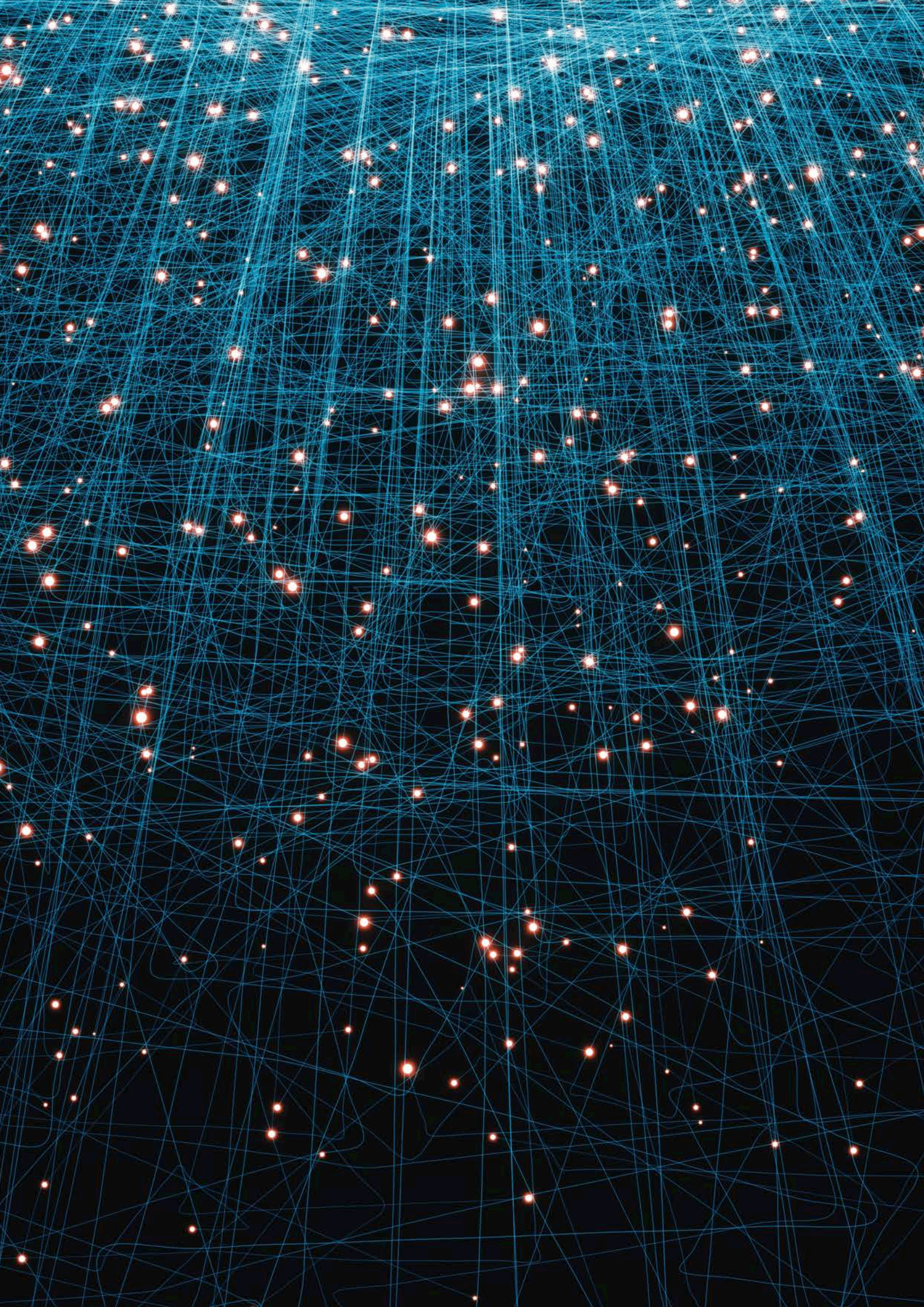
Cybersecurity Trends

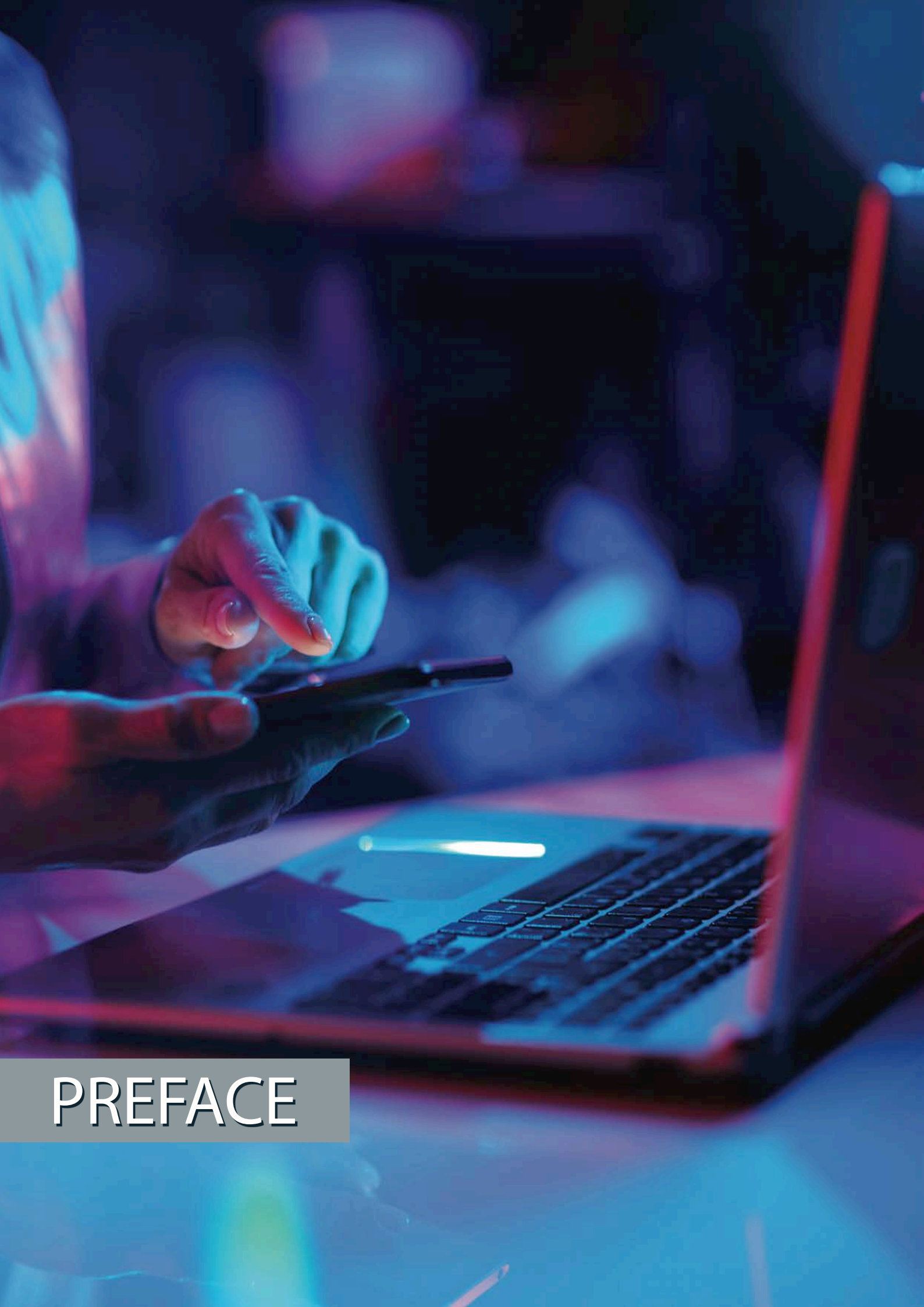
Special Edition. 2025

Artificial Intelligence, Metaverse and Quantum Computing: their hidden side, risks and challenges

This volume was realised with the support of POLICE GENEVE



Index



PREFACE

Preface.

Author: Monica Bonfanti, Commander of the Geneva State Police

Prevention

With the fourth industrial revolution having been completed three years ago, we are now witnessing a frantic succession of new technologies, all based on Artificial Intelligence (AI), which is developing second by second thanks to quantum computers, enabling it to integrate millions of data points, measurable in tera-, peta- and zettabits.

Everyone, whether they like it or not, now has a virtual alter ego that goes beyond and encompasses their physical existence.

BIO

Monica Bonfanti graduated in forensic science from the Institute of Forensic Science and Criminology (IPSC) at the University of Lausanne, where she obtained a PhD in 2001. Between 1993 and 1998, she worked as a teacher at the IPSC. From September 1993 to January 2000, she was an expert for the Swiss and French courts, specializing in the field of firearms and tool marks. From September 1993 to January 2000, she was responsible for teaching and research in the field of firearms, gunshot residue and tool marks, as well as handgun training at the IPSC. In February 2000, Monica Bonfanti joined the police force of the State of Geneva as technical head of the Technical and Forensic Police Brigade (BPTS). In August 2006, she was promoted to the highest position in the Geneva State police force, becoming Commander.

With the Metaverse, and the most complete of these 'alter egos', the avatar's critical threshold has now been reached.

Now, more than ever, the Geneva State Police have given themselves the means to keep pace with these technologies so that they can carry out their mission.

And for good reason: there can be no prevention, action or repression without mastering the basics of the cutting-edge technologies now used by most of us. So, the Geneva Cantonal Police, drawing on its experience and collaboration with the INTERPOL Metaverse Expert Group (i-MEG), has developed a vision for 2030+.

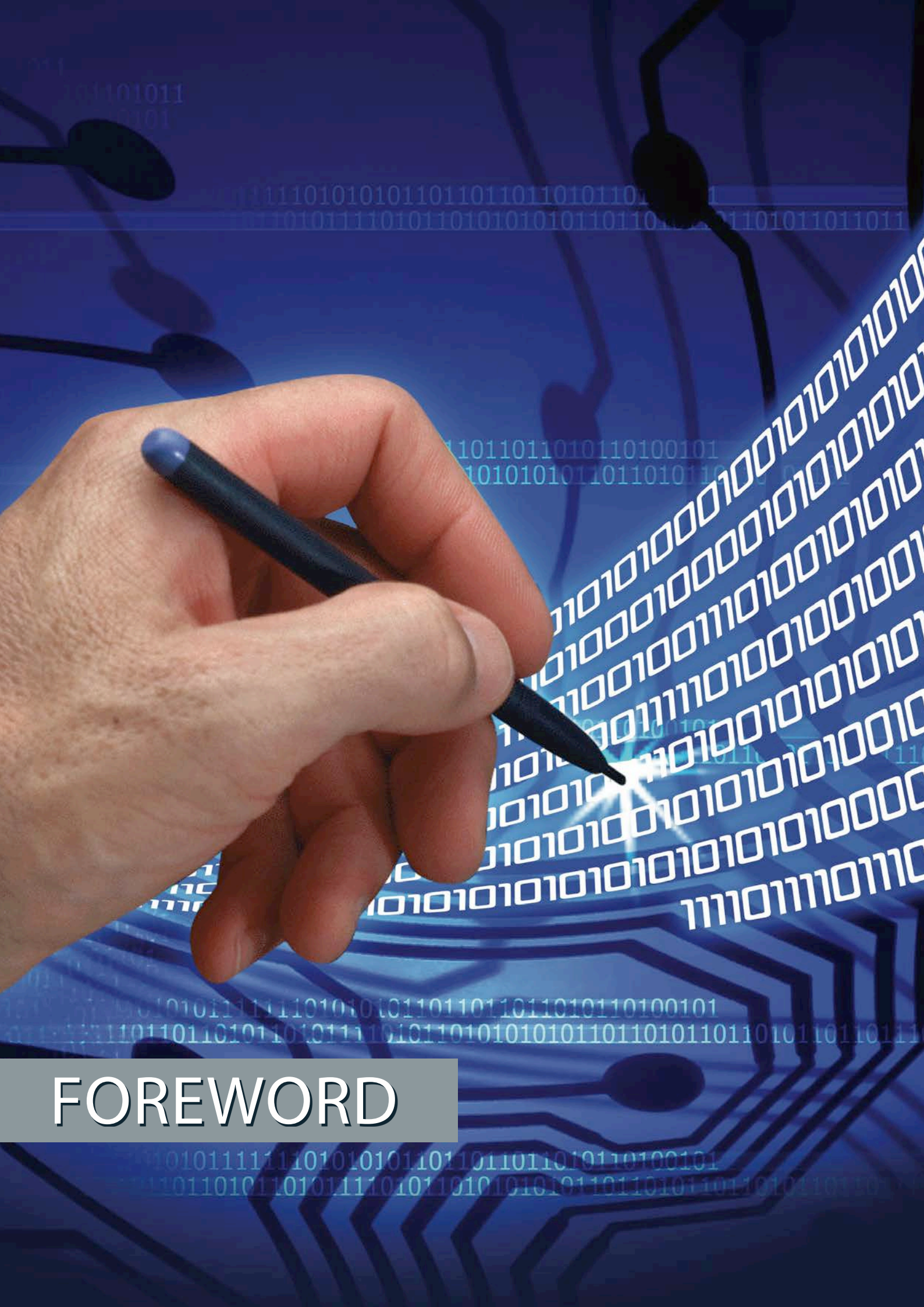
The aim of this strategy is to give police officers the knowledge they need to prevent crime and assist victims, but also, for the forensic core of the Geneva State Police, to be a visible player in the virtual world, particularly in the Metaverse.

Then, it's a matter of gathering clues that can be used as evidence for everything to do with 'virtual' crimes, i.e. scams, fraud, smartphone theft and the use of their tools - financial and personal - right up to the worst-case scenario, that of a real crime whose perpetrator has preventively 'protected' himself with a cohort of virtual alibis.

In the vast virtual world of AI, quantum computing and the Metaverse represent an evolution that we should not judge, but to which we must adapt, as simple users or as representatives of the forces of law and order.

The Geneva State Police were therefore delighted to contribute to this volume, which is rich in viewpoints and lessons learned from specialists from different professional backgrounds and countries. Useful or not, ethical or not, legal or not, these three pillars of the digital world are dissected so that their 'dark side' is better known by users, who can then use them wisely, while remaining in control of their most vital data.

We would like to thank the authors and publishers for this collective effort, which is as necessary as it is welcome. ■



FOREWORD

Foreword.

IN CYBER FORUM

Author: General (ret.) Marc Watin-Augouard

"Trains will prevent cows from having milk, electricity will be fatal to humans and animals", while the danger posed by the motor car was such that, according to the Locomotive Act (1865), it had to be preceded in town by a man waving a red flag 50 metres ahead to warn of the threat.

These examples may make you smile, but they illustrate the fears aroused by new technologies. Ignorance is a source of fear. Digital innovations are no exception to the rule. They are all the more worrying because they are based on dematerialisation, on an abstraction whose effects we know without always mastering its causes. By opening the bonnet of a car, we can see, touch and understand how each element of the engine or transmission enables movement, thanks to a fuel that can be perceived by sight and smell.

© Petar Milošević, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons

The digital metamorphosis escapes our senses, our scale of reference points in time, space and size; it is all at once infinitely fast (1 quintillion operations per second for an exaflop computer), infinitely small (10 billionths of a metre for the fineness of a 10 nanometre engraving on a chip) and infinitely large (a thousand billion terabytes of data created in 2030). Quantum physics contradicts Newtonian principles.

The real world, made up of land, sea and air environments, has been joined by an invisible digital substratum, imperceptible because it is immaterial. Our traditional society was characterised by the unity of time, place and action. This trilogy is now being undermined by a digital metamorphosis that is fuelling fears and fantasies. What cannot be easily conceived, cannot be clearly stated.

AI and the metaverse are part of this sometimes agonising questioning, whenever humans devise a technology that could lead to their downfall.

BIO

- Founder and Co-Director, Forum InCyber
- Former Director, Research Centre of the National Gendarmerie School (CREGN)
- Former Inspector General of the Armed Forces
- Army General (2008)
- Lieutenant General (2007)
- Major General (2006)
- Commander, Gendarmerie for the Northern Defence Zone (2005 - 2008)
- Commander, Nord-Pas-de-Calais Gendarmerie Region
- Advisor for the Gendarmerie, Cabinet of Dominique de Villepin (2004 - 2005)
- Brigadier General (2003)
- Security Advisor, Cabinet of Nicolas Sarkozy (2002 - 2004)
- Commander of the Champagne-Ardenne Departmental Gendarmerie Legion (2000 - 2002)
- TEACHING: Lecturer in law, Universities Panthéon-Assas (Paris II), René Descartes (Paris V) and Aix-Marseille III-Méditerranée

Raymond Aron once said, with reference to nuclear weapons, that humanity now had the means to destroy itself. A similar view is held by some scientists, who fear a general AI that surpasses human intelligence and takes decisions could enslave humans. In 1863, Samuel Butler said that *"Man will become to machines what the horse and the dog are to Man"*. Where will the separation between humans and machine be? As for the metaverse, a space projected from our physical space by a bijective relationship, not devoid of any link with AI, at what moment will its level of veracity lead us to no longer know where we stand, to mistake the virtual for the real and vice versa? Where will the separation between the true and the false be?

© Rehab Center Parus. <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons

Yes, the technologies emerging today are truly disruptive, because they force us to rethink humanity by asking existential questions.

AI, in its primitive conception, is already ancient. In 1943, Warren McCulloch and Walter Pitts imagined a cognitive system in the form of a network of artificial neurons in their prophetic article *A logical calculus of the ideas immanent in nervous activity*. In 1950, Alan Turing wrote an article in *Mind* entitled *Machines of Computation and Intelligence*, which heralded the work of 1956 at Dartmouth College. Two years later, in 1958, Frank Rosenblatt produced the first connectionist automatic learning system (based on the connection of artificial neurons), the Perceptron.

From symbolists to connectionists, AI, despite its 'winters', continued to evolve in the secrecy of insiders, until the release of generative AI on 30 November 2022 democratised the technology with 1 million users in less than five days (compared with 3.5 years for a million Netflix subscribers). It is likely that Deepseek achieved this result in just a few hours.

Lifting the lid on AI! In the beginning, there was data, the raw material of the digital metamorphosis. This data, which is now produced in profusion - and will be even more so tomorrow with the development of the IOT - feeds LLMs (Large Language Models) and is 'scraped' and 'crawled' in the immense reservoir of the web. As Tariq Krim points out, *"LLMs have taken global culture, all the things that make up our culture, be they photos, images, videos and so on. We've scanned all of YouTube, all of Wikipedia. Most platforms sucked up all the content by pirating and did the same thing with all the world's literature. It's called Common Crawl. All the world's culture can be phagocytosed, digested"*. So, like oil, data is extracted, refined and delivered to the engine of algorithms.

Blogtrepreneur, CC BY 2.0 <https://creativecommons.org/licenses/by/2.0>, via Wikimedia Commons

But do we have the right to extract? Is the data protected by law (personal data, national defence secrets, business secrets, intellectual property, professional secrets)? Is it the result of illegal action, such as data leaks? This is the first question that springs to mind. If we look simply at the creative process, how many AI products disregard the rights of artists and writers? The current lawsuits against ChatGPT and others bear witness to the shameless use of cultural assets belonging to others for profit. If the courts do not take significant corrective action, AI could spell the death of creators. Are the algorithms sufficiently controlled and transparent to transform data into predictions, text, images, sound and action (agentic AI) without risk? This is far from certain, as evidenced by the many calls for transparency issued by international organisations and bodies (UN, OECD, Council of Europe, EU, OSCE, Asilomar Conference, Montreal Declaration, Global Partnership on AI, etc.). This transparency is one of the requirements of the European regulation on artificial intelligence (AI Act - June 2024), which deals in particular with generative AI and imposes traceability obligations. Biases and hallucinations are likely to multiply if the product cannot be traced back to its source. This is one of the major dangers if the use of incomplete, biased or false data is accepted as true.

While uses follow the emergence of new technologies, misuses appear more quickly, because fraudsters and predators are always more active, more reactive: the sword is more agile than the shield, in particular because the shield must apply legal rules and respect procedures. AI-generated images of naked young girls posted on social networks precede the articles of the criminal code that punish sexual content. The use of AI to facilitate cyber attacks (phishing, off-the-shelf malware) adds to the sense of alarm and reinforces the temptation to over-regulate at the risk of undermining innovation.

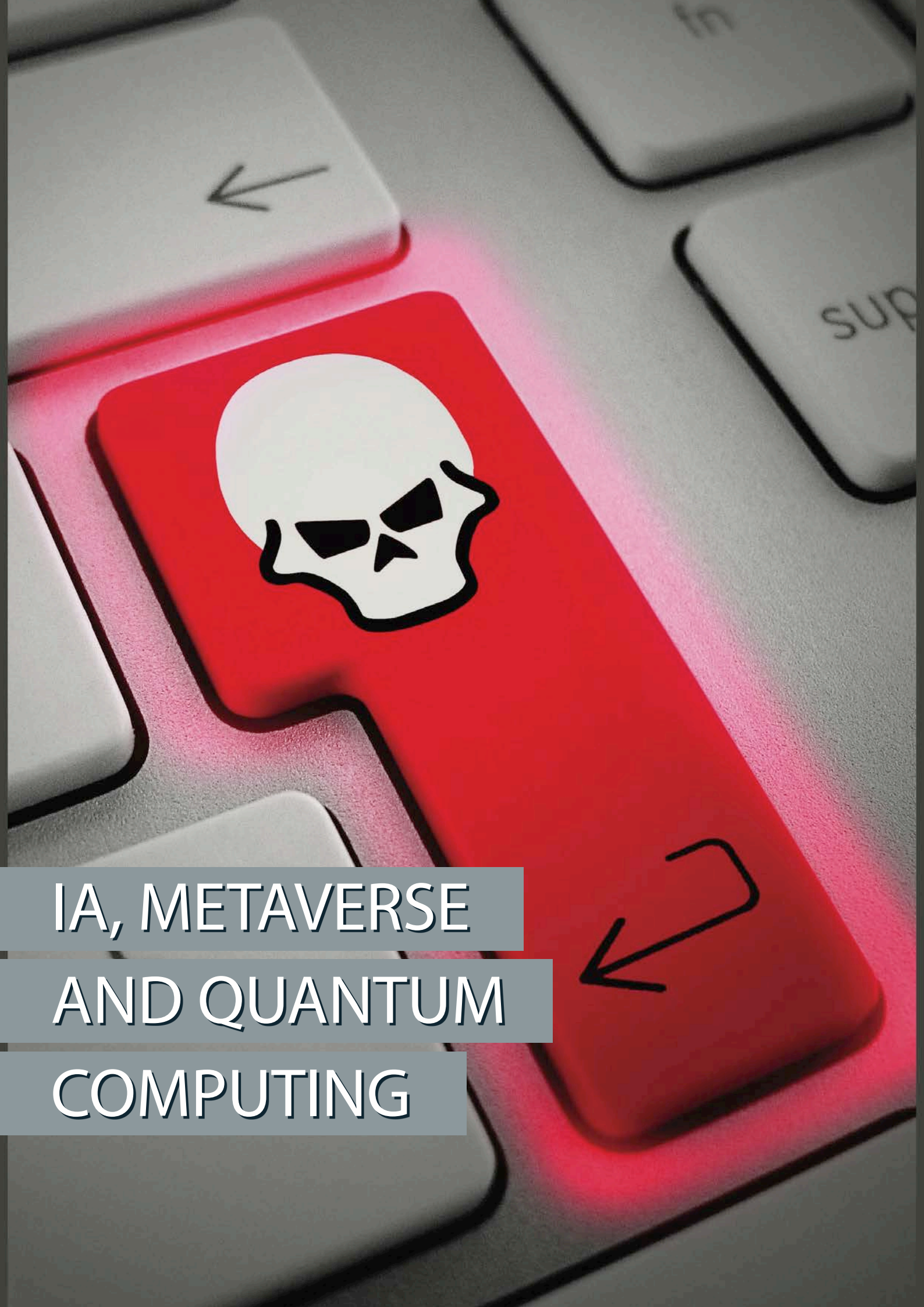
The manipulation of information is also a danger for democracy, businesses and individuals, whose reputations can be put to the test. An MIT study showed that false information takes six times less time to reach 1,500 people than true information. Distorted information benefits from the virality of networks, while restoring the truth is much more laborious.

