Important concepts & everyday business life:

Large language models & Zero Trust in practical use.





Welcome back, dear curious readers!

hen Alexander von Humabout this in our article from page 26 and the boldt set off on his most fainterview with Executive Manager Jan Ciupka mous expedition, the Amerfrom page 42. ican voyage, in 1799, it was And as usual, you can also get to know us preceded by years of preparation. He spent better! In the cabin talks on pages 24 and 46, we many years acquiring scientific training, seintroduce you to two of our exciting colleagues. curing funds (much of it from his own fortune) Our travel report is rounded off with an and collecting the necessary equipment. His Al playlist that you are guaranteed never to thorough preparation also enabled him to perhave seen before and a text by our namesake form a large number of scientific observations Humboldt that simply couldn't be better suited and measurements that were invaluable for to today's world. subsequent research.

The editorial group behind our travel report is now called the Humboldt Group, because we have always been truly captivated by the exploratory spirit of our namesake. And also by the way he conducted his work and used it to derive practical benefits for science.

Similarly, we have spent years preparing ALAN, our own large language model. Our initial thoughts around artificial intelligence for our clients in 2011 resulted in what we consider to be the most pragmatic system for European customers. Read about how this came about and learn more about what is important to us from page 6 in our look behind the scenes.

The introduction of Zero Trust in IT is also more of a journey than a short-term action. Almost everyone has already heard about this topic, but just what exactly is behind it and what can be done pragmatically? We explain more





With this edition, we hope that we can once again bring you closer to our engagement with the world, science, business and new topics and give you valuable impetus for your daily life. Please feel free to send us your feedback:

redaktion@humboldt-gruppe.com.

Because we also want to keep on learning. Just like Humboldt: "The most dangerous worldviews are the worldviews of those who have never viewed the world."

Warmly yours

Andry Fischer

Dr. Andrej Fischer **Executive Management**



Contents



12 -

22 -

31 -

-Gruppe

o6 —	The journey to our own artificial intelligence
	Comma Soft presents ALAN
12 —	From fun to day-to-day business: Why GenAl is not finding its way into the company
14 —	ALAN: A solution made for our customers' everyday lives
22 —	GenAl in use: An excerpt from our customer projects
	NI 1 I . I
26 —	No trust please Zero Trust and its actual use
31 —	The basic principles of Zero Trust
32 —	How to build a Zero Trust architecture
40 —	Zero Trust is the future
42 —	Is Zero Trust only a corporate matter?
	Interview with Dr. Jan Ciupka
	Also in this issue:
24 —	Start-up enthusiasm with Dr. Hua-jing Han
46 —	The cabin talk with COO Benjamin Schulte
48 —	Peterberger Gespräche 2023: A showroom for business, technology and science
62 —	From Humboldt's "Cosmos: A Sketch of a Physical Description of the Universe"
64 —	Al tools to try out
66 —	Imprint



The journey to our own artificial intelligence **Comma Soft presents ALAN**

s if it had fallen from the sky - this is how the rapid development of generative AI and large language models appears to some. Although it was way back in 2011 that we started to look at the impact of the topic and, above all, its benefits. Like the eponym of our editorial group,

Alexander von Humboldt, we refined our tools after the initial theses, developed our own approaches and prepared ourselves for the great departure that was to come in 2023. But you can read about it for yourself - a look behind the scenes of ALAN, our practical AI for European companies.

2011 - 2018

Where the journey began...

A data state of mind. Between a cool head and a hot heart." - this was the motto of the Petersberg Talks 2011¹.

It formed the basis for all the data and AI developments that are still driving us today. With cool, analytical minds and hearts on fire with enthusiasm, we have dedicated ourselves to the topics of data science and AI ever since. Because one thing was already clear to us at the time: artificial intelligence has an unimagined transformative power that will have an impact on companies, science and society - and is

already having an impact, as we are seeing today with the rapid development of generative Al. Alongside technology, social and ethical issues have always been part of the journey for us. In 2017, a policy paper² and a discussion forum³ on AI were drawn up, which - again as part of the Petersberg Talks - focused on the issue of dealing with AI in a responsible and beneficial way.

AI is starting to grow out of its fledgling stage

In the first ten years of our Al journey, the gap between AI hype and reality in companies was often still striking: machine learning models encountered legacy IT infrastructures, PoCs encountered IT systems, Python faced mainframe.

"The old myth that 70-80% of data is not used, but has to be searched for, collected and cleaned up, was indeed often the case. Most of the time, the missing or inappropriate data was the sticking point," remembers our LLM evangelist



Dr. Andrej Fischer recalling the first years in Data Science Consulting. From the beginning, however, NLP (natural language processing) projects were particularly rewarding, mostly in the context of customer communication.

