UNITED NATIONS CONFERENCE ON TRADE AND DEVELOPMENT

OCTOBER 2017

UNCTAD Research Paper No. 10
UNCTAD/SER.RP/2017/10/Rev.1

Digital tools for foresight

Martin Hilbert
Department of Communication, University of California, Davis
hilbert@ucdavis.edu

Abstract

The report illustrates how digital tools, namely big data and computer simulations, can be used to enhance foresight exercises. By using concrete case studies it also shows the feasibility of using these computational methods in developing countries and for the implementation of the 2030 Agenda for Sustainable Development. The report concludes by discussing the considerations that need to be taken into account when doing foresight exercises of a digital future in the context of the implementation of the sustainable development Goals.

Key words: Foresight, sustainable development, sustainable development goals, science, technology and innovation, big data, computer simulations

Contents

Acknowledgements

This research paper was prepared by a team comprising Martin Hilbert, Department of Communication; University of California, Davis (with the assistance of Suzanne Cheng and Shengling (Wilburn) Wang, University of California, Davis), Bob Bell, Katalin Bokor, and Claudia Contreras (UNCTAD), under the guidance of Dong Wu (Chief, Science, Technology and Innovation Policy Section, UNCTAD). Comments from Professor Ozcan Saritas on an earlier version of this paper are gratefully acknowledged.

Introduction: Foresight for the Sustainable Development Goals in a digital age

The Sustainable Development Goals (SDGs) are 17 sets of goals with 169 targets to be achieved by 2030. As political aspirations, they aim at a future that is yet uncertain. In other words, this future is not deterministically given to us, but has yet to be socially constructed. Therefore, concrete policy tools are needed to deal with and shape this uncertain future. This report argues that the natural tool for doing so is to use foresight methods. Foresight involves bringing together key agents of change and sources of knowledge, to develop strategic visions and anticipatory intelligence (Miles et al., 2008: 11), to shape the future, and it is often executed through participatory consultations. Foresight not only provides approaches and methods about scanning issues that can be measured today (i.e. trends), but also informs policy-makers about future issues or wild cards that are not yet considered in policy design but must be tackled today if we are to develop our societies in a sustainable way. Foresight makes particularly sense in addressing sustainable development challenges. It is vital for any forward planning or policy activity to be able to meet future challenges proactively.

One of the main challenges of current foresight exercises is that we live in a period of technological change. Technological change is exponentially fast-paced, all-embracing, and global in nature. An important driver of this current phase of technological change is digitalization. The digitalization of the entire stockpile of technologically mediated information has taken less than 30 years, as less than one percent was in digital format in the mid-1980s (the rest in analogue format on paper, tape, vinyl, etc.), and more than 99 % in digital format today (extrapolated from (Hilbert & López, 2011)). The exponentially fast innovation cycles of digital technology create high uncertainty; its general purpose applicability embraces all sectors, and its inherent defiance of national borders intermingles the most diverse aspects of a heterogeneous world.

All of this leaves us with the aspiration to shape an extremely uncertain and fast paced future. This paper presents examples of how digital tools can help to implement future foresight (as means to help achieving the SDGs). Digital big data footprint can be used to detect empirical realities, artificial intelligence to extract insights from data, and computer simulations to explore future scenarios that are different from today's reality; to explore the world as we would like it to be; the world where the SDGs are the reality. The paper shows that these tools can be extremely useful to foster foresight studies in developing countries. They are cost-effective, scalable, and sensitive to local contexts. The paper explores various practical applications that use these computational tools in order to implement Sustainable Development Goals.

After this Introduction, Chapter 1 focuses on digital tools, namely big data and computer simulations, which can be used to enhance foresight exercises. It also reviews methods from the field of what is nowadays known as 'computational (social) science'. Chapter 2 presents concrete case studies that show the feasibility of using these computational methods in developing countries and for the implementation of the SDGs. Chapter 3 discusses the considerations that need to be taken into account when doing foresight exercises of a digital future. It places foresight efforts into the challenging context of our global reality, especially focusing on the particularities for implementing the 2030 Agenda for Sustainable Development in developing countries, and addresses a caveat about the role, scope, and limits of foresight exercises.

1. Digital tools for Foresight

Information and communication technologies (ICT) are currently transforming the way information is obtained, knowledge is created, and insights are derived. This revolutionizes the way research is being done. The applications reach from inductive empirical inquiry (e.g. 'big data'), to deductive theoretical scholarship (e.g. 'computer simulations'). The resulting opportunities are very broadly labelled as "computational social science" (Cioffi-Revilla, 2014; Conte et al., 2012; Lazer et al., 2009).

Computational science tools have given new wind to a variety of mixed method approaches used in foresight exercises (e.g. Alexander & Maiden, 2005; Haegeman, Marinelli, Scapolo, Ricci, & Sokolov, 2013; Hansen, Rasmussen, & Jacobsen, 2016; Kwakkel & Pruyt, 2013). Empirical evidence allows to ground foresight into reality, while theoretical models allow to explore futures that are desired, but for which no empirical data exists. With this paper we argue that it is useful to make an explicit effort to adopt digital methods as an integral part of foresight methods. This does not argue to completely replace any human involvement in foresight exercises with algorithms, but to complement human planning with data-driven decision making and the formal development of scenarios.

Given that ever more of human conduct is taking place in digital networks, and given that digital conduct inevitably leaves a digital footprint, the social sciences currently have access to an unprecedented amount of data on the most diverse aspects of our reality and its dynamics (Lazer et al., 2009). The catch-phrase here became 'big data' (Manyika et al., 2011; Mayer-Schönberger and Cukier, 2013). Big data considers the accumulation and analysis of greatly increased information resources, beyond the storage and analytical capacity of earlier hardware and software resources. Big data is usually used to describe a massive volume of both structure and unstructured data that it is difficult to process with traditional database and software techniques due to its magnitude (UN Global Pulse, 2012: 13). It is made possible by increases in both data storage capacity and the range of available data sources.[1] Its impact on the social sciences has been compared with the impact of the invention of the telescope for astronomy and the invention of the microscope for biology (providing an unprecedented level of detail about the system of interest). Confronted with such increase in the level of perceivable granularity in social dynamics, analysts and policy makers have an inevitable obligation to make use of it to inform future planning.

This digital deluge is nowadays not only used to predict stock market behaviour, commercial consumption patterns, and traffic jams, but also to predict epidemics, medical necessities, environmental disasters, and poverty levels (Hilbert, 2016a; Manyika et al., 2011; Mayer-Schönberger and Cukier, 2013). For example, using simple metadata digital footprints like call duration and call frequency, it has been shown that one can predict socioeconomic, demographic, and other behavioural trades with 80-85% accuracy (J. E. Blumenstock & Eagle, 2012; J. E. Blumenstock, Gillick, & Eagle, 2010; Frias-Martinez & Frias-Martinez, 2014; Frias-Martinez, Frias-Martinez, & Oliver, 2010; Raento, Oulasvirta, & Eagle, 2009). This basically allows to reverse-engineer much of a traditional household survey or census (Frias-Martinez and Virseda, 2013), which is information often lacking in least developed countries. In a combination of machine learning and publicly available data, including high-resolution daytime and night-time satellite imagery, researchers recently were able to explain up to 75% of the variation in local-level economic outcomes in Nigeria, Tanzania, Uganda, Malawi, and Rwanda (Jean et al., 2016). In other words, they were able to predict the distribution of poverty better than any other existing approach.

It seems not only natural to take advantage of these computational science methods for foresight purposes, but also imperative. This is especially relevant in the light of the lack of existing data to inform specific development goals in developing countries.

For practical purposes, it is important to recognize that the big data paradigm includes a heavy focus on analyses, often involving machine learning, but also other forms of artificial intelligence. For one, the catch-phrase 'big data' refers to new sources of data. The digital footprint created with each digital communication and transaction can replace traditional data sources (like surveys) with proxy indicators that correlate with the variable of interest. The benefit is the low cost and real-time availability of the digital proxy indicator. The epitome is Google's illustrious use of 50 million most-common search terms as a proxy for the spread of the

[1] Report of the Secretary General on Information and Communication technologies for inclusive social and economic development, p.8

seasonal flu (Ginsberg et al., 2009). Secondly the notion of big data goes beyond data itself and focuses on methods of data analytics to inform intelligent decisions. Independent from the specific giga-, tera-, peta-, or exabyte scale, the big data paradigm argues to systematically place analytic treatment of data at the forefront of intelligent decision-making. The process can be seen as the natural next step in the evolution from the "Information Age" and "Information Societies" to "Knowledge Societies" (for more see Hilbert, 2013). Building on the digital infrastructure that led to vast increases in information during recent decades, the big data paradigm focuses on converting this digital information into knowledge that informs intelligent decisions. Continuing with the previous example, Google processed an impressive 450 million different mathematical models in order to test for correlations between online search terms and flu outbreaks reported by official data. Eventually, 45 search terms were identified that outperformed traditional models of flu outbreak with real-time predictions (Ginsberg et al., 2009).

While the opportunities of big data are enormous, especially for developing countries in which traditional statistics are scarce (Hilbert, 2016; Letouzé, 2012; WEF, 2012), the very same Google flu trend case also exemplifies the ultimate limitation of (big) data analytics: the fact that all data is from the past. The application of the same Google flu trend algorithm became increasingly out of sync with actual flu epidemics over time (Lazer et al., 2014). The reason is straightforward: reality (in this case search behaviour) is changing over time, which requires a different input and therefore changes the results if the same algorithm is used. The big data algorithm was optimized for a reality that does not exist anymore, and applying it to the new reality will be deceptive.

What the best big data analytics can do is to update estimations in 'real-time': so-called 'nowcasting'. Both the terms are actually misnomers, since the very act of recording converts it into 'past-time' data. Data is always from the past, per definition. Therefore, it can only detect patterns that have occurred in the past. When the past and the future follow the same stationary logic, data analytics is extremely useful in predicting future patterns. In other words, if the generative mechanism that produced the past data continues to be valid in the future, then the past data can tell us something about the future. However, if significant changes occur in the system's dynamic, empirical statistics based on past data are at best limited to predicting the future, if not deceiving (Hilbert, 2014a, 2015b). In a technical sense, the key concept here is the 'stationarity' of the data source. This is usually in the fine-print of most empirical data analysis (big or small data) and refers basically to this idea: if the generative mechanism is 'stationary', then past data is useful to predict the future. If not, mere descriptive statistics cannot help us, because they were produced from a different reality. Most statistical forecasting methods are based on the assumption that the time series can be rendered approximately stationary through the use of mathematical transformations.