The metaverse is defined by the interministerial report of the French mission on the development of metaverse (2022) as "an online service giving access to shared and persistent real-time 3D space simulations, in which we can live immersive experiences together". Its link with data and AI is obvious; it allows us to move on from video games to a universe of stunning realism in which our bodily identity is transformed into an avatar. Projection onto this immaterial representation raises unprecedented questions: can our avatar commit sexual assault on another avatar? The metaverse transforms abstract concepts into tangible experiences. Beyond the control of the imaginary, there is a risk of losing the imaginary, of losing our relationship with the real, which is modified and corrected. You can't see what the sensors don't see; it's the virtual that's right. It's perversion through images. The boundary between the real and the imaginary becomes porous, leading to a hybridisation between the human and his avatar.

An inventory of all the risks associated with these technologies could lead to their being strictly regulated, or even banned. That would be a mistake! The truth is that their deployment requires us to protect the most extraordinary automated data processing system, which today has no firewall or antivirus: our brains.

© Wikimedia Commons

The answer lies in rediscovering the critical spirit, in the general culture that has been all too abandoned by our education system. The "ultimate battle" has begun. The one that will enable human intelligence to dominate artificial intelligence. This is not an option, but an absolute necessity if we are to save humanity from digital slavery. ■



IA, METAVERSE AND QUANTUM COMPUTING

The Digital Pandora's Box: AI, Quantum Computing, and the Metaverse Unleashed.

Author: Marco Essomba

In the race towards technological supremacy, humanity stands at a precipice.

Imagine a scenario where quantum-powered AI systems operate within the Metaverse, creating a hyper-realistic virtual world indistinguishable from reality.

Harry Bates - Pandora - Tate Britain Sep 2010 detail of right hand and box front (5131762464).png. via Wikimedia Commons

BIO

Marco Essomba is the Founder & CTO of BlockAPT – a UK based innovative cybersecurity company. An influential thought leader in cybersecurity with almost 2 decades of working with some of the largest and well known institutions. Marco's passion, expertise and knowledge has culminated in the design of the unique central management, command and control BlockAPT platform which allows businesses to stay ahead of cyber threats 24/7. Marco is often called upon as a panellist at cybersecurity conferences and has been a host ambassador at CyberTalks, one of London's largest cybersecurity events. With 16,000+ followers on LinkedIn and 35,000+ on Twitter, he is sought after for his quick problem-solving approach and helping businesses futureproof their security infrastructure. To find out more about BlockAPT, please visit : https://www.blockapt.com
To find out more about Marco Essomba,please visit https://www.linkedin.com/in/marcoessomba/ or https://twitter.com/marcoessomba

Artificial Intelligence, Quantum Computing, and the Metaverse promise to revolutionise our world, offering unparalleled advancements in science, medicine, communication and so much more. Yet, as we eagerly embrace these innovations, we risk opening a Pandora's box of unprecedented threats to our privacy, security, and our very way of life.

Consider the dark potential of AI in the hands of malicious actors. In this environment, sophisticated deepfakes could be generated in real-time, making it impossible to tell the truth from fiction. Advanced machine learning algorithms can now create deepfakes indistinguishable from reality, enabling widespread disinformation campaigns. We have seen recent examples where it has already allegedly swayed elections or incited social unrest.

AI-powered autonomous weapons systems raise ethical concerns about machines making life-or-death decisions on the battlefield. And as AI becomes more sophisticated, there's a growing risk of it being used to manipulate and exploit human psychology on a massive scale.

State actors could use this technology to create elaborate false narratives, manipulating public opinion and rewriting history on an unprecedented scale.

The Metaverse, touted as the next evolution of the internet, presents its own set of troubling implications. As we increasingly live our lives in virtual worlds, the line between reality and simulation blurs. Addiction to these immersive environments could lead to widespread social isolation and mental health issues. Moreover, the vast amounts of personal data generated in the Metaverse – from our movements to our social interactions – create a treasure trove for surveillance capitalism and authoritarian control.

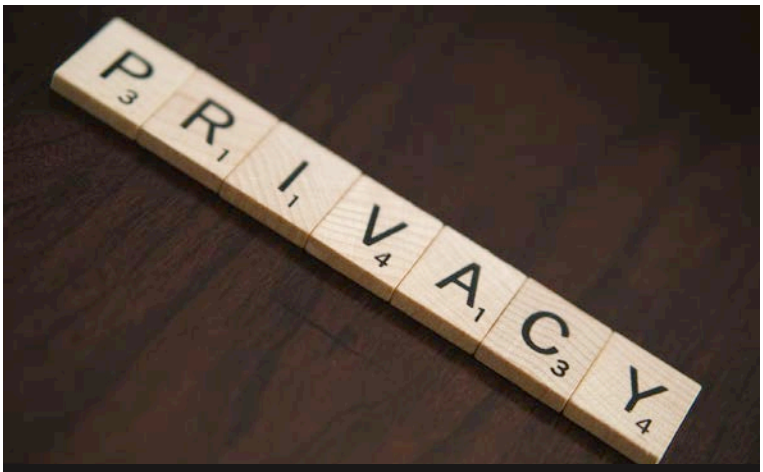
Quantum Computing, while still in its infancy, looms as an existential threat to our **current** cybersecurity infrastructure which relies on data privacy and integrity. Once fully realised, Quantum Computers will be able to break most modern encryption methods in seconds, rendering our digital communications, financial systems, and critical infrastructure vulnerable to attack.

Moreover, the integration of AI and Quantum Computing in the Metaverse could lead to the creation of virtual entities that surpass human intelligence. These digital 'beings', unconstrained by physical limitations, could evolve and multiply at an exponential rate, potentially overwhelming human users and even challenging our dominance in both virtual and real worlds. The convergence of these technologies compounds their potential for serious misuse.

The economic implications of these technologies are equally staggering. As AI and Quantum Computing advance, entire industries may become obsolete overnight, leading to massive job displacement. The Metaverse could create new virtual economies that rival or surpass real-world GDPs, raising complex questions about taxation, regulation, and wealth distribution across digital and physical realms.

Another concerning aspect is the potential for these technologies to exacerbate existing social inequalities. Access to advanced AI, Quantum Computing, and fully immersive Metaverse experiences may become a privilege of the wealthy, creating a new digital divide that further separates the haves from the have-nots. This could lead to a two-tiered society where those with access to these technologies have insurmountable advantages in education, career opportunities, and quality of life.

Privacy, once considered a fundamental human right, is under siege from all sides. AI-powered facial recognition and behaviour prediction algorithms track our every move in the physical world. Quantum Computing threatens to lay bare our digital lives. And the Metaverse promises to map our innermost thoughts and desires. In this brave new



Owen Moore, CC BY 2.0 <https://creativecommons.org/licenses/by/2.0>, via Wikimedia Commons

world, the concept of personal privacy may become nothing more than a quaint relic of the past.

The environmental impact of these technologies also warrants serious consideration. The massive energy requirements for training large AI models, operating Quantum Computers, and maintaining the infrastructure for a global Metaverse could significantly contribute to climate change. Balancing technological progress with environmental sustainability will be a critical challenge in the coming decades.

As we grapple with these challenges, it's crucial to recognise that the development of AI, Quantum Computing, and the Metaverse is not occurring in a vacuum. Geopolitical tensions and the race for technological supremacy among nations add another layer of complexity to the situation. The first country to achieve quantum domination or develop artificial general intelligence could gain an insurmountable advantage in global affairs, potentially destabilising the current world order.

In light of these multifaceted challenges, a coordinated global response is imperative. International cooperation on AI ethics, quantum security standards, and Metaverse governance will be essential to mitigate risks and ensure these technologies benefit humanity as a whole.

Educational initiatives will also play a crucial role in preparing society for this new era, fostering digital literacy and critical thinking skills to navigate an increasingly complex and interconnected world.

Ultimately, the story of AI, Quantum Computing, and the Metaverse is still being written.

As we stand on the brink of this technological revolution, we must ask ourselves: Are we prepared for the consequences? Can we harness the power of AI, Quantum Computing, and the Metaverse while safeguarding our fundamental rights and values? Or will we unwittingly usher in a dystopian future where privacy is extinct, reality is malleable, and humanity is at the mercy of the very machines we created?

The digital Pandora's box has been opened - but the choice is ours and the window for action is rapidly closing. ■

Artificial Intelligence, Metaverse and Quantum Computing: Power Issues and Economic Warfare.

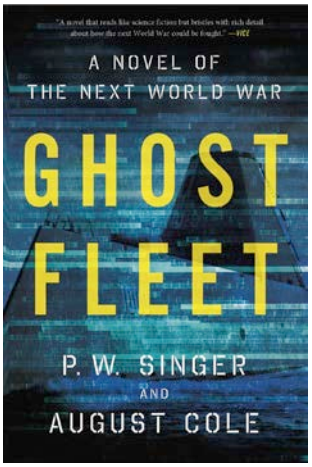
Author: Stéphane MORTIER

BIO

Stéphane holds a PhD in management science (Paris 1 Panthéon-Sorbonne), a degree in political science, sociology and international politics from the Université libre de Bruxelles, and a degree in strategic management from the École de Guerre Économique. He is a research associate at the Université Gustave Eiffel and a lecturer at the École de Guerre Économique, the École Supérieure de Gestion and the University of Likasi (Democratic Republic of Congo).
He spent 20 years with the French Ministry of the Interior, including almost 15 as deputy director of the economic intelligence service. In this capacity, he represented the French Gendarmerie Nationale on public economic security policy bodies. He is currently an independent consultant and Vice-Chairman of Oplaa.Tech (Office of Prospective and Logical Adaptative and Anticipation), a consortium specialising in services integrating business intelligence and artificial intelligence.

Introduction

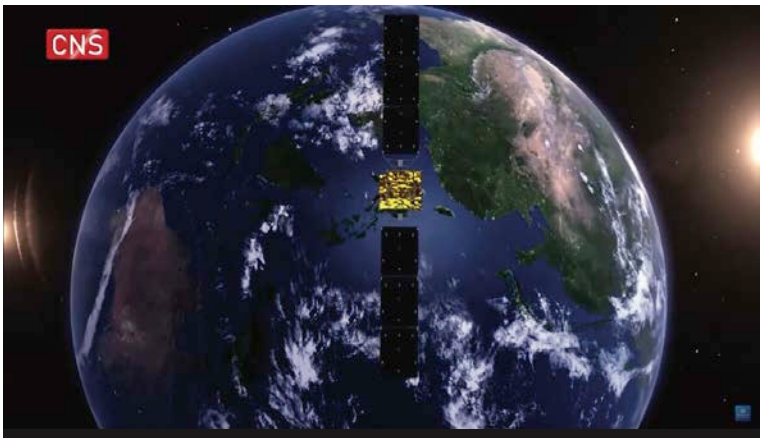
Technology has always been a battleground for rivalry between nations and territorial claims: from coal to steel, from steel to railways, from steam to electricity. Times have changed but the stakes remain the same: dominating markets, establishing power, imposing a vision of the world. It is through these constants that economic warfare has refocused on high technologies such as artificial intelligence and the quantum computer, not forgetting their entire value chain, from strategic minerals to algorithms and connected objects.



In 2015, Peter W. Singer and August Cole published a novel entitled *Ghost Fleet: A Novel of the Next Worls War*[1], a fictional account of a China that is technologically advanced and even dominant in certain critical areas, a post-communist power that is still strongly nationalist and authoritarian. The latter, allied with Russia, opposes and attacks the United States... It is doing so by combining disruptive innovations, in laser weapons to neutralise American satellites and to control access to and use of outer space, in military robotics with aerial drones, in detection systems enabling them to know the position of American strategic assets (economic warfare), in cyberspace, with its army of hackers and the "backdoors" introduced into American weapons and command systems through hacking or the inclusion of compromised Chinese electronic components in the supply chains of major American arms manufacturers (economic warfare)[2]. All the

technologies used in this fiction are in fact very real. The authors' intention was to describe a plausible future and to encourage the digital superpowers to discuss the issue[3]. However, 10 years on, these discussions have still not taken place and the competition is raging with ever-increasing tensions.

In recent years, a number of events have corroborated the predictions of this novel, including the Chinese remote sensing satellite Qimingxing 1, whose camera was controlled by an artificial intelligence, and more recently the launch of Deepseek, as well as power strategies in the metaverse market and a veritable race for quantum technology.



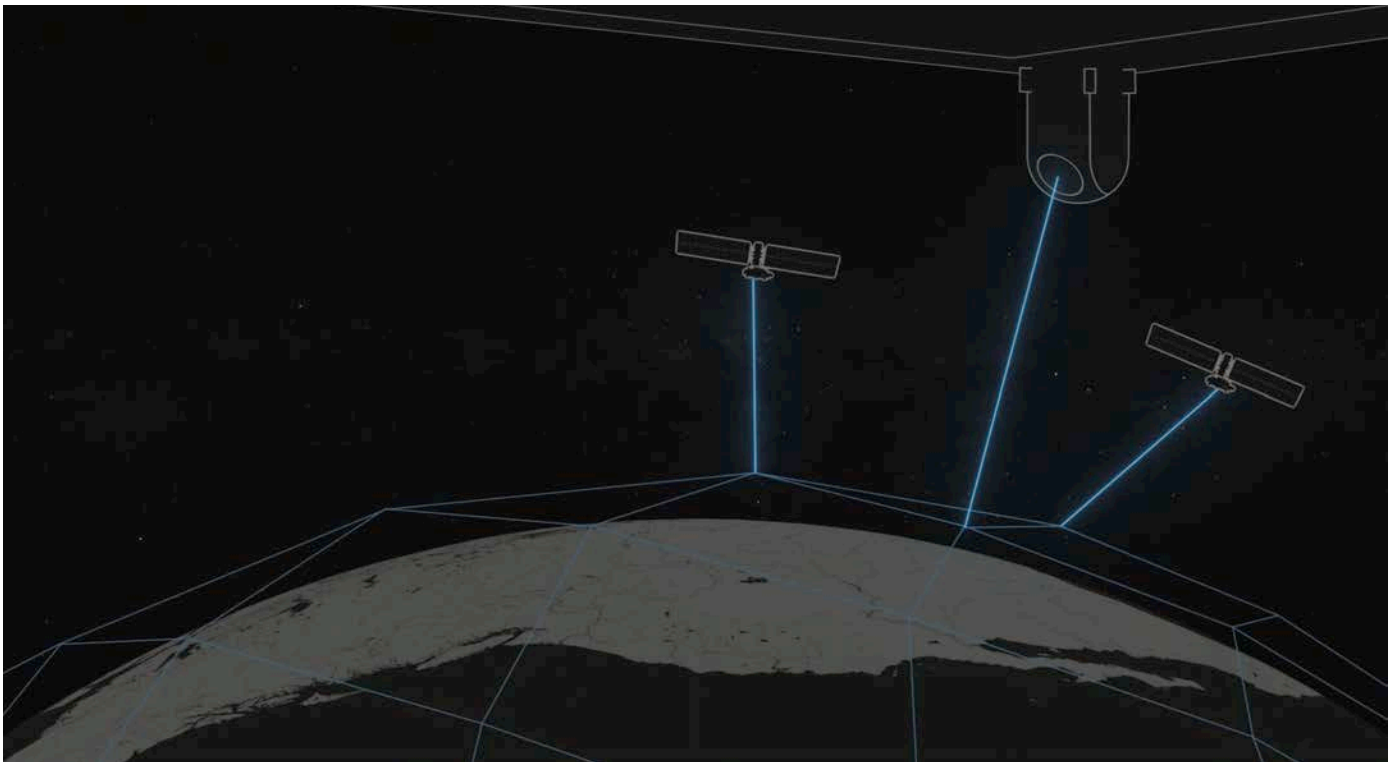
China News Service, CC BY 3.0 <https://creativecommons.org/licenses/by/3.0>,via Wikimedia Commons

Artificial intelligence and power issues

In early 2023, the Chinese satellite Qimingxing 1, was completely controlled by an artificial intelligence system, without any commands, tasks or human intervention[4]. The researchers (Wuhan University) were surprised by the ground targets that the AI chose to observe. In particular, it was interested in the port of Osaka in Japan, an occasional passage and logistics point for US Navy ships operating in the Pacific. Under the control of artificial intelligence, the satellite's camera also focused on the Indian city of Patna. The researchers speculate that these choices may have been dictated by recent events, which have seen China and India clash in 2020 over a border dispute and, of course, the strategic rivalry between China and the United States. The military theme could therefore be one of the criteria chosen by the AI, although we do not know exactly how it came to this decision. The researchers explained that the ultimate aim of this work was to one day be able to use AI to make more effective use of China's 260 remote sensing satellites, which are currently under-utilised. One of the avenues envisaged would be to put AI in charge of national defence surveillance to detect military activities.[5]

Meanwhile, in the United States, SpaceX is developing a constellation of satellites equipped with artificial intelligence tools for military applications (Starshield) in collaboration with the US Defence and its National Reconnaissance Office (NRO). These satellites would be equipped with imaging systems to monitor ground movements anywhere on the planet. The satellites will also be able to share their observations in real time with all the US intelligence services.

Clearly, observation capabilities, and therefore intelligence, are strategic priorities for the major powers, and artificial intelligence is at the heart of this.



More recently still, the Chinese start-up Hangzhou DeepSeek Artificial Intelligence shook up the world of artificial intelligence when, on 20 January 2025, it published its freely usable *Large Language Models* (LLMs), which it claimed could rival those of major American companies (OpenAI, Google Gemini, etc.). What's more, they will cost less in terms of the IT resources required to provide results for user queries[6]. This is a technological and economic disruption that calls into question the entire American investment model. Indeed, Deepseek is an AI model that is equivalent to or even more powerful than OpenAI and Meta, but developed for just USD 6 million, with an application programming interface that is around 96.4% cheaper than its American rivals, all within an innovative opensource approach.

The announcement of this news provoked surprising reactions with considerable economic impact. The financial markets went into a panic, and here are just a few examples:

- Nasdaq falls sharply (-3.07%);
- Nvidia lost 13% and its position as the world's largest market capitalisation (a loss of $590 billion in value), while Taiwan Semiconductors lost 17%;
- The *Magnificent* 7 (Nvidia, Apple, Microsoft, Amazon, Alphabet, Meta, Tesla) are all down sharply on the stock market;
- 28% jump in the US financial market volatility index VIX (often called the stock market fear index, this index generally rises when there is turbulence and prices fall).

Beyond artificial intelligence and Deepseek, a simple announcement can send financial markets into a tailspin. This information strategy is a veritable weapon of economic warfare, and the Deepseek 'affair' illustrates this perfectly, even if the surprise effect could have been avoided with high-quality competitive intelligence and anticipation strategies.

The race for Metaverse territories

As for the metaverse, whose experts estimate a global market of around USD 82 billion in 2023, and growth to USD 936 billion by 2030, it has not escaped the power strategies of the major state players, China and the United States.

In October 2023, Microsoft completed its merger with Activision-Blizzard, the largest American video games publisher, for the sum of $69 billion, which suddenly positions Microsoft as the second world leader in video games after the Chinese company Tencent. Immersive and interactive platforms use the full range of high technology in terms of computing power and networking. We can therefore speak of an economic war around these technologies, particularly in the video game sector, given the influence it has on individuals throughout the world.

More or less at the same time, in June 2023 to be precise, the three major state-owned mobile operators, China Mobile, China Unicom and China Telecom, established an industrial alliance for metaverse, with an initial partnership between 24 operators and companies all along the industrial chain, including Huawei, Xiaomi and companies specialising in the development of artificial intelligence. The Chinese metaverse industry is dedicated more specifically to the development of mixed reality (MR), immersive experiences in which digital and physical objects interact. Above all, this approach supports the real economy, the acceleration of work in manufacturing and industrial design. As a result, Chinese metaverse industries and technologies are in a frantic race to register trademarks and patents. In September 2023, five Chinese ministries published a three-year Joint Action Plan for the Innovative Development of the Metaverse Industry 2023-2025. It focuses on promoting three to five globally influential companies to form a metaverse industry ecosystem, which will carry this sector forward as a growth pole of the digital economy and a global technological powerhouse.[7]

Here too, China and the United States are vying with each other, both in terms of ingenuity and strategy, to occupy the leading position in the metaverse ecosystem, and it is the companies (state-owned or linked in one way or another to the state) that are the main players, engaging in ultra-competition.

Quantum: who will calculate the fastest?

In addition to the scientific advances, they will make possible, quantum technologies represent a direct threat to current cryptographic security systems. In other words, the security of financial transactions, communications and critical infrastructures. Quantum and post-quantum encryption have therefore become major strategic issues in the technological rivalry between the West and China. While quantum computers promise to revolutionise cryptography by rendering some of today's systems obsolete, the race to master these technologies is dividing the world's powers.[8]

As early as 2021, IBM was selling a quantum computer to a German research institute, and a second was sold to a medical research centre in the United States in 2023[9]. Although a reality, these two sales are more in the realm of IBM R&D than operational applications. However, the process has begun and is moving very quickly.

In December 2024, Google unveiled 'Willow', a superconducting chip made up of 105 qubits. In less than five minutes, this chip performed a calculation that would take 10 septillion years (i.e. year1025) for one of today's most powerful supercomputers[10] . At the same time, China unveiled Tianyan-504[11], a quantum computer with unprecedented computing power. It is based on a superconducting chip called Xiaohong-504, capable, as its name suggests, of handling 504 qubits. The characteristics of this chip, such as the lifetime of the qubits, the fidelity of the logic gates and the depth of the quantum circuit, are comparable to those of leading international platforms, notably those developed by IBM or Willow (Google).

However, all of them still have to develop algorithms that are inaccessible to conventional computers, while still being useful for concrete problems. The first to develop concrete applications of quantum computing will be the master of the world for as long as it takes its competitors to catch up technologically... In short, a more than strategic objective.

Conclusion

Whatever the technology or technological field concerned, two major players are battling it out in an unprecedented race for innovation: the United States and China. In this confrontation, no holds are barred, and the weapon of information is being widely used. Need we remind you of US President Donald Trump's declaration that the United States would take a 50% stake in TikTok? Or China's announcement of Deepseek?

Information is therefore at the heart of both countries' power strategies. And this information has to be sought where it can be found by the technological means at our disposal, artificial intelligence in particular. The economic aspects of power are systematically present: R&D budgets, mergers and acquisitions, strategic development plans, etc. Consequently, technological development (artificial intelligence, quantum computing, metaverse, chips, etc.) is inseparable from a desire for economic hegemony over its geopolitical competitors.

The geo-economy of technological developments is shaping the world of tomorrow. In this world of the near future, many seem destined to remain on the sidelines and submit, with no other alternative, to the most powerful. Today, Russia, the European Union, India and others are no more than secondary players who are doing their best to put on a brave face, but who in reality have already lost the technological battle. Will an artificial intelligence summit like the one held in Paris in February 2025, at which the French President called on Europeans to take a leap forward to catch up in artificial intelligence, suggesting that we act in *Notre-Dame mode*, in reference to the reconstruction of the famous Paris cathedral be enough? It's all very puzzling, and looks like a last stand... ■

1 Singer P.W. & Cole A., Ghost Fleet: a novel of the next world war, Eamon Dolan/Mariner Books, May 24, 2016
2 https://defenseidentity.fr/2020/11/14/ghost-fleet-la-somme-de-toutes-les-peurs/
3 Spatola N., L'intelligence artificielle. De la révolution technologique à la révolution sociale, L'Opportune, 2018, p.39.
4 https://www.agenzianova.com/fr/news/aerospace-china-place-l%27intelligence-artificielle-dans-le-contr%C3%B4le-total-d%27un-satellite-pendant-24-heures/
5 https://www.cnetfrance.fr/news/la-chine-a-laisse-une-ia-controler-un-satellite-39957400.htm
6 https://www.lemonde.fr/pixels/article/2025/01/28/deepseek-la-reponse-chinoise-a-chatgpt-expliquee-en-six-questions_6520488_4408996.html
7 See Sanchez, Maria José, "Brève histoire des politiques su Métavers en Chine", IRIS - Asia Focus, n°217, April 2024, 16p.
8 Le Coguic, Erwan, "Guerre économique et informatique quantique", École de Guerre Economique, 20 January 2025. https://www.ege.fr/infoguerre/guerre-economique-et-informatique-quantique
9 Branco, Adrian, "For IBM, the quantum revolution could begin as early as 2024", 10 April 2023, https://www.01net.com/actualites/pour-ibm-la-revolution-quantique-pourrait-commencer-des-2024.html
10 Guerrini, Yannick, "Google unveils Willow, its super-powerful quantum chip: 10 septillion years of computing reduced to 5 minutes", 9 December 2024, https://www.01net.com/actualites/google-devoile-willow-puce-quantique-surpuissante-10-septillions-annees-calcul-reduites-5-minutes.html
11 Guerrini, Yannick, "504 qubits: China beats a new record in quantum computing and defies the West", 12 December 2024, https://www.01net.com/actualites/504-qubits-chine-bat-nouveau-record-informatique-quantique-defie-occident.html

The Invisible Battlefield: How AI, Quantum Computing and the Metaverse are Reshaping Espionage and Warfare.