There were usually more than enough of them. And companies were desperate to find ways to master the flood of communication, direct customer inquiries to the right departments or find out more about customers and their attitude to service and products. A customer complaint should be recognized and dealt with



directly as such and not get bogged down in departmental ping-pong. There is a lot of relevant information in the spoken (or written) word. And to this day, making them practicable, usable and analyzable is of great value to companies.

At the end of the 2010s, a new movement took off in the NLP universe - very slowly at

¹ https://petersberger-gespraeche.de/jahr/2011/

² https://www.petersberger-gespraeche.de/petersbergererklaerung/

³ https://www.petersberger-gespraeche.de/thesen/

first, then increasingly faster. There were ever more attempts to model text semantically, i.e. to capture the meaning of text, as AI was now also able to independently find patterns and correlations in data. Special applications from the field of deep learning opened up completely new possibilities for our projects, such as finetuning to industry-specific texts. "At the time, we built a customized deep learning model for

an insurer that was trained on tens of thousands of documents from the input management system. When it came to classifying documents, these models were unbeatable," says machine learning expert Dr. Lars Flöer, recalling the early forays into this new world.

When AI feels with us

n retrospect, the decisive moment for the development of large language models as we know them today was certainly the publication of the paper "Attention is All You Need" in 2017. Here, a team of Google experts presented the so-called "transformer architecture". This specialist machine learning technology is fundamental to today's language models and aroused our interest at the time. They were impressive in their ability not only to recognize words, but also to use the context of a word. A classic example: sentiment analysis, with which emotions could suddenly be identified "between the lines", e.g. irony, joy or anger. A real benefit for companies that receive tens of thousands of messages from customers every day and can now automatically prioritize them according to "emotional urgency".

The biggest benefit was that these models no longer had to be taught to understand language from scratch. They already possessed a basic understanding of language and could therefore be deployed much more easily in new areas comparable to a person who learns a foreign language more easily if they have already mastered their mother tongue. If they are not good at their first language, learning other languages will be much more difficult. However, we and the community were slow to realize that these models also enabled a completely new mode of interaction.

GenAI-Glossar

Large language model - LLM - Generative Al - Prompting: The terminology can be confusing. If you are interested in delving deeper into the topic, you can find more details on the most important terms here.



2018 - 2022

Talk to me! In dialog with the LLM

It was in 2018 that the first forerunners of generative language models emerged, where the focus was on producing language (e.g. in response to a question). The models became progressively larger and were trained with an increasing amount of data. It became clear during the further development of GPT-1 (2018) to GPT-3 (2020) that bigger really is better. Trillions of text tokens were used to adjust trillions of parameters in the AI models - and thousands of graphics cards were brought to glowing point in the process. The amazing thing: the ability of these models not only to generate grammatically and syntactically accurate text, but also to reproduce facts, argue logically, understand irony and analyze increased with each successive iteration. "We have discussed many different ideas with our customers on how this could be used in practice: chatbots were naturally an obvious choice, but so were automated email replies

Page 10

and better, more semantically adept analysis and research assistants. However, the implementation in projects was very challenging, as there were simply no usable ecosystems for this yet," explains Dr. Sebastian Schönnenbeck, who, in addition to developing AI solutions, also frequently discusses this topic with customers through AI training.

At the end of 2022, OpenAl then published ChatGPT, a simple chat interface for GPT-3.5 on the web, free for the whole world. The ChatGPT hype was born and it changed the way people interact with machines forever. Real, powerful Al was suddenly accessible - "at your fingertips" - could answer questions and even engage in real dialog with humans. This epoch-making development made Al available to everyone in their everyday lives, at work or in their private lives.

From leisure time fun to day-to-day business

Everyday work is also the key word. Because our claim at Comma Soft is clear: technology must benefit companies and people, not the other way around. ChatGPT as the leading standard solution and the first open source alternatives were both entertaining and useful in private, but hardly applicable in a business context - especially for European companies.

Why many GenAl tools do NOT make it into the company:

- Data security: One of the main reasons for opening ChatGPT to the 01 general public was certainly to collect new training data. This is problematic if you then enter your intellectual property in the form of software in order to get tips on the code, as Samsung engineers have done.⁴
- Superficiality: The language models were only aware of the world in 02 line with their training with data from the web. But the most relevant knowledge for companies is usually available internally: manuals, experience reports, repair and test reports, etc. This data was not part of the pre-training and will not be in the future.
- **Language problems:** Although the early GPT versions had a very good command of English, other languages were not mastered to the same extent, as they were significantly less represented in the training data. This is no longer the case for ChatGPT, but it is still a major issue for many open source models.
- 04 **Blinkers:** Most commercial solutions are designed for individual use: a person in dialog with a machine. However, a company is above all a network of people who want to help each other and share knowledge and solutions. Solutions such as today's GPT store must therefore be developed further for companies and sectors.



