Summary report 18th Security Talk: Generative AI - Powerful tool or fundamental security risk?

Presentation Katharina Fulterer

Opportunities, challenges and potential applications for artificial intelligence (AI)



What are the actual effects of AI on people? There are statements such as that of Bill Gates, who says that AI is our new mobile phone. Or Sundar Pichai from Google compares it to innovations such as fire and electricity. However, there are also voices such as Jeff Bezos, who says that we are only at the beginning of this development. So where are the potential areas of application, where can we as a company, as a public institution, really utilise AI in a meaningful way? The topic of route optimisation is certainly a big evergreen. Our navigation apps and fitness trackers help us in various areas of our lives. Large companies such as Swiss Post also use AI to optimise routes when delivering parcels. Various data, such as weather data, traffic volumes or parcel volumes, are taken into account to ensure the efficient use of resources. AI also makes a valuable contribution in the fight against fraud and is used by banks to check billions of transactions every day.

Generative AI as a new player

There are now many examples, e.g. in image generation, **where AI actually comes close to or even surpasses the performance of humans**. In the context of Generative AI, it is worth differentiating once again what exactly we are talking about. Ultimately, we are talking about intelligent machines or intelligent computer programmes - artificial intelligence. Various experts have been researching this topic since the 1950s. Be it in the field of machine learning or, in recent years, increasingly on neural networks, which fall into the area of deep learning.

If we take a closer look at deep learning, we can see two areas that show development potential today. On the one hand, there is the area of predictive AI, classic AI. This involves recognising patterns based on historical data and predicting future events. Examples of this include weather forecasts or route optimisation. In contrast, since the publication of ChatGPT, we have seen models that fall into the field of generative artificial intelligence. These involve generating completely new outputs on the basis of training data. These can be generated in textual form, as a voice, in images and many other modalities. **Generative artificial intelligence**

is a medium that allows us to enter into a dialogue with artificial intelligence. It is, so to speak, a dialogue-capable interface with which we humans can converse in our daily lives.

There are also applications that work much better today in combination with predictive and generative AI than they did a few years ago. This includes, for example, the translation of languages, testing, correction and generation of software code, as well as the automated creation of documents. And then there are areas where Gen-AI enables outputs that were not possible before. For example, personalized customer interaction or the topic of deepfakes.

Risks in the implementation of AI

In addition to a multitude of opportunities, the use of AI also harbours many risks that need to be addressed. These can be divided into three areas. There are risks that affect us as humans, e.g. if an AI produces results that contain a certain bias or reveal personal data. There are risks that affect companies, e.g. when data is reused that is subject to legal requirements or when employees are dissatisfied because they have to hand over certain tasks to an AI. There are also risks that affect society, such as the impact of artificial intelligence on Switzerland's security of supply. Ultimately, the aim is to develop trustworthy AI solutions and put them into practice.

Applications of Generative AI in Swiss companies

Where are applications with Generative AI being used by Swiss companies in the Swiss market today? The public discourse is not always transparent about how demanding the use of AI actually is in a corporate context. According to the OECD, only 8% of companies in the OECD area used AI in 2023. This is due to the complexity of implementation. After a successful PoC (proof of concept), companies often fail to really scale the project and transfer it to a productive setting. In the real business world, completely new influencing factors are added and the new AI solutions have to be connected to existing systems. Such projects often fail because certain security and governance issues were not sufficiently analysed and considered in an initial lean PoC. As a company, we therefore try to shed sufficient light on these issues and think about how we can design secure and trustworthy systems from the outset. One example of such a solution is the recording and collection of relevant information.

Generative AI in the area of Knowledge Management

Generative AI has very broad fields of application, but the one that is currently being pursued most actively by many companies in the Swiss market is in the context of knowledge management. How can we make accessible data available to our employees, and perhaps in a second step also to our customers, as efficiently and usefully as possible? Behind this is the ability to analyse and synthesise documents and create good summaries from them. There are very similar use cases in police work, especially when it comes to recording incidents. This is a relatively tedious process, which we are currently trying to automate in a PoC by using an

AI to record spoken language, create a protocol and summarise the important points again and transfer them to the existing systems.