Author: **Raj Meghani**

In the shadowy world of international espionage and cyber warfare, a new arms race is underway. The weapons? Artificial Intelligence, Quantum Computing, and the Metaverse. These cutting-edge technologies are not just transforming our daily lives; they're revolutionising the way nations spy, sabotage, and wage war in the 21st century.

At the forefront of this revolution is Artificial Intelligence. State-sponsored hackers are leveraging AI to create more sophisticated and elusive cyber-attacks. Machine learning algorithms can now analyse vast amounts of data to identify vulnerabilities in enemy systems, automate the creation of malware, and even mimic human behaviour to bypass security measures. AI-powered bots can flood social media with disinformation at an unprecedented scale, swaying public opinion and destabilising democracies from within.

But the true game-changer in cyber espionage may be Quantum Computing. As these incredibly powerful machines become a reality, they threaten to render current encryption methods obsolete. A quantum computer could potentially break the most advanced encryption in mere hours, exposing sensitive government communications, military secrets, and critical infrastructure to foreign adversaries. The nation that achieves "quantum supremacy" first could gain an insurmountable advantage in the global intelligence game.

The Metaverse, which many would argue is still in its infancy, presents new frontiers for espionage and conflict. As more of our lives, work, and social interactions move into virtual spaces, these digital realms become rich targets for intelligence gathering. State actors could create elaborate fake identities and virtual environments to lure and manipulate targets, extracting sensitive information without ever setting foot in the physical world. The lines between reality and deception blur even further in this new landscape.

One of the most concerning developments is the convergence of these technologies. Imagine AI-powered autonomous agents operating in a

quantum-secured Metaverse, carrying out complex espionage operations with minimal human oversight. These digital spies could infiltrate virtual boardrooms, government meetings, and military simulations, gathering intelligence and potentially even influencing decisions in real-time.

The implications for privacy in this new paradigm are staggering.

As AI systems become more adept at analysing human behaviour, and Quantum Computers make data encryption less reliable, the concept of personal privacy may become obsolete. We are moving more and more towards an environment where our behaviours are constantly monitored, analysed, and influenced by intelligent machines.

State surveillance could become ubiquitous and nearly undetectable, monitoring not just our digital communications but our virtual lives and even our thought patterns as we interact with AI-driven interfaces. State-sponsored hackers armed with quantum capabilities could engage in espionage on an unprecedented scale, stealing trade secrets, military intelligence, and personal data with impunity.

The ethical concerns in this new era of digital espionage are also taxing. The use of AI in intelligence gathering raises questions about accountability and human rights. Who is responsible when an AI-driven operation leads to civilian casualties or diplomatic incidents? How can we ensure that AI systems don't perpetuate biases or make critical decisions based on flawed data? The potential for abuse is enormous, and international laws and norms have yet to catch up with these technological realities.

The rise of these technologies is also blurring the lines between peacetime and war. Cyber-attacks powered by AI and Quantum Computing can cause massive disruptions to a nation's infrastructure, economy, and society without a single shot being fired. The Metaverse could become a new theatre of war, where battles are fought over virtual resources and territory with real-world consequences. This ambiguity makes deterrence and conflict resolution far more challenging in the digital age.

Moreover, the democratisation of these technologies means that non-state actors, including terrorist groups and criminal organisations, can now access capabilities once reserved for world superpowers. A small team of skilled hackers armed with AI tools and quantum resources could potentially wreak havoc on a global scale, challenging traditional notions of national security and power dynamics.

BIO

Internationally recognised thought leader and cybersecurity influencer, Raj Meghani is the Co-Founder & Chief Marketing Officer at BlockAPT. A leading edge, highly acclaimed, innovative cybersecurity business, empowering organisations with a centrally managed, command and control single platform experience. Passionate about turning the complex into something simple in cybersecurity, technology and digital transformation, Raj has over 25+ years' experience in FTSE 100/250 to high growth ventures helping businesses across financial services, IT and professional services with their business strategy, digital transformation, growth and retention plans. She's esteemed as a successful brand builder, business growth hacker and judges for 2 cybersecurity awards. Her unique expertise in scaling start-ups and disrupting markets with new tech has earned her recognition as a "One in a Million" female founder by The Entrepreneur's Network and placed her in the Top 44 "Cyber Power Women" by Top Cyber News Magazine. Raj is also Non-Executive Director on the Board of Money Matters Community Bank.
LinkedIn - https://www.linkedin.com/in/raj-meghani/
Twitter: https://twitter.com/blockapt
Company website: https://www.blockapt.com

As nations race to harness these technologies for espionage and warfare, an arms control crisis looms. How can we verify compliance with treaties limiting AI weapons or quantum capabilities when these technologies are largely invisible and rapidly evolving? The risk of miscalculation and escalation in this new landscape is frighteningly high.

The human factor in intelligence work is also being transformed. As AI systems become more capable of analysing intelligence and predicting outcomes, the role of human analysts and decision-makers may diminish. In my eyes, this raises concerns about maintaining human judgment and intuition in critical national security decisions, especially as AI becomes cloudier and more difficult for humans to understand.

Education and workforce development in the intelligence community must evolve rapidly to keep pace with these changes. Nations will need to cultivate a new generation of cyber warriors, quantum interpreters, and virtual world operatives to maintain their competitive edge in this high-tech espionage landscape.

Despite these challenges, these same technologies also offer new tools for **counter-espionage and defence**. AI can be used to detect and respond to cyber-attacks in real-time, quantum secure technologies could create unbreakable codes, and the Metaverse might provide secure virtual spaces for sensitive communications and training simulations.

As we stand on the brink of this new era in espionage and warfare, the stakes could not be higher.

The nations and alliances that master the integration of AI, Quantum Computing, and the Metaverse will likely dominate the global stage in the coming decades.

However, with this power comes great responsibility. The international community must work together and partner with organisations and institutions to establish norms, treaties, and ethical guidelines for the use of these technologies in espionage and conflict.

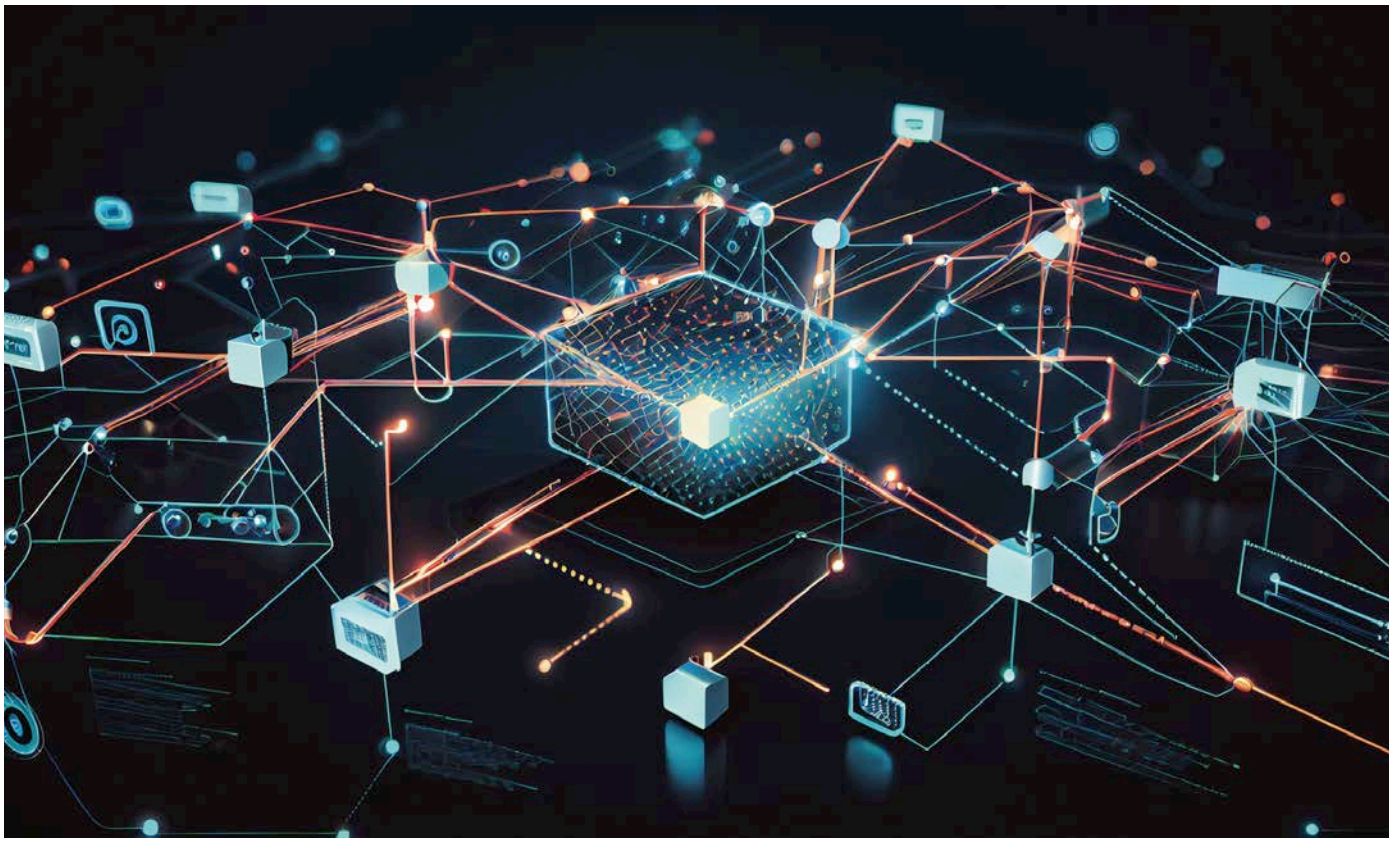
The invisible battlefield of the 21st century is taking shape around us, fought with algorithms, qubits, and avatars rather than the bullets and bombs we are used to. Players on this field will need to carefully navigate this new reality, remain vigilant and be open to being adaptive in both the physical and digital worlds.

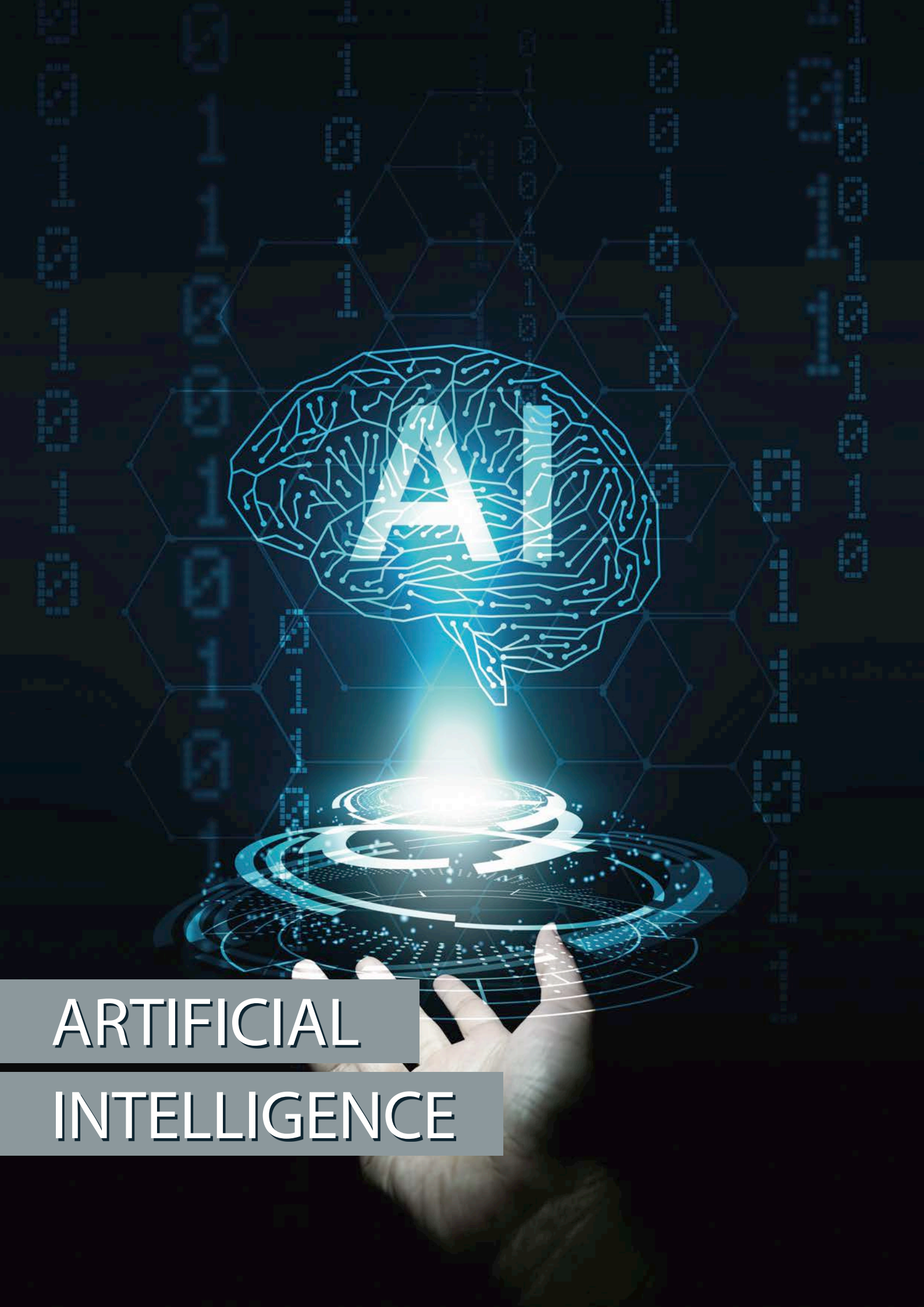
The future of global power and peace may well depend on how we manage these powerful, transformative technologies.

In this brave new world of digital espionage and virtual warfare, we stand at a crossroads.

Will we harness these formidable technologies to level the playing field peacefully in a controlled manner or will we unleash a state of uncontrollable conflict that may ultimately spell the end of privacy, autonomy, and perhaps even humanity as we know it?

Time will tell but the one thing which is a fact is that the clock is ticking… ■



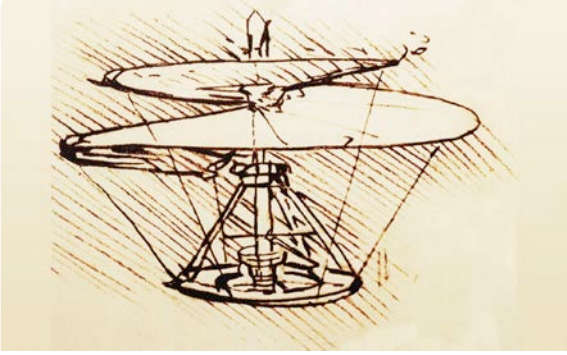


ARTIFICIAL INTELLIGENCE

AI and a Sign of the Times.

Author: Mika Lauhde

Leonardo da Vinci was inventing the helicopter, airplane and many other futures looking technical devices, but they were not really flying (literally) and there was good reason. The technology and financial ecosystem around them were missing. Society would not have been ready for such innovation and Leonardo would have most probably been sued instantly due to the witchcraft like Galileo in 1633.

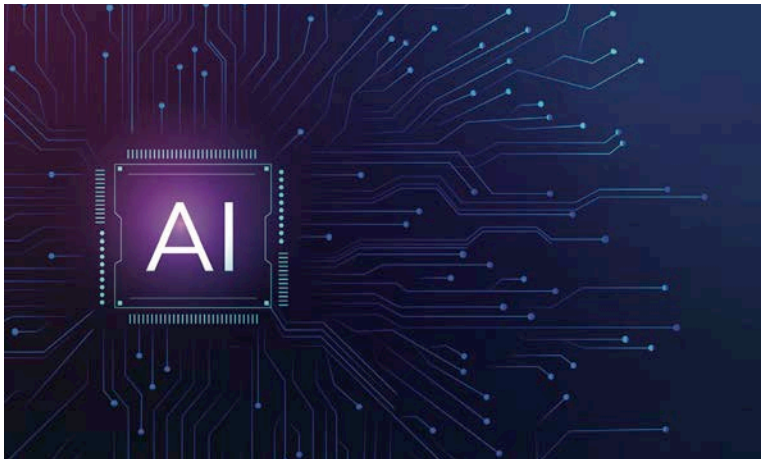


Today, we are getting super excited about every technology step which the technology ecosystem around us is bringing: Blockchain, Zero Trust, Quantum Computing, etc. This is due to the marketing and social pressure which will create a very short hype Similar in kind to the social media hype itself.

However, we often forget how society, ecosystems etc. are being impacted by the new technology as it is accepted, as it behaves and the way it is perceived.

So, let's look at the technology called AI (Artificial Intelligence) and how the current environment where it is landing could be impact how it will be appearing and how it will be used.

Let's start from today's situation. Today, 85% of the browsers used in EU are provided by two vendors. 99% of mobile operating systems are provided by two vendors, 90% of desktop operating systems are provided by two vendors, 90% of the global search is done by one company search engine etc.



So, what can be perhaps predicted, is that after few years, you can could be selecting perhaps from 2 different company's who's AI you will be using in your daily life and who's LLM (Large Language Model) you will be contributing to. Perhaps there will be third provider, but it might be totally unacceptable to use it, it is not coming from a "like minded country" and is therefore very dangerous to use.

For military purposes, there will be some specific national vendors, who are working with their own AI and LLM but naturally they will be not available for large audiences. Perhaps some other small scale AI's are also existing, for special purposes, and are serving some smaller user groups, which cannot use the big companies due to the neutrality, security, privacy, etc. reasons

The other prediction is what can be done now with the quality of AI generated content. So far, all the AI systems have been using "human generated data" to build their LLM systems. But soon, we can be sure, that ever more powerful AI systems will be harvesting all available "human generated" data. At the same time, AI productivity to generate own data will be far beyond a humans' contribution, AI will be moving to eat their own data. One engineer described this process with words "Shit in, shit out". With other words, quantity of the data will be increasing, but the quality will be lower and the human touch will finally disappear. Or to be precise, the new human touch is defined by AI. So, do we need AI to find good data and separate that from bulk data?

The third prediction on what can easily be done now is also related to innovation. The EU was super successful with GDPR but with the EU AI Act regulation, the target is regulating innovation, not so much the result. It is not the same issue. When looking at the world today, super innovations in

BIO

Mika Lauhde is working at the International Committee of Red Cross as Head of Technology in Delegation of Cyberspace and Global Cyber Hub leading ICRC's R&D for cyber and telecom solution development; having +30 years of experience working in Cyber security and privacy field, he had lead world biggest telecom companies security functions, worked in EU and outside EU in several key positions defining the future of directives, regulations, strategies of the critical cyber and privacy functions. As Director of Business Security and Continuity, Mika was in charge of global cyber security efforts at Nokia Telecommunication. Among other things, he handled cyber-related government relations and Nokia's crisis management, as well as security related to handsets and other manufacturing operations worldwide. As Global Vice President of Cyber Security and Privacy at Huawei Technologies, Mika Lauhde was advising the company's top executives on policy, law, regulations, technology, and broad cyber security trends and was leading governmental, and other stakeholder relations world wide. Before joining Huawei, he was VP of Government Relations and Business Development with SSH Communications Security (inventor of SSH protocol), where he advised governments and other stakeholders on security and privacy issues including critical infrastructure protection, compliance, software assurance, and risk and identity management globally. Mika served for 11 years as a Member of ENISA (the European Network and Information Security Agency) and advises Europol on cyber security and privacy. He served as a Member of the British government's critical infrastructure protection group CPNI from 2005 to 2009. He is currently a Senior Fellow at the Maastricht University Faculty of Law's Centre of Data Protection and Cyber Security and a Fellow at the United Kingdom's Institution of Engineering and Technology (FIET) and working in International Committee of Red Cross as Head of Technology in Delegation of Cyberspace and Global Cyber Hub leading ICRC's R&D for cyber and telecom solution development.

AI area are happening by big giants, military and criminals. The common denominator with all of these is, that they do not follow the rules and regulations. Either they are creating those, or they will and can ignore those. These will define real *"de facto"* regulation for the years to come. Until regulatory frameworks start to be really advanced, AI manufacturers as well users won't be able to make sense of what they should and should not do.

Last but not least is the approvals. There can be a long road for SME (Small and Mediums Enterprises) companies to cross the desert for success, given approvals for AI will have a huge fragmentation impact. So far, we have had rather unified approval schemes between Europe, US and China. EU with EC ("European Compliance"), US with FCC "Federal Communications Commission" and Chinese MIIT (Ministry of Industry and Information Technology). However now, thanks to the ever-increasing geo polarization and high-tech war, we might be losing this, not perfect but still working, relationship as well.

EU has already put the EU AI Act in place, but no real existing and credible AI shipset nor SW (Software) house have invested much financing to put to the challenge. The US has very little regulation but lot of investments from start-ups and the few significant giants in this area.

And China, they did the first AI regulation before EU! But they did that differently. 2021 China was regulating recommendations from algorithms, 2022 rules for deep synthesis and 2023 rules for generative AI. All these are being aimed to propel China into being the future global leader in governance and regulation of AI.

And where is this leading us?

China will have its own AI law which needs to be followed if you would like (and can/is allowed/is permitted etc.) to ship your AI to China, where you will be having have lots of competition, investment, local demand, and a desire to export AI to all countries which allowing that. The EU will have their own regulation, which will allow usage in the EU area, but lack not support in relation to the R&D or putting real money for development. Finally, the US will still have less regulation but entering into the US market is commercially very difficult for other countries due to the big giants protecting their market, national security and import restrictions.

So, coming back to Leonardo da Vinci. We do not seem to have a society which would like to welcome AI with open hands, but societies which would like to burn all the wrong thinkers with their AI and just leave their own AI which will be representing their ethics, values, and views on what that local society is seeing as being beneficial.

Not a big leap in the last 500 years. ■

Generative AI, Legitimate or Not? From Knowledge Creation to New Venture Creation.