This argument is crucial for our purposes since the 2030 Agenda for Sustainable Development[2] is an aspiration that explicitly aims at creating a world that is different from today's reality. A world with no poverty (SDG 1), zero hunger (SDG 2), clean water and sanitation (SDG 6) and affordable and clean energy (SDG 7) is decisively different from the world in which we lived in 2015 when the goals were adopted. It is the definition of development goals to change the 'stationarity of the system'. It is the ambition to replace the dynamic that produces poverty with a new one that does not. The data footprint left behind by the old mechanism can only help us to understand the old mechanism. This is a useful first step, but it cannot automatically tell us what will happen if we change it. We have never lived in a world with zero hunger and affordable and clean energy. Who could tell us what it would be like and how to get there? Data from the past alone cannot.

In order to predict a future that has never been, theory-driven models are necessary (Hilbert, 2015b). These allow variables to be adjusted with values that have never existed in statistically observable reality. Digital tools also act as a game changer in this challenge. Computational simulations allow to setting up theory-driven models that greatly expand the scope and level of sophistication of traditional 'paper-and-pen' models. While traditional models are only able to handle a very limited number of variables (at most a dozen or so), today's computational power allows creating mathematically formalized models with thousands and even millions of dynamic variables. Such computer simulations of artificial societies have no conceptual

[2] See A/RES/70/1.

limitations on the achievable level of detail and accuracy (in practice they do face computational limits). Most recent simulations are based on individual agents ('agent-based models') (Epstein, 2005; Epstein and Axtell, 1996; Helbing and Balietti, 2011; Miller and Page, 2007; Schelling, 2006; Wilensky and Rand, 2015), which allow to simulate the interactions of individuals (not merely aggregated factors or variables). In essence, they are similar in nature to popular computer strategy games, such as SimCity. The researcher sets up a reality (often based on existing modules of behaviour and structure) and then lets the model run in order to observe how the myriad of interactions often produce surprising and counterintuitive outcomes. Agent-based computer simulations are essentially equation based, just that differential equations are both programmed into the flexible behaviour of the agents and result from their interaction. Compared to traditional equation-based model, "the agent-based model is more versatile and therefore preferable" (Sukumar and Nutaro, 2012: 6).

These models can and should be calibrated by empirical data. "A good complex systems model both begins and ends with data: Low-level data [from the micro-scale] is used to formulate the assumptions about the building blocks of the model, and both high [macro-scale] and low-level data is also used to test whether the resulting emergent phenomena properly correspond to those observed in the real world" (Farmer, 2012). However, they go beyond mere data, and require theoretical assumptions of behaviour and interactions. This also allows transferring insights from one case to the other, often with surprising results. Theoretical models allow us to transfer local and small-scale evidence into different contexts. A developed Africa will not simply be a statistically extrapolated version of Europe's development trajectory. The statistical model of a fitted curve to some data points cannot give us an in-depth insight here. However, we can take bits of empirical evidence from different instances and piece together an assumed theoretical model of Africa's development trajectory. This would show us how Africa would develop 'in theory'. Both theory and empirical data are needed. The combination of both empirical (big) data analysis of the digital footprint and theoretical computer simulation models of possible scenarios provide a wide variety of opportunities to inform foresight exercises.

2. Examples of how Big data and computer simulations can inform foresight studies

Several UN agencies (including the UN Statistical Commission and UN Global Pulse) have recognized that the potential of big data is an important dimension of ICTs' contribution to the monitoring and assessment of SDGs. In the following, we will review several case studies that provide evidence on how big data methods can be used to produce intelligence to inform foresight studies. The examples presented as well as the Sustainable Development Goals that they are focused on were selected with a view to illustrate our argument and does not mean that these Goals are more relevant than others. For additional cases showing the use of big data for development see for instance Hilbert (2016a) and Letouzé (2012).

SDG2: Sustainable food production in Colombia

Focus area: SDGs 2: End hunger, achieve food security and improved nutrition, and promote sustainable agriculture, Target 4: By 2030, ensure sustainable food production systems and implement resilient agricultural practices that increase productivity and production, that help maintain ecosystems, that strengthen capacity for adaptation to climate change, extreme weather, drought, flooding and other disasters and that progressively improve land and soil quality.

The International Center for Tropical Agriculture is an organization that strengthens agricultural technologies, innovations, and new knowledge that helps small farm owners improve their crop yields, incomes, and usage of natural resources. Scientists collaborated with the Colombian Government, Agriculture and Food Security, and Colombia's National Federation of Rice Growers to collect a big volume of weather and crops data in last decade in Colombia. The initiative predicted upcoming climate changes in Córdoba, a major rice-growing area in Colombia. The results are highly localized. In the town of Saldaña, for example, the analysis showed that rice yields were limited mainly because of solar radiation during the grain-ripening stage. Meanwhile, in the town of Espinal, the team found that it suffered from sensitivity to warm nights. Solutions do not have to be costly, as such information can help farmers to avoid losses simply by sowing crops in right period of time. The climate change foresight helped 170 farmers in Córdoba avoid direct economic losses of an estimated $ 3.6 million and potentially improve the productivity of rice by 1 to 3 tons per hectare. To achieve this, different data sources were analysed in a complementary fashion to provide a more complete profile of climate change. So-called 'data fusion' is a typical big data technique. Additionally, analytical algorithms were adopted and modified from other disciplines, such as biology and neuroscience, and were used to run statistical models and compare with weather records. With support from national and international organizations such as the World Bank and the Fund for Irrigated Rice in Latin America, the initiative has started to approach rice growers associations in other countries, including Nicaragua, Peru, Argentina and Uruguay.[3]

SDG 8: Global networks against forced labour

Impact area: SDGs 8: Decent work and economic growth, Target 7: Take immediate and effective measures to eradicate forced labour, end modern slavery and human trafficking and secure the prohibition and elimination of the worst forms of child labour, including recruitment and use of child soldiers, and by 2025 end child labour in all its forms.

Made in a Free World is a non-profit organization that tracks and identifies raw materials and goods associated with slave labour, using data uploaded by the companies. Using digital footprints, the organization's marketing analysis program FRDM (Forced Labor Risk Determination & Mitigation) helps businesses and organizations to look into their supply chains to identify slave and child labour, or products that come from conflict zones. The database contains over 54,000 products, services, and commodities from the United Nations Standard Products and Services Code. Business products and services have been broken down into material and parts, and traced to their origin. Country risk is generated using internationally recognized and peer-reviewed reports. The organization also partnered with Ariba Network, the largest business-to-business trading platform that connects more than two million companies with their global supply chains. The tool is sponsored by the US federal government and used by federal contractors and independent companies. When a company uploads purchased data, the software compares the data to its Global Slavery database. An accompanying action plan aims at strengthening vendor agreements, improving supplier code of conduct, and providing suggestions to companies on effectively getting rid of slavery in their supply chains.[4]

SDG 3: Big Data for ending Tuberculosis in Brazil

Impact area: SDG 3: Ensure healthy lives and promote well-being for all at all ages. Target 3: By 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne diseases and other communicable diseases.

The Brazilian Cadastro Único (Unified Registry in English) is a database that aims at identifying all existing low-income families for inclusion in social welfare and income redistribution programs. Through a computerized system, the federal government of Brazil consolidates the data collected in the database to formulate and implement specific policies that contribute to the reduction of social vulnerabilities that these families are exposed to. The system is a gateway to 20 public policies. To register, family members must have a monthly income of up to half the minimum wage per person (less than R$ 394, or about USD 125) per person. As the social program expanded, extensive efforts were made to ensure all potential beneficiaries were included and to verify their information –with over 1200 crews visiting homes across the country and

[3] Sources: (The Guardian, 2014; CCAFS, 2015).
[4] Sources: (The Australian Business Review, 2015; How Big Data is driving sustainability, 2016; Business Wire, 2015)

even trekking deep into the Amazon by canoe. As of 2016, the Cadastro Único was tracking the information of over 103 million people.[5] The system also captures tuberculosis cases and collects information that distinguishes people that have received cash transfers from those who have not. The information contained in this massive database has helped to forecast the prevalence of tuberculosis in the country until 2025. It is estimated that if the related social programs were to stop today, the Tuberculosis prevalence in 2025 would be approximately 6% higher than if the programs were to continue. By contrast, an expansion of the cash transfers to all tuberculosis cases would contribute to a faster decrease in the prevalence of the disease, with a 33% lower prevalence in 2025. Policy makers and administrators are better informed by these foresights and can make more effective policies related to tuberculosis, regulate respective health and pharmaceutical markets more efficiently, and use these insights to explore the side-effects of tuberculosis on other aspects of the Brazilian economy and society, including family policies, labour rights, and living conditions. [6]

SDG 4: Universal education in India

Impact area: SDG 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all; Target 1: By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes.

Akshara Foundation is a Non-Government Organization in India that aims at narrowing the gaps in the universalization of pre-school and primary education in India. The organization has been collecting primary data from 40,000 schools and over one million students since 2006. The organization itself did not have sufficient capacities to fully make use of these data. Motivated by the abundance of primary data available, Hewlett Packard approached Akshara Foundation with a fresh analytical perspective to turn the data into useful information identifying the links between education resources, facilities and Indian children's education experiences. Data scientist from HP's Analytics and Data Management Practice helped to clean and organize previously messy and incomplete dataset of Karnataka, India. Together they created a custom dashboard that provides insights of the ratio of students and teachers and the adequate number of books per student. Machine learning algorithms were used to look for hidden patterns in the data. One surprising finding is that the introduction of separate bathrooms reduces drop-out rates in school. Such insights allow administrators and policy makers to set any foresight and policy planning exercise on a more comprehensive outlook of reality. Sources: (Gutierrez, 2015; *Datanami*, 2015).

SDG 6: Water management in Taiwan province of China

Impact area: Goal 6: Ensure availability and sustainable management of water and sanitation for all; Target 5: By 2030, implement integrated water resources management at all levels, including through transboundary cooperation as appropriate.

Despite the Province's high annual rainfall, only about 20% of the rain can be used as a water resource due to its mountainous terrain, population strips and other factors. The Taiwan Water Corporation (TWC) is a state-owned water company that provides 8.6 million cubic meters of high-quality potable water per day. TaKaDu is an Israeli organization that optimizes operational efficiency and improves water management holistically by using big data analytics, advanced algorithms, and Internet of Things (IoT). In 2016, both organizations partnered to find a solution for the effective water management system, enabling the buildout of a solid water management network and save water resources. TWC used a cloud-based smart system by TaKaDu to monitor parts of the 3000 km pipe networks across Taiwan. The TaKaDu water management system applies big data techniques such as data fusion, real-time detection, and machine learning to achieve reduced leakages and effective water management. The results provided early warning of incidences, such as assets, leaks, bursts, water pressure issues, water quality, and operational faults.[7]

[5] https://www.riopluscentre.org/blogs/big-data-predicting-future-tuberculosis-rates-in-brazils-uncertain-times
[6] Sources: (Rasella et al., 2016; Rio + Centre, 2016).
[7] Sources: (www.ChinaPost.com.tw, 2016; Israël Science Info, 2016).