Generative AI with chat- and voicebots

Chatbots and voicebots are another use case for AI. In their current version, these are often a real nuisance for customers. Generative AI is now able to understand the language spoken, not just High German but also the various Swiss dialects, recognise the customer's intention and derive the appropriate actions from this.

Generative AI in a military context

We also see many potential applications for generative AI in a military context. This could be the simulation of training scenarios, for example. Today, news articles, social media posts and Twitter messages are painstakingly analysed by hand to make simulations and training appear as realistic as possible. This is precisely where Generative AI can make an appropriate contribution with its creative power.

Presentation Stefan Preuss

Generative AI for marketing purposes



Generative AI is the easiest for marketing purposes. The technology works excellently. You can do so much with it. AI is everywhere. The next wave of text-to-video or text-to-image generation is currently underway. There are no guardrails content policies that dictate what is allowed and what is not. This also gives you the opportunity to generate problematic videos. Perhaps someone knows the tool Flux from Black Forest, a German company. It is one of the fastest image generators currently on the market. If you tell Flux to show a picture of a doctor, you get a white male doctor my age. So there's a bias in this tool because it can't interpret this task at all. If you compare it with a Google search and ask the same question, the result was the same five years ago, but today it's a bit more mixed. These tools therefore have their limits.

Failed AI applications

'The best AI fails 2020-2024": We started collecting major AI fails around 2020. A good example is a chess robot that is able to play chess and operate several boards simultaneously. However, it was not trained to allow the little boy it was playing against to simply reach into the playing field. The robot arm should then have stopped and walked out. However, it did not recognise the situation, grabbed and broke the boy's finger. A somewhat more recent case from McDonalds is about a relatively large project that had to be cancelled. They wanted to use generative AI in a drive-through to automate the ordering process. As a result, 260 iced teas were ordered and various parallel orders from other locations were included. We need to learn from these examples how errors occur and how they can be avoided.

Trust in AI as an essential component

Why should we trust AI? After the Wall of Shame, we initially have doubts about this technology. We use technology a thousand times in our daily lives and trust it a thousand times accordingly. We travel by train; we trust the train driver and SBB that the train will arrive. We trust the pilot flying our aeroplane that he will land it safely. In the end, **AI is also about trust**. There was a very interesting article on this by Bruce Schneier, a cyber security specialist. He defined four dimensions for when we can trust technology. Firstly, when the technology itself implies security mechanisms and is fully developed. If there are values within the company offering the technology that lie not only in maximising profits, but also in providing a product that works. Or also the awareness of one's own reputation, which is associated with this product. And finally, sensible regulation that also takes into account a system of penalties for misbehaviour and failure.

The question is, do I trust the technology and would I use it? Schneier has said that social trust characteristics are relevant for secure technology and regulation, while interpersonal trust characteristics are relevant for reputation and value. And when we talk about generative AI, we see a lot of humanity in this technology. AI responds in the same way that we speak. And that's why we are placing more emphasis on interpersonal trust and neglecting social trust a little too much.

Autonomous driving as an example

I take autonomous driving as an illustrative example, but this is not a specific Gen-AI topic. Car manufacturers all say that the technology is ready and working. As a critical observer, I would say that if autonomous driving is widespread, then the number of lorry drivers should also fall, which is not the case according to the figures in the USA and Europe. So this is completely at odds with the technology and the promise of the technology. Why is that? The technology is there. We have radar, GPS, 4G, sensors in modern vehicles that measure every possible value from 2 cm to 250 metres and calculate every driving scenario 10 seconds in advance. Today's modern vehicles are data centres on wheels. So the technology is ready. Is it a question of regulation? In Switzerland, there is now at least a draft for the regulation of autonomous driving. In Germany, autonomous driving is possible at level 5, up to 60 km/h on

the motorway (!). Neither the technology, the regulation nor the users are the problem. If we look at San Francisco, the largest experiment in autonomous driving is currently underway. There are around 1600 vehicles in use. If you look at the driving reports, it is statistically the safest way to drive. There have been no fatal accidents since the Uber accident six years ago. The technology is mature, but challenges remain.