Authors: Kevin Dumoux, Jun Zhou

March 12, 2025

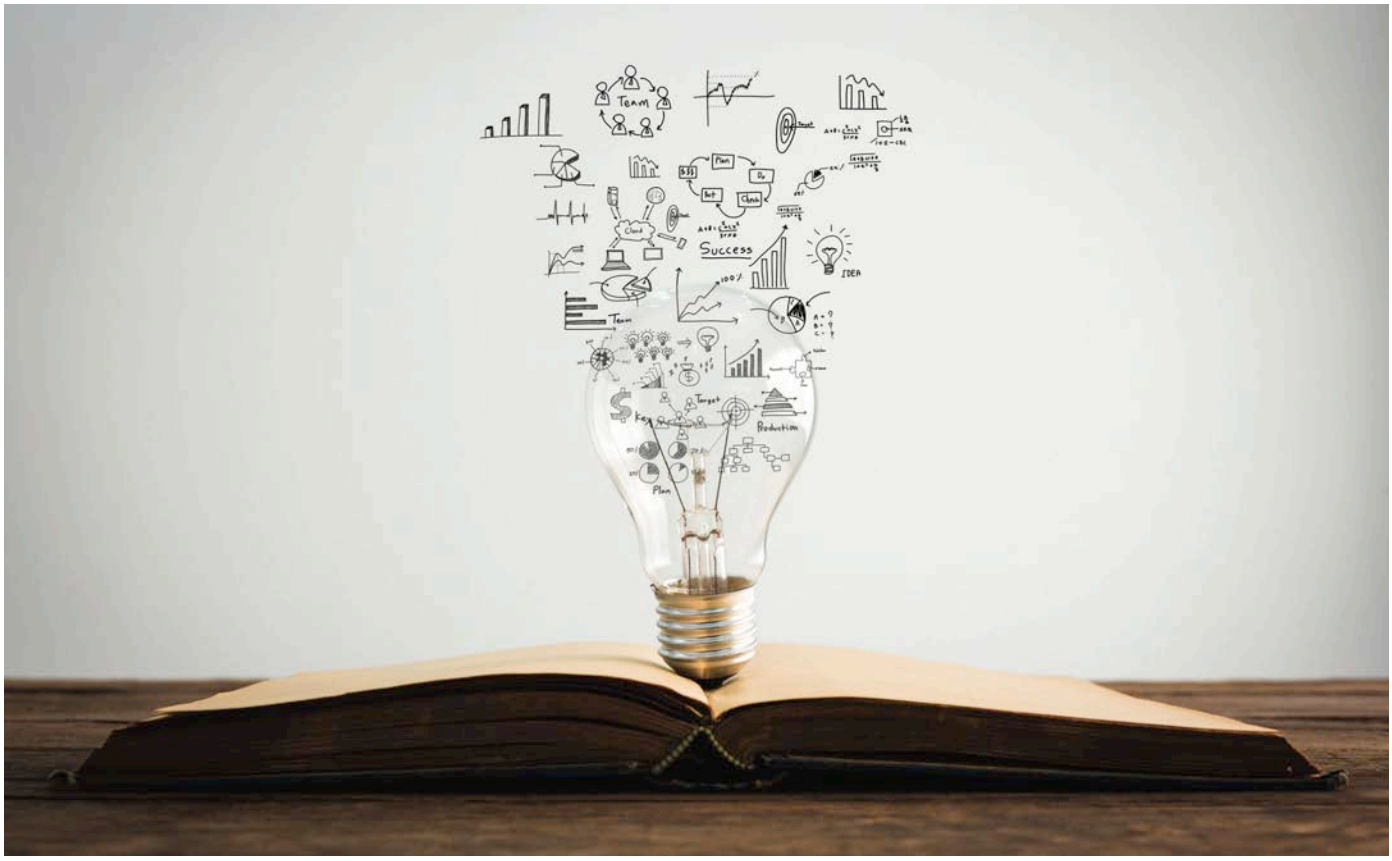
Since the modern era, from East to West, knowledge and entrepreneurship have been two critical pillars contributing to progressive development in society and humanity. There hasn't been a proof of shortcut to advance such development without collective efforts from humans. However, the introduction of generative AI, which transforms how machines interact with and understand humans, has raised questions regarding its ability to outperform humans and replace humans in the creation process.

Generative Artificial Intelligence (Generative AI) refers to a type of artificial intelligence system that is capable of resembling human creation in generating text, images, audio, and videos, given predefined prompts. While many AI systems primarily target tasks such as automation and decision-making, generative AI focuses on generating new content based on existing data. Generative AI models, particularly Large Language Models (LLMs) such as ChatGPT, rely on self-supervised learning techniques to learn the patterns and structures through massive training data and generate coherent new data or predictions that share similar characteristics (Bender et al., 2021). LLMs leverage the underlying statistical framework of language to fundamentally generate a "reasonable continuation" of preceding texts (Wolfram, 2023). Since the booming of generative AI in the 2020s,

generative AI applications have become a great heat across a wide range of fields. Despite its widespread adoption, the legitimacy of generative AI in knowledge creation and new venture creation might appear too good to be true.

Knowledge Creation

Historically, knowledge is created through a blend of practices that are primarily influenced by cultural, philosophical, and religious contexts. In modern society, the knowledge creation heavily relies on research and educational institutions preserving, advancing, and disseminating across various disciplines. Scholars in various disciplines may apply different methods and reasoning logics yet all follow a systematic and rigorous approach to create new knowledge, in know-what and know-how, for advancing our understanding of the world. While it is important to recognize



the promise of generative AI for assisting the entire value chain of knowledge production from synthesis, to creation, to evaluation and translation (Bartunek, Rynes, & Daft, 2001; Kilduff, Mehra, & Dunn, 2011; Van De Ven & Johnson, 2006), it is also critical to not overlook the potential perils of doing so (Grimes et al., 2023).

This goes back to the mechanisms of generative AI in content generation. Despite the models being self-supervised learning techniques, the current tools such as ChatGPT are trained on a massive dataset from a wide range of sources, including newspapers, social media posts, and blogs, posing issues regarding the traceability of the source and its reliability. Moreover, the generating process is often seen as a "black box" that we don't know how these tools come to conclude the new content even though they try to mimic the human creation. It raises questions about whether the mechanism follows any systematic steps of approach or any particular type of reasoning logic. These are the critical issues closely linked with transparency and reliability, the essential elements of being rigorous in academic research. Without transparency and reliability, it is difficult to validate the steps in data collection and data analysis, thus, hardly validating the process of knowledge creation and its results.

Another issue with generative AI models is that, unlike human's cognitive ability to understand and interpret the contexts of the situation or phenomena, they are blind to the personal, social, and cultural circumstances we inhabit (Lindebaum & Fleming, 2023). Human creation always follows reflexivity that we problematize the world in which we live and challenge the taken-for-granted perception of experience (Hibbert & Cunliffe, 2015). The system of generative AI unfortunately is an acontextual algorithm focusing on calculability rather than understanding the situation and its environment. The generative AI models cannot distinguish 'word form' from meaning that they lack the ability to identify the reality and the interpretations of content they are training on (Bender et al, 2021) and, thus, are liable to hallucinate (i.e., making up 'facts') (Kulkarni et al, 2023). The generating results would be merely considered poor knowledge or even deepfakes.

New Venture Creation

In this fast-changing world, competing for speed and efficiency is vital for many businesses' survival and thrive. Generative AI has been a great tool to enhance productivity. For example, adopting generative AI in customer service, results in a 14% increase in the number of chats an agent can successfully resolve per hour (Brynjolfsson et al., 2023). Generative AI can also increase the consultants' performance without a substantial organizational or technological investment, especially in tasks such as writing and programming (Dell'Acqua et al., 2023; Noy and Zhang, 2023; Peng et al., 2023). For new venture creation, many entrepreneurs have also embraced the benefits of generative AI in different applications during their entrepreneurial journey. Some might leverage it as a great tool to have guidance simply as how to make a website, or what strategies to launch social media campaigns,

while others might integrate it into their business offerings such as creating content for their marketing brochures, or supporting customer service as chatbots. With no doubt, generative AI has boosted the speed and efficiency of entrepreneurs launching new projects and creating new businesses. However, despite the considerable promise generative AI has brought to the business world, the peril of the same coin cannot be overlooked.

For entrepreneurs, gaining a competitive advantage means acquiring key resources that are valuable, rare, inimitable, and non-substitutable (Barney, 1991). In that, the differentiations of resources and capabilities play a critical role in enabling new ventures within the same industry to outperform others (Barney, 1991; Dierickx & Cool, 1989). Although generative AI might be helpful to a certain extent to speed up processes for productivity, it is under question whether this benefit enabled by generative AI is considered the key resource in gaining competitive advantages. Generative AI models such as ChatGPT nowadays are commonly introduced to the public and share open access with everyone, thus, the capabilities they offer are not seen as rare, inimitable, or non-substitutable, adding no extra value for gaining competitive advantages. Besides, generative AI fastens the new venture creation process collectively for all entrepreneurs, thus, by relativity, it promotes no unique advantages that could be differentiated from others. Instead, the real alert to entrepreneurs and managers is that the productivity impacts of generative AI are highly uneven. While less-skilled and less-experienced workers can improve significantly, minimal impacts are shown on the productivity of more-experienced or more-skilled workers (Brynjolfsson et al., 2023). From an effective human resources perspective, it is highly questionable to argue that a productivity increase by generative AI indicates a positive outcome.

Another alert for entrepreneurs and organizations is that due to the opacity characteristic of LLMs, generative AI models tend to produce incorrect, but plausible results (hallucinations or confabulations), even including performing math or providing citations (Dell'Acqua et al., 2023). This implies the potential pitfall for applying generative AI models to certain tasks that require high accuracy. As mentioned earlier, the process of LLMs generating results is rather a black box that lacks transparency on the training process and explainability behind predictions, making it difficult for entrepreneurs and managers to grasp errors and verify them. It is irresponsible to assume that generative AI applications in business are reliable and solely advantageous.

Last but not least, overly relying on generative AI makes entrepreneurs and managers path-dependent and hinders their ability to think, reflect, and innovate. Nowadays challenges faced by entrepreneurs are the innovation related to both tangible and intangible assets, the resources and capabilities encompassing management skills, organizational processes, routines, and knowledge. To be innovative in the new venture creation, entrepreneurs need to leverage their imagination, vision, knowledge and experience. They have to go through a reflexivity process to better understand their intrinsic motivation, and assess the market and the value proposition. The convenient fast guidance and strategies from generative AI models cannot be solely depended on for new venture creation, because the true power of creation is still with entrepreneurs.

To Sum Up

The journey of creation, whether in knowledge creation or new venture creation, embodies an understanding

BIO

Graduated from EDHEC Business School (MSc in Finance), Kevin Dumoux is the co-creator of Cercle K2, an international think tank that develops various networking and dialogue platforms. Over the past 12 years, the Institution has brought together 3,000 members across 30 countries. Professionally, Kevin leads a consulting firm specializing in Strategy, business development, and digital transformation. In 2024, he was appointed as a "Conseiller du Commerce Extérieur de la France" (Foreign Trade Advisor for France) by decree of the French Prime Minister.

BIO

Jun Zhou is a PhD researcher in entrepreneurship at ESCP Business School, and member of Cercle K2. With a background in business and innovation, Jun Zhou holds a master degree from HEC Paris and has accumulated eight years of corporate experience, along with two years as an entrepreneur in China.

of oneself and the surroundings, moreover, a continuing reflectivity by interrogating our experiences and questioning our relationships with our social world. Yet generative AI is unable to fully capture the creator's mind, personal experience, and social and cultural backgrounds. In Kant's work, "transcendental grammar" signifies that the universal form of human language has its ground in the universal structure of thought (D'Agostino, 2023). Generative AI, particularly Large Language Models, however, still lacks the ability to articulate and construct ideas in the same way. Lastly, we need to be aware that the fast speed of content generation by generative AI models promotes content homogeneity, in the sense that they cannot keep up with the speed of how human ability responds to radical changes in perception and understanding when crafting authentic, original and creative work, which heavily relies on reflexivity that entails thinking processes with new, unexpected, and unanticipated information.

REFERENCES

Barney, J. B. 1991. Firm resources and sustained competitive advantage. *Journal of Management*, 17: 99-120.

Bartunek, J. M., Rynes, S. L., & Daft, R. L. 2001. Across the great divide: Knowledge creation and transfer between practitioners and academics. *Academy of Management Journal,* 44: 340–355.

Bender EM, Gebru T, McMillan-Major A, et al. (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Virtual Event, Canada, pp.610-623. Association for Computing Machinery.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at Work* (Working Paper 31161). National Bureau of Economic Research. https://doi.org/10.3386/w31161

D'Agostino, P. (2023). Kant's Transcendental Theory of Universal Grammar. *The Cognitive Foundation of the Structure of Language.* Kant Yearbook, 15(1), 1-24. https://doi.org/10.1515/kantyb-2023-0001

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.4573321

Deresiewicz, W. (2023). Why AI Will Never Rival Human Creativity - Predictive mechanisms preclude the originality needed for true art. Retrieved from https://www.persuasion.community/p/why-ai-will-never-rival-human-creativity on 23 June 2023.

Dierickx, I., & Cool, K. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science,* 35: 1504-1511.

Grimes, M., von Krogh, G., Feuerriegel, S., Rink, F., & Gruber, M. (2023). From Scarcity to Abundance: Scholars and Scholarship in an Age of Generative Artificial Intelligence. *Academy of Management Journal,* 66(6), 1617–1624. https://doi.org/10.5465/amj.2023.4006

Hibbert, P., & Cunliffe, A. (2015). Responsible Management: Engaging Moral Reflexive Practice Through Threshold Concepts. *Journal of Business Ethics,* 127(1), 177–188. https://doi.org/10.1007/s10551-013-1993-7

Kilduff, M., Mehra, A., & Dunn, M. B. 2011. From blue sky research to problem solving: A philosophy of science theory of new knowledge production. *Academy of Management Review,* 36: 297–317.

Kulkarni, M., Mantere, S., Vaara, E., van den Broek, E., Pachidi, S., Glaser, V. L., Gehman, J., Petriglieri, G., Lindebaum, D., Cameron, L. D., Rahman, H. A., Islam, G., & Greenwood, M. (2024). The Future of Research in an Artificial Intelligence-Driven World. *Journal of Management Inquiry.* https://doi-org.revproxy.escpeurope.eu/10.1177/10564926231219622

Lindebaum, D., & Fleming, P. (2023). ChatGPT undermines human reflexivity, scientific responsibility and responsible management research. *British Journal of Management.* 35: 566-575. https://doi-org.revproxy.escpeurope.eu/10.1111/1467-8551.12781

Noy, Shakked and Whitney Zhang, *"Experimental evidence on the productivity effects of generative artificial intelligence,"* 2023. Available at SSRN: https: // ssrn. com/ abstract= 4375283 .

Peng, S., E. Kalliamvakou, P. Cihon, and M. Demirer, *"The impact of ai on developer productivity: Evidence from github copilot,"* arXiv preprint arXiv:2302.06590, 2023.

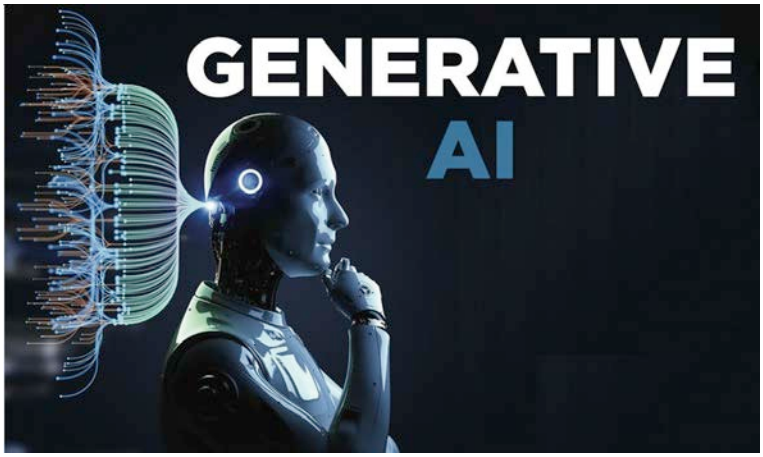
Van De Ven, A. H., & Johnson, P. E. 2006. Knowledge for theory and practice. *Academy of Management Review,* 31: 802–821.

Wolfram S (2023) *What Is ChatGPT Doing ... and Why Does It Work?* Retrieved from https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/ ■

Hallucinatory Artificial Intelligence and Autointoxication.

Author: Christine Dugoin-Clément

In November 2022, the first version of ChatGPT made its appearance in a rather positive atmosphere. This generative Artificial Intelligence (AI) won over a large number of users by translating texts, providing articulate and credible, if not truthful, answers. It was also in February 2022 that Russia launched a massive invasion of Ukraine, which was accompanied by multiple waves of disinformation (Dugoin-Clément 2022) in a variety of content and media, following on from the actions taken in the wake of the war in the Donbass that began in 2014 (Jaitner and Geers 2015, Dugoin-Clément 2019). While these waves of information worried governments and institutions about their impact on public opinion and democratic processes, ChatGPT and, more generally, LLM generative AI, made many professions concerned about their potential disappearance, but also the teaching profession, which feared an exposure to cheating, an impoverishment of thought and, in general, the possibility of weakening discrimination between truthful content and content appearing to be truthful.



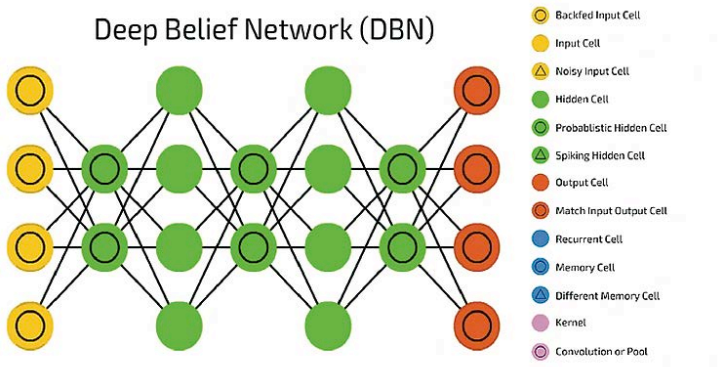
Generative AI, a new technological era?

Before giving birth to ChatGPT, Gemini, Claude, Poe, Bard and other generative AIs, AI underwent numerous developments and pauses in its non-linear growth, constrained in particular by the computing capacity of computers and the amount of data available. The 50s and 60s were marked by the development of the General Problem Solver (Newell, Shaw and Simon 2000) and ELIZA (Weizenbaum 1966) programmes, and it was not until 2015 that AlphaGo (Chen 2016, Granter, Beck and Papke Jr 2017) signalled the return in force of the AI known to the general public during the Go games played against Fan Hui and Lee Sedol. An unsupervised machine learning generative AI, ChatGPT belongs to a class of language models called Generative Pre-trained Transformers (GPTs). It follows on from

BIO

Christine Dugoin-Clément is a researcher at the Risks Chair of IAE Paris-Sorbonne, at the Artificial Intelligence Observatory of Paris 1 Panthéon-Sorbonne, at the research centre of the Saint Cyr Coëtquidan schools (CREC) and at the Research Centre of the National Gendarmerie (CRGN). A former IHEDN auditor, his work specialises in strategies, particularly influence, cyber and AI.

other models such as Deep Belief Networks and Deep Boltzmann Machines (Salakhutdinov and Larochelle 2010) which still suffered from gaps in the generalisation of their activities (Hinton, Osindero and Teh 2006, Pan, Yu et al. 2019).



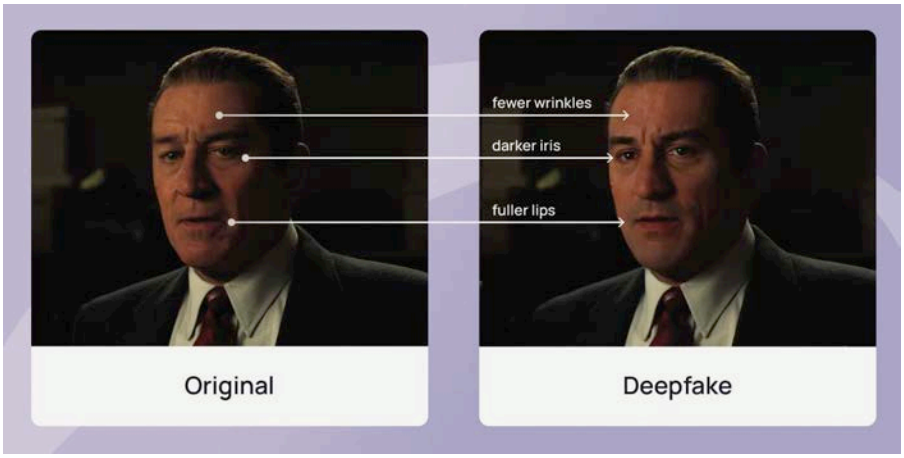
Transformers are based on attention mechanisms (Vaswani, Shazeer et al. 2017) in which the model focuses on different parts of the input sequence while generating the output sequence. This attention allows for more finesse in text generation. However, based on Gans, Adversarial Neural Networks (Goodfellow, Pouget-Abadie et al. 2020), in which two neural networks, a generator and a discriminator, compete to produce content that is impossible to discriminate from authentic products. We are therefore entering a new era of generative AI in which texts ranging from poems (Köbis and Mossink 2021) to political speeches (Bullock and Luengo-Oroz 2019) video with DeepBrain, images with MidJourney or even audio can be produced and difficult to differentiate from content produced by humans.

A number of risks induced by the development of these technologies were quickly highlighted, while, echoing the issues mentioned above. A number of ethical points were put forward, such as the possibility of cheating, counterfeiting and theft of intellectual property (Gillotte 2019) and the issue of personal data (Pucheral, Rallet et al. 2016, Siau and Wang 2020, Fui-Hoon Nah, Zheng et al. 2023). But in light of the increase in misinformation, the malicious use that can be made of deepfakes and content produced by generative AI for the purposes of influence has become a subject in its own right (Whittaker, Kietzmann et al. 2020, Dugoin-Clément 2022, Dugoin-Clément 2022).

Feeding generative AI

In addition to transformers, generative AI, particularly for text, works on a set of underlying concepts and model architectures that are commonly known as Large Language Models (LLMs). These are probabilistic models used to produce text or images. To do this, these models will be pre-trained on massive quantities of data, especially as these models can accept different inputs, for example image and text inputs to generate text outputs, as does ChatGPT (Achiam, Adler et al. 2023). As they are statistical models, these AIs simply imitate human-type responses, but have no understanding or perception of the underlying meaning of their output (Shukla Shubhendu and Vijay 2013), nor do they have any intention to transmit: when an AI starts to produce a text, it has no intention of its purpose; the text is written by probability without any intention.

One of the limitations often highlighted with regard to generative AI models is that they are designed on the basis of historical data, which therefore has a specific cut-off date. (Feuerriegel, Hartmann et al. 2024). In addition, several models do not store information prior to this training, while



others compress content, preventing these algorithms from remembering what was seen during their training (Chiang 2023).

To get around these limitations, recent models have been equipped with a real-time search for information on the web. While this approach solves the problem of obsolescence, it weakens the models in terms of the possibility of hallucinations.

Hallucinations and self-trolling

As we can see, the aim of generative AI is not to be accurate, but to create a feeling of authenticity. Under these conditions, one of the issues is the possible hallucination of systems. Here, hallucination is the phenomenon that leads to the production of content that is absurd or untrue to the given input source (Ji, Lee et al. 2023, Rawte, Sheth and Das 2023). We will then be faced with content that looks realistic, but makes no sense, or is totally unrealistic, fuelling disinformation processes (Dwivedi, Kshetri et al. 2023, Susarla, Gopal et al. 2023). Generative AI is being used more and more widely to produce a variety of inaccurate content, even content citing non-existent sources, and this content is being projected onto the Internet, often in the same movement that gave rise to its creation. To create the content that will satisfy the person who ordered it, the algorithms will search the Internet for data. As a result, the

hallucinations of generative AIs may run the risk of suffering from a form of autointoxication or autotrolling that will become tricky to distinguish in the midst of growing infobesity.

Under these conditions, the proliferation of generative AIs makes it even more difficult for public authorities to combat disinformation (Matasick, Alfonsi and Bellantoni 2020, Pherson and Heuer Jr 2020, Vériter, Bjola and Koops 2020). In addition, discrimination may be made increasingly difficult due to the possibility of AIs becoming self-poisoned, including discrimination by hallucinations, or even by products designed for disinformation purposes or to increase a flow making authentic facts statistically less representative than certain counterfeit products. However, this issue is being taken into account by the digital giants, who are trying to protect their generative AI against the possibility of hallucination. However, this will not prevent the deliberately deleterious practices implemented by actors whose aim is disinformation, whether to provoke a specific action in the real world, to sow chaos or to induce perpetual doubt, thereby fragmenting society into a series of small groups that are unable to reach a critical mass that would enable a major organisation, whether an institution or a government, to take a decision.

Bibliography

Achiam, J., et al. (2023). "Gpt-4 technical report." arXiv preprint arXiv:2303.08774.

Bullock, J. and M. Luengo-Oroz (2019). "Automated speech generation from UN General Assembly statements: Mapping risks in AI generated texts." arXiv preprint arXiv:1906.01946.