SDG 7: International energy efficiency

Impact area: SDG 7: Ensure access to affordable, reliable, sustainable and modern energy for all; Target 3: By 2030, double the global rate of improvement in energy efficiency.

Opower is an energy efficiency cloud service company headquartered in Virginia, USA, dedicated to reducing energy consumption for their customers. The company works with more than 100 utility services (such as PG&E, and National Grid) and serves more than 60 million homes in several countries, including Australia, Canada, France, Japan, New Zealand, UK, USA. The software plugs into the IT system of the clients and utility company, unifying customer information in a single data warehouse. The software analyses consumer's utility energy data, collected data from customer interactions, and along with third-party data such as customers' income, geographic location, and weather reports, to give consumers personalized suggestions to save energy and money. The system conducts data analysis of energy usage information and overlying weather patterns to estimate energy used for adjusting room temperature versus other needs. Opower's engagement platform and solution suite provide data insight in supporting utilities in customer outreach activities, such as improve the customer experience, and involving their customers in energy and cost efficiency goals programs at the same time. Since its launching in 2007, the company estimates to have saved over 11 terawatt-hours (TWh) of energy, enough energy to power over 1 million U.S. homes for a full year. The average customer receiving the Opower platform has cut energy usage by more than 2.5% per year and prevented emissions of over 2.23 million tons of CO2 each year.[8]

SDG 11: Waste Management from South Korea to Colombia

Impact area: SDG 11: Make cities and human settlements inclusive, safe, resilient and sustainable; Target 6: By 2030, reduce the adverse per capita environmental impact of cities, including by paying special attention to air quality and municipal and other waste management.

Waste collection is usually performed using static routes and static schedules. As a result, some trash sites are overflowing with waste causing unnecessary clean-up costs, while others are nearly empty making waste collection unnecessary. This type of inefficiency arises from a pure informational problem that wastes both time and money, and is harmful to the environment. The E-CUBE LAB from Seoul, South Korea focuses on the development of smart trash bins and city waste management solutions. The organization has developed advanced sensors to cut down operational costs and increase the effectiveness of waste management. The E-CUBE LAB had rich experiences of waste management in South Korea and other developed countries. In 2015, E-CUBE LAB introduced their advanced sensors (Clean CAP) to two cities in Colombia, Ibague and Santa Marta. These sensors can monitor how much waste has been accumulated and send the information to the waste control centre, so they these two cities are equipped with the ability to access data and monitor waste in real time. Such insights are not only essential in order to foresight city planning in overcrowded metropolis areas, but the real time analysis has already proven to optimize waste and reduce costs by avoiding unnecessary pick-ups and dynamically calculating waste collection routes in a flexible manner. According to E-CUBE LAB, this may reduce operational costs of waste management up to 80%.[9]

SDG 15: Big data for detecting illegal timber trade in Brazil

Impact area: SDG 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss; Target 1: By 2020, ensure the conservation, restoration and sustainable use of

[8] Sources: (Fischer and Shilts, 2014; Business Wire, 2016).
[9] Sources: (Ecube Labs, 2015; The Guardian, 2016).

terrestrial and inland freshwater ecosystems and their services, in particular forests, wetlands, mountains and drylands, in line with obligations under international agreements.

Bolsa Verde do Rio de Janeiro, known as BVRio, is a non-profit organization in Rio de Janeiro, Brazil. The organization collaborates with Rio De Janeiro State Government, the Municipality of Rio de Janeiro, and the Brazilian Amazonian states to prevent illegal timber trades in Brazil. BVRio adopts a computational system that includes analytical tools for due diligence and risk assessment system. The system consists of four components: document analysis, remote sensing (i.e.: satellite imagery), field audits, and external databases. BVRio's analysis found that between 2007 and 2015, more than 40 percent of the forest management operations in the Brazilian states of Pará and Mato Grosso were likely to violate severe breaches of the law. Those areas are responsible for more than 70 percent of Brazil's timber production. Different from conventional databases, which usually only include logging permits and sawmill operating licenses to detect illegality of timber, the system collects more diverse data sources from legal records of forest owners, loggers, and forest engineers. Algorithms that allow to scan data in real time and to find hidden correlations in datasets make modelling and predictive analysis more accurate and efficient. As a result, administrators and policy makers can identify inconsistencies and illegal behaviours and respond to them quickly. At the same time, buyers can inform themselves about their timber sources and protect themselves from purchasing illegal ones.[10]

SDG 17: Public-private partnerships in Brazil

Impact area: SDG 17: Strengthen the means of implementation and revitalize the global partnership for sustainable development; Target 17: Encourage and promote effective public, public-private and civil society partnerships, building on the experience and resourcing strategies of partnerships.

In a public-private sector collaboration, the Rio de Janeiro city government and IBM have created an integrated operations system across many sectors to achieve real-time city and smart urbanism. More than thirty public agencies contributed data streams, including traffic and public transport data, municipal and utility services, emergency services, weather information, and call details records. Citizens interact on social media platforms, contributing real time data, such as tweets about traffic jams. As a result, decentralized data and information go into one centralized command centre, consisting of both private sectors and public sectors. With data from different sources, policy makers and administrators have a better understanding of city dynamics and find hidden correlations. Combined with advanced algorithms, predictive models are created to develop scenarios for natural disasters and crimes in a real time. Due to this big data project, it is reported that the city of Rio de Janeiro emergency response is 30 percent faster than before. Rio Operations Center is one of the seven pilot projects (other pilots include 3 US, 1 Polish, 1 Vietnamese, and 1 Chinese cities) that are part of IBM's Smarter Cities Challenge. Since its launch in 2010, IBM has advised (and in some cases awarded grants) over 100 cities in how they can address some of their more challenging issues while also doing this in a way that is more sustainable through the usage of these smarter technologies.[11]

Future scenarios: computer simulations

The first existing social science computer simulation models were inspired by the work of engineers, who used computers to simulate buildings and bridges. Similar to our policy problem of building a future that never existed before in reality, engineers are frequently confronted with the task of building something very specific and local for which no adequate data exists. The main building blocks of the first social science models were the same differential equations and macro-level factors that traditionally make up paper-and-pen-based models, such as common in so-called systems theory, including variables like capital, labour, and

[10] Sources: (Mongabay, 2016; Responsible timber I BVRio Timber Exchange, 2016)
[11] Sources: (Singer, 2012; Urenio, 2015; Vineetadurani, 2011).

income (Gilbert and Troitzsch, 2005). The increase in computational power over the subsequent decades allowed to fine-grain those approaches, which led to an increasing transition of simulations "from factors to actors" (Macy and Willer, 2002). Instead of modelling factors of abstract aggregate variables, these models simulate individual actors. This allows for the inclusion of much diversity in models, as each actor is a multidimensional variable itself. For example, instead of working with an average income per capita in the model, one can work with a heterogeneous distribution of income per person, which is more realistic, as in reality, few people have exactly the same aggregate income. It can be deceiving to neglect such diversity, as diversity leads to complexity, and often to unexpected emergent phenomena (Page, 2010). These more recent models are often referred to as agent-based models (ABMs), also multi-agent (Epstein, 2005; Epstein and Axtell, 1996; Helbing and Balietti, 2011; Miller and Page, 2007; Schelling, 2006; Wilensky and Rand, 2015). Their signature is the possibility to study the emergence of non-linear macro patterns that arise out of a multiplicity of dynamical micro-interactions (Schelling, 2006).

Agent-based computer simulation models have a number of characteristics that make them not only very useful, but also cost-effective and scalable for developing countries. Some of the features include (Hilbert, 2015a):

1. They enable hypothetical what-if questions, by exploring theoretical scenarios without empirical precedent (much in line with the discussion above).

2. They build on modules of algorithms, which allows for scalable adjustment to local conditions and contexts in a cost-effective manner by building on algorithmic modules developed for other purposes. For example, it might be costly to program a module that simulates the acceleration and deceleration patterns of different kinds of vehicles, but once the library of these lines of code exists, they can easily be 'copy-pasted' to simulate all kinds of local traffic conditions in specific cities around the world. In the same sense, it might be costly to develop hurricane simulations for coastal areas in the US, but once they exist, parts of the generic parts of the code can be adopted and fine-tuned to the specific context in developing regions. This modular approach provides a cost-effective way to break away from 'one-size-fits-all' models and to invest scarce resources into the adoption to local contexts.

3. They are a constant and humbling reminder of the probabilistic nature of all models. Computer simulations are usually not run once, but many different times, with varying initial conditions. This considers the effects of path-dependencies which are omnipresent in dynamic systems. Each run will result in a more or less different result, which provides a probability distribution (called the invariant distribution). Any conclusion from these kinds of numerical simulations can only state that "with x % probability, y will happen" (following the probability distribution of obtained outcomes). Never a certain result. This helps to avoid falling for the false believe that a policy provides for a certain outcome.

4. Computer simulations visualize the systematic reasoning with which a model carries out the interactive dynamic that arises from model assumptions and therefore is often more intuitive than a set of equations. This helps in communicating sophisticated and complex findings to policy-makers and the general audience by interacting with the model similar to interactions with a video game. This last aspect is especially important for foresight exercises, as it allows for the intuitive involvement of stakeholders, which would usually not be convinced or could not follow formal mathematical arguments or models. Picking up on this idea, several initiatives have recently developed so-called 'playable models' or 'serious games', which basically are game-like computer simulations open for public exploration.

In the following, we will review several case studies that provide evidence on how these kinds of techniques can be used to produce intelligence that could potentially inform foresight studies.

SDG 3: Simulating malaria mosquitos in Niger

Impact area: SDG 3: Ensure healthy lives and promote well-being for all at all ages; Target 3: By 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne diseases and other communicable diseases.