- **05** Unreliability: Each subsequent version of the language model tried to reconcile even more skills with so-called alignment - i.e. desirable behavior. However, this sometimes leads to undesirable (and above all unforeseen) effects and, to this day, to fluctuating quality. This is a no-go for productively used services.
- **Costs:** A Plus or Enterprise account with ChatGPT or Copilot licenses quickly become expensive if tens of thousands of employees are going to use the technology on a daily basis or if the models are used hundreds of thousands of times in volume processes. Monopolies are never good for prices.

⁴https://www.golem.de/news/kuenstliche-intelligenz-samsung-ingenieure-leaken-internedaten-an-chatgpt-2304-173220.html

Gruppe



2023 - today

ALAN

A solution made for our customers' everyday lives

ust like many other companies later (and often more publicly), we built our own "CommaGPT" based on ChatGPT in early 2023. This in-house version enabled us to use language models for a wide range of applications: to help with coding, structuring documents, translating text and much more. We were not only power users ourselves, but also helped our first customers to use this technology. In spring 2023, another opportunity presented itself to us: open source language models had their breakthrough. The first real and serious alternatives to ChatGPT were the openly available models, e.g. Meta and Mistral.

"For us, this was confirmation that we were on the right track with the development of our own secure LLM, which was already underway at the time. So we developed our own solution further

and were subsequently able to offer it to our customers: ALAN was born. It is our response to our customers' need for greater autonomy and safety when using this key technology," explains Michael Tannenbaum, who heads up product development for our LLM.

Beneath the surface – a lot to do...

What impressed many of the first users of ChatGPT and Co. was certainly its ease of use. It was suddenly possible to use AI without any tech know-how or programming skills. To ensure that this could also work with the open source models, there were many problems to solve under the surface. So our Comma Soft development team began working with project practitioners to train our own LLM for use in companies. Day and night, our LLM went through the following learning process:

Lesson 1 - German

Most of the powerful openly available language models were mainly trained on English material. German, for example, accounted for only 0.17% of the training material in Meta's Llama 2 model. As a result, they have little or no German language skills. So German lessons were the first thing on the agenda. This was no easy task, as it was not just a question of training the models in a new language, but also of grasping the nuances and peculiarities of the German language.

So we collected and curated documents, contracts, communications, data and lots of other information. However, all these materials could not simply be used for training. Dr. Sebastian Schönenbeck, who took care of exactly that, remembers very clearly just what a challenge it was: "The open source models were specially trained for chat interactions. This is precisely why they also needed training material in dialog format. For example, if we had used a lengthy documentation for training, the model would have become poorer because it would have learned to give answers in a documentary style. We therefore first had to convert our materials into a question-and-answer format - which we did using a specially developed LLM." Only after this sub-LLM had processed the information could it be used for ALAN's German training.

Lesson 2 - European context

Besides language, it was also critical to embed European contextual knowledge and specific industry knowledge into the models, which were often underrepresented in the original training data. We used European legal texts and other public sources, for example. The challenging part here was that the basic model was already tuned "ex works" for chatting, i.e. for question and answer mode. So here, too, we had to take the detour via a second LLM. In order for our LLM to be able to provide answers in the right context, it first needed context to which questions and answers had to be generated - a bit like Jeopardy.

Lesson 3 - Weight loss and training

Whereas the early models were relatively small and manageable, today's state-of-the-art large language models have become extremely large indeed. The 70B in Llama 2 stands for 70 billion numerical parameters, each of which occupies 16 bits. With Grok-1 from xAI, there are even 314 billion parameters. This is larger than the memory of any graphics card available on the market. In order to be able to run ALAN costefficiently on-premises on a single graphics card, our LLM had to go through a "weight loss program". Today, our customers can actually





run it on individual A100 graphics cards - and there is even room left over for fine-tuning. This made it possible for our customers to deploy the system on-premises and to use our scalable asa-service model. To ensure data protection and data security for the latter, we deliberately opted for the Telekom Cloud to ensure that customer data remains in Germany.

Lesson 4 - Integration

Jiři

Many companies have either dealt with the pre-training of foundation models or the hosting of open source models, but have not dealt with the operational running in companies. Our vision was and is to develop a product that companies can use productively. To achieve this, we had to ensure that our large language model was not only powerful, but also intuitive to use. It had to be user-friendly on the one hand and integrate well with the company's systems on the other.

To achieve this, we have developed a chat interface that contains the tried-and-tested el-

Gruppe

ements of familiar chat tools and expanded it for use in companies. We have also developed a comprehensive technical interface (API) that companies can use to flexibly connect their systems and applications.