Chen, J. X. (2016). "The evolution of computing: AlphaGo." Computing in Science & 18(4): 4-7.

Chiang, T. (2023). "ChatGPT is a blurry JPEG of the web." The New 9: 2023.

Dugoin-Clément, C. (2022). Artificial Intelligence (AI) and deepfake in disinformation: a tool under-exploited? MENACIS 2022.

Dugoin-Clément, C. (2022). DeepFakes: what immunity for law enforcement? (work in progress). AIM Conference.

Dugoin-Clément, C. (2022). Infleunce et manipulation, des conflits armées modernes en ukraine aux guerres économiques.

Dugoin-Clément, C. (2019). «The Use of Cyber Activities as a Weapon: The Empirical Case of Ukraine.» The journal of intelligence and cyber security.

Dwivedi, Y. K., et al (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy." International Journal of Information 71: 102642.

Feuerriegel, S., et al. (2024). "Generative ai." Business & Information Systems 66(1): 111-126.

Fui-Hoon Nah, F., et al. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration, Taylor & Francis. 25: 277-304.

Gillotte, J. L. (2019). "Copyright infringement in ai-generated artworks." UC Davis L. Rev53: 2655.

Goodfellow, I., et al. (2020). "Generative adversarial networks." Communications of the 63(11): 139-144.

Granter, S. R., et al. (2017). "AlphaGo, deep learning, and the future of the human microscopist." Archives of pathology & laboratory 141(5): 619-621.

Hinton, G. E., et al. (2006). "A fast learning algorithm for deep belief nets." Neural 18(7): 1527-1554.

Jaitner, M. and K. Geers (2015). "Russian information warfare: Lessons from Ukraine." Cyber war in perspective: Russian aggression against Ukraine.: 87-94.

Ji, Z., et al. (2023). "Survey of hallucination in natural language generation." ACM Computing 55(12): 1-38.

Köbis, N. and L. D. Mossink (2021). "Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry." Computers in human 114: 106553.

Matasick, C., et al. (2020). "Public governance measures in the face of misinformation."

Newell, A., et al. (2000) "Report on a general problem solving program." Artificial intelligence: critical 2: 69-87.

Pan, Z., et al. (2019). "Recent progress on generative adversarial networks (GANs): A survey." Ieee 7: 36322-36333.

Pherson, R. H. and R. J. Heuer Jr (2020). Structured analytic techniques for intelligence analysis, Cq Press.

Pucheral, P., et al. (2016). "Privacy by design: a false good solution to the personal data protection issues raised by Open data and connected objects?" Legicom(1): 89-99.

Rawte, V., et al. (2023). "A survey of hallucination in large foundation models." arXiv preprint arXiv:2309.05922.

Salakhutdinov, R. and H. Larochelle (2010). Efficient learning of deep Boltzmann machines. Proceedings of the thirteenth international conference on artificial intelligence and statistics, JMLR Workshop and Conference Proceedings.

Shukla Shubhendu, S. and J. Vijay (2013). "Applicability of artificial intelligence in different fields of life." International Journal of Scientific Engineering and 1(1): 28-35.

Siau, K. and W. Wang (2020). "Artificial intelligence (AI) ethics: ethics of AI and ethical AI." Journal of Database Management (JDM31(2): 74-87.

Susarla, A., et al. (2023). "The Janus effect of generative AI: Charting the path for responsible conduct of scholarly activities in information systems." Information Systems 34(2): 399-408.

Vaswani, A., et al. (2017). "Attention is all you need." Advances in neural information processing systems .30

Vériter, S. L., et al. (2020). "Tackling COVID-19 disinformation: Internal and external challenges for the European Union." The Hague Journal of 15(4): 569-582.

Weizenbaum, J. (1966). "ELIZA-a computer program for the study of natural language communication between man and machine." Communications of the 9(1): 36-45.

Whittaker, L., et al. (2020). "All around me are synthetic faces": the mad world of AI-generated media." IT 22(5): 90-99. ■

Changing Security: the AI effect.

Authors: Gérald Vernez, Diego Kuonen

** A first version of this article entitled "AI and its consequences for the security professions" was published in January 2024 in the journal of the Police Academy at Savatan in Switzerland.*

The phenomenon

On 30 November 2022, OpenAI released ChatGPT, the first conversational agent based on generative AI[1], free and accessible to anyone with an Internet connection. It was an earthquake that surprised many informed observers and its own inventors. ChatGPT popularised AI and triggered a groundswell and unprecedented proliferation of tools and services using AI-based technological building blocks. There is clearly a before and after ChatGPT.

Since then, the development of competing generative AI algorithms has been supported. Among the best known, OpenAI is now in its versionGPT-4o. Google has designed PaLM for complex natural language processing tasks. Anthropic has developed Claude for aligning with human intentions and reducing undesirable behaviour. The Chinese firm Baidu has developed ERNIE and Meta's model is LLaMA. Recently, the Chinese algorithm DeepSeek-R1 was announced, raising concerns about the competitiveness of today's technology giants. Nvidia, which specialises in graphics cards (GPUs) and AI, and is a leader in high-performance computing and AI chips, has seen its market capitalisation plummet by almost $590 billion. However, investors re-evaluated their positions and the initial volatility quickly subsided. These examples show how rapidly the situation is evolving, but also how unstable it is.

Beyond these developments, it is crucial to understand the basic workings of generative AI algorithms. First and foremost, they are made up of fuel, i.e. data. Their second key component is the brain or algorithms of these LLMs, or Large Language Models[2]. LLMs are trained on vast quantities of data from

BIO

Gérald Vernez studied geology, meteorology and safety policy. He began his career as an engineer in special works and risk management. After joining the Armed Forces General Staff, he prepared and trained the Confederation's crisis staff for the Year 2000, then planned the Armed Forces' command structures and processes and developed the Armed Forces' information operations. He then served as Chief of Staff of the Armed Forces Joint Staff and Deputy Director Project for the development of the first national cyber security strategy. As the Chief of the Armed Forces' delegate for cyber defence, he planned and coordinated the development of this area within the Armed Forces and then the Department of Defence. In 2020, he founded digiVolution, a think tank for strategic action on digital change, and more recently dVCyberGroup SA, a consultancy and operational support company 360° for our digitally changing security society.

which they establish a statistical model enabling them to suggest, with a high probability, which word comes after which other in an answer to a question, or to generate new ideas from the information they possess.

These key elements - data and algorithms - explain why AI is still largely perfectible. All it takes is an error in one of them for the result to be directly affected. "Garbage in! Garbage out! " And this is true right from the early stages of development and training of each of these algorithms. You poison the input, you will have nothing better at the output. Data quality explains why AI systems produce numerous biases and are sometimes even guilty of racism or discrimination, even going so far as to invent facts. This is known as "hallucination". If the question (which is data) is badly posed at the input, the tool will produce an answer inadequate (data) at the output. And if the recipient does not know how to interpret the result or check it, erroneous facts will be taken as truths. These limits are also complicated when politics get involved. With DeepSeek, for example, there is no point in asking a question about the events in Tiananmen Square in 1989, as the algorithm simply refuses to answer it. Quite apart from the controversy over the censorship expressed in this way, the legitimate question is "What other subjects are being deliberately hidden or modified to meet the interests that escape us? " It would be naive to believe that only the Chinese impose restrictions on the circulation of information.

Uses of AI

It would be a monumental mistake to consider AI's teething problems as definitive and to condemn it ad aeternam, or to think that it is limited to generative AI and conversational such as ChatGPT agents and similar systems (such as Copilot or Gemini). AI is in fact already established in many activities and products everyday, summarised by a few application examples in the table below.

The many technologies behind the acronym AI are at the heart of our daily lives, and the development of AI long predates the appearance of ChatGPT. AI is at the centre of digital mutation. As in biology, "mutation" means that it is an irreversible process. Whether we like

Areas of employment	Examples
Machine Learning	- Image recognition, medical diagnostics. - Property price prediction. - Weather forecasts. - Analysis of genomic data - Recommendation systems. - Artistic creation (paintings, music, fashion, architecture, media). - Cyber security.
Automatic Natural Language Processing (NLP)	- Machine translation, text generation. - Social media monitoring, customer services. - Virtual assistants, automatic transcription (dictation). - Content creation, automatic summaries.
Vision	- Video surveillance. - Autonomous vehicles. - Facial recognition. - Medical diagnosis by imaging. - Industrial inspection. - Safety monitoring. - Sports analysis.
Recommendation systems	- Streaming platforms, e-commerce, advertising. - Predictive analysis (e.g. in machine maintenance). - Education professions.
Robotics	- Delivery and construction robots. - Autonomous vehicles. - Industrial automation (production, quality control). - Robotic prostheses, surgical robotics and assistance for operations. - Rescue and disaster relief. - Automated response to cyber attacks.
Expert systems	- Support for clinical decision-making and personalisation of precision medical treatments. - Stock market forecasts. - Detection of financial frauds to insurances. - Agriculture (disease prevention, crop development, crop improvement and selection, pests' detection and treatment). - Managing and optimising energy resources and flows. - Research and development. - Optimising logistics and processes. - Network monitoring and detection of intrusions and cyber attacks.

BIO

Prof. Dr. sc. Diego Kuonen, PhD in Statistics and CStat PStat, founded Statoo Consulting in 2001 and advises companies and authorities in Switzerland and throughout Europe at operational, tactical and strategic levels in the fields of Analytics, Statistics and Data Science, including AI and Machine Learning. With over 24 years of professional practice, he has developed in-depth expertise and a proven track record in these areas. In addition, since 2016 he has been Professor of Data at the Science Geneva School of Economics and Management (GSEM) at the University of Geneva, Switzerland, and Founding Director of new programme. GSEM's Master of Science in Business Analytics Finally, he recognised the need for a data culture in society and co-initiated a call in Switzerland for a national data literacy campaign in 2020. He was also co-initiator and co-author of the Swiss Data Literacy Charter (published in 2024 by the Swiss Academies of Sciences; see akademien-schweiz.ch/fr/themen/culture-scientifique/data-literacy-charta/), to promote a fundamental cultural change in society with regard to data and its use. Since 2021, he has also been a member of the Advisory Board of the digiVolution Foundation.

it or not, with AI society is moving towards ever greater automation[3] and "augmented intelligence."

The benefits of AI are undeniable, but every coin has a second side, and the list of risks and malicious uses is as long as the list of benefits. AI made a remarkable entry into the political arena with the 2016 US presidential election, and is now omnipresent in levers that are the driving force in operations, radicalisation and conspiracy theories in particular. Hardly a day goes by without AI being associated with blunders, ongoing conflicts or cyber attacks. Criminals and fraudsters are often quicker than many businesses and public services to adopt technologies such as AI to boost their activities. AI is also omnipresent in armaments, and debates about killer robots are legion, albeit without any results really tangible to date. And that's not to mention all the legal problems, intellectual property issues and deepfakes. The list is endless. And among the next acceleration factors, it is quantum computing that will boost AI.

The future of AI and superintelligence

AI is neither good nor bad, but it arouses both fascination and fear. The argument between the pros and cons is far from exhausted. A responsible development policy such as that promoted by the Swiss Confederation[4], which places human beings and their control of technology at the heart of processes, is all the more necessary given that the future will see an inexorable increase in the proportion of AI systems. This will be particularly the case in complex areas or tasks with high added value, such as strategic planning or decision-making. While the public increasingly have the impression of conversing and interacting with real intelligence, in reality they are interacting with AI systems whose intelligence resides still in statistics. While may may quadruped robots resemble dogs from a distance, and bipedal metal robots be given humanoid characteristics, they are still frustrating AI systems.

However, there are a number of signs that major progress is imminent. We may be approaching "general artificial intelligence" and its derivatives more quickly than anticipated. The key term is the "technological singularity", the moment when technology has its own consciousness and autonomy, enabling it to improve itself and design new generations of increasingly intelligent technology ever more rapidly. This would be an "intelligence explosion" that would far exceed human intelligence. Ray Kurzweil, the inventor of the concept, has predicted that this will happen in 2045. However, this development and this date are the subject of bitter debate and speculation. Some believe, for example, that recent technological developments would make it possible as early as 2028, or even earlier. The social and political consequences will be major, but there is no consensus.[5]

AI fascinates as much as it frightens, and there is intense activity[6] to try and control its development. But the problem is Kafkaesque. On the one hand, there is the risk of not imposing rules soon enough that are enough powerful and therefore ineffective, leading to a loss of control over AI and its uses. On the other hand, there is the risk of over-legislating and stifling developments at home while others make unimpeded progress and reap major strategic, military and benefits. Fear of the monster or fear for business? How can we regulate something we don't understand and on which the experts themselves disagree? The answer is called the "precautionary principle" and vigilance (hence intelligence), illustrated opposite by a symbol in this field (produced by an AI), provided that everyone plays by the same rules. However, as the US vice-president recently stated at the summit in Paris Alon 10 and 11 February 2025, the new US administration does not want a framework that restricts its industry.[7]

AI and safety

How should security professionals approach the challenges posed by AI? The world is becoming ever more complex and managers can no longer "just do what they did before". There is no simple answer. They need to adapt, and here are five recommendations from digiVolution that the authors of this article are practising and implementing in their consulting and operational support activities.

1 **Information and training** - Whether we like it or not, technology is moving forward. Managers and employees need to be informed and trained in the use of technology, so that they can take informed action in the face of the risks (intended or otherwise) of AI and the opportunities it opens up. Interacting with AI is now a skill essential. As the fuel of AI is data, as are the results produced by AI-based tools it is imperative to promote data literacy. This is an essential condition for an informed society[8]. It must

also be a requirement for all security professions. Every individual must acquire a "Data and AI Literacy for Everyone! Leave No One Behind!".[9]

2 **The key lies in governance** - Decision-makers must at all costs avoid lapsing into technophile angelism or technophobic paranoia. The monitoring of technological developments must be integrated into their strategic risk management so that they can take timely action to control disruptive developments linked to AI or to take advantage of its contributions. The various departments must (and this is not an option) put in place a strategy that enables managers to take advantage of advances in AI and control the consequences that flow from them.

3 **Strong policy data management** - Data is the fuel for decision-making. Every company or organisation needs to identify its data, manage it, protect it and add value to it through a dedicated process. It's not just a question of storing data - protected where necessary - but, within a specific organisation, making it usable for all those who could benefit from it and putting it to work to produce additional knowledge.

4 **Technology integration strategy** - Enlightened use of AI as a "cognitive co-pilot" should enable organisations/companies to generate significant added value and savings in repetitive, low added-value tasks. These savings should enable employees to concentrate on what they do best. With the volume, speed and volatility of information increasing massively, controlling its integrity is taking on major importance. This means investing in tasks to verify the authenticity of information before incorporating it into the decision-making process. Will AI destroy jobs? Let's start by transforming them. It's not AI that's going to take jobs, but the people who work with your AI.

5 **Staying secure and resilient at all costs** - The promises of technology may be immense, but they are far from infallible. So, what about accidents, breakdowns and vulnerabilities? Taking care of your technological affinity, yes, without a doubt, but guarding against any systemic dependency and being ready to manage crisis situations is also key.

These recommendations and the facts set out above illustrate the turning point at which finds society itself. Ignoring this situation is tantamount to falling behind an inevitable development, and weighted risks making it irrecoverable. In an interview at the end of 2023, Peter Brabeck-Letmathe said that "we have lost control of time and change". We invite our readers to reflect on this.

At digiVolution we share this view, which is why we are committed to developing solutions that strengthen Switzerland's digital confidence, resilience and sovereignty. And at the heart of our actions are the principles of anticipation and precaution. ■

---

1 Generative AI" is a broad concept that refers to AI systems trained on the basis of large quantities of data from the physical and virtual world to generate data autonomously (texts, images, sound recordings, videos, simulations, codes, etc.). They are often multimodal, with input and/or output in one or more modalities (text, image, video, etc.).
2 By integrating additional modalities into LLMs, it is possible to create LMMs (Large Multimodal Models); see https://huyenchip.com/2023/10/10/multimodal.html
3 The automation and rationalisation of repetitive tasks involving large quantities of data should take place in a stable situation, where the rules apply today and tomorrow, where the future resembles the past and where no one can break the rules. Many areas of policy or activity are characterised by situations where uncertainty reigns and where the rules cannot be applied as they stand. This is the case in the kitchen, for example, when it comes to responding to the individual wishes of guests or reacting correctly to unexpected emergencies during the preparation of a meal, such as the absence of essential ingredients or utensils. In his book Klick, published in 2021, the German psychologist Gerd Gigerenzer describes this uncertainty by talking about how difficult it would be to play chess "if the king could break the rules on a whim and the queen could leave the board in protest after setting fire to the rooks". (Source: column „Gesucht: Kochroboter/in - Gefunden: Schachroboter/in" [Diego Kuonen, Walliser Bote, 18 November 2021; see www.data-literacy.ch/_files/ugd/09ac11_8c69c58c404349d185646df86a6127b5.pdf). AI is only promising in stable situations. Where there is uncertainty, however, it can provide assistance to human decision-makers.
4 The Confederation's "Code of good practice for human-centred and trustworthy data science and AI"; see www.bfs.admin.ch/bfs/fr/home/dscc/dscc.assetdetail.29325685.html
5 https://arxiv.org/abs/2401.02843
6 These include the International AI Safety Report (https://www.gov.uk/government/publications/international-ai-safety-report-2025), the European Artificial Intelligence Law (https://artificialintelligenceact.eu/fr/), UNIDIR's Artificial Intelligence Policy Portal (https://parispeaceforum.org/fr/projets/artificial-intelligence-policy-portal) and the Framework to Advance AI Governance and Risk Management in National Security (https://ai.gov/wp-content/uploads/2024/10/NSM-Framework-to-Advance-AI-Governance-and-Risk-Management-in-National-Security.pdf).
7 See https://www.politico.eu/article
8 https://fr.data-literacy.ch/
9 " Sans littératie des données, pas de littératie en IA "www.ictjournal.ch/articles/2024-09-06/sans-litteratie-des-donnees-pas-de-litteratie-en-ia

The Evolving Malware Menace: When AI Strikes at the Heart of a Small Business.

Author: Raj Meghani

"I don't understand, how could this have happened?" Jack slumped back in his chair, staring at the ransomware message cynically taunting him from his computer screen.

His business partner, Emma, shook her head in disbelief. "The malware snuck past our security systems undetected and has encrypted all our critical files, design schematics, client orders - you name it. They're demanding 25 bitcoins, over $200,000, to be paid within 48 hours or everything gets corrupted permanently."

Jack read the demand again - $200,000 in cryptocurrency in 48 hours or all their data would be permanently corrupted. As a small manufacturing company, that exorbitant ransom could bankrupt them.

"This is crazy. We had standard cybersecurity protections like firewalls and antivirus in place, the whole nine yards! How did some malware slip in and bring us to our knees like this?"

Emma's expression turned grim. "I think I know...remember that cybersecurity report last quarter? It warned about the alarming rise of AI-powered malware that can dynamically evade detection."

Realisation hit Jack like a brick. "You think this is one of those advanced, adaptive, self-learning viruses that AI cybercriminals are deploying? Ones that can evade traditional defences?"

BIO

Internationally recognised thought leader and cybersecurity influencer, Raj Meghani is the Co-Founder & Chief Marketing Officer at BlockAPT. A leading edge, highly acclaimed, innovative cybersecurity business, empowering organisations with a centrally managed, command and control single platform experience. Passionate about turning the complex into something simple in cybersecurity, technology and digital transformation, Raj has over 25+ years' experience in FTSE 100/250 to high growth ventures helping businesses across financial services, IT and professional services with their business strategy, digital transformation, growth and retention plans. She's esteemed as a successful brand builder, business growth hacker and judges for 2 cybersecurity awards. Her unique expertise in scaling start-ups and disrupting markets with new tech has earned her recognition as a "One in a Million" female founder by The Entrepreneur's Network and placed her in the Top 44 "Cyber Power Women" by Top Cyber News Magazine. Raj is also Non-Executive Director on the Board of Money Matters Community Bank.
LinkedIn - https://www.linkedin.com/in/raj-meghani/
Twitter: https://twitter.com/blockapt
Company website: https://www.blockapt.com

"It would explain how it bypassed our safeguards," Emma said grimly. "The military-grade encryption keeps morphing itself repeatedly to bypass our security and the malware likely probed our system for vulnerabilities using machine learning reconnaissance techniques."

The true ingenuity and evolution of their cyber adversary weighed heavily on Jack. This was no ordinary malware or virus - it had thought for itself, learned, and adapted to infiltrate their networks and the company.

"What do we do?" Jack ran his hands through his hair anxiously. "Paying ransom demands like these only funds future attacks and emboldens the criminals further. But if we don't pay..."

Emma shook her head sombrely and looked at him with a mix of determination and regret. "We'd likely lose everything - years of proprietary data, product designs, customer records. We'd be forced into a devastating rebuild from scratch."

Looking into their cyber insurance policy yielded no relief. The fine print excluded "ransomware or malicious software attacks utilising artificial intelligence capabilities" from coverage.

Emma continued gravely, "In hindsight, we clearly should have implemented more robust AI cybersecurity monitoring of cyber threats. Secured our networks with behavioural analytics that use defensive AI to detect these advanced threats in real-time."

Jack sighed heavily. "We skimped on upgrading to behavioral analytics that use defensive AI algorithms to monitor networks and flag anomalies."

"Not to mention sorely lacking proactive vulnerability management," Emma added ruefully. "AI malware is brilliant at sniffing out obscure security flaws surgically and exploiting them. If we'd continuously patched and updated everything religiously, it may never have found a way in."

A humourless laugh escaped his lips. "Talk about being a few steps behind. The crooks had ruthless AI fighting for them, while we feebly guarded the fort with checklists software scans, signatures and rules."

The stark truth settled over them - their limited security stack was simply outmatched and outwitted by offensive AI systems meticulously designed to adapt, infiltrate, and proliferate. It ruthlessly leveraged their vulnerabilities against them.

"We have an agonising decision to make," Emma said finally. "Cripple the business by draining our funds to pay the ransomware or dig our own grave by refusing – we are destroyed either way."

An uneasy silence settled over them, the gravity of their missed opportunities to bolster their cyber defenses against AI threats weighing heavy.

In the end, their small business was simply outmatched and outsmarted by systems using AI in unprecedentedly malicious ways.

After an agonisng debate, they settled on not paying, hoping to rebuild while heavily investing in advanced AI cybersecurity countermeasures. AI had won this battle but had not won the war yet.

Rebuilding from rubble was grueling, but they emerged stronger - prioritising AI-driven security monitoring, continuous vulnerability assessments, and developing ransomware resilience capabilities.

Moral of the Story:

Their devastating encounter with AI-powered malware was a harsh wake-up call - one that forced them to embrace advanced cybersecurity measures or risk obsolescence.

Businesses and individuals should adopt a new mantra – **proactive cyber resilience**. The only way to neutralise AI threats is to fight fire with fire - combat advanced AI with advanced AI defences.

Those who turn a blind eye to AI's darker potentials are doomed to fall prey to them. ■

The Role of Intelligence in a Cybersecurity Strategy.

Author: Nicola Sotira

Cybersecurity has become a crucial component of protecting information and technology infrastructure in the digital age. As cyber threats increase, the need for advanced cybersecurity strategies is more pressing. Intelligence, or the collection, analysis, and interpretation of information, plays a central role in formulating and implementing these strategies.

Cyber threats are constantly evolving, becoming increasingly sophisticated and difficult to detect. Attacks can come from various sources, including hackers, organized criminal groups, and even government actors. These attacks can have serious consequences, such as theft of sensitive data, operational disruptions, and reputational damage.

Purpose and Components of Intelligence in Cybersecurity

Intelligence in cybersecurity refers to collecting, analyzing, and using information to anticipate, identify, and mitigate cyber threats. The goal is to provide security professionals with a thorough understanding of the threat landscape, enabling them to make informed decisions and implement effective preventive measures. Intelligence in cybersecurity can be broken down into several key components, each with a specific role in protecting digital assets.

Information Gathering

Information gathering is the first step in the intelligence process. This step involves gathering data from a variety of sources, including:

- Open sources: Publicly available information, such as news stories, blogs, forums, and social media.