The Hydrology, Entomology, and Malaria Transmission Simulator (HYDREMATS) is a coupled hydrology and entomology model that has been developed to simulate the links between environmental variables and malaria transmission with high temporal and spatial resolution. In a simple form, it is a spatial simulation model of malaria transmission and rainfall. The hydrology model uses satellite-derived vegetation and soil parameters. Excess precipitation is routed over this rough, infiltrating surface to simulate the formation and persistence of pools that allow malaria mosquito breeding. An agent-based entomology model simulating known behaviour relevant to malaria transmission constitutes the link between hydrology and malaria. In this way, environmental determinants of malaria transmission are simulated in detail. It simulates individual mosquitoes as "agents" while building spatial relationships between vectors and hosts. More specifically, the model is used to simulate the life cycle and movement of individual malaria vector mosquitoes in Banizoumbou (a Niger Sahel village), taking interactions with the physical environment and humans into account. With the help of the simulation, scientists found that one of the pools in the village was highly related to numbers of mosquitoes and potential transmission of malaria, so they suggested targeting the pool by using larvicide. It turned out that a 49% decrease in village adult mosquito abundance was observed with the application of the targeted growth-inhibiting larvicide. [12]

SDG 10: Reduce inequality within and among nations

The question of income inequality has been studied with agent-based computer simulations since their pioneering large-scale application. One of the foundational simulations (which is used in many university courses) is the so-called Sugarscape, an agent-based social simulation developed by Epstein and Axtell in their book Growing Artificial Societies in the mid-1990s. In the model citizens with different skills and abilities wander randomly around a landscape and feed off sugar (resources) that they find in different locations. The model starts very simple and one of the first surprising facts about is it how easy it is to grow the same Pareto-law distribution of sugar that is usually found for income distribution. While this extremely unequal distribution of resources grows almost naturally in this very simplistic society, the computer simulation then allows to test how it can be changed (something not as easily testable in real societies). Computer simulations allow to defy 'natural orders' and look for creative ways to create realities that are different. Two quick findings are the importance of a well-functioning lending and borrowing. Mechanism among agents, and the sensitivity of the accompanying interest rate. Another one is the role of inheritance to foster social inequality further (the model simulation is publicly available through the user-friendly software NetLogo). Typical for computer simulations, over the years the Sugarscape model has been expanded to test a myriad of hypotheses, including simulations with over a million agents at interactive rates (the so-called 'Sugarscape on steroids' model).

Other models started to explore the nature of wealth inequality instead of income. One model by Harckbart was implemented using open software R Studio and includes variables like net worth (current and maximum), annual return drawn, capital gains tax rate, a total of capital gains tax collected in a year by the government, and a 5% return rate per year. Several hypotheses were tested around the effects of these variables on wealth inequality for different societal conditions. It was found that long-run inequality is greatly impacted by the volatility of returns and capital gains tax rates. The model shows stability even in simulations with high unequal stating wealth distributions, where inequality appears to move towards levels determined by the economy's volatility and tax rate. The process of reduced inequality can be observed after a couple of decades or centuries in all simulations, the pace determined by initial inequality and the long run inequality

[12] Sources: (Bomblies, 2014; Bomblies et al., 2008; Yamana and Eltahir, 2010).

level expected given taxes and volatility. An interesting observation from the results showed that even moderate capital gains tax rates with equal redistribution of wealth are enough to prevent inequality. As with all models, the author notes that the results might depend on the inclusion and assumption of different additional variables might affect the results (such as equality of education, skills, and opportunity, which were assumed in the model).[13]

SDG 6: Participatory water management in Tunisia

Impact area: SDG 6: Ensure availability and sustainable management of water and sanitation for all; Target 5: By 2030, implement integrated water resources management at all levels, including through transboundary cooperation as appropriate.

Tunisia has a long history of coping with water scarcity. Until the 2000s, water resources in Tunisia were managed solely by government bodies without any consultation with water users. At the local level, Regional Agricultural Development Committees (CRDAs), which represented the government bodies, managed water resources. The Tunisian Government then decided to involve farmers in the local management of water resources by creating Agricultural Development Associations (GDAs). Water resource management by CRDAs must now take into account water demand from farmers belonging to GDAs and, in return, farmers must acknowledge water scarcity and accept the need for resource-sharing. How can CRDAs and GDAs work towards the goal of cooperative water management? A traditional continuous foresight exercise could have been the answer to this question. However, instead, at the request by officials at Kairouan CRDA (Regional Agricultural Development Committees), a group of researchers developed a simulation game, AquaFej, aimed at involving and raising farmers' awareness of water management and collaborations between farmers and officials in central Tunisia. The simulation game model is based on a hydrological model of the watershed, which allows to calculate dam and groundwater recharge according to rainfall and artificial reload of groundwater and water flow in rivers. In the simulation game, the dam manager and farmers are the main agents in a scenario focusing on the management of water allocation from the dam. Based on winter rainfall, the hydrological model indicates the water available for storage and groundwater recharge. Once the decision on water allocation has been taken, the hydrological model determines the amount of water available for each type of water access. From the cropping plans defined (winter and spring-summer crops), the technical-economic model evaluates the impact of irrigation on groundwater levels and calculates farm incomes according to yields and market prices. This game allows decisions by stakeholders to be taken into account irrespective of whether they are made collectively or as a result of negotiation. AquaFej built a bridge between farmers and policy makers and stimulated them to achieve collaborative water management. On one side, AquaFej makes farmers more aware of water resource-sharing, scarcity, and the impact of overexploitation. On the other side, the game informs CRDA officials regarding water demands from farmers and other sectors. For example, by running the simulation, farmers can notice the consequences of individual cropping choices with considering economic and environmental factors.[14]

SDG 17: Playable models for stable budgets in the United States

Impact area: SDG 17: Strengthen the means of implementation and revitalize the global partnership for sustainable development; Target 13: Enhance global macroeconomic stability, including through policy coordination and policy coherence.

Budget Hero is a computer simulation game to invite players to enact different financial policies to make an ideal budget decision in a playable platform. The game is built on the Congressional Budget Office model, which includes advantages and disadvantages for each policy. The game was developed in May 2008, relaunched in September 2012 by David Rejeski, and retired in August 2014 by the Woodrow Wilson International Center for Scholars, which is a United States Presidential Memorial funded by the Congress, in

[13] Sources: (D'Souza et al., 2007; Epstein and Axtell, 1996; Harckbart, 2015; Wilensky and Rand, 2015).

[14] Sources: (Le Bars et al., 2014).

collaboration with the American Public Media. The game is described as a 'playable model', as it emphasizes to deal with a complex system through scaling general public engagement by exploring effects of causal factors affect other parts and the whole. A playable model is a form of computer simulation model that allows researchers to better understand the different processing of complex questions. In other words, the game creates an interactive and exploratory platform to help people test their hypotheses and come up with their own solutions. Such engagement aims to involve traditional media (journalists as gatekeepers) to the audience who are able to understand headlines by themselves. In Budget Hero players will attempt to balance the budget by choosing from more than 70 policies, provided with the advantages and disadvantages for each policy. The gamers have opportunities to avoid worsening the debt or crash the economy in the game by better understanding the relationships between financial policies and budgetary issues. Budget Hero has a record of 2 million players, and all of the data in the model and decisions were recorded. The advantages of a playable model include having the potential to be scaled to a large population, allowing users to learn through their experiences, increasing public engagement around public issues, the ability to experiment with different options and to see the consequences of their choices, are all helpful in improving understanding and developing theories of the situations. Thus, playable models like the budget simulation game can be used as both a strategic (predictive planning), as well as a participatory foresight tools.[15]

SDG 1: Ending poverty with decision-oriented insights

Impact area: SDG 1: End poverty in all its forms everywhere; Target 1: By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than $1.25 a day.

The non-profit organization Live 58 has launched a simulation game in 2012, known as Survive 125. The purpose of the simulation game is to make gamers experience living situations, decision-makings, and trade-off when they only have $ 1.25 per day. In particular, this simulation game provides decision-making scenarios that truly happened in the real world. The gamers are required to think critically to determine the survival strategies of his or her family. For example, the player is a 26-year-old woman living in India with four children, works as a bricklayer and makes $1.25 a day. In order to survive, the player is forced to make difficult decisions constantly, including whether to send her teenage daughter to work at a factory with a presence of sex trafficking in order to make more money and whether to stay home with a sick child, which would require missing one day's worth of wages. The goal of the game is to survive as long as possible with $1.25 per day. The experimental experience helps policymakers and the general public better understand extreme poverty and people who suffer from extreme poverty, instead from watching video material and the media. The computer simulation makes the complex ideas more self-explanatory while providing opportunities to explore "what if… then" scenarios. It invites people to imagine what would happen with more vivid experience. It has also shown to motivate players to think about and try to come up with potential solutions, which would feed into the action oriented aspect of foresight exercises.[16]

SDG 9: Computational simulations to foster innovation in policy design

Impact area: SDG 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation.

Using SpatialOS computational platform, developers can create large-scale complex models, real-time simulations of large and detailed virtual worlds (such as an entire ecosystem). Because the system works as a cloud-based infrastructure service that uses thousands of servers, the platform can be scaled to millions and is able to continue stimulating entities even when the users are logged off. These worlds contain millions of individually simulated entities, the core building blocks of a SpatialOS simulation. Entities can be non-player characters (NPCs), animals or pedestrians depending on simulation settings. Each entity is defined by

[15] Sources: (Games for Change, 2011; Rejeski, Chaplin, Olson, & Scholars, 2015).
[16] Sources: (Khoo, 2014; Mazarakis, 2012).

a set of Components. Components are the specified behaviours through defined properties, events and commands. In a simulated ecology, a fox is an entity with a hunting component, which has health and hunger properties. As entities interact overtime properties will change, in the fox example foxes' health properties will decrease after a fight, and hunt other animals when hungry. SpatialOS can also be used for city management, economics and defence. The Improbable City simulation model includes several elements such as the interaction of traffic patterns, energy consumption, and waste management by combining existing simulation models such as models of traffic, pedestrians, and pollution. It can stimulate situations such as the impact on traffic of a broken down delivery truck, the effects it has on deliveries for the remainder of the day, and even the impact on the wider economy. SpatialOS is the first product of Improbable, a start-up company from London. The company has raised $20 million from Andreessen Horowitz in 2015, and is also funded by from a group of technical companies' entrepreneurs and investors.[17]

Maybe the boldest project developed in terms of using computer simulations for the design of public policies is the FuturICT Knowledge Accelerator and Crisis-Relief System (FuturICT). It is a 10-year visionary project designed in response to the European Commission Flagship call. Its core component is the Living Earth Simulator Project (LES), a proposed computer simulation system that can simulate everything on Earth as a crystal ball of the world, estimated to cost EUR 900 million. By using a massive scale of data to calibrate the simulation (coming from sensor information through digital media, social networks, and public infrastructure), the LES can model human activities and their impact on global political stability, the environment and financial markets. The resulting simulations aim at predicting unexpected and complex shifts in collective opinions and social attitudes, changes in consumer behaviour, emergence of tensions in communities, demographics, migration, mobility patterns, or health trends. Hailed by Scientific American as "The Machine That Would Predict the Future", the project originally received funding from the European Union Seventh Framework Programme till 2014, and was ranked as the top of the six finalists in competition for EUR 1 billion European Commission funding, but lost to the Human Brain project in the final round of application. It is currently maintained as the research initiative of ETH Zurich. While still not at its full potential, the initiative outlines the possible scope of computer simulations.[18]