Lesson 5 - Reading and arithmetic

An important lesson learned from the first customer projects involved the integration of internal company data. While tools such as ChatGPT have broad knowledge of the world through the internet, they lack specific knowledge of individual companies and their internal affairs. To do this, our LLM had to learn how to handle company-specific knowledge and data. The solution we found: Retrieval augmented generation (RAG) and plugins that allow internal documents and data to be linked directly to our language models. With these solutions, users can today access their company's entire wealth of knowledge without the need for time-consuming retraining of the LLM for this specific case. The best thing about it: company knowledge



in LLM is always up-to-date and the generated results contain a source reference to the information used, in which the individual access rights are taken 100% into account. This means that sensitive information remains protected

Independence and practical relevance – a GenAI startup "the Comma Soft way"

any large providers of generative AI operate as tech startups with large investments from various financial backers.

And for the pre-training of large foundation models (the glowing graphics cards ...) this is also necessary. However, this was and is not an option for us, as our CEO Benjamin Schulte emphasizes: "As a family business, our independence is extremely important to us. That is



Behind the scenes: the ALAN product team during development

- and the results of the LLM can be transparently verified if required.
- That was just a small excerpt from our LLM curriculum. The first MVP was ready to take on the big boys in the fall of 2023.

why we are developing our LLM entirely with our own financial resources. This is the only way we can operate independently of third parties and develop GenAl in such a way that the focus is always on customer benefit."

Our initial tech team was built up with these in-house resources. It still consists of experienced data science and machine learning experts such as Michael Tannenbaum, Dr. Lars Flöer and many other esteemed colleagues.



The authors:







Dr. Hua-Jing Han



Dr. Carsten Draschner



Dr. Andrei Fischer

Alongside machine learning and MLOps, we also needed a front end, business development, marketing, partner management, contracting, legal etc. etc. What had we let ourselves in for?

The fact that most of the people involved were still working on real customer projects at the same time brought a lot of practical relevance and pragmatism, but also required a lot of discipline in planning and implementation. What motivated us was the vision behind ALAN: each day we could see this LLM taking shape and becoming a reality. With hindsight, it is fair to say that such a development, achieved in this very special way, was probably only ever possible thanks to the special framework conditions of an independent family business with a courageous entrepreneurial family and an inquisitive team.

Today, we look forward to Michael Tannenbaum's "Weekly Product Update" every Monday and are always amazed at what is possible in a week. Our internal feedback and sharing channels are now ablaze. Of course, we are still power users ourselves and have countless ideas on how we want to develop ALAN further.

Meet ALAN!

Would you like to get to know our LLM? We would be delighted to show you in a personal live demo. You can find more information and book your demo appointment here.



What's next?

Today, our LLM is already being used by many of our customers. But the journey is only just getting started: in the near future, we will be training and offering special industry models and equipping our LLM with even more skills.

It will be used in the development and operation of software, in communication with com-

Gruppe

pany data platforms and also in a community platform for LLM workflows - and will of course have a multimodal extension to image, video and sound. This is how AI can open up unimagined opportunities for creativity and make work easier for all people and all companies. We're paving the way for this.

GenAI in action

We have implemented various GenAI scenarios with our customers on our journey. Here you will find three examples that simultaneously reflect the development from ChatGPT to ALAN:

Baloise

From secure ChatGPT use to GenAl literacy

The Swiss subsidiary of insurer Baloise decided to familiarize its employees with GenAl at the beginning of 2023 and offer them a secure way to use it. There is a lot of potential here, especially in knowledge management: employees could obtain consolidated answers to their questions there instead of having to manually search through countless documents and databases in various systems. However, the direct use of ChatGPT is not possible due to regulatory and compliance requirements. The use of Microsoft Azure OpenAI Services and clear specifications for approved use cases offered a way out. If employees use ChatGPT via this access, the data entered is not used to train ChatGPT. Employees can now chat securely with the tool. Over time, further use cases were created that make work easier for employees in regular processes. Other national companies will now also be granted access in stages. What stood out when planning the introduction of GenAl: new technology on its own is not enough. Employees also need to be involved. Through special training and change programs, they were sensitized to the safe handling of generative AI, making GenAl literacy part of the corporate culture.

BITMARCK GenAl helping the helpdesk

BITMARCK supports statutory health insurance companies with digitalization. With more and more health insurance companies showing interest in generative AI, the managed service provider is helping them to gain access to this key technology while complying with the strict guidelines of the healthcare market and protecting patient data. Currently, employees are testing out using ALAN in the helpdesk themselves. In future, support requests from health insurance companies will be analyzed here and forwarded to the right contact person or department. They will then also receive suggested solutions for the tickets, including relevant articles from the BITMARCK knowledge database. The LLM can also help maintain this database and generate new entries by extracting information from new ticket cases. This would also be helpful when onboarding new colleagues. After all, having access to a well-maintained and updated knowledge database requires less training time - and the health insurance companies would in turn receive competent information on their concerns more guickly. As soon as the solution has been rolled out, additional sources will be connected and further possible uses for the LLM will be explored.