- Closed sources: Data obtained from private networks or collaboration with other organizations.
- Dark web: Hidden forums and marketplaces where information about vulnerabilities and attacks is exchanged.
- Threat intelligence feeds Services that provide up-to-date data on known and emerging threats.

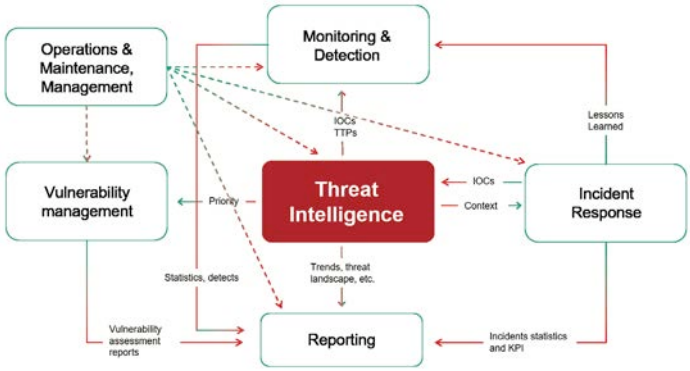
BIO

Head of CERT and security operation in Poste Italiane, Nicola Sotira works in information security and network with over twenty years of experience gained in international environments. He was involved in encryption design and network security, working in complex infrastructures like mobile and 3G networks. He has collaborated with several computer industry magazines as a journalist, contributing to disseminating security and legal and technical aspects. He has taught security at the Master in Network Security of Sapienza University and LUISS since 2005. Member of the Association for Computing Machinery (ACM) since 2004 and promoter of technological innovation, collaborates with several start-ups in Italy and abroad. Member of the board of start-up companies where he participated in developing and designing services in the mobile sector.

Information Analysis

Once collected, information must be analyzed to identify patterns, trends, and potential threats. This phase uses advanced data analysis techniques and machine learning tools to extract useful information from the raw data.

- Pattern analysis: Identification of recurring patterns in suspicious activities.
- Data correlation: Linking different events to identify complex attacks.
- Threat profiling: involves creating detailed profiles of threat actors, including their techniques, tactics, and procedures (TTP).



Contextualization of Threats

Intelligence must be contextualized to be useful. This means understanding how a specific threat may affect the organization and which assets are most at risk. This process involves:

- Impact assessment: Analysis of the potential consequences of an attack.
- Threat prioritization: Determining the most critical threats based on their likelihood and potential impact.
- Asset mapping: Identification of the organization's critical assets and their vulnerability.

Communication and reporting

Intelligence must be communicated clearly and effectively to decision-makers. This involves creating detailed reports and regular updates and responding quickly to requests for information.

- Intelligence reports: Documents that summarize key threats and recommend specific actions.
- Security Dashboards: Visual tools that provide a real-time overview of security status.
- Security Briefings: Regular meetings with organizational leaders to discuss threats and responses.

Applications of Intelligence in Cybersecurity

Intelligence can be applied in several areas of cybersecurity to improve the protection of digital assets and the organization's overall resilience.

Threat Detection

Intelligence is critical for early threat detection. Analyzing real-time data and identifying anomalous behavior can detect suspicious activities before they become security incidents.

© Freepik.com

- Intrusion detection systems: Use intelligence to identify and report anomalous activity.
- Continuous monitoring: Constant surveillance of networks and systems to detect emerging threats.

Incident Response

Intelligence is essential for rapid and effective response to security incidents. By providing detailed information about the nature and extent of the attack, intelligence enables security teams to respond in a targeted manner.

- Incident Response Plans: Detailed plans that use intelligence to guide incident response actions.
- Forensic Analysis: Use of intelligence to collect and analyze evidence after an incident.

Threat Prevention

Intelligence can be used to prevent attacks by identifying and mitigating vulnerabilities before they are exploited.

- Patch management: Identification of vulnerabilities and timely deployment of security patches.
- Systems hardening: Implementation of security measures to reduce the attack surface.

Training and Awareness

Intelligence can be used to educate employees about cybersecurity threats and best practices for security.

- Training programs: Training sessions that use real-world case studies to teach employees how to recognize and respond to threats.
- Awareness campaigns: Initiatives to raise awareness of cyber risks and promote safe behaviors.

Compliance and Audits

Intelligence can help organizations meet compliance requirements and prepare for security audits.

- Risk Assessments: Threat analysis to identify areas of risk and implement appropriate controls.
- Documentation: Creation of detailed reports and records to demonstrate compliance with security regulations.

Main advantages

Proactivity

Intelligence enables organizations to be proactive rather than reactive. By identifying threats before they materialize, preventive measures can be implemented, reducing the risk of security incidents.

Reducing Response Time

Intelligence provides real-time information that enables security teams to respond quickly to

incidents. This reduces the time attackers stay in systems and minimizes damage.

Resource Optimization

Using intelligence, organizations can optimize the allocation of security resources. With a clear understanding of the most critical threats, efforts can be focused on the highest-risk areas.

Improved Awareness

Intelligence increases threat awareness within the organization by educating employees and leaders about risks and security best practices.

Decision Support

Intelligence provides decision-makers with the information they need to make informed decisions about their security strategy. This includes prioritizing security investments and adopting new technologies.

Challenges

Despite its many benefits, integrating intelligence into cybersecurity presents some challenges.

Data Complexity

The amount of data collected can be enormous and complex to manage. Advanced tools and specialized skills are needed to analyze and interpret the data properly.

Accuracy of Intelligence

Intelligence must be accurate and timely to be effective. Incomplete or outdated data can lead to incorrect decisions and increase the risk of incidents.

Privacy and Compliance

Data collection and analysis can raise privacy and regulatory compliance concerns. Ensuring intelligence activities comply with data protection laws and regulations is important.

Resources and Costs

Implementing and maintaining an effective intelligence program can be costly and require significant resources. Organizations must carefully weigh the costs and benefits.

Threat Evasion

Threat actors can develop techniques to evade intelligence-based detection systems. Tools and strategies must be updated to remain effective against new attack techniques.

© GarryKillian / Freepik

Integrating Intelligence Processes into Cybersecurity

Organizations should follow some best practices to maximize the benefits of intelligence in cybersecurity.

Develop an Intelligence Strategy

It is essential to develop a clear and consistent strategy for using intelligence. This includes defining objectives, selecting data sources, and planning analysis activities.

Use Advanced Tools

Organizations should use advanced data analysis and machine learning tools to manage and interpret intelligence data. This includes threat intelligence platforms, pattern analysis tools, and intrusion detection systems.

Collaborate with Other Organizations

Collaboration with other organizations can improve intelligence effectiveness. Sharing threat intelligence and best practices with partners, vendors, and industry consortia can provide a more comprehensive view of the threat landscape.

Training Personnel

It is important to provide ongoing training to security personnel to ensure they have the necessary skills to use intelligence effectively. This includes training courses on data analysis techniques, security tools, and incident response best practices.

Monitor and Evaluate

Organizations should continuously monitor the effectiveness of their intelligence activities and make improvements where necessary. This includes regularly reviewing data collection and analysis processes, updating security tools, and evaluating overall performance.

Technologies Supporting Intelligence in Cybersecurity.

The effectiveness of intelligence in cybersecurity is strongly influenced by the technologies used to collect, analyze, and interpret data. The following are some of the key technologies that support intelligence in cybersecurity.

Machine Learning and Artificial Intelligence

Machine learning and artificial intelligence (AI) are critical for advanced security data analysis. These technologies can identify complex patterns and anomalies in data, improving threat detection and attack prediction.

- Supervised Learning: Used to train patterns on labeled data and detect known threats.
- Unsupervised Learning: Used to identify unknown threats by analyzing patterns and anomalies in unlabeled data.

Big Data Analytics

Big data analytics technologies enable managing and analyzing large volumes of security data from various sources. This includes:

- Data Warehousing: Centralized storage of security data for efficient analysis.

- Data Mining: Extracting useful information from raw data through advanced analysis techniques.

Threat Intelligence Platforms (TIP)

Threat intelligence platforms centralize intelligence data collection, analysis, and distribution. These platforms integrate data from diverse sources and provide advanced threat analysis and visualization tools.

- Data Integration: Data is collected from internal and external sources.
- Analysis and Correlation: Tools to correlate events and identify attack patterns.
- Information Distribution: Sharing relevant information with security teams.

Future Trends in Intelligence for Cybersecurity

The cybersecurity landscape constantly changes, and intelligence plays a crucial role in adapting to these transformations. Below are some of the future trends that will influence the use of intelligence in cybersecurity.

Artificial Intelligence

Artificial intelligence will continue to evolve, offering new tools and techniques for threat analysis. This includes the development of more sophisticated algorithms for threat detection and attack prediction.

Multilevel Integration

Multilevel integration of intelligence information will enable a more complete and detailed view of the threat landscape. This includes integrating data from diverse sources such as security sensors, IoT devices, and cloud platforms.

Collaboration and Sharing

Collaboration and information sharing among organizations will become increasingly important. Security alliances and consortia will provide platforms for sharing threat information and best practices.

Critical Infrastructure Security

Protecting critical infrastructure, such as energy networks, transportation systems, and health services, will become an increasing priority. Intelligence will play a key role in protecting these infrastructures from cyber-attacks.

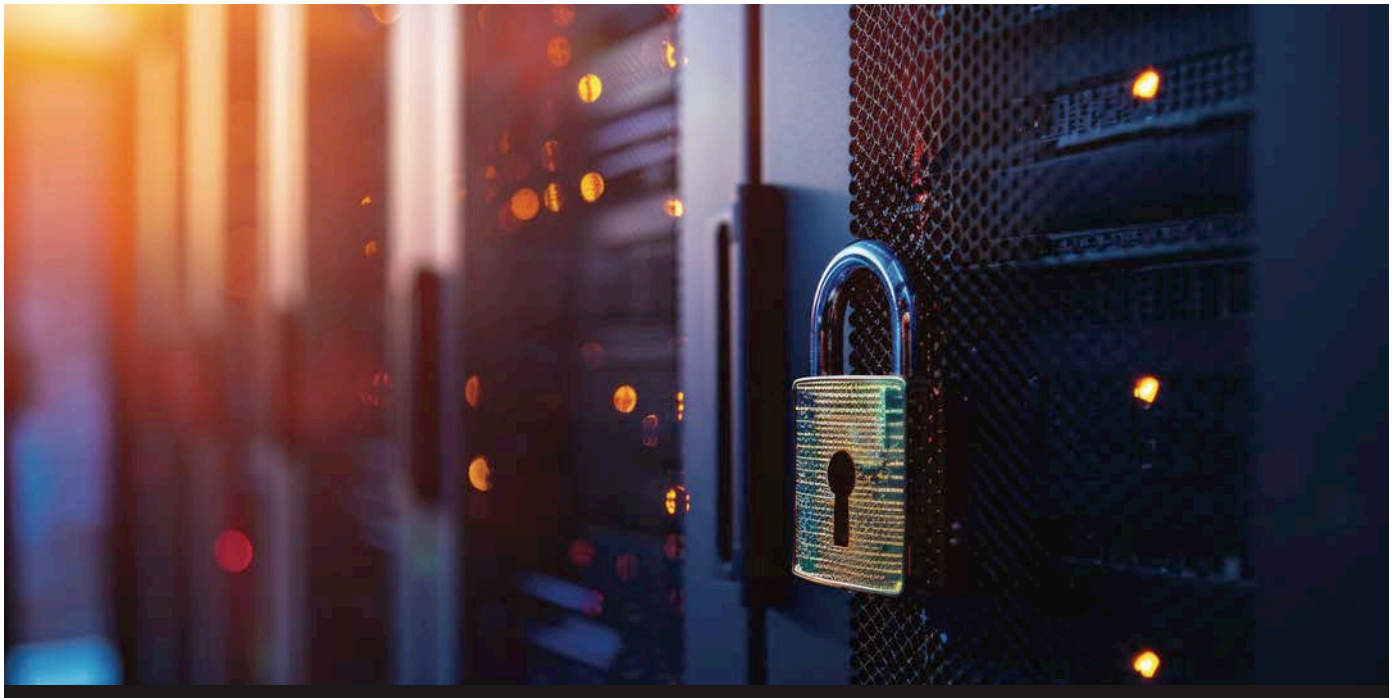
Regulation and Compliance

Cybersecurity regulations will continue to evolve, and intelligence will be essential to ensure compliance. Organizations will need to adapt to new regulations and demonstrate their adherence to security best practices.

Conclusion

Intelligence plays a crucial role in an effective cybersecurity strategy. By collecting, analyzing, and interpreting information, intelligence helps organizations predict, identify, and address cyber threats. Despite the challenges, integrating intelligence into cybersecurity offers several benefits, including increased proactivity, reduced response time, and improved threat awareness.

To fully leverage intelligence's benefits, organizations need to develop clear strategies, utilize advanced technologies, collaborate with other entities, and provide ongoing training to security personnel. Looking ahead, intelligence will continue to evolve in response to changes in the threat landscape, playing a key role in safeguarding digital assets and critical infrastructure. ■



METAVERSE

The 2030+ vision of the Geneva State Police in Metaverse.

How the Geneva Police is Preparing for the Challenges of Virtual Worlds.

Author: Patrick Ghion

What is the Metaverse?

The Metaverse is a term used to describe all the immersive, persistent and interconnected virtual worlds accessible through technologies such as virtual reality (VR), augmented reality (AR) and mixed reality (MR).

The word Metaverse is a contraction of meta (beyond) and universe (universe), and was popularised by Neal Stephenson's 1992 science fiction novel *Snow Crash*.

Since then, the concept has evolved in line with technological advances and Internet surfers' uses: today it includes platforms such as Second Life, Fortnite, Roblox, Decentraland and Facebook Horizon.

The Metaverse is a space for entertainment, creation, socialisation, education, work and business, where users can interact with digital content, virtual objects, avatars and intelligent agents.

The history of the Metaverse is rich and fascinating, and has been influenced by many cultural, technological and social events and developments. Here are some of the key moments in this history:

- **1982:** The film "Tron" is released, presenting a futuristic vision of a virtual world in which users can interact with computer programmes.
- **1992:** The science fiction novel "Snow Crash" by Neal Stephenson popularises the term "metaverse" to describe an interconnected virtual universe.
- **1999:** The film "The Matrix" is released, presenting a dystopian vision of a virtual world controlled by machines.
- **2003:** "Second Life" is launched, offering users the chance to create and explore a persistent virtual world.
- **2011:** The novel "Ready Player One" by Ernest Cline is published, presenting a vision of a near future in which most people spend their time in a virtual world called the OASIS. A movie adaptation has been released in cinemas in 2018.
- **2021:** Facebook changes its name to Meta, reflecting its commitment to building and developing the Metaverse.

These and other events have helped shape our understanding and vision of the Metaverse, and it's exciting to see how this story will continue to unfold in the future.

What are the financial and criminal stakes in the Metaverse?

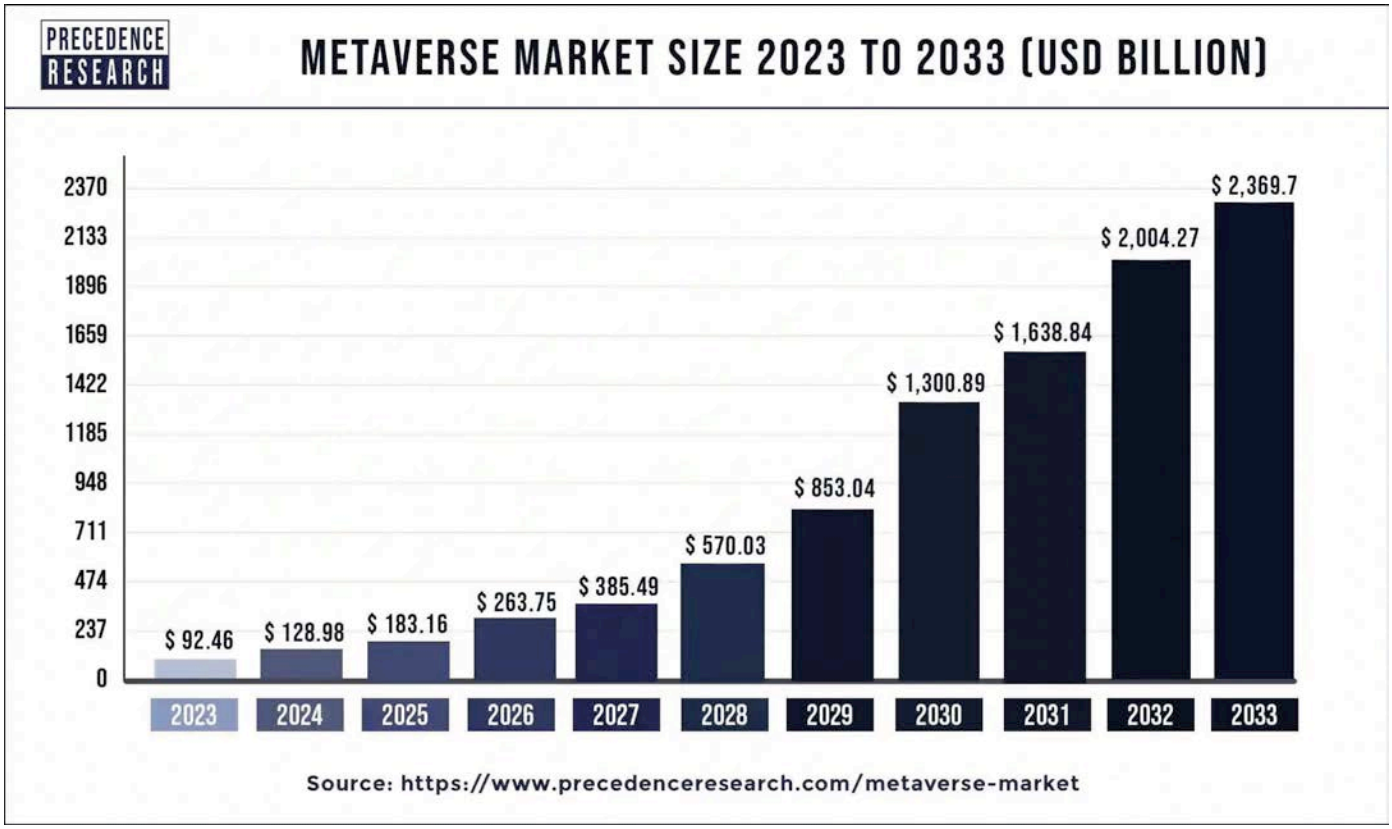
The Metaverse represents a colossal potential market, estimated to be worth several hundred billion dollars over the next few years. Players in the sector, including digital giants, video game publishers, entertainment companies and innovative start-ups, are in fierce competition to attract and retain users, and to create business ecosystems around their platforms.

The Metaverse is based on cutting-edge technologies such as blockchain, cryptocurrencies, NFTs (Non-Fungible Tokens), the Internet of Behaviour (IoB) and Artificial Intelligence (AI), which offer both opportunities and risks. These technologies can be misused for malicious purposes such as money laundering, fraud, theft, hacking, extortion, harassment, propaganda, terrorism and child pornography.

The Metaverse also poses legal, ethical and social challenges, such as the protection of personal data, respect for privacy, content regulation, the responsibility of players, the impact on mental health and the digital divide.

What is the 2030+ vision of the Geneva State Police in the Metaverse?

Facing these challenges, the Geneva State Police has drawn up a 2030+ vision to adapt its missions, skills and resources within the context of the Metaverse. This vision is based on three strategic pillars:



- Police officer training: this involves equipping police officers with the knowledge and skills they need to understand how the Metaverse works, its uses and risks, and to use the appropriate tools to intervene. The training includes theoretical, practical and immersive modules, as well as ongoing personalised follow-ups. But it will also involve exploiting virtual reality opportunities to increase the level of training by practising situations that would not be possible in real life.
- Prevention: the aim is to make Metaverse users aware of good practices, rights and duties, dangers and the reflexes to adopt in the event of a problem. Prevention involves information campaigns, partnerships with players in the sector, outreach and mediation activities, and an active and visible police presence in virtual worlds.
- Forensic aspects: this involves collecting, analysing and exploiting digital evidence relating to offences committed in or in connection with the Metaverse. The forensic aspects involve technical, legal and operational skills, as well as close collaboration with the relevant national and international authorities.

Against this backdrop, the Geneva State Police worked with the INTERPOL Metaverse Expert Group (i-MEG), contributing to the drafting of the document "METAVERSE: A Law Enforcement Perspective"[1], which was published in January 2024.

What is the future potential of the Metaverse?

The Metaverse is a fast-growing phenomenon that will continue to develop and diversify in line with technological advances and user expectations. The Metaverse will also generate new forms of data, which can be used for a variety of purposes, including personalisation, recommendation, prediction, simulation and creation.

This data can come from the ocular cameras in virtual reality headsets, which capture the gaze, facial expressions or emotions of users, or from brain sensors, which measure the electrical activity of the brain and interpret the intentions, preferences or mental states of users.

This data can have positive applications, such as improving the experience, optimising learning or increasing empathy, but it can also have negative applications, such as manipulation, spying, profiling or control.

BIO

Captain Patrick Ghion joined the Geneva Cantonal Police in 1997 after five years' experience in two Swiss banks. His career as a criminal police Detective has led him to conduct investigations in the fields of narcotics, burglaries, and the Juvenile Division. He joined the Computer Crime Division when it was created in 2003, where he remained until 2013, when he became Head of the Division until 2015. In 2015, he joined the Criminal Police Management Board, where he became Head of the Department of Forensic Sciences, comprising six Divisions, including the Cyber Triptych (Forensics, Investigations and Intelligence), the Homicide Division and the Crime Scene Investigation Division (CSI). There, he developed the cyber-Triptych, comprising analysts, investigators, and forensic specialists. This long-term strategy led him to implement and head the Regional Cyber Competence Center for Western Switzerland (RC3) from 2019. Since September 2023, he has held the new position of Chief Cyber Strategy Officer at the Geneva Cantonal Police, while retaining his duties within the RC3, among others. He is now responsible for the strategy and operational implementation of the Geneva Police's "Vision 2030+" in the Metaverse and for establishing partnerships between police organizations and the public, private and academic sectors at national and international level.

The Metaverse therefore raises crucial questions about data sovereignty and security, which require constant vigilance and regulation.

Conclusion

The Metaverse is a fascinating phenomenon, opening up unprecedented horizons for humanity, but also posing major challenges for society, democracy and the rule of law. The Geneva State Police is well aware of these challenges and has set itself an ambitious and realistic vision for 2030+, to ensure that it fulfils its public service mission in virtual worlds. The Geneva State Police is convinced that the Metaverse can be a place of freedom, creativity, solidarity and progress, provided that it is governed by clear rules that are respected and applied adhered to by all those involved. ■

1 https://www.interpol.int/content/download/20828/file/Metaverse%20-%20a%20law%20enforcement%20perspective.pdf

The power of platforms in the great agora of the 'metaverse'.

Interview with Luciano Violante.

Author: Massimiliano Cannata

** A first version of this article has been published in Cybersecurity Trends Italia 2:2024.*

"It is still too early to say what scenarios open up with the development of the metaverse. McKinsey estimates two to three trillion dollars in investments in this field, within this decade. Avatars will soon also have touch and taste. It is difficult to make reliable evaluations, our effort must tend towards the construction of a digital civilisation, characterised by the autonomy and dominance of the person over the algorithm, without which there can be no future.' Luciano Violante, jurist and magistrate, president of the Leonardo Foundation – a defence, airspace and security giant – engaged in defining the ethical and legal status of what is the last stage of the digital revolution, invites us to be cautious when dealing with a multifaceted topic such as the metaverse.

Mr. President, huge public and private players are trying to improve network* governance*, in an attempt to ensure the security of communications and transactions, the respect for privacy and the inviolable personal rights, and transparency. How should this challenge be addressed?