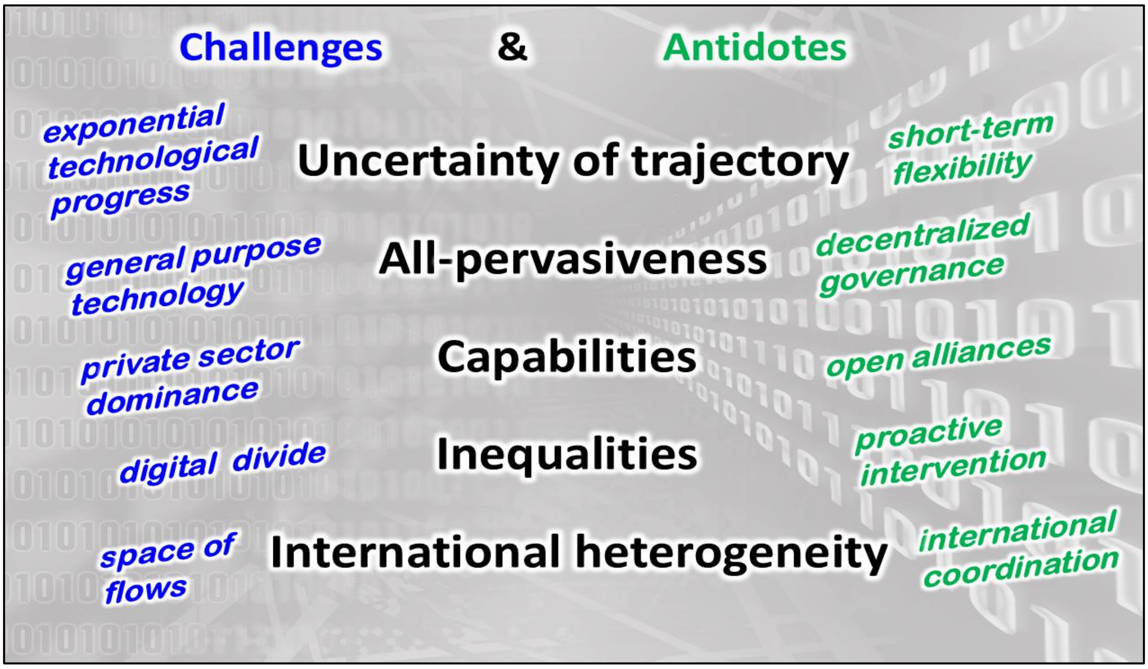
3. Considerations of foresight exercises in the digital age

The previous section explored how digital tools can enhance foresight studies. In this section, we turn the perspective around, and ask what foresight exercises have to consider in our digital age. Doing foresight for international development in our times faces certain particularities, which need to be addressed when implementing these exercises. Some of them cannot be remedied easily with computational science tools, but are rather the result of the digital age. In this section, we will review five of them and will propose strategies to address them. The list is not exhaustive, but touches on some important aspects. This aims at embedding modern tools into an adequate institutional context.

For each of the following particularities, we will discuss the related challenge and one possible antidote to confront this challenge. This aims at providing concrete policy recommendations for the implementation and potential institutionalization of foresight exercises for development purposes. Summing up, short-term flexibility, decentralized governance in the public sector, open alliances with the private sector, proactive policies for the reduction of inequalities, and international coordination are some of the institutional factors that should be considered by any foresight exercise in the digital age.

[17] Sources: (Franklin-Wallis, 2015; Johnson, 2016; Olson, 2015).
[18] Sources: (Abbott and Schiermeier, 2013; Helbing et al., 2011; The Machine That Would Predict the Future, 2011)

Figure 1
Particularities and respective antidotes of foresight exercises in a digital age



Uncertainty of trajectory

First of all, in a field of technological progress that has shown exponential growth for the last few decades, uncertainty is the name of the game. Innovation cycles become shorter and shorter. It took humankind some 150,000 years to make the innovative step from the domestication of fire (maintaining fire found elsewhere) to the proactive starting of fire (through sparks or heat), but only 66 years from the first human flight of 40 meters to a manned flight to the moon. It took radio technology 55 years to reach universal access, but the mobile phone only 20. It took humankind the 572 years between 1440 (printing press) and 2012 to store 2.5 zettabytes of information in some kind of media, but only 2.5 years to obtain just as much information storage capacity between 2012 and 2014 (see Figure 1).

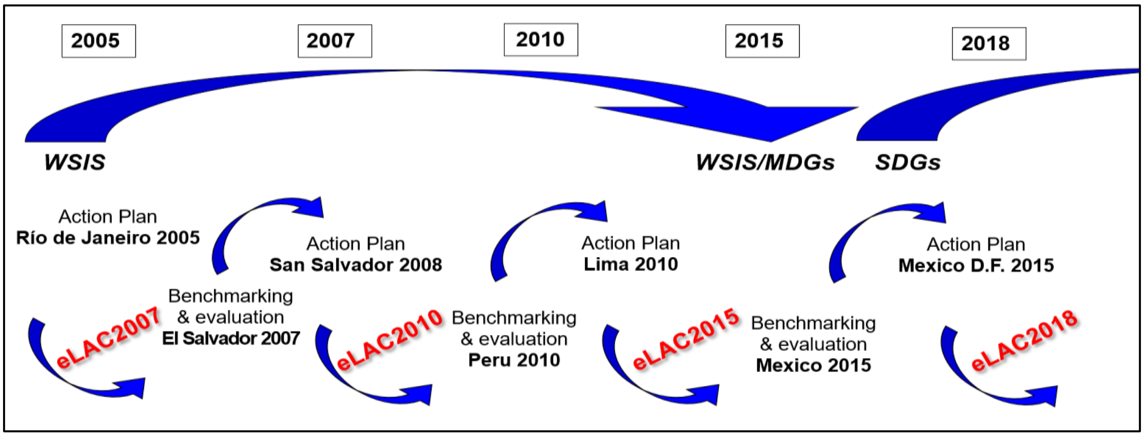
The accelerating nature of progress is not only an empirical observation that holds generally, but stems from the nature of what innovation is (Arthur, 2009). The foundational innovation theorist Joseph Schumpeter defined "that innovation combines factors in a new way, or that it consists in carrying out new combinations" (Schumpeter, 1939). Mathematically, combinations are defined through a factorial logic (i.e. "n choose k"), which results in the fact the number of possible combinations of things grows exponentially with the linear growths of inputs. While not all combinations of previous building blocks are useful or desirable, the current literature uses this as the basis for more intricate innovation models (often through so-called Urn-models, e.g. Dosi, Ermoliev, & Kaniovski, 1994; Tria, Loreto, Servedio, & Strogatz, 2014; Youn, Strumsky, Bettencourt, & Lobo, 2015). The exponential power of combinatorics is, therefore, the very definition of what innovation consists of (new combinations). It is, therefore, the nature of technology that results in this persistently exponential nature of technological progress in general (specific technologies might naturally have divergent paths, or even stop in their evolution).

As Ray Kurzweil has often pointed out on repeated occasions, "when people think of a future period, they intuitively assume that the current rate of progress will continue for future periods. However, careful consideration of the pace of technology shows that the rate of progress is not constant, but it is human nature to adapt to the changing pace, so the intuitive view is that the pace will continue at the current rate. Even for those of us who have been around long enough to experience how the pace increases over time, our unexamined intuition nonetheless provides the impression that progress changes at the rate that we have experienced recently. From a mathematician's perspective, a primary reason for this is that an exponential curve approximates a straight line when viewed for a brief duration." (Kurzweil, 2001).

For example, it is not intuitive for us to understand that if it took 40 years to add 1 billion transistors to a microchip, it might only take another two years to add another billion. Exponential doubling implies that one makes as much progress in the next step as in all previous periods combined. Some 40 years ago, the ICT paradigm was dominated by the first Xerox computers, with black and white text-processors developed by the Palo Alto Research Centre, while we currently see virtual and augmented reality on the go. An exponential doubling period suggests that during the next period we make just as much progress as has been achieved since the beginning. It is extremely difficult to develop foresight in a field of such dizzying growth, as the rapid doubling of performance is counterintuitive to the human imagination.

One way to confront the uncertainty is by combining short-term flexibility with long-term goals. This assures that policies can react to new development, while social, economic and political actors are still provided with sufficient planning security. One example of this approach the Latin American and Caribbean Action Plan for Information Society development, eLAC.[19] At the global level, the challenges and goals of Information Society development have been discussed at the highest possible political level, during the two phases of the WSIS in Geneva in 2003 and Tunis in 2005. The eLAC process fits into this global political stream as a regional complementarity to the challenge. In contrast to the 167 global action points ambitiously delineated by WSIS for the year 2015, the basic idea of eLAC2007 was to identify the most urgent and important short-term goals for the region. While the global plan had to find a compromise between the most and least developed countries, Latin America and the Caribbean was able to zoom in on their more particular problems. The result was a selection of 30 goals and 70 concrete activities to be implemented during 2005-2007. eLAC2007 is, therefore, a regional action plan whose purpose is to mediate between the ambitious goals of the global political stream and the local demands of individual countries by identifying common regional priorities. It recognizes the eminently dynamic and short-lived character of digital development, the need to be realistic and the importance of taking firm steps towards the long-term aims of the Millennium Development Goals and WSIS, which converged in 2015.

Figure 1
Long-term vision and short-term actions



[19] For more information about eLac see https://www.cepal.org/es/elac2018/antecedentes

The short-term nature of the eLAC process gave the region the opportunity to review progress towards its goals and reformulate the targets along the way, by strengthening provisions or making significant changes as part of this long-term challenge. As such, the Action Plan was evaluated through a benchmarking report, which provided empirical evidence about the advancement of each goal in a 144 -page report (Hilbert, 2012; OSILAC, 2007). Simultaneously, an eLAC Policy Priority Delphi was carried out.[20] The countries gathered in El Salvador to evaluate the results, and to subsequently approve a new plan. During the 2.5 years of its life, the first round of Delphi foresight, participants suggested to maintain 23 of the 30 goals of eLAC2007 for a new plan, while they suggested 7 completely new goals (Hilbert et al., 2009). This implies that in only 2.5 years, the larger public esteemed that about a quarter of the policy goals had already become obsolete. This shows the necessity to revisit policy plans every few years.

Table 1
Final monitoring of progress of the eLAC2007 Action Plan (Hilbert, 2012; OSILAC, 2007)

Area	Goal	Amount of progress
A. Digital access and inclusion	1 Regional infrastructure	Progress
	2 Community centres	Strong progress
	3 Online schools and libraries	Progress
	4 Online health centres	No progress
	5 Employment	Moderate progress
	6 Local government	Strong progress
B. Capacity-building and knowledge creation	7 Alternative technologies	Moderate progress
	8 Software	Moderate progress
	9 Training	Progress
	10 Research and education networks	Strong progress
	11 Science and technology	No progress
	12 Businesses	Progress
	13 Creative and content industries	Progress
	14 Internet governance	Progress
C. Governmental transparency and efficiency	15 e-Government	Progress
	16 e-Education	Strong progress
	17 e-Health	No progress
	18 Disasters	No progress
	19 e-Justice	Moderate progress
	20 Environmental protection	Moderate progress
	21 Public information and cultural patrimony	Progress
D. Policy instruments	22 National strategies	Progress
	23 Financing	No progress
	24 Universal access policies	No progress
	25 Legislative framework	No progress
	26 Indicators and measurement	Strong progress
E. Empowering environment	27 Monitoring of the World Summit and execution of eLAC2007	Strong progress

As shown in Figure 1, this dual design of long-term vision and short-term policy action was considered successful by member countries and the process is currently in its fourth phase: eLAC2007, eLAC2010, eLAC2015, and currently eLAC2018.