Let's take a look at what can go wrong. Option one, the telephone network fails. Number two, the police stop an autonomous vehicle, the vehicle brakes briefly, the police get out, walk over, the vehicle accelerates and drives on, only to stop again 200 metres later. Number three, teenage gangs have fun covering the sensors with traffic cones, creating an artificial traffic jam. Number four, a human-driven vehicle hits a woman, the woman is thrown under an autonomous vehicle, the autonomous vehicle is not prepared for this situation. The autonomous vehicle stopped because it had recognised a stop sign.

AI works 95-99 per cent of the time, never 100 per cent

There are situations that you can't foresee. I have called it the 'fly-and-cow dilemma' because nobody expects a cow to fly through the neighbourhood and the vehicle to have to avoid it. So AI will work 95 or 99 per cent of the time, but it will never work 100 per cent of the time. That's why we have to be able to cushion this one per cent in such a way that no damage occurs and, in particular, no people are harmed.

To summarise, reality writes its own laws, even in connection with AI. Even if we implement all the dimensions described as well as possible and try to master them as well as possible. We have the feeling that we have to do everything with AI today, but not every problem can be solved with AI. It can perhaps be solved differently. Nevertheless, it is important to consider when AI and when human power or alternative solutions are suitable.

Presentation Dr. Thomas Rothacher

Technological development in the last century



Over the last 100 years, we have seen an exponential increase in technology transfer in our environment. In the past, the military was still a technology driver. Today, however, it is the civilian world that drives technology. The financial situation has changed. Furthermore, technological change has much more influence on our culture than we realise. The iPhone, which has been in use for 14 years, has disruptively changed the world. Google has given knowledge a different value and artificial intelligence will make this fact even clearer to us. Another phenomenon is the penetration of this technology, which is happening ever faster and more widely. It took Netflix three and a half years to reach one million subscribers/users. ChatGPT, on the other hand, took just five days to reach this milestone. Another graph shows how long it took for an algorithm to perform a certain task better than the average person. In the case of handwriting recognition, this took over ten years. If we go to 2018, it only took two years for speech recognition and understanding. **So there has been a huge acceleration in these skill developments**.

Impact of generative AI on cyber security

When we come to armasuisse, we have a major advantage over private-sector companies. If they want to sell a product, they have to be able to rule out all eventualities. You heard earlier that it could be difficult to achieve the last one per cent until AI is fully controlled. This is a little easier for S&T (armasuisse's Science and Technology department). S&T is concerned with the question of how we can bring new gadgets to the troops or into use while they are still at the experimental or test stage.

We asked ourselves where we should use this technology. There are certain competitions in the army where good plays against evil. As Switzerland, we took part in an international competition and finished in the bottom third in 2017. We then collected several years' worth of data and used it to programme and train algorithms. Thanks to the help of these algorithms, we finished in the top third of the Block Shields 19 competition. This form of AI has been in operational use in the army's systems since 2020.

AI's learning ability

There are already war games that can be played using artificial intelligence. For example, we can have two artificial intelligences play against each other. The aim here is to equip troops with different technologies and then position them against each other. The amazing thing is that the solution an artificial intelligence proposes is different from the one we would choose. Although artificial intelligence may have a bias, the bias of humans is normally even more pronounced. We have applied the technology in the Air Force's new tactical simulator. The artificial intelligence uses a large amount of data to calculate the best way for the enemy to attack our positions. From this, we in turn learn what this means for our systems - how they need to be designed and how we can best organise them.

Drones controlled by AI

Drones are one of the big issues that have been fuelled or accelerated by artificial intelligence. A Ucrainian officer who is in charge of these drone operations told us that in Ucraine they currently need a million drones every three months. Furthermore, they are destroying 95% of the protected vehicles with these drones because they do not have enough artillery ammunition. The drones are put into the air and then 'fire and forget' applies. This means they are no longer under human control but select their targets themselves using artificial intelligence and are constantly being developed further. Ucraine is the first country to introduce an unmanned unit as a separate species. This is a topic that we are already discussing in some areas of the army today.

Switzerland is also a leading nation in the field of AI and drones. Our universities are recognised as part of the world's largest talent pool. We have also been number one in the global innovation index for several years. Armasuisse has therefore worked to establish the drone task force in the last two months. The aim is to network the individual players in Switzerland and make a contribution to our security.

'It is not the most intelligent species that will survive, but the one that can adapt best.'