Technological development can certainly not be stopped, just as human creativity cannot be stopped. But one can think of addressing and constructing a governance of technologies together with technologists, knowing that the economic and social spheres that will be most affected by this new revolution are the same as those of Web 2.0: human relations, marketing, communication, finance and economics, education, health. But with one important difference: the power of mutation in this case will be immeasurably greater than in other paradigm 'leaps' we have experienced. There is one thing I would like to point out right away: the construction of the autonomy and domination of the person over the algorithm is the objective to be set in the present, in order to be free in the future. It is our responsibility to be able to realise it.

In your official speech at the Warrant of Privacy on the occasion of the European Data Protection Day, you dwelt on a neologism:* 'figital'. *Can you help us explain what it means?

In ordinary language, the term indicates a new paradigm affecting every aspect of living, living, learning, doing, from the environment to the city, from gaming to education, from ethics to work, from the seabed to space. Just one example: the Chinese chain Kentucky Fried Chicken has installed smart screens in its fast-food restaurants capable of exploiting facial recognition and artificial intelligence to propose special offers to customers.

From the* 'homo videns' *thematised by Giovanni Sartori to the* 'homo distans', *which is substantiated in a dimension of the self that is* 'distanced from its own body'. *Should we be afraid of this condition of being that we still know little about?

I would speak of *homo distans*, with reference to the possibility of socialising between people who are distant and converse through their avatars. There is a risk, which must be addressed and tempered, of a further loosening of

social and interpersonal bonds. A risk that we cannot underestimate and which we must deal with, with the utmost responsibility.

The risks of reintermediation

The oligopoly of platforms puts democratic freedoms at risk. Are there effective countermeasures?

In our society, the most dangerous deception is disintermediation. Mediators are not being erased: they are being replaced. The old mediators - party, association, church, family, trade union - presented themselves as such on the public stage, were scalable, had knowable statutes. The new mediators do not present themselves as such, are not scalable, have no visible statutes. Microsoft, Amazon, Google give us the services we need at an acceptable cost and with efficiency. In return we freely and deliberately hand over all our data to them. If the same data were demanded from us by the State, protests and press campaigns would start. In reality, what is underway is a process of reintermediation, which must be studied very carefully.

What in particular are you referring to?

The platforms, referred to earlier, guide our daily lives to a greater extent than traditional mediators. In the past, one knew the address, the telephone number, the leaders and workers of one's party or trade union, of one's parish. One could question the leadership of the party and trade union and be a candidate to be elected at their place. In contrast, today we do not have the address and phone number of Amazon, let alone the ability to climb in its structure. The risks are obvious. For covert brokers there are neither rules

nor countervailing powers; without appropriate countermeasures they are destined to wield infinite power over our lives. The streams of thought driven through social media count for more than individual intelligence. Those who govern the digital environment have the ability to decide not only what we buy, but what we think and how we orient ourselves in the world.

Where, in your opinion, does this immense power come from?

Microsoft, Google, Amazon, the largest platforms, control 64% of the cloud infrastructure market. Microsoft has about 90% of the operating systems for servers and PCs and runs Office, which is the most popular software package in the world. 92% of our mailboxes are managed by Microsoft, Apple and Google. Again, to define the weight of digital companies, I recall that in 1990, the top five US companies in terms of capitalisation were IBM, Exxon, General Electric, AT&T, Philip Morris. In 2020 they were Apple, Microsoft, Amazon, Alphabet, Facebook. In the same year, each of the five largest technology companies was worth more than the 76 largest energy

companies combined. We live in an oligopoly. The oligopolists have people, companies, states in their hands. If they flipped the switch, the world would grind to a halt. This is their strength with which we will have to learn to measure ourselves. ■

BIO

Luciano Violante
President of the Leonardo Foundation, Luciano Violante teaches Public Law at 'La Sapienza' University in Rome and at the University of Valle d'Aosta. A jurist and magistrate, during the 'years of lead' he held the delicate post of examining magistrate in Turin, where he immediately distinguished himself for his systematic commitment and action against the phenomenon of terrorism. A parliamentarian with a long militancy in the ranks of the PCI, PDS and DS, he was president of the Anti-Mafia Parliamentary Commission (1992-1994) and of the Chamber of Deputies (1996-2001) and was a member of the commission of '10 wise' men, appointed by the President Emeritus of the Republic Giorgio Napolitano, with the aim of defining the institutional and economic-social reforms necessary for the country's development. He is currently at the head of 'Italiadecide' (cf. www.italiadecide.it). The purpose of this association, whose founding members include top-rank personalities, is to contribute to the improvement of the quality of public policies and to enhance the different levels of territorial governance through a precise and rigorous application of the logic of subsidiarity. He is the author of several books, on geopolitics, law, justice, and State institutions.

The Metaverse Will Impose a New Way of Being on the Planet.

Interview with Mattia Fantinati.

Author: Massimiliano Cannata

"Communication is increasingly preponderant in our lives, which are now themselves sometimes transformed into medium, sometimes into message and are sometimes embodying both. The Internet reduced the space between people and accelerated the phenomenon. We organise sessions, meetings and seminars mainly to provide tools to fight fake news and deepfake that put our democracies at risk. Human needs and rights must always be at the centre of every internet governance. This is the mission of IGF Italy, the Internet Governance Forum, an organisation that belongs to the UN and supports through its commitment the ethical democratisation of the governance of Internet. In short, we must arrive at a body of rules that protects online rights just as it does offline". In this interview, Mattia Fantinati, president of IGF Italy, outlines the delicate relationship between law, ethics and the evolutionary rhythms of technological progresses, which are becoming ever tighter and more disruptive.

BIO

Mattia Fantinati is the president of IGF Italia (Internet Governance Forum) since 2021 and is involved in the organisation of EuroDIG since 2019. At the United Nations, he represents Italy at the HLPDC (High Level Panel of Digital Cooperation) roundtables. He has been a Member of the Italian Parliament, Chamber of Deputies and Undersecretary of State of the Public Function. His main areas of interest are Digital Transformation in Governments and Innovation Management in the Public and Private sectors. He was appointed by the Secretary General of the United Nations as a member of the MAG (Multistakeholder Advisory Group) to advise the Secretary General on the IGF (Internet Governance Forum). He regularly participates in conferences and seminars on Digitalisation held by the most important public and private agencies (ITU, WSIS, CES,E-Leaders, OECD).

Mr. President, Stefano Rodotà made an allusion to a Constitution for the Internet, emphasising the necessary global vision and real cooperation between States needed to regulate an instrument that is by definition borderless. Is the great jurist's intuition destined to remain utopia?

We see that today states do not have identical approach to human rights. This is also the case in the digital realm. Perhaps it is utopian to imagine an equal law for all, but at the UN or EU level, coordination between nations that sign agreements and abide by them is already a good starting point. IGF has mapped all 38 inter-stakeholder agreements worldwide and we observe two recurring themes: respect for human rights and the importance of digital cooperation between countries. Of course, then these states have

to put these principles into practice, but awareness is growing and we are playing our part. After all, it is up to supranational entities to balance the weight of the big platforms and network giants.

The second half of the Internet

Scholars speak of a second time of the Internet. After the first phase that opened up a new universe for the world of communication and production, social networking, the IOT, now we have 'that fourth revolution', well described by Luciano Floridi, in which the real and the virtual intermingle, giving birth to the 'on life' dimension in which we are all immersed.

How do you regulate this dual reality?

The prompt given by the question leads me to talk about the metaverse, a young technology that, however, already hints at its future implications, in which we move from 2D to 3D and the person is projected into a virtual world. Each of us can do many things in the metaverse. One can attend meetings, train or work remotely with great benefits because the experience is made more immersive by the shift from pixels to voxels, a three-dimensional version of the pixel itself. One can learn to use objects or do training in many areas without risking some inherent offline dangers. Of course, the metaverse cannot replace reality, the emotions of a journey for example. But it is wrong to be frightened and legislators should not chase the changing world.

What concrete strategies should be adopted?

I believe it is right for a multi-stakeholder regulator to sit down with the creators of these platforms to try to outline a scenario starting with some clear and stable legislative *pillars*, in order to leave then the regulatory framework unchanged for as long as possible. Indeed, we must avoid the regulatory schizophrenia that generates only uncertainty and excessive bureaucracy. Obviously, the centrality of human beings, education, respect for privacy and individual identity, and finally the improvement of cybersecurity safeguards are among the fundamental guidelines. Basically, it works like mobility rules and the highway code: I rely on certain binding conventions, such as traffic light colours, because I know that more or less everyone respects them. The same thing must happen when I use a connected device: I need to know that there is a set of rules protecting me.

Metaverse, democracy and the Internet

Democracy and the Internet are at a crucial junction. In a recent essay, Michele Mezza denounced the oligopoly of platforms, which put fundamental freedoms at risk. Where do we stand on this delicate issue that has to do with that surveillance capitalism, theorised by Shoshana Zuboff, which has generated so many questions?

These platforms have turnovers equal to the GDP of many states. So an individual nation cannot balance them alone, while unions of countries, by creating a critical mass, can limit their abuse, for instance in terms of data usage. The EU alone has 600 million inhabitants and

can therefore have a say. It is no coincidence that the EU Commission with measures such as the Digital Markets Act (DMA) and the Digital Services Act (DSA) has a strong and integrated regulatory approach on platforms. As Roberto Viola, general director of DG Connect of the same Commission, said, platforms are important, but they are not States, they have no independent legislation. On the contrary, they must respect the rules of a healthy market. And then there is an indispensable ethical plan: the big digital players have made huge profits and now have a duty to reinvest to make the internet an increasingly safe and profitable place.

You mentioned the metaverse perspective. On the occasion of the European Data Protection Day, the Warrant of Privacy stated: **'It is not so much the conceptual novelty that should shock us - Pierre Levy, Negroponte and the early theorists of the Net have spoken to us about virtuality - as the issues related to the impact that this new digital 'habitat' will have on social relations and on the behaviour of individuals, on their freedoms and rights, as well as on the decision-making processes of the community.** ***We are facing a new test for jurists, politicians and men of institutions.***

What is your opinion on this?

We have already addressed the legal aspect in this conversation. On the economic side, it is clear that the race for the new 'mode of being' may cause imbalances, bubbles and speculative risks. Some will set off in search of a new Eldorado, hoping to gain large slices of an unprecedented market and consumption pie, so it is clear that regulation will also have an impact on economic, financial activities, real estate, etc, and not only on the various 'virtual' and de-centralised markets that may emerge. However, as the digital philosopher Cosimo Accoto argues, virtuality is probably only one of the possible outputs of the metaverse, while the immersive internet will also be embedded in objects, in our bodies themselves, in our environments, starting with augmented reality, and we will live within the many simulations that are transforming the world and generating a new 'terraforming', which imposes a new way of being on the planet.

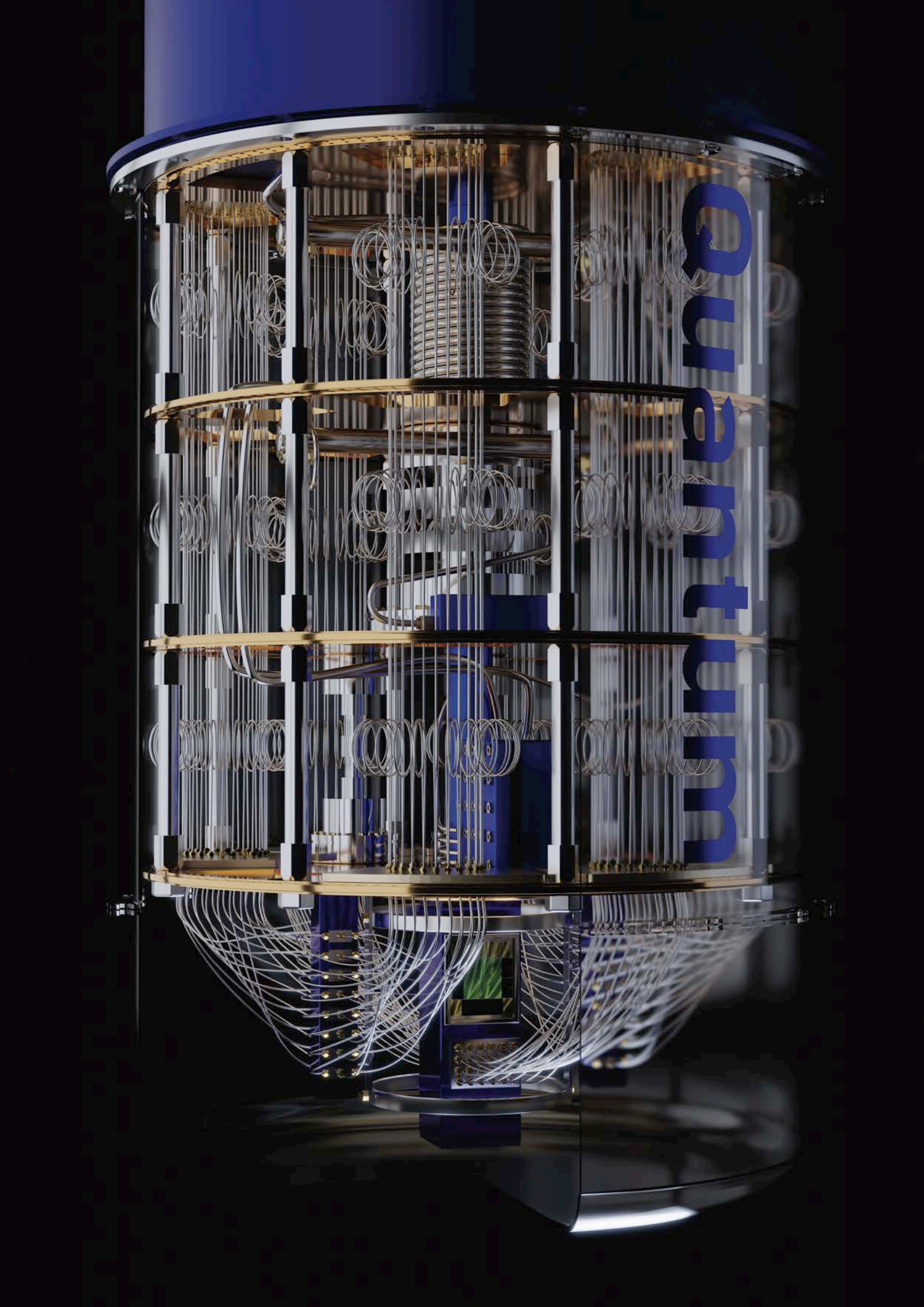
In this scenario, there is naturally great attention and great disquiet surrounding the issues of cybersecurity, a now crucial terrain on which criminal battles are fought for economic profit, but also wars of an allegedly 'ethical' nature led by transnational groups and by militant activists of the most diverse causes. To these are added the skirmishes between sovereign states for global geopolitical supremacy. What can you say about this?

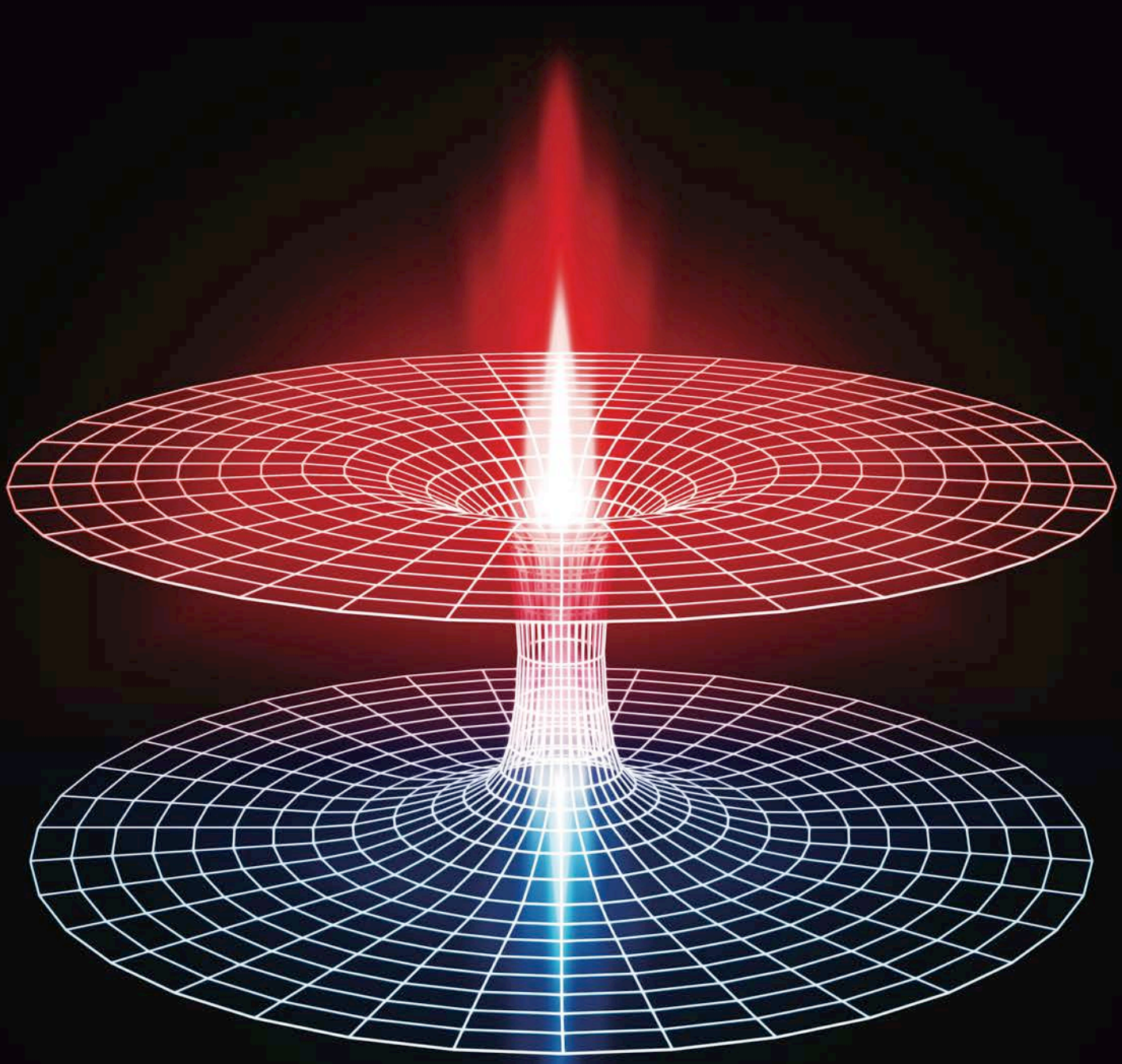
Europe is taking power away from platforms and, at the same time, investing in digital sovereignty and in the defence of institutions and citizens. The old IT market rules are calibrated on a 50-year-old reality and therefore need to be updated. Today, we have to look after data management and precisely after the platforms: the legislation has hence to be rewritten in an innovative way, as is happening with the aforementioned DSA and DMA rules. I have just come back from a professional trip to Israel and over there, they are very attentive to the advancement of technology from a military perspective.

What can the State of Israel teach us about a sensitive subject like cybersecurity?

They are very advanced on both hardware and software. They invest a lot in training: a sector that is subject to constant change and needs robust schooling, especially secondary schooling. We also need to be able to push on the ability of young people to do innovative business, creating a favourable environment for start-ups. In Tel Aviv, they are very good and we, despite some excellence, have a lot to learn.

On cybersecurity, in fact, awareness in Italy is not yet very high, but the IGF has recently joined *'RepubblicaDigitale',* the strategic initiative of the Department for Digital Transformation of the Presidency of the Council of Ministers, precisely to support the strengthening of digital skills. Finally, we are running an *Internet governance* school to teach businesses, organisations and associations what different technologies we have are and how they work. We celebrated the launching event in Ancona last November and we will definitely continue along this path. ■


quantum



QUANTUM COMPUTING

Will it still be possible to protect secrecy in a post-quantum world?

Author: Franck DeCloquement

The probable deployment of quantum technology, and its correlative mastery, will have the reasoned objective of significantly increasing the nation's strategic autonomy, and guaranteeing the operational superiority of our forces and our State. On the other hand, what about the deleterious effects of future quantum resources, extending their offensive capabilities operated by determined attackers, compared with defenders who currently only have traditional computers, equipped with conventional encryption?

Vladimir Putin's famous statement, which has since become emblematic, is often quoted as illustrating the decisive strategic stakes involved in the speed race between the major powers vying for absolute mastery of AI: *«Whoever becomes leader in this field will be master of the world»*. But as one train can very often hide another, the probable advent of quantum computers remains one of the major challenges today for guaranteeing the integrity and security of our data stored across the world's infrastructures, which we all know are essentially based on the principle of encryption.

In the beginning was the number.

In principle, all our top-quality mobile phone messaging services are encrypted «end-to-end». All the hard disks in the world that are in contention, and in place in the servers of the mega data farms, containing the bulk of our personal data - medical for example - or sensitive data, are drastically encrypted. Just as the secrets held by our armies and our various intelligence centres are also encrypted. From our online payments (which are made via all our payment methods), our electronic signatures, our energy expenditure records (recorded via smart meters hosted in our own homes), our on-board computers in our current (and soon driverless) vehicles, and even the biometric chips embedded in our passports, all depend directly

BIO

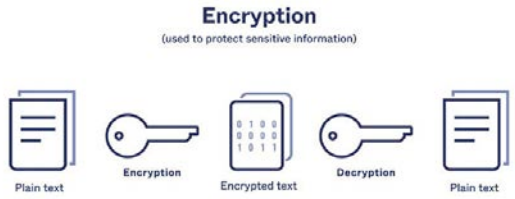
Former student of the École de Guerre Économique (EGE), expert practitioner in strategic intelligence, member of the CEPS and of the scientific council of the Institut d'étude de géopolitique appliquée (IEGA). He is also the author of the chapter «Cyber, quantique, intelligence artificielle: vers une bascule majeure géopolitique? in the collective work «Géopolitique du XXI siècle», published in April 2024 by Ellipses.

on complex algorithms developed in the 1970s, which transform easy-to-read data into encrypted messages accessible only to those who have the 'key' to unlock them. These complex algorithms, in turn, depend on mathematical functions that are simple to use to create encrypted messages, but particularly difficult to reverse to find the source message, this time in clear text, if you don't have the right key.

Having said that, we regularly hear the term 'encryption' used in everyday language, but the word has no meaning whatsoever. Here's a quick lesson: the word 'encrypt' entered common parlance a long time ago, often incorrectly. The media have also played a big part in popularising the misuse of the term, for example by regularly referring to «encrypted channels», as well as a mistranslation of the English word «to encrypt». But the word has enough to drive even the most fussy cryptologists mad. And with good reason, because once you understand the origin of the word a little better, you immediately realise that it has no meaning for our current uses.

The principle of encryption.

There are many situations in which we want to make data or messages «unintelligible» to anyone who might intercept them improperly. Whether it's a secret message or simple personal data stored in third-party data centres.



Throughout history, numerous methods have been devised to «encode» messages. Some are very simple, but others are much more complex. Among the best known is Caesar's code, which simply involved replacing the letters in the message with the next letter of the alphabet (or with a fixed offset): A becomes B, B becomes C, etc. There are, of course, many other methods, infinitely more effective and subtle. Nazi Germany, for example, used the famous «Enigma» machine (recently popularised in the cinema in 2014 by the film directed by Morten Tyldum, «Imitation Game», and starring Benedict Cumberbatch), during the Second World War, to encrypt messages of strategic interest to the Reich. All by substituting one letter for another. But with the addition of a regular substitution rotation, to avoid always encoding the same character in the same way. This parameter for switching from a «clear» message to a «coded» message is called an «encryption key». It can be very simple («move one letter forward in the alphabet» is an easy algorithm to use, for example), or very complex (as the current methods generally used on the Web, including the «Advanced Encryption Standard» - or AES - for example).

This is how most of our data is encrypted when it travels through certain email systems, or when it is stored on most of the servers of the main web platforms. But also some of our personal objects (IOT), like our indispensable smartphones.

Deciphering or decrypting?