[20] For additional information on the eLAC Policy Delphi see Hilbert et al. (2009).

Policy-making in the context pervasiveness of ICTs

Secondly, ICTs are general purpose technologies and, like others before them, such as electricity or motorized transportation, they are at the heart of a multitude of transformations of the social and economic organization that are leading entire societies through a paradigm change (Helpman, 1998; Perez, 1983; Schumpeter, 1939). The cross-cutting nature of the transformations does not fit in with the traditional organization of political institutions, such as the ministries of education, transport, health, trade and public administration, and so forth, nor does it consider the leadership of the private sector in the development and implementation of many of the required methods.

In governments, there are authorities for infrastructure, science and technology, and often planning, but connectivity is only the first step, while the transition toward a digital age requires the expertise of each of their authorities and actors. This has severe consequences for the building of public policy agendas. One of these is that the insights are distributed among the most diverse institutions and organizations. This spans education, health, entrepreneurship, legal challenges, cultural heritage, disaster management, and national security, among many others.

This is also reflected in the respective budget allocations. For example, a study in Chile found that the US 205 million of public ICT spending in 2003 was distributed among 22 public authorities (DIPRES, 2005; Hilbert, 2011a). The Ministries of Finance, Education, and Defence all spend about 15 % each, while the other half was distributed among other authorities. The thematic experts in the subject matter, the telecommunication authority, merely counted with a telecommunication development of US$ 5 million, which is a small fraction of the total public spending. This can be seen as positive, since the digitalization of health should be in the hands of those that understand the health sector, and the digitalization of education, agriculture, justice and international relations likewise. These are the agents that have the required insights. Therefore, forecasting and planning for solutions for a digital future is not one task that can be assigned to one single authority.

One way to confront the all-pervasiveness of digital transformations is the creation of multi-agency agendas. This implies close coordination among authorities, in order to identify potential synergies, to avoid double efforts, and shared lessons learned. In the field of digital development, this is a commonly adopted truth, as many countries around the world have long set up multi-agent e-strategies (Adamali et al., 2006; Hilbert et al., 2005). It is natural that foresight exercises for a digital future have to follow this decentralized approach, and that it would be counterproductive to assign this task to one single authority.

Availability of foresight capabilities

The need for decentralization and openness of foresight exercises goes beyond a decentralized approach in the public sector, and has to consider the capabilities of the private sector, including crucial resources for modern foresight exercises, like skills and data.

We have seen in many of our case studies that the private sector often takes the lead in the use of digital tools for the generation of empirical insights and theoretical scenarios. Companies are often the ones with access to updated data, and have the skill set to implement digital foresight aids. Undoubtedly, the largest hosts of data are companies from the private sector that specialize on its curation and analysis, like Google, Facebook, and other specialized companies, like Premise, a San Francisco based data company that specializes on producing data for the “world’s hardest-to-see-places”, operating on more than 80 cities in 30 countries (*The New Yorker*, 2013). Their participation in foresight exercises through private-public sector partnerships seems very productive, not only to inform public decision-making, but also for the long-term sustainability of the private enterprise through close coordination with the respective authorities and the prevailing policy agenda. This last point can become a slippery slope, as more and more calls have accumulated that accuse private sector data monopolies to use their insights for private interest.

For example, the peer-review principle in academic research usually assures public access to databases on which academic arguments are formed. In a recent study published by the highest ranked Journal of the world's largest general scientific society (AAAS's Science[21]), a group of researchers from Facebook examined how 10.1 million U.S. Facebook users interacted in response to the rising critique that Facebook's algorithmic news ranking exposes users to less diverse and more polarized news content. The authors concluded that "compared with algorithmic ranking, individuals' choices played a stronger role in limiting exposure to cross-cutting content" (Bakshy et al., 2015: 1130), which means that the longstanding insight that people tend to read what they like and know (ideological homophily) is a much more important force in biased content selection, or more succinctly: "it's not our fault" (*WIRED*, 2015). This conclusion made wide-spread news. Naturally, given the proprietary nature of their massive dataset, nobody else had access to the opaque and self-reportedly "large, comprehensive dataset". Given that 10 million users is an only small fraction of the at least 200 million American Facebook users, commentators were left with open questions about the representativeness, and details of the dataset (including the self-reported nature of the ideological affiliation of the users, etc.) (*WIRED*, 2015). Nevertheless, the study was published in the highest academic journal, which shows that even the academic world has started to render to the monopoly of private sector data dominance for the sake of obtaining interesting new insights. The involvement of private sector data providers in foresight exercises confronts a dilemma of a potential conflict of interest that can have more tangible consequences, as in this case these very same actors would be involved in shaping future policies.

While it is true that the private sector hosts important and very insightful data sources, in contrast to previous technological revolutions, in the case of the data revolution, the distribution of capabilities situation is actually less skewed. For 2010, it was estimated that each organization of the U.S. government to host some 1.3 Petabytes of data, compared with a national organizational mean of 0.7 PB, while the government itself hosts around 12 % of the nationally stored data, and the public sector related sectors of education, health care, and transportation another 13 % (Manyika et al., 2011). In many areas, the public sector has a natural data monopoly and the digital footprint has undoubtedly provided more data on many important issues that can possibly be analysed in the near future. The public sector, therefore, comes to any potential public-private-sector partnership with an important weight.

Additionally, both the collection and analysis of further information, as well as the possession of the required skills and computational power by the public sector is rather a question of public sector priorities. Governments around the world are leaders in cutting-edge data analysis and complex simulations when it comes to security and national intelligence issues. Defence and national security agencies are some of the key innovators in the field of computational science. This does not mean that one could easily use these massive datasets and skills for foresight purposes, but it shows that it is not only theoretically, but also practically possible that the public sector gets to the forefront at the data revolution. For other countries around the world, this shows that the employment of qualified ICT personal in the public sector is rather a question of public sector priorities, than a structural problem of the economy or the labour market.

Of course, the uniquely massive datasets managed by national security agencies are often classified to different degrees and their access is regulated. Together with the foregoing argument about possible conflicts of interests in public-private-sector partnerships, one reasonable way forward is to pursue open data alliances for foresight purposes. Over recent years the so-called "open data" movement has gained much momentum, both in the public and private sector (for more see Huijboom & Van den Broek, 2011; Janssen, Charalabidis, & Zuiderwijk, 2012). First results are impressive, as, for example, the above-mentioned outstandingly successful identification of poverty levels through satellite imagery and other databases was exclusively based on open data (Jean et al., 2016). This shows that open data can be very useful to inform foresight exercises.

New digital and capabilities divides

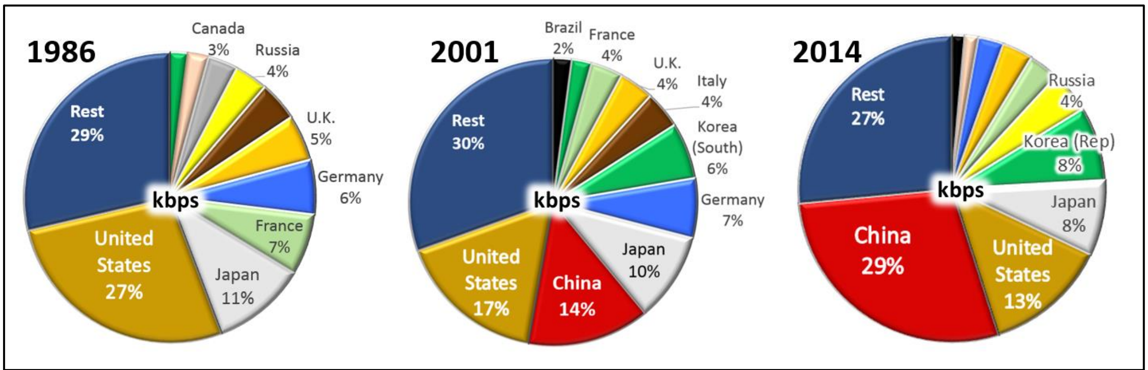
[21] American Association for the Advancement of Science.

Both the process of technology diffusion and the economic incentives inherent to the information economy increasingly concentrate information and computational capacities. On the one hand, technology does not fall uniformly distributed from heaven but diffuses through social networks over time.[22] During this diffusion process, a divide between the 'have' and 'have-nots' is unavoidable (Hilbert, 2011b). The rising challenge consists in achieving both a universal diffusion of the technology throughout the network and to assure that there is no socially damaging inequality in the different flavours of a certain technology. On the other hand, basic economic principles, such as economies of scale in information storage and short product lifecycles, lead to a natural concentration in the digital economy (Shapiro and Varian, 1998).

The digital divide is evolving from a divide that centres on the question of access to technology, toward a divide in terms of access to information (Hilbert, 2014b, 2016b). The result is a continuous divide in terms of bandwidth, data, and effective computational power.

While in 1986, 20 % of the world's largest storage technologies were able to hold 75% of society's technologically stored information, this share grew to 93 % by 2007. The domination of the top-20 % of the world's general-purpose computers grew from 65 % in 1986 to 94 % two decades later (Hilbert, 2014c). Additionally, Figure 2 shows that installed bandwidth potential on a national level is extremely concentrated. Over recent decades, the share of the top-10 countries in terms of installed bandwidth has stayed stable at 70 % of global bandwidth (Hilbert, 2016b). This means that historically, a very small group of countries dominates global bandwidth. It is interesting to notice that the top-3 countries usually captured a share of about 40 – 45 %, but expanded its influence to 50 % in recent years. In 2014, only three countries contributed the globally installed bandwidth potential for every second telecommunicated bit of information. The Figure shows clearly that this persistence in concentration is despite the rearrangement of important players. Back in 1986, a group of traditionally highly developed countries was found in the global top ranks. The U.S., Japan, France and Germany hosted more than half of global bandwidth. During the second half of the 2000s, three Asian countries, Japan, South Korea and China occupied ranks 2, 3 and 4. Starting in 2011 China has taken the global lead in terms of installed bandwidth capacity, hosting some 1.30 Petabytes per second, versus 1.26 in the U.S. By 2014, China hosts almost twice as much bandwidth potential than the U.S.

Figure 2
Global concentration of the installed bandwidth potential per country in kbps (Hilbert, 2016b)



[22] Roger (2003, p.5) defines diffusion as ''the process by which an innovation is communicated through certain channels over time among the members of a social system''.