My conclusion: the speed of the developments mentioned is increasing exponentially, especially in the area of security. And we are now seeing revolutions in warfare that go beyond any disruptions. Furthermore, the nations around us are all arming themselves enormously. This is not my opinion as Head of S&T, but as Deputy Head of Armaments. The Swiss are in a bubble and are not aware of what is happening around us. However, we must orientate ourselves to our environment. After all, it won't be the most intelligent species that survives in the end, but the one that can adapt best.

Panel discussion

In addition to the two speakers, Stefan Preuss and Dr Thomas Rothacher, other experts took part in the panel: Jennifer Scurrell, PhD student at the Center for Security Studies at ETH Zurich, Patrick Fontana, Digital & App Innovation Specialist at Microsoft and Dr Peter Friedli, partner at Eraneos Switzerland. The panel was moderated by Fredy Müller, Managing Director of FORUM SICHERHEIT SCHWEIZ.



AI as fundamental change

Fredy Müller: Stefan Preuss, is the technological revolution we are experiencing with artificial intelligence a similarly big step in development as reading, writing or arithmetic?

Stefan Preuss: In my opinion, definitely yes. If you look at developments since the industrial revolution, from the steam engine to electricity, information technology and now artificial intelligence, it's definitely a revolutionary step, because there's data behind it. And we now have this data available to analyse, we can use it. If you look not only at generative AI, but also at image recognition processes in the medical field, it gives us an incredible boost.

Freddy Müller: Patrick Fontana, you have been working in the tech sector for 20 years. What has changed at Microsoft during this time?

Patrick Fontana: A lot. Firstly, we have taken the step from a pure sales and licensing organisation to a technology development and adaptation company. Artificial intelligence is the next step. We no longer simply talk about bringing out a product, but instead provide most technologies as services. A single company often no longer has the resources to do everything

independently. This always brings a twist to the company, because you have to think carefully about what the use case behind the AI applications is. All failed implementation attempts did not have a clearly defined use case.

Freddy Müller: So the aim is no longer to bring a product per se to the market, but a service or a complete system?

Patrick Fontana: On the one hand certainly a service, but on the other hand also an ecosystem.

Freddy Müller: Peter Friedli, you have been working in the security consulting sector for many years. How have you experienced this technological change?

Dr. Peter Friedli: I sense a great need for counselling services. As Patrick Fontana said, we should actually consider what problems we actually have and how we can solve them. It remains to be seen whether AI or generative AI is the solution.

Freddy Müller: Jennifer, you belong to the younger generation and are right in the middle of it with your research. How do you experience the hype surrounding AI and generative AI?

Jennifer Scurrell: I wouldn't describe AI as hype, but as a revolution. Especially given my background in human-air interaction. You can already see the impact that AI is having today. Be it chatbots, which are getting smarter and smarter, but these bots can also do a lot of good in the field of medicine or psychotherapy. I've been working with machine learning, deep learning and AI for around eight years and I'd like to say once again that this technology is absolutely revolutionary.

Understanding how AI works - not a prerequisite for trust

Freddy Müller: How well do people know AI and generative AI anyway?

Jennifer Scurrell: Surprisingly, a lot of people really do use ChatGPT, including my mum. I wouldn't say she knows how it works. But does she really need to know? It's more important to acquire general skills in dealing with technologies such as the internet, social media and AI. Critical thinking is essential when dealing with any technology.

Freddy Müller: Stefan Preuss, how much do you need to know about the technology from a business perspective? Should you just use it and not scrutinise it too much as a user?

Stefan Preuss: From the insurance company's point of view, this is perhaps not the best way to go, because you are then playing very heavily with customer trust. I would recommend a cautious and conscious use of such technologies, trying them out and comparing them. It becomes obvious how much a technology, for example in battlefield development, completely overturns a previously familiar behaviour and changes things. But of course this does not fit into every company scenario.

The six principles for the use of Responsible AI

Freddy Müller: Patrick Fontana, we know the six principles that are needed to use AI responsibly. Dealing with these principles is certainly a constant topic at Microsoft.