Once the message has been «encrypted», however, it is vital to be able to retrieve its content intact and in its entirety, one way or another. When you send a message to a friend using an application such as «Signal» or «WhatsApp», for example, you don't want a third party to be able to read your private exchanges in clear text if they are intercepted. But you naturally want your correspondent to be able to read it clearly. Your correspondent must therefore be able to «decrypt» the packet of data received. Depending on the method used, decryption can take several forms. With a symmetric encryption system, the key used to decrypt the message is identical to the one used to encrypt it. Asymmetric systems are a little more complicated, since each party has two keys: a public key and a private key, both of which are mathematically linked. The sender uses the recipient's public key to encrypt his message, while the recipient uses his private key to decrypt it. However, we will not go into further detail here, as cryptographic methods are not the main subject of this article.

How do you decrypt a message?

There are, however, scenarios that go beyond the simple transfer between two consenting individuals. Hackers and certain governments, for example, seek to take advantage of the security loopholes inherent in these devices to recover data in circulation. But as mentioned above, this data is usually encrypted and they do not have the key to decrypt it. Then begins the tedious task of «breaking» the code. The main method is to «brute force», trying every possible combination, one by one, until the result is conclusive. This is an extremely long and time-consuming process, which can take several years, and in some cases even longer. Without a key, it's not a matter of «decrypting», but of «decrypting».

To sum up:

- Encrypting a message involves encoding it with a key.
- Decrypting a message means decoding it using a key.
- Decrypting a message means decoding it without a key.

As you can see, if we follow this logic, «encrypting» something would consist of encoding it without having a key. However, as soon as we use a method to modify a message in order to make it unreadable, even in a totally random way, we create at the same time the «key» enabling the process to be reversed. It is therefore impossible to «encrypt» a message.

This is why, in theory, if we follow the lexicon of «cryptography» to the letter, we should not talk about «encryption». However, it should be remembered that French is a living language, which evolves with its various usages, even when these are abused. And as we mentioned earlier, the term 'crypter' has become part of everyday language. It may be fundamentally wrong, but for many people it is a way of making themselves understood by others, and of putting across an idea that will be understood by both experienced cryptographers (who may - admittedly - grit their teeth a little) and newcomers to the subject. The fact remains that, from now on, our dear readers will be able to explain why it is preferable to use the word «encrypt». It might come in handy at the next family dinner...

When custom defines the norm.

In view of the decisive contributions made by technology, all the regalian law enforcement agencies of the states throughout the world (including France and the United States) are increasingly in active demand of «backdoors»

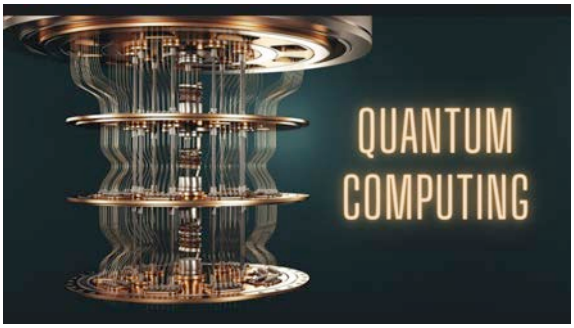
included in the encryption systems that protect our personal data and our telephone exchanges. But also those of the countless criminals at work in our computer ecosystems, safe from the prying eyes and ears of government services. Arguing that the national security of countries and their citizens is at stake. Which is not untrue. And there are new indications that government agencies already have very active methods of action at their disposal, as well as specific, high-performance intrusive tools which, for better or worse, enable them to access the data on locked smartphones at will, thanks - moreover - to weaknesses in the Android and iOS security systems themselves.

Although the protections currently in place for our smartphones are relatively adequate to deal with a number of threat models (or potential aggressive attacks), computer researchers have concluded that they do not answer the question of specialised forensic tools that government agencies can easily acquire or implement as part of their police investigations or intelligence work. The soap opera of the media adventures of Pegasus spyware, based on zero-day vulnerabilities and flaws, has undoubtedly demonstrated this to the world over the last few years.

A few years ago, a report by researchers from the non-profit organisation Upturn found nearly 50,000 examples of US police in all 50 states using legal computer forensic tools that can be deployed on mobile devices to access smartphone data between 2015 and 2019. And while citizens in many countries may still think it unlikely that their computing devices will be specifically subject to this type of search, mobile surveillance remains ubiquitous in many parts of the world. And at an ever-increasing number of border crossings...

When the advent of quantum is just around the corner.

However, if quantum computers are ever developed, these problems, which are particularly complex to solve in the present day, will undoubtedly become almost «child's play» for the malicious intelligences that are at work and that are destined to appropriate them in the future (hackers, mafias and criminal circles, warring states, etc.). To decrypt a message protected by the 'RSA' protocol, for example (a system that also allows encryption keys to be shared), a conventional computer today would take a time comparable to the lifetime of the Universe itself. But the researchers estimate that a future quantum computer could perfectly well do the same job in just eight hours.



The Diffie-Hellman key exchange (a widely used cryptographic method named after its two inventors) would also be easily reversed by a quantum machine. On the other hand, another type of protocol, such as AES (Advanced Encryption Standard), would not be considered to be directly threatened by advances in computer technology, but it is often used in combination with the previous methods, and therefore cannot entirely replace them.

Quantum technologies that will enable us to perform astronomical calculations in the future. But also within the reach of our adversaries...

The fields of application envisaged are already very vast: whereas conventional computers work with ordinary digital bits made up of «1» and «0», quantum computers are based schematically on quantum bits, or «qubits». These units take advantage of a specific property of quantum physics: «superposition». This allows a qubit to be, for example, 70% «1» and 30% «0» at the same time. The ability to be in many states - «at once» - means that quantum computers can perform complex mathematical operations much more quickly than even the most powerful classical computers today. The time saving is such that certain astronomical calculations that would be unimaginable today would be immediately within the reach of scientists. And even if such computers are not produced for at least another twenty years, according to some futurists, the problem is already a matter of urgency, according to the many researchers consulted. Because our data could already be in immediate and irreparable danger. And that's even if it hasn't yet been produced, according to Dustin Moody, a renowned mathematician who also heads the NIST's «post-quantum cryptography» project.

Post-quantum? Or how to guarantee secrecy in a world that has become quantum.

Existing quantum computers contain a few hundred qubits at most, and ultimately offer fairly limited performance. IBM, for example, presented a chip with 1,121 qubits at the end of 2023, and claims that it will have a computer with more than 4,000 qubits by 2025. Scientists at Google and the Swedish Communications Security Authority estimated in 2021 that 20 million qubits would be needed to break a 2048-bit RSA key. A very commonly used key length.

There is no longer any doubt that quantum computing, if it comes to pass, will drastically change the way computing is used in the future. And vital questions are being asked about the security associated with this inevitable development. It's a safe bet that cyber-criminals of all stripes, and the predatory states that often employ them as their 'hunting team', are currently working hard to prepare themselves to break this emerging revolutionary architecture. We can therefore legitimately assume that certain intelligence centres and the clandestine actions associated with them around the world could already be collecting strategic encrypted data, identified as particularly sensitive or decisive, while waiting patiently for the technology to emerge so that

they can eventually read it «with an open heart». Some specialists rightly believe that this is what the great powers in global rivalry, such as the United States and China to name but two, are doing. And with the dizzying prospect of the shock that the probable arrival of quantum computers will produce in our ecosystems of life, cryptographers and standardisation bodies the world over are undoubtedly striving to develop a set of encryption devices and techniques that these machines will find as hard to attack as our «classic» computers of today. With this in mind, many researchers are exploring and testing new, much more robust algorithms.

Post-quantum in a nutshell.

Faced with the security threat posed by quantum computers, the world of cryptography is racing against time to develop new algorithms that can be used as effective countermeasures.

Faced with these machines of a new kind, which will be able to carry out calculations at a speed infinitely faster than current computing resources (and will therefore be able to decrypt encrypted data fairly quickly), post-quantum cryptography ultimately refers to the frantic search for new encryption algorithms capable of theoretically withstanding the onslaught of the quantum computer.

What's more, their operation does not require the use of a quantum computer itself, as they can already be implemented on all our current machines. In short, post-quantum cryptography can be defined as a specific branch of cryptography designed to guarantee the security of information against attackers with quantum computers «in the future». This discipline is distinct from classical quantum cryptography, which aims to build cryptographic algorithms that use physical, rather than mathematical, properties to guarantee security. Shor's, Grover's and Simon's quantum algorithms extend the capabilities of an attacker using only a conventional computer. While there are currently no quantum computers, strictly speaking, representing an immediately concrete threat to the security of cryptosystems deployed to date, these algorithms conceptually make it possible to solve certain complex computational problems on which several popular primitive values are based. A primitive value or «primitive data structure « is data that is not an object and has no method. In computing, a «primitive function» can refer to a basic function provided by a software layer such as BIOS, just above the hardware architecture of a computer. These primitives are generally provided by a programming interface. They are often faster and more efficient than their equivalent high-level programmed versions, because they are optimised for the hardware being driven, by managing I/O. Both symmetric and asymmetric encryption will undoubtedly be affected by this likely emergence. Symmetric encryption

uses a single key to encrypt and decrypt protected data, which must therefore be communicated to everyone who needs access to it. The solution for this is fairly simple. All that needs to be done to protect the data is to increase the size of the key without changing the algorithm. This is what is known as «asymmetric» encryption, which will be particularly vulnerable to quantum computers. It is based on two keys: one «private» and the other «public». The recipient's public key is used to encrypt the data sent to him, and only his private key is used to decrypt it. This method is used for secure exchanges, including instant messaging.

A brief overview of the different types of post-quantum cryptographic algorithms.

So-called «post-quantum» cryptographic algorithms can be broadly divided into several categories. Each has its own strengths and weaknesses.

- « Lattice-based » cryptography is based on the mathematical properties of lattices, which are also structured collections of points in a multi-dimensional space. The hardness of certain computational problems associated with «lattices» forms the basis of encryption and key exchange algorithms.
- « Multi-variable » cryptography involves solving non-linear equations to obtain secure encryption. The complexity of these equations makes it particularly difficult for a hacker to recover the plaintext from the ciphertext without the appropriate decryption key.
- «Code-based » cryptography uses error-correcting codes to guarantee data security. Thanks to these codes, it is much more difficult for a hacker to extract significant information from the ciphertext, even if he has unlimited computing power at his disposal.

We are obliged to fight the future now.

The message was clear: *«Tomorrow, a sufficiently powerful quantum computer will be able to break all cryptographic algorithms and decrypt our messages. To counter this threat, developing post-quantum encryption technologies is a strategic challenge. And here we are»,* stated the French head of state, Emmanuel Macron, in a tweet dated 1 December 2022. Quantum computing, which will probably eventually offer the possibility of solving infinitely complex problems hitherto inaccessible to our contemporary means of countermeasures (while at the same time posing a major threat to the solidity and integrity of many of the digital safeguards currently in use, which we have hitherto taken for granted), is forcing us... A cold shower guaranteed for the flippant.

Pushing the implementation of these new sets of algorithms to the limit in order to counter the future power of quantum decryption has become the imperative objective of multiple multi-year competitions organised across the planet, in order to develop a robust post-quantum cryptography strategy, where only the strongest will survive. This is because certain approaches that initially appear promising will logically wither over time. This is forcing researchers around the world to come up with new ideas at an ever-increasing pace. For example, as reported in specialist scientific journals, out of 69 proposed algorithms, selected at the end of 2017 in one of these competitions, between 25 and 30 have been completely «broken» (or have undergone very significant fracturing attacks). At the end of August 2023, for example, the National Institute of Standards and Technology (NIST), an agency of the US Department of Commerce whose aim is to promote the economy by developing technologies, metrology and standards in conjunction with industry, had published draft standards for three robust algorithms. The agency planned to finalise the standards during 2024. As the field continues to advance, key innovations such as quantum key distribution and quantum shock-resistant digital signatures are coming on stream, offering promising solutions to the challenges posed by quantum computing. But ultimately, the future of post-quantum cryptography, sometimes also called «resistant cryptography» or «quantum security», which aims as we have discovered to develop encryption methods that will remain secure (even against attacks from super-powerful quantum computers), will definitely be what we make of it.

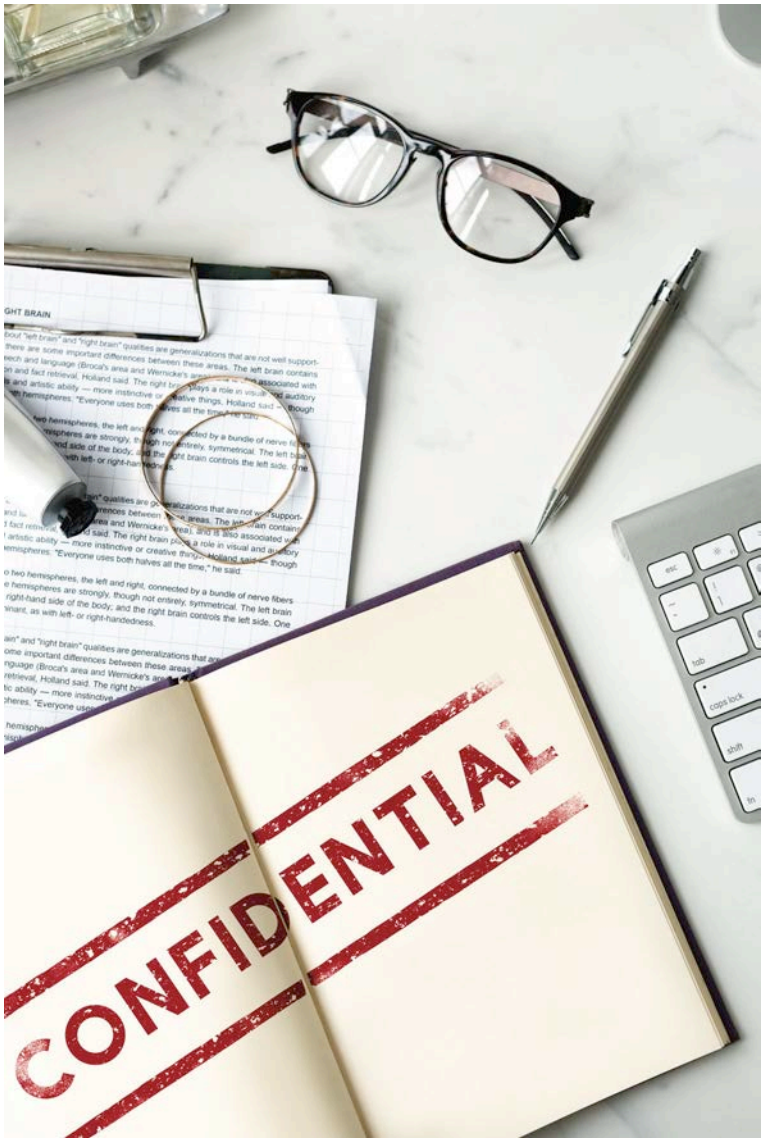
By exploring new mathematical problems and developing quantum-resistant algorithms, researchers are striving to guarantee the confidentiality and integrity of sensitive information in the post-quantum era that lies ahead. We can only hope that this future will be kind to us... ■

This text is freely inspired - for certain parts only - by one of the author's publications on encryption for Atlantico, published in 2021, and by the article Keeping secrets in a quantum world, published by Nature in November 2023, which has been retranscribed, taken up and adapted many times by various scientific publications and media.

POST SCRIPTUM

Artificial Intelligence, between Great Promise and Reality with Multiple Risks.

Author: Laurent Chrzanovski



Artificial Intelligence, with its impressive number of applications, has become indispensable. While there is no doubt about its ability to accelerate and improve research in many scientific fields, when skillfully supervised by humans, its use across the board in the tertiary and private sectors raises many questions and just as many problems.

The consequences of ill-considered adoption of AI: the end of confidentiality

The world is now divided into two categories: companies and individuals who have massively adopted AI products and those who have not yet taken this step.

The first category is particularly exposed to danger. Indeed, like many others before them, these tools store everything in their memory, and using them for all your personal and professional activities means handing over a huge amount of data, often some of the most confidential in nature.

And with good reason: within a company, as well as asking chatbots all sorts of questions, employees can now write projects and professional documents at lightning speed, improving and correcting them, summarising them and formatting them with images and graphics. AI also plans the group sessions at which these documents will be presented, and takes the minutes. What's more,

the AI can be entrusted not only with managing the company's projects, marketing and social networks, but also with customer service, organising emails, appointments and job interviews.

In a way, the entire strategy and actions of a company hyper-connected to AI - in short, how it works, its data and its entire intellectual property - are now known by its own suppliers of AI products. A treasure trove of data that enables AI to generate, on its own, comparative plans to respond better and enhance similar requests. Ultimately, this could well lead to the delivery - anonymised - of one company's business plan to another.

So it's hardly surprising that more and more large corporations, as well as players in the most sensitive fields (both public and private), are banning the use of AI from their companies altogether.

The abolition of intellectual property

But the thirst of AI players for access to confidential or protected data does not stop there.

On 13 March, OpenAI asked the US government to lift the ban on feeding its products with copyrighted data.

Whatever the federal government's response to this request, the approach is unambiguous: research, patents and publications are the last link that AI still needs to provide answers worthy of those of the best specialists in each field.

There can be no doubt that it is only a matter of time before the giants of the Big Tech get their way on this issue.

So, thanks to all the data collected in open access, as well as that obtained by the users themselves, this latest area of knowledge will complete the squaring of the circle for AI players in their quest for omnipotence over virtually all digital information.

This subject is also central to a number of studies by Luciano Floridi[1], who wonders how we will ever be able to protect our own copyright and, vice versa, asks whether the results obtained by AI tools will not themselves become texts in the future, with copyright belonging to the companies that generated them.

The «chaos» that AI could produce: a healthy mistrust on the part of citizens.

A few weeks ago, the US intelligence services revealed that several states and their «troll factories» were flooding the web and AI products with fake news, untruths and subversive material.

Even in the long term, it will become increasingly difficult for designers of AI products, and *chatbots* in particular, not to mix verified information with intoxication, delivering users a 'magma' where facts and 'fakes' are mixed together, going well beyond what is highlighted in *Christine Dugoin-Clément's article in this book*.

In this sense, the biggest change we can see is in people's perception of the use of artificial intelligence.

Indeed, in 2019, more than a third of Europeans would have preferred to replace their governments with AI, against a backdrop of hope based on a technology that is still almost unknown, coupled with a growing distrust of the political class.[2]

BIO

Laurent Chrzanovski holds a doctorate in Roman archaeology from the University of Lausanne, a post-doctoral research diploma in history and sociology from the Romanian Academy and a doctorate in history and related sciences. He is a professor at the doctoral school of «Lucian Blaga» University in Sibiu. He is the author/publisher of 42 books, over 150 scientific articles and as many articles for the general public.

In the field of cybersecurity, Laurent Chrzanovski is a member and contract consultant of the ITU expert group (UN-Geneva). He founded and directs the annual Cybersecurity Dialogues public-private partnership conferences. In the same spirit of public-private partnership, he is the co-founder and editor-in-chief of the only free quarterly cybersecurity awareness magazine, Cybersecurity Trends, published in Romanian, French, English and Italian. His main areas of research focus on the human relationship with the digital world, as well as the risky behaviour that results from a lack of awareness of the dangers of the virtual world; he also works on finding the right balance between the security and privacy of 'digital' citizens.

Since March 2022 and the launch of ChatGPT 3, accessible to the general public, distrust and scepticism have reigned, particularly with regard to the manipulation of data (text, images, videos) and therefore the veracity of the information accessible.

Citizens are increasingly fearful of attacks on their privacy and democracy. Paradoxically, they are very much in favour of the use of AI tools by their own state, provided they are informed[3], while more and more businesses and citizens are using AI products to increase and improve their professional productivity, or to facilitate many aspects of their private lives.

The inevitable convergence towards monopolies - of surveillance?

The American example[4] demonstrates that the last thirty years has seen a reduction in the number of economic players working in each area of the economy from 20 to 2 or 3, i.e. the emergence of virtual monopolies in almost every professional sector. This strategy of the «big» companies gradually absorbing the «small» ones has long been adopted by the web giants, especially GAFAM and BATX. As soon as a product or app became interesting, they bought it and integrated it into their range of services. Examples include Microsoft's takeover of Skype and Facebook's (Meta's) acquisition of Instagram and WhatsApp, not forgetting Twitter, now X, by the corporation headed by Elon Musk.

AI development has evolved from its nascent stages to full flourishing. In the future, after this period of market entry, a few major players will maintain their autonomy, the weaker ones will disappear, many others will

be relentlessly bought out, leading to the monopolies *described by Mika Lauhde in this volume*.

In view of what we have described above, this is all the more worrying because Big Tech players are «our best spies», acting with our own consent, according to Shoshana Zhuboff's masterpiece[5], constituting a capitalism of surveillance - with its backdoors open to certain States. This situation, in which we were already living long before AI, is therefore likely to continue, with a few powerful players who will have in their hands, thanks to AI, an exponential volume of confidential data on every company and every individual.

The winners of the moment: the major cybercrime groups

We have already mentioned the 'troll factories' generating a magma of fake news, untruths and subversive material, contaminating the AI world on a daily basis. But this is just the tip of the iceberg.

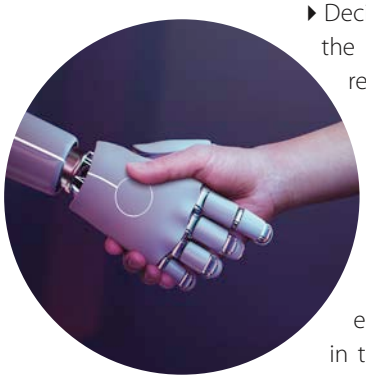
For more than two years now, the most sophisticated cyber attacks have all been designed using AI. As a result, in addition to the 'usual' type of attack, the weekly speciality reports show a veritable flood of exploits against the most robust systems, *in primis* cybersecurity tools and... AI tools (on this subject, we should mention the theft of hundreds of thousands of ChatGPT user data records, on several occasions).

Against this backdrop, the security market is evolving rapidly, but the latest-generation products - essential for insurance cover in the USA and the UK - are becoming increasingly expensive. Without them, it's impossible to cope with an AI attack, or to apply for insurance, as *Raj Meghani explains in this volume*.

Conclusion

AI is here to stay. The trick is to keep its everyday use in proportion, and only use it as a tool to complement (rather than replace) routine activities, taking care to keep the amount of private data it contains to a minimum.

For the rest, AI is simply adding fuel to the fire on the most sensitive safety points in our digital ecosystem. These are always the same, as highlighted in the 1,200 or so articles and interviews published by our magazine since it was founded ten years ago:



- Decision-makers' poor understanding of the difference between cybersecurity and resilience (of which cybersecurity is only the fundamental pillar).
- The lack of implementation of concrete policies for the convergence of all connected tools within a facility.
- The lack of investment in resources, people and continuing education in the field of cybersecurity in the broadest sense, both in the public sector and in the private sector, not to mention ordinary and vulnerable citizens.
- The difficulty of sorting and then protecting confidential data, both personal and professional, within the great magma known as «private data».

In addition, this will require implementing rigorous regulatory frameworks on a company-by-company basis, clearly delineating which AI applications will be permitted and which will be prohibited… ■

1 Luciano Floridi, The Ethics of Artificial Intelligence-Principles, Challenges, and Opportunities, Oxford University Press 2023
Luciano Floridi, The Ethics of Artificial Intelligence: exacerbated problems, renewed problems, unprecedented problems - Introduction to the Special Issue of the American Philosophical Quarterly dedicated to The Ethics of AI (April 20, 2024). Available at: http://dx.doi.org/10.2139/ssrn.4801799
2 EUROPEAN TECH INSIGHTS 2019 (available at: https://docs.ie.edu/cgc/European-Tech-Insights-2019.pdf/)
3 EUROPEAN TECH INSIGHTS 2024 (available at: https://static.ie.edu/CGC/European%20Tech%20Insights%202024%20-%20IE%20CGC.pdf/)
4 Jonathan Tepper, Denise Hearn, The Myth of Capitalism: Monopolies and the Death of Competition, 2018
5 Shoshana Zuboff, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, 2019

A publication

web for business
swiss webacademy

Redaction:

Laurent Chrzanovski and Romulus Maier (†)

Our deepest thanks to Raj Meghani for the English proofreading.

ISSN 2559 - 6136

ISSN-L 2559 - 6136

https://swissacademy.eu/


AI