In addition to this international data divide, the domestic digital divide “in terms of bandwidth does not show any clear monotone pattern. It fluctuates up and down over the decades as the result of an intricate interplay between incessant technological progress and diffusion of technology” (Hilbert, 2016b: 567). Far from being continuously closed, the digital data divide seems to be a moving target. It incessantly evolves through an ever changing heterogeneous set of innovations in the existing telecom bandwidth capacities (Hilbert, 2014b).

But the digital divide does not only affects access to information capacities. It can also bias the digital footprint on which the big data paradigm is based. First and foremost, the lack of infrastructure access and usage limits the availability of any digital footprint. “Twitter does not represent ‘all people’, and it is an error to assume ‘people’ and ‘Twitter users’ are synonymous” (boyd and Crawford, 2012). Not surprisingly, it turns out that the question of sample representativeness is closely linked to the degree of digital inequality. Blumenstock et al. worked with mobile phone data from Rwanda during 2005-2009 (Blumenstock and Eagle, 2012; Blumenstock et al., 2012), when mobile phone penetration was between 2% and 20%. They found that “phones are disproportionately owned and used by the privileged strata of Rwandan society” (Blumenstock et al., 2012). Frias-Martinez et al. worked with mobile phone big data from a more advanced “emerging economy in Latin America” (Frias-Martinez & Virseda, 2013) with a mobile phone penetration of around 60-80%. The big data sample matched the social stratification of the available census data impressively well (Hilbert, 2016a).

The digital divide does not stop with access but also includes human resource and skills challenges: Hal Varian, chief economist at Google, emblematically stated: “the sexy job in the next ten years will be statisticians… And I’m not kidding” (*The New York Times*, 2009). The same counts for programmers and computer scientists. In order to inform foresight exercises, they are not only needed for data analytics of the digital footprint, but also to set up more intricate computer simulation. It is predicted that in the near future even the job magnet United States will face a shortage of some 160,000 professionals with deep analytical skills (of a total of 450,000 in demand), as well as a shortage of 1.5 million data managers that are able to make informed decisions based on analytic findings (of a total of 4 million in demand) (Manyika et al., 2011). Interestingly, some less developed countries, such as countries from the former Soviet bloc (e.g. Romania, Poland, and Bulgaria), as well as the world’s large developing BRIC countries (Brazil, Russia, India and China) produce an important part of the global professionals with deep analytical skills (Hilbert, 2016a).

The challenge of technology and skill gaps is not new. A recent report of the Secretary-General highlighted the importance of “the basic requirements to take advantage of information and communications technologies are still pertinent” (ECOSOC, 2016: 13). The solution is also not new and consists of a large variety of proactive policy interventions and adequate institutional and regulatory arrangements. This reaches from basic telecommunication access policies, over the constant updating to state-of-the-art infrastructures (i.e. in terms of bandwidth), to education and skill building policies, and incentives for highly skilled labour to join the public sector (i.e. for policy foresight). These policies are more relevant than ever.[23] Several decades into the digital revolution the resulting divides are by far not closed. They constantly evolve. Any attempt to make use of computational science tools for foresight exercises will inevitably be impacted by the se divides.

Importance of international coordination

The digital paradigm is inherently global. Therefore, foresight exercises of a digital future have to consider the resulting international interdependencies of a world that is becoming more global and local (Wilkinson, 2016). In order to explain the implications of this fact, leading sociologist Manuel Castells has coined the

[23] For a detailed discussion see "Report of the Secretary General on foresight for digital development" (E/CN.16/2016/3). Available at http://unctad.org/meetings/en/SessionalDocuments/ecn162016d3_en.pdf

concept of the "space of flow" (Castells, 1999). According to this author, what matters in the digital age is not geographical space and distance, but the underlying network structure of digital networks.

International coordination can help to address this challenge. International foresight exercises contribute to make policy-making in developing countries more participatory, and can help foster transparency and accountability of public decision-making (Hilbert et al., 2009:880). Moreover, research has shown that international cooperation provides an adequate platform and sufficient scale for developing countries to adjust their policy actions to permanent technological change (Hilbert et al., 2009:893). Coordination does not imply homogenization, as the consideration of local contexts is essential for the production of relevant foresight studies, both for accurate forecasting and the creation of ownership in the creation of adequate solutions. Local context and skills are fundamental to maximize the benefits of information and communications technologies (ECOSOC, 2016: 14). Therefore, policy makers and practitioners could consider designing different levels of foresight exercises addressing global and local opportunities and threats with a view to design agendas aimed at exploiting the strengths of cooperation and strategic planning on their respective level. At the national and local levels, there are a myriad of ongoing policy agendas, including from federal governments, regions, municipalities, communities, and individual organizations (such as companies and universities). At the global level, there are long-term outlooks are general goals.

Moreover, international coordination can also help to take advantage of potential economies of scale in the development of foresight exercises. Because the variable reproduction costs are negligible ("copy-paste"), the cost structure of digital information almost entire consists of fixed costs. This leads to infinite economies of scale (Shapiro and Varian, 1998) and creates the technical feasibility of sharing sources between and within levels of analysis (e.g. regions, sectors, countries). The same accounts for the reuse of computer code for simulation models. However, just because resources can in principle be shared, does not imply that they would be. It is important to set up alliances to exploit these potential economies of scale in the development of foresight exercises that take advantage from global insights, while considering local contexts.

For instance, the innovative initiative on big data created in 2009 by the United Nations Secretary-General, Global Pulse, is currently helping to recognize and leverage the opportunities that big data offers for sustainable development and humanitarian action. This initiative was established based on a recognition that digital data offers the opportunity to gain a better understanding of changes in human well-being, and to get real-time feedback on how well policy responses are working.[24] Global Pulse works by public-private data sharing partnerships, generating high-impact analytical tools and approaches through its network of Pulse Labs, and by promoting broad adoption of useful innovations across the UN System. Several collaborative research projects, prototypes and experiments analysing data big data have been conducted by this initiative in diverse sustainable development areas including climate change, macroeconomics, and natural disasters. In addition, events that bring together scholars, practitioners and policy makers also provide the platform for networking and knowledge sharing. According to the UNDP Global Centre for Public Service Excellence (GCPSE), when properly facilitated, collaboration with independent foresight academics, practitioners and institutions could potentially help policy-makers in low-capacities countries to develop long term sustainable development policies (UNDP, 2014: 4).

Box 1
The Pulse Labs, UN Global Pulse

Pulse Labs bring together government experts, UN agencies, academia and the private sector to pioneer new methods and frameworks for using Big Data to support development goals. In order to gain insights that are actionable for global development policymakers, data needs to be understood in context. Therefore, Global Pulse is organised into a network of Pulse Labs in different regions of the world with a headquarters lab in New York, a lab in Jakarta, Indonesia, and a lab in Kampala, Uganda. Pulse Labs tap into local knowledge and innovation, establish key partnerships, test and pilot real-time monitoring approaches at the country

[24] http://www.unglobalpulse.org/about-new

level, and support the adoption of proven approaches. They design, scope, and co-create projects with UN Agencies and public sector institutions who provide sectoral expertise, and with private sector or academic partners who often provide access to data or analytical and engineering tools.

Global Pulse has identified regional vulnerability to global crises; existing real-time data collection initiatives; a nascent open source technology community; a vibrant local business environment; accelerating mobile phone coverage; and a track record of grassroots innovation as factors that are conducive to the success of in-country Pulse Labs. By sharing their breakthroughs, the labs help to establish and mainstream global best practices of how new sources of digital data and emerging technologies may be used to help policy makers understand, in real time, what is happening to vulnerable populations.[25]

Foresight in policy making: final considerations

We will end with a final caveat about the role of foresight in policy making, especially about the role of both computational methods and human outlooks on policy options. Computational numerical methods aim at providing us more detailed insight into reality. Like all science, computational methods aim at getting closer to 'the truth'. On the contrary, "Politics concerns itself only with those realms where truth is not —or is not yet— known. We do not vote for the best polio vaccines or conduct surveys on the ideal space shuttle, not has Boolean algebra been subject to electoral testing. But Laetrile and genetic engineering, while they belong formally to the domain of science, have aroused sufficient conflict among scientists to throw them into the political domain—and rightly so. Where consensus stops, politics start" (Barber, 1984: 129).

We have presented several examples where computational science methods allow us to get closer to 'the truth'. We have seen that they allow to identify poverty levels better than ever before, reveal hidden structures in global supply chains, monitor and simulate diseases with unprecedented accuracy, and allow to expose the interactions of taxes and inequality.

However, first, just because there are numbers and algorithms involved does not necessarily mean that we are dealing with something useful. Numerical algorithms can even be programmed to propagate lies, which would take us further from 'the truth'.

Second, no model can ever be 'true', because everything excluded from our limited model is seen as uncertain "noise". No model can ever get to 'the truth', as they are inevitably wrong. Reality is too complex therefore we never know all conditions and ongoing dynamics. Models abstract from this reality. The only accurate simulation would be a 1-to-1 simulation, which would imply to reconstruct reality itself, and would therefore likely be as big as reality itself. The only working model of the universe is the universe. Therefore: "models must be intrinsically probabilistic in order to specify both predictions and noise-related deviations from those predictions" (Gell-Mann, 1995: 49). In other words, any kind of model, be it a mental model used by an expert or layperson, a formal analytical model, or a numerical computational model, can only be probabilistic. Probabilities imply that the truth is not known, but that selected aspects have a certain likelihood of occurring. This leads to the well-known insight that "all models are wrong, but some are useful" (Box and Draper, 1987). This is a fact sometimes forgotten when working with the impressively detailed big data footprint or the surprising accuracy of today's algorithms, often accompanied by a certain degree of number-fascination or numerical overconfidence. It is important that numbers require us to be more specific and precise in the description of the input and output variables of the models, but they do not assure that the data is unbiased (some things are inevitably left out), or that our assumptions are valid (some dynamics are not or wrongly considered). Just because there is a number or an algorithm, it does not make models true. The benefits of using quantitative models in foresight studies lie more in their facilitation of a comparison

[25] Source: http://unglobalpulse.org/pulse-labs

between decision alternatives relating to specific scenarios than in their absolute values (Hansen et al., 2016). In order to counteract the pitfall of number-fascination, it is important that practitioners clearly understand the complementary purposes of using computational methods for foresight exercises.

References

Adamali A, Coffey JO and Safdar Z (2006). Trends in National E-Strategies: A Review of 40 Countries. *Information and Communications for Development: Global Trends and Policies in 2006*. World Bank. Washington D.C.

Alexander IF and Maiden N (2005). *Scenarios,?Stories, Use Cases: Through the Systems Development Life-Cycle*. John Wiley & Sons.

Arthur WB (2009). *The Nature of Technology: What It Is and How It Evolves*. Free Press.

Bakshy E, Messing S and Adamic LA (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*. 348(6239):1130–1132.