APPL

It also works without chat

Using LLMs beyond chats - how does it work? One example is GenAl at APPL. This group of companies from the printing industry has integrated ALAN into its ticket system for machine faults. Shift supervisors record faults or defects directly with their smartphones in production. Information on the machines, fault types, fault descriptions and images of the faulty machine parts can then be automatically sent to the LLM at the touch of a button. The respective machine manuals are also stored there as a knowledge database. If a machine malfunction then occurs, employees receive suggestions for solutions directly at the machine. Another practical feature is that image recognition allows errors to be analyzed and diagnosed directly in the ticket system. This means that machine faults can be rectified more quickly and downtimes minimized.

Gruppe



Below deck Start-up enthusiasm with strategist Dr. Hua-jing Han, also known as Jingjing

Jingjing, at Comma Soft you are responsible for strategy topics and work with clients on their corporate strategy, which is closely linked to a digital and IT strategy. Now you've hinted that you also do this in your private life. How might one imagine that?

At university, I was already involved in innovation together with my partner: we supported startups with their establishment and their first steps out of academia. The main focus here was on offering exchange formats, inspiration and impetus - and of course sparring, i.e. the exchange of ideas and the illumination of certain aspects from different perspectives. And brainstorming, thinking about new ideas together. The development of this initiative was also a bit like a mini start-up: you build something from scratch and then there's something that wasn't there before, with people who are suddenly working on a shared vision. How do we do marketing to reach more people, how do we find sponsors, how do we find interesting speakers for our events?

If friends and acquaintances and alumni are still involved with start-ups today, are about to found a start-up or have already founded one, then I am still more than happy to offer support.

What fascinates you about start-ups?

Start-ups are a passion of mine. There is something magical about building something new when something new is created. A bit like Herman Hesse's "a magic dwells in each beginning". I just find it so fascinating how perceived boundaries shift when you develop and establish a start-up and suddenly things that you thought were impossible actually work. This way of doing things differently – and of course this kind of scalability – has such a different aesthetic and directness to it that I find its creation super exciting. I can definitely sense the spark inside you by the way you talk about it. Are there any other things you like to do to clear your head? After all, in your job you're challenged mentally the entire day.

Sometimes I feel a bit jealous of people who have had this ONE hobby for years. I like to say that I am an empiricist, since I also come from empirical behavioral research, and my life currently consists more of sampling: I usually try something new every few months. At the moment, for example, I'm crocheting (laughs) ... I'm not very good at it yet though.

Nevertheless, there are some constants that are super diverse and changeable. For instance, I'm a total foodie and like to go the extra mile



Dr. Hua-Jing Han, Senior Consultant Data Business for the Insurance & Banking sectors

Gruppe

for good food. I love music and enjoy going to concerts, especially classical ones. And I love planning trips!

And is everything planned out and does it work according to a certain strategy?

On the one hand, I am aware that I generally think more strategically in terms of my thought patterns and also approach things in life quite strategically. But I try to stop strategizing during breaks and after work. My dog helps me with this, because he comes racing up to me and insists that we should do something that is completely unstrategic. His name is Poffertjes, or Poffi for short.

So another new task for your leisure time?

"Task" sounds like a to-do list, and we probably all have enough of those... For me, what's more important here is what I get back, and that's another family member. I always say that dogs are good for the soul. Incidentally, Poffi is a Shiba Inu, a national treasure in Japan and genetically one of the dog breeds closest to the wolf. Impatience is not a helpful quality when it comes to his training, he has a mind of his own. Stability and guidance – in other words, everything that constitutes a good strategy – are on the agenda for me here. He has also definitely made me more relaxed and patient.

Thank you for the insights and the little tour of your cabin, Jingjing!



No trust please

Zero Trust and its real-life use

rust is the glue that holds our society together. Trust is important in there is one area that must forget about trust and that is the complex IT system of companies. Because even the slightest mistake may jeopardize the security of the entire organisation. Nothing new so far - we all know that. But how can this knowledge be used systematically? The Zero Trust approach reveals it to us - and uses the example of Siemens' Zero Trust journey to show you how it can succeed. In an interview with Executive Manager Jan Ciupka, you can find out how the topic can be pragmatically implemented in SMEs and what the first steps are.

"It was a QR code that caused my company to go bankrupt!" - Sounds far-fetched, but it's not at all unrealistic.