Patrick Fontana: The six principles for the use of responsible AI are definitely an issue, but ultimately they go right back to the topic we heard earlier. Because when we use a technology, we have to trust that the output is correct. In terms of principles, AI must be transparent, secure and traceable in order to prevent biases. If a company can stand behind it, then that creates trust. The question remains as to whether this needs to be laid down in law so that everyone adheres to it, or whether a self-declaration is sufficient without any consequences. This is the question behind the six principles of Responsible AI.

Regulation of AI

Freddy Müller: Is a self-declaration sufficient or do we need legal provisions for AI to prevent abuse?

Dr. Thomas Rothacher: In the military world, it's a bit difficult with regulation. I'm someone who doesn't really believe in this regulation. Because all those developments and functions that are possible will also materialise sooner or later. We need to think more about how we want to deal with this and not regulate something that is going to happen anyway. The EU's approach to such regulation is the worst possible way.

Jennifer Scurrell: I also agree from a scientific point of view. Sensible regulation is fine, but I find the EU's approach to the issue problematic. Progress and creativity are being hindered, which also inhibits Europe as a potential leader in the AI sector. That's why we need sensible but minor regulation.

Freddy Müller: It's a balancing act. Stefan Preuss, you showed it on a slide, it needs regulation.

Stefan Preuss: Two hearts beat in my chest in this regard. On the one hand, I believe that regulation is absolutely essential for a disruptive technology like AI. We missed the opportunity to regulate the internet and social media, and today we are painstakingly trying to close Pandora's books again. In this respect, I was glad that the EU AI Act was passed. However, it will take a few more years to define the objects that we want to regulate and to understand what we are talking about. In this respect, I am very ambivalent. From my perspective, I am

of course happy to have a target that I can use as a guide. But it still takes a lot of time to implement.

Freddy Müller: Patrick Fontana, do we need global or local regulation? It has been said that a lot has probably been missed in the area of social media. Microsoft is often criticised in this context.

Patrick Fontana: Of course, it would be much easier for us if everything was regulated globally somewhere. We could then harmonise all of our compliance processes with this regulation and have less work to do. The fact is that every single country and every organisation is currently trying to introduce such regulations. We have to prove ourselves to these regulations and comply with them. We are happy to do this because it promotes trust. We are a little ambivalent: Regulation yes, but only in such a way that innovation and development are not inhibited.

Social and political changes achieved through (generative) AI

Freddy Müller: Next, we will talk about the social and political changes generated by AI or generative AI. This is primarily the research area of Jennifer Scurrell.

Jennifer Scurrell: I'm mainly looking at how chatbots in social media can influence political opinion. The topic first hit the headlines in 2016 when it was suspected that Russian bots had influenced the US elections. In this case, scientists were divided. Some studies have shown that the bots actually influenced the results due to their large numbers. However, other studies show that these bots were not smart enough to effectively influence political opinion. With ChatGPT, we have a completely different situation because the bots are getting better and smarter. People sometimes can no longer distinguish bots from humans. In recent months, many studies have been published on this topic. It has been shown that these bots can influence other people at least as well as humans with propaganda. We already have a major problem here.

Dealing with bots

Freddy Müller: Patrick Fontana, how does a company like Microsoft deal with this bot issue?

Patrick Fontana: As Microsoft, we don't offer a bot that can do everything. Instead, we offer clear reference architectures and support customers directly. When using OpenAI, we have a two-zone security system. We have an automated grid recognition system for questions that should not be asked, such as questions about suicide or questions of a manipulative nature. For data protection reasons, this human component can also be removed. In this case, however, we as a company must protect ourselves, as we are responsible for ensuring that certain questions cannot be answered. One example of this is a proof of concept that we carried out with the army.

AI and the young generation

Freddy Müller: Peter Friedli, how does a consulting firm deal with this challenge?

Dr. Peter Friedli: We are almost all digital natives, or at least have become so. However, my 9-year-old daughter is entering a world in which everything digital is already normal. How do we get this generation to be able to deal with it? To a certain extent, it's a train-the-trainer principle, as we know it from the army. Are we as trainers prepared to pass this on? Are the teachers prepared to pass it on? I think these are also questions that can be transferred to companies.



Jennifer Scurrell: Many people, whether students or adults, are affected but not prepared for it. We are influenced by social media on a daily basis and bombarded with information. When ChatGPT came out, people wanted to ban it in schools and universities. But that's the wrong approach. We need to teach children from an early age how to use these technologies. This could be addressed once a week in a 'technology' subject. The aim would be to playfully try out which questions generate which answers, which data is used and how it is combined?