Barber BR (1984). *Strong Democracy: Participatory Politics for a New Age*. University of California Press.

Big Data, Big Impact: New Possibilities for International Development (2012). WEF (World Economic Forum) and Vital Wave Consulting. See http://www.weforum.org/reports/big-data-big-impact-new-possibilities-international-development (accessed 24 August 2012).

Blumenstock J, Eagle N and Fafchamps M (2012). Risk and Reciprocity Over the Mobile Phone Network: Evidence from Rwanda.

Blumenstock JE and Eagle N (2012). Divided We Call: Disparities in Access and Use of Mobile Phones in Rwanda. *Information Technologies & International Development*. 8(2):1–16.

Blumenstock JE, Gillick D and Eagle N (2010). Who's Calling? Demographics of Mobile Phone Use in Rwanda. *AAAI Spring Symposium: Artificial Intelligence for Development*. AAAI.

Box GEP and Draper NR (1987). *Empirical Model-Building and Response Surfaces*. Wiley.

boyd danah and Crawford K (2012). Critical Questions for Big Data. *Information, Communication & Society*. 15(5):662–679.

Castells M (1999). *The Information Age, Volumes 1-3: Economy, Society and Culture*. Wiley-Blackwell. Cambridge (Mass.); Oxford.

CCAFS (2015). Big Data for climate-smart agriculture. International Center for Tropical Agriculture (CIAT). C.

Cioffi-Revilla C (2014). *Introduction to Computational Social Science*. Springer. London.

Conte R et al. (2012). Manifesto of computational social science. *European Physical Journal - Special Topics*. 214325–346.

Datanami (2015). How Big Data Is Helping Indian Schoolchildren. 13 March.

DIPRES (2005). Informe de cuantificacion del gasto en gobierno electronico, Ano 2003. Direccion de Presupuesto, Ministerio de Hacienda, Gobierno de Chile.

Dosi G, Ermoliev Y and Kaniovski Y (1994). Generalized urn schemes and technological dynamics. *Journal of Mathematical Economics*. 23(1):1–19.

ECOSOC (2016). Foresight for digital development: Report of the Secretary-General. Nineteenth session. Commission on Science and Technology for Development. Geneva.

Epstein JM (2005). Remarks on the Foundations of Agent-Based Generative Social Science. Available at http://www.santafe.edu/media/workingpapers/05-06-024.pdf (accessed 27 March 2013).

Epstein JM and Axtell RL (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. A Bradford Book.

Farmer JD (2012). Economics needs to treat the economy as a complex system. Complexity Research Initiative for Systemic instabilities (CRISIS). Berlin.

Frias-Martinez V and Frias-Martinez E (2014). Spectral clustering for sensing urban land use using Twitter activity. *Engineering Applications of Artificial Intelligence*. 35237–245.

Frias-Martinez V, Frias-Martinez E and Oliver N (2010). A Gender-centric Analysis of Calling Behavior in a Developing Economy Using Call Detail Records. *AAAI 2010 Spring Symposia Artifical Intelligence for Development*. .

Frias-Martinez V and Virseda J (2013). Cell Phone Analytics: Scaling Human Behavior Studies into the Millions. *Information Technologies & International Development*. 9(2):35–50.

Gell-Mann M (1995). *The Quark and the Jaguar: Adventures in the Simple and the Complex*. St. Martin's Griffin. New York.

Gilbert N and Troitzsch K (2005). *Simulation for the Social Scientist*. Open University Press. Maidenhead, England ; New York, NY.

Ginsberg J et al. (2009). Detecting influenza epidemics using search engine query data. *Nature*. 457(7232):1012–1014.

Gutierrez D (2015). HP Data Scientists Help Akshara Foundation Transform Education for Indian Students March. Available at http://insidebigdata.com/2015/03/11/hp-data-scientists-help-akshara-foundation-transform-education-indian-students/ (accessed 25 August 2016).

Haegeman K, Marinelli E, Scapolo F, Ricci A and Sokolov A (2013). Quantitative and qualitative approaches in Future-oriented Technology Analysis (FTA): From combination to integration? *Technological Forecasting and Social Change*. Future-Oriented Technology Analysis. 80(3):386–397.

Hansen MS, Rasmussen LB and Jacobsen P (2016). Interactive foresight simulation. *Technological Forecasting and Social Change*. 103214–227.

Helbing D and Balietti S (2011). How to Do Agent-Based Simulations in the Future: From Modeling Social Mechanisms to Emergent Phenomena and Interactive Systems Design. Available at http://www.santafe.edu/research/working-papers/abstract/51b331dfecab44d50dc35fed2c6bbd7b/.

Helpman E (1998). *General Purpose Technologies and Economic Growth*. MIT Press.

Hilbert M (2011a). The end justifies the definition: The manifold outlooks on the digital divide and their practical usefulness for policy-making. *Telecommunications Policy*. 35(8):715–736.

Hilbert M (2011b). The end justifies the definition: The manifold outlooks on the digital divide and their practical usefulness for policy-making. *Telecommunications Policy*. 35(8):715–736.

Hilbert M (2012). Towards a Conceptual Framework for ICT for Development: Lessons Learned from the Latin American "Cube Framework." *Information Technologies & International Development*. 8(4, Winter; Special issue: ICT4D in Latin America):243–259 (Spanish version: 261–280).

Hilbert M (2013). Big Data for Development: From Information - to Knowledge Societies. SSRN Scholarly Paper No. ID 2205145. Social Science Research Network. Rochester, NY.

Hilbert M (2014a). *Big Data Requires Big Visions for Big Change*. TEDxUCL, x=independently organized TED talks. London.

Hilbert M (2014b). Technological information inequality as an incessantly moving target: The redistribution of information and communication capacities between 1986 and 2010. *Journal of the Association for Information Science and Technology*. 65(4):821–835.

Hilbert M (2014c). How much of the global information and communication explosion is driven by more, and how much by better technology? *Journal of the Association for Information Science and Technology*. 65(4):856–861.

Hilbert M (2015a). e-Science for Digital Development: ICT4ICT4D. *Centre for Development Informatics, SEED, University of Manchester*. Development Informatics Working Paper Series.

Hilbert M (2015b). ICT4ICTD: Computational Social Science for Digital Development. *2015 48th Hawaii International Conference on System Sciences (HICSS)*. IEEE Computer Society: 2145–2157.

Hilbert M (2016a). Big Data for Development: A Review of Promises and Challenges. *Development Policy Review*. 34(1):135–174.

Hilbert M (2016b). The bad news is that the digital access divide is here to stay: Domestically installed bandwidths among 172 countries for 1986–2014. *Telecommunications Policy*. 40(6):567–581.

Hilbert M, Bustos S and Ferraz JC (2005). Estrategias nacionales para la sociedad de la información en América Latina y el Caribe. Information Society Programme No. LC/W.17. United Nations ECLAC. Santiago.

Hilbert M and López P (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*. 332(6025):60–65.

Hilbert M, Miles I and Othmer J (2009). Foresight tools for participative policy-making in inter-governmental processes in developing countries: Lessons learned from the eLAC Policy Priorities Delphi. *Technological Forecasting and Social Change*. 76(7):880–896.

Huijboom N and Van den Broek T (2011). Open data: an international comparison of strategies. *European journal of ePractice*. 12(1):4–16.

Janssen M, Charalabidis Y and Zuiderwijk A (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*. 29(4):258–268.

Jean N et al. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*. 353(6301):790–794.

Kurzweil R (2001). The Law of Accelerating Returns. Available at http://www.kurzweilai.net/the-law-of-accelerating-returns.

Kwakkel JH and Pruyt E (2013). Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change*. Future-Oriented Technology Analysis. 80(3):419–431.

Lazer D et al. (2009). Computational Social Science. *Science*. 323(5915):721–723.

Lazer D, Kennedy R, King G and Vespignani A (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*. 343(6176):1203–1205.

Le Bars M, Le Grusse P and Albouchi L (2014). AquaFej : a simulation game for planning water management : an experiment in central Tunisia. *International Journal of Sustainable Development*. 17(3):242–260.

Letouzé E (2012). Big Data for Development: Opportunities and Challenges. White p. United Nations Global Pulse. New York.

Macy MW and Willer R (2002). From Factors to Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology*. 28(1):143–166.

Manyika J et al. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. McKinsey & Company.

Mayer-Schönberger V and Cukier K (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think*. John Murray. London.

Miles I, Harper JC, Georghiou L, Keenan M and Popper R (2008). The many faces of foresight. *The handbook of technology foresight: Concepts and practice*. 3–23.

Miller JH and Page SE (2007). *Complex Adaptive Systems*. Princeton University Press. Princeton, New Jersey.

OSILAC (2007). Monitoring eLAC2007:Progress and current state of development of Latin American and Caribbean information societies. Information Society Programme No. LC/W.151. United Nations ECLAC Observatory for the Information Society in Latin America and the Caribbean. Santiago, 143.

Page SE (2010). *Diversity and Complexity*. Princeton University Press.

Perez C (1983). Structural change and assimilation of new technologies in the economic and social systems. *Futures*. 15(5):357–375.

Raento M, Oulasvirta A and Eagle N (2009). Smartphones: An Emerging Tool for Social Scientists. *Sociological Methods & Research*. 37(3):426–454.

Schelling TC (2006). *Micromotives and Macrobehavior*. W. W. Norton & Company.

Schumpeter J (1939). *Business Cycles: A Theoretical, Historical, And Statistical Analysis of the Capitalist Process*. McGraw-Hill. New York.

Shapiro C and Varian HR (1998). *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Press.

Sukumar SR and Nutaro JJ (2012). Agent-Based vs. Equation-Based Epidemiological Models: A Model Selection Case Study. IEEE: 74–79.

The Guardian (2014). Colombia rice growers saved from ruin after being told not to plant their crop. 30 September.

The New York Times (2009). For Today's Graduate, Just One Word: Statistics. 6 August.

The New Yorker (2013). A New Premise to Capture Prices Around the World. 14 October.

Tria F, Loreto V, Servedio VDP and Strogatz SH (2014). The dynamics of correlated novelties. *Scientific Reports*. 4.

UN Global Pulse (2012). Big data for development: Challenges & opportunities. *United Nations, New York*.

UNDP G (2014). Foresight as a Strategic Long-Term Planning Tool for Developing Countries.

Wilensky U and Rand W (2015). *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. The MIT Press. Cambridge, Massachusetts.

Wilkinson A (2016). Using strategic foresight methods to anticipate and prepare for the jobs-scarce economy. *European Journal of Futures Research*. 4(1):12.

WIRED (2015). What Facebook's 'It's Not Our Fault' Study Really Means. 8 March.

Youn H, Strumsky D, Bettencourt LMA and Lobo J (2015). Invention as a combinatorial process: evidence from US patents. *Journal of The Royal Society Interface*. 12(106):20150272.