Gruppe

How does this come about?

- business. We all want to trust. But 1. A QR code at a trade fair does not actually link you to the site map, but to an app containing malware.
 - 2. The app accesses the user's login data the next time they log in to a company website.
 - 3. The compromised account can now be used to take over other accounts, including an admin account, via shared file storage and social engineering.
 - The admin account, in turn, can now be 4. used to encrypt the company's infrastructure, including email, ERP, production-related systems and accounting.
 - 5. Recovery takes weeks and the company's liquidity collapses due to a lack of production and invoicing.

In order to counteract such a scenario, a modern and effective IT security approach is necessary: ► Zero Trust. Zero Trust is "not a technology that you can simply introduce or migrate to at short notice"¹ says Thomas Müller-Lynch, Global Director Digital Identities at Siemens. There needs to be a rethink and a paradigm shift. Rather, Zero Trust is - as Siemens puts it -"not a destination, it is ultimately a journey"². We embarked on this Zero Trust journey together with Siemens and learned a lot in the process.

You can read the most important findings and insights from this ongoing journey below between the green markers. <

High castle walls were yesterday

Before Zero Trust, the perimeter approach dominated the security sector. This can be explained simply using the principle of a castle: the enemy 3.



Thomas Müller Lynch, Global Director Digital Identities at Siemens

is expected outside the walls, so they are built high and thick, and everything contained within these walls is considered safe.

The origins of the Zero Trust principle date back to the early 2000s³. Alongside the professionalization and industrialization of the cybercrime industry and the increasing accessibility of IT security information thanks to the internet, three facts led to a rethink in the area of cyber security at this time:

- The way in which IT resources and services were being consumed and from where was undergoing a fundamental change even then. Access from anywhere and via a wide variety of devices had to be possible or had to become possible.
- 2. Through cooperation, collaboration and service providers, "external parties" have needed access to a company's IT resources ever since.
- Companies began to realize that IT security is more than just an annoying evil; it is essential for survival.

Since then, the simple perimeter approach has become outdated and the same destiny awaits it as once awaited the castle walls. Relying solely on the perimeter approach today is like using a castle from the Hundred Years' War as a defense against modern equipment such as helicopters, missiles and drones – a futile endeavor.

¹https://comma-soft.com/blog/deep-dive-mit-siemenschancen-challenges-von-zero-trust/, 30.01.2024 ²https://blog.siemens.com/de/2023/10/erster-schritt-vonsiemens-in-die-zero-trust-zukunft-fur-itot/, 08.01.2024 ³ JERICHO: https://www.blackhat.com/presentations/bhusa-04/bh-us-04-simmonds.pdf, 01.03.2024



What "Zero Trust" actually means in cyber security

he term "Zero Trust" probably first appeared in the cyber security context in 2010 in a Forrester⁴ model paper.

The core idea that a corporate network must not be regarded as intrinsically secure is stated clearly there. Shortly before this, the so-called Operation Aurora took place, in which Chinese hackers carried out major attacks on large US companies, including Google. Large amounts of data were stolen in this attack.

In response, Google published a series of white papers in 2014 and the following years and, us-

ing BeyondCorp⁵ as an example, implemented a model that does not grant implicit trust to any asset such as a user, device or account and makes all access to any type of data or resources dependent on the trust status of the asset something that is checked regularly. This is the model that Google itself uses as a benchmark for combating attacks and phishing.

 ⁴ https://www.forrester.com/report/Build-Security-Into-Your-Networks-DNA-The-Zero-Trust-Network-Architecture/ RES57047 (01.03.2024, Aktualisierte Version von 2012)
⁵ https://beyondcorp.com/ • At Siemens too, IT in particular has long been undergoing a transformation process in which security is one of the key objectives. Siemens, as a technology company, has long been conscious of the far-reaching consequences of a successful cyber attack, as well as the fact that, as a globally known and operating company, it is itself continually the target of cybercrime. For this security to have a holistic effect in the company, it is necessary for tools and processes to be integrated into this transformation, particularly within the individual specialist areas.

• The groundwork for the transformation was laid back in 2013, when Siemens began consolidating its identity and authentication solution. Originally distributed across many different instances, it was migrated to a standardized operating and security model. These efforts were the foundation for the later Zero Trust initiative at Siemens. This is because standardization and a central collection of configurations and logs are essential in order to be able to make dynamic decisions about access at a later date.