Freddy Müller: Should society, primarily in companies, be trained and made aware of what we are facing here?

Stefan Preuss: It's almost impossible to do this from a parent's perspective. When you see what is being communicated to children via social media and that AI is amplifying a lot of things, it's a very difficult battle. On the other hand, the technology impact assessment is exciting. For the first time, we are confronted with a technology that forces us to think about our basic ethical attitudes. Do we want to retain human-in-the-loop for weapons systems? Or what does it mean to integrate social scoring systems into society, not to use them in Switzerland, but to sell them?

Freddy Müller: Thomas Rothacher, what can we do to avoid being fooled by deepfakes, whether in civil society or in the military world?

Dr. Thomas Rothacher: I don't really understand the technology behind AI, but what I have learnt is to make certain plausibility considerations. With deepfake videos, it helps to think about how realistic what is being shown is. In my view, this critical thinking and critical discussion should be encouraged in a more targeted way.

Companies need to adapt the way they deal with AI

Freddy Müller: Patrick Fontana, I assume that Microsoft is in great demand at schools due to its expertise. Do you have the opportunity to appear there?

Patrick Fontana: We actively make ourselves available. However, demand is very low, especially at primary school level, as well as in the area of basic teaching or the eighth and ninth grades. Universities, on the other hand, are actively addressing the issue. Even when we are out and about in companies, we always say that an adaptation programme is needed. The first step is to learn how to use this technology and to properly validate the information provided.

Freddy Müller: Is there a need for a more professional approach to the internet, social media and AI?

Dr. Peter Friedli: In my view, there should be one where there is not yet a need. We often write AI strategies for companies that are out of the ideation phase. AI strategies already contain a technical part. However, it is more about the operational aspects, about processes and considerations regarding data outflow. A typical use case: I have to write a reply to an email. To do this, I copy the content, call up ChatGPT and have replies created. A large amount of data flows out during this process. No firewall can help against this. Finally, there is a kind of portfolio management and prioritisation of AI activity. The technology is therefore only a small part of these processes.

Successful application of AI in companies and public authorities

Freddy Müller: Let's move on to the third topic block, the use of AI in companies and by authorities. Stefan Preuss, you have vividly illustrated what can go wrong when using AI. What are the three most important elements for the successful use of AI?

Stefan Preuss: As already mentioned, AI is based on company data. This is the only way to generate added value for the company-specific context. The **information must be available in a suitable quality and quantity**. This is probably the biggest challenge for most companies. In addition, you should allow as many smaller scenarios as possible within a **strategy scenario before taking the big leap**. Because even we at Mobiliar cannot afford 99 per cent accuracy, we need 100 per cent. So before you take this step, you need a lot of experience

Freddy Müller: At this point, we would like to ask the audience: Do you use AI, do you already have experience? Who has a chatbot in use?

Georg Kaufmann: I am Chief of Staff in the Armed Forces Personnel Division. We are currently in the process of introducing a chatbot. We currently have a hotline that answers 30,000-40,000 enquiries a year. 60 per cent of these are by phone and 40 per cent by email. The head of the army now wants to integrate a new project, first-level support. This would require an additional 1.6 million staff, which is out of the question. That's why we ended up with the chatbot and are now in the MVP. We're currently carrying out troop trials. I will then present the project to the army leadership. They will then decide whether the chatbot will be used, as this also requires financial and human resources.

Patrick Fontana: The Training Command has been running a project for some time. We often get in touch with AI via the support channel. In this area, ChatGPT has introduced the component of speech-to-text understanding, which allows us to actively communicate with our data.

Generative AI is currently missing a relatively important component, namely the ability to gain an overview of the entire case and find different variants. We are limited here at the moment. That would be the next development step. Today, we are stabilising the entire AI phase. Do we even want to take the next step? First of all, ethical issues need to be clarified before the AI can act in a quasi-autonomous way, take stock of the situation and find variant.

Freddy Müller: We are talking about AGI, Artificial General Intelligence. What is your assessment, when will it arrive?