In Comma Soft, Siemens found a partner who was able to analyze and understand the complex environment and develop a comprehensive concept for the future technology stack, infrastructure and operations - including all the technological and organizational peculiarities and characteristics that can be found in a company of this size. П



Basic principles of Zero Trust

The basic principles of Zero Trust

But first back to what Zero Trust is. A NIST paper, which was published in its final version in 2020, describes Zero Trust in a somewhat more abstract way. In this paper, this concept is described as a collection of various guidelines for IT operations and processes as well as for the system architecture itself with the aim of improving the security of corporate IT. Three core ideas can be summarized as follows:

01 Assume breach – Design and treat your IT systems under the assumption that they tain processes, such as a dual control principle for system-critical activities

2 Least privilege – For each task to be performed with an IT system, only the authorizations that are required to perform the task should be assigned. When applied to an everyday example, this means that people who only need read-only access to a directory, for example, should also only be given read permissions. Access rights must be evaluated before any access is granted.

Dynamic access control – The decision whether or not to grant access must be able to adapt dynamically based on data available at the time. This automatically means that every access must be checked. If, for example, the device is no longer recognized as trustworthy between two accesses from an end device, subsequent access should be blocked.

In 2017, three years before the publication of the final paper, Siemens took another major step towards its own IT security when the company introduced M365 services globally. This made it possible to use behavior-based risk assessments of user activities in real time and thus also to dynamically control access. This enabled further technological innovations to be implemented more and more quickly, something that was specifically accelerated two years later in large-scale projects to improve M365 and Azure AD security.

The continuous support of Comma Soft's experts helped with the evaluation and implementation of security features, especially in the Microsoft cloud, which were and still are being developed at a rapid pace.

Gruppe

have already been infiltrated by an attacker. This already leads to conclusions for cer-



How to build a Zero Trust architecture

hen introducing Zero Trust in a company, it makes sense to divide the analysis into five sub-areas: identity, devices, network, applications, and finally monitoring and automation. It may seem overwhelming at first glance, so let's delve into a brief tour of these areas.

Identity: Are you real?

Identity is one of the most important points in the implementation of Zero Trust, as it involves permanent identification: a change in the security paradigm away from a corporate network that is assumed to be secure and towards an identity-centric security approach. Proof of identity is therefore a basic requirement for correct authorization. In this context, proof of natural identity is important (e.g. checking an ID card), as is the legitimate use of a user account, which is achieved in the Zero Trust approach, for example, through multi-factor authentication (MFA).

Devices: Everything - and a lot more besides

When it comes to devices used in the company, the focus is on identifying them and evaluating their health status. Critical assets can only be protected if users access them from known and verifiably securely configured end devices.

Siemens recognized this basic prerequisite and, together with Comma Soft, developed a solution that focuses primarily on systems outside of Microsoft Windows. The fact that we were allowed to delve into the depths of operating systems and compare long, green strings of numbers on two black screens in order to understand the mechanisms was all part of the Photo: David Clode

fun. But what does this have to do with the identification of devices in companies?

An identity that is physically tied to the device and cannot simply be transferred to other devices is ideal. There are indeed approaches that make use of hardware security modules such as TPM chips and store (private) cryptographic keys there. However, validating that such a chip has been used is a major challenge for processes and technical implementation. The alternative is to trust that such a chip has been used - an obvious discrepancy to the basic idea of the concept. And this dilemma is just the beginning, as the next question that arises is how to integrate it into existing applications. This is because once the key has been stored in the chip, you also need the corresponding interfaces to be able to use it. A truly comprehensive solution for devices involves an enormous amount of additional work. Companies therefore need to



5 pillars of Zero Trust

think carefully about where they really need keys stored in chips and where a software-based key in conjunction with measures from other Zero Trust pillars is sufficient or even more sensible.

Network: More than an intranet

A common misunderstanding arises regarding the role of the network in Zero Trust. Some simply equate Zero Trust with the abolition of the internal company network and hosting in the cloud, but this does not quite do it justice. Segmentation in the network certainly helps to make a company's IT more secure. However, the mere fact that a person is in a network may no longer be sufficient to justify access. Furthermore, access to certain segments can be dynamically blocked at network level if information collected on user behavior reveals suspicious activity. One thing to bear in mind here is that the other sub-areas also need to be considered. A straightforward segmentation using an overlay network, for example, i.e. a logical separation on one and the same physical network, and no further verification of access would only shift the intrinsic trust to the overlay and not actually comply with the Zero Trust concept.

Applications: Authenticate, authenticate, authenticate!

The previous considerations relate primarily to the underlying infrastructure, but applications and resources themselves also play an important role in Zero Trust architecture and must be able to evaluate trust-relevant information. Even if access has already been approved by the application, the application itself must also check whether the user has valid authentication and whether the user is authorized to access certain information based on the information in the application. This in turn means that many applications that cannot perform such an evaluation in their current state will have to be reconfigured or, in the worst case, even reprogrammed or replaced. > Together with Siemens, we have developed best practices for secure configuration for developers and app owners and automated the monitoring of compliance with these in the back-end systems.