Patrick Fontana: As always, Microsoft plans a maximum of two years into the future with such roadmaps and then looks further ahead. Bill Gates said in his last podcast, The Next Big Thing, that he expects 10-15 years. I would agree with this opinion.

AI in a military context

Freddy Müller: We now come to the fourth thematic block, the military context. Thomas Rothacher, you said that Switzerland can play in the top league. What do we need to do to maintain our position behind the USA and China?

Dr. Thomas Rothacher: We are in third place in terms of research and innovation, but implementation is a little lacking. So how do we manage to manufacture important products in Switzerland, especially under more difficult conditions, for example when the borders are no longer so open in the event of a conflict? In my opinion, this is the greater challenge.



Dr. Peter Friedli: Funnily enough, two studies have just been published today by armasuisse on how Switzerland should be promoted as a centre of technology. One of the key points was to create transparency. I think that armasuisse and Swiss Innovation Forces are endeavouring to bring the start-up ecosystem in Switzerland closer to the defence industry and break down resistance there. Above all, the benefits for the armed forces and defence should become clear.

Patrick Fontana: As the highest-ranking officer at Microsoft Switzerland, I can also comment on this. The defence ecosystem works. The question is always how regional an action has to be. For example, do I trust an American company when I am in a conflict or does everything have to happen within national borders? We are in active dialogue with our American colleagues at DARPA (Defense Advanced Research Projects Agency), but we also implement corresponding use cases here in Switzerland, which in turn supports the local ecosystem.

Freddy Müller: How can we optimise the use of AI and generative AI in the police, justice and intelligence services?

Dr. Thomas Rothacher: There are already some examples of this. We use AI in the area of staff exercises, for example. How do I feed the system with my own data? We now use AI to develop scenarios, as well as to analyse data for fused images. AI is also used today to analyse text and images. The initial analyses are often already being carried out by AI and in this way AI helps those responsible to make decisions. So humans are still in the loop.

Critical questioning replaces knowledge

Nathalie Gratzer (Question from the audience): Is there a danger of artificial intelligence making us stupid? As mentioned earlier, AI extends my skills to a certain extent and provides me with solutions quickly without me having to think, reason and use my brain. If ChatGPT had existed when I was at school, I would probably have done very few tasks myself and only learnt half of them.

Thomas Rothacher: The relationship has changed completely and we have to deal with certain things differently today. I crammed knowledge at school. That has a performance value for me. Today, it's no longer difficult to acquire knowledge. The difficulty is evaluating this knowledge and making it plausible? I have no idea how to teach my children to distinguish truths from untruths. Dealing with the subject is a challenge.

Stefan Preuss: I see the point absolutely, the ability to scrutinise things critically is becoming increasingly important. When we look at a medical assistance system, an experienced doctor is able to judge whether its statements are meaningful and correct. Critical questioning is an important skill for us humans. The question is whether we can maintain this competence?

In my everyday life, it looks like this. I have three different chatbots open and ask all three bots the same question. I then compare the answers and continue working with the one that seems most plausible to me. I try to scrutinise the answers critically and always ask for the sources and links so that I can verify the information. This is a question of method, of methodological expertise.

Freddy Müller: Jennifer Scurrell, how much do you rely on AI and generative AI?

Jennifer Scurrell: I don't use it in my research and I don't take a possible answer that chat GPT gives me to my research question as the truth. To avoid becoming stupid, it is important to learn how to deal with such information and to scrutinise the knowledge that is spouted. We must also learn to ask the right questions. If we succeed in doing this, working with AI has great potential.

Patrick Fontana: We are now realising that we need courses to actively evaluate the relevant information. A chatbot and an AI assistant are helpful for making statements and gaining a few insights. However, we also need to be aware of the consequences behind this. This is a skill that we need to develop in the future.

Conclusion: Lisa Kontratieva



To summarise, we have seen that AI, and generative AI in particular, offers a great deal of potential. There are an infinite number of possible applications and not just chatbots, but also multimodal LLMs, e.g. for image or video generation. We can automate processes. The possible applications are very diverse. And yet there is a risk that something will not work or that we will be deceived by very truthful content. So there is a security risk, an ethical risk, a political risk and technological challenges. Switzerland now has the opportunity to take a leading role in the field of AI. We can achieve this if we also actively manage these risks and challenges.