Monitoring and automation: Safeguarding security

interaction between the individual components, central monitoring is needed for all information collected, e.g. failed login attempts, configuration changes or resource access. Since the volume of information becomes unmanageable for individuals and the manual response to a cyber attack takes too long, the automation of processes is a crucial element in a Zero Trust architecture.

While automation is often fairly easy to implement afterwards, the greater challenge lies in collecting the data centrally in a monitoring system and deciding which data is actually valuable for evaluation. In any case, Zero Trust-relevant components must be monitored for function and changes.

Around the end of 2020, and therefore even before the US President issued his executive order to implement Zero Trust in government networks, Siemens had already launched its official Zero Trust program, which reflected exactly the structure described above in its organization. The remarkable thing about it: Siemens not only concentrated on the technical implementation details, but also established communication with the workforce as an essential part of the program. Only by getting everyone involved was it possible to make this mammoth project a success.



As Zero Trust involves stronger networking and

Humboldt

Gruppe

Zero Tust with Comma Soft

If you would like to find out more about our expertise in the area of Zero Trust or gain further insights

into the concept structure and its implementation at Siemens, scan this QR code.



No trust, but still a risk

ven with the introduction of Zero Trust, however, not all risks will disappear at once (especially since there is not just one solution, and Zero Trust is an ongoing process). On the contrary, the Zero Trust concept creates new risks, some of which are still better than the alternatives, but of which we should be aware.

- First of all, there is the risk of a denial of service. The fact that there must be a central engine that constantly evaluates whether access is legal also means that this system is an attractive target for attackers. These could try to flood the engine with requests, making it impossible for other users to continue working.
- Furthermore, there is a risk that the nesting of many components in the automation system could lead to an unforeseen cascade of dependencies, in which the failure of a seemingly unimportant system could lead to major IT failures like a snowball system.
- · Finally, false-positives and false-negatives

can occur in the engine during the access decision. False positives in particular can also lead to a denial of service.

When Zero Trust is introduced, the integration of the systems may well lead to a situation where one becomes dependent on a single provider. This is not necessarily a bad thing, but it is something that is good to be aware of. However, if one relies on a large number of different providers to counteract this, one runs the risk of the complexity of the entire system getting out of control.

This decision must be weighed up carefully and requires experience. Partners with experience in technology, architecture and individual implementation are helpful in this context, says Thomas Müller-Lynch. "Because Zero Trust and cyber security as a whole should always be introduced and maintained individually."⁶

⁶ https://comma-soft.com/blog/deep-dive-mit-siemenschancen-challenges-von-zero-trust/, 08.01.2023



Becoming an unattractive target

utomation is an integral part of Zero Trust. Policies stored centrally as code make an attractive target for attackers. Knowledge of the company's architecture alone can be valuable to attackers. If the policies are stored with a service provider, this provider could become the target of an attack instead. It is important to remain vigilant here so that the trust problem is not shifted to the third-party provider.

The accounts used for automation are now also more likely to become a target for attackers, as they often have far-reaching authorizations and cannot rely on dynamic MFA mechanisms such as an app notification for authentication like real individuals can.

Recently, there has been another publication⁷ by Google on the BeyondCorp model, which deals with special cases in the introduction of Zero Trust. As with the experiences at Siemens, the introduction of Zero Trust requires resources that cannot be migrated in the short term. <

In such cases, it is important to name a person responsible for an exception granted and to set a time limit for the exception so that the progress of the migration can be ensured at regular intervals.

7 https://research.google/pubs/beyondcorp-and-the-longtail-of-zero-trust/

It is now clear that Zero Trust as a concept in a company is never really complete. There will always be technological developments that make it necessary to adapt existing mechanisms.

Poisoned AI

One innovation that has already become widespread is the use of generative AI and machine learning in general. The use of ML models Zero Trust the code to assess the risk of user behavior is already Generative AI is producing an ever increasing standard today and forms the basis of many amount of software code. These program and Zero Trust policy engines that decide on access. script frameworks must always be checked However, there is a risk that companies will rely before use. The idea of "always verify" also aptoo much on automation and that attackers will plies here, as developers should not rely on the use AI tools to continuously feed data into the fact that code generated by AI is automatically ML system, ultimately leading to the wrong desecure. And the ever-increasing networking cision being made and the attackers gaining of devices outside of IT will also keep compaaccess. This is referred to as data poisoning nies busy with plenty to do. Even something as Nowadays, freely available tools for generating seemingly simple as the smart heating thermoimages and texts are used by cyber criminals to stat in the office requires a secure integration carry out phishing campaigns. The quality of the option for these devices. fakes is so high that even sensitive users often no longer detect the attack in their busy everyday



Example of the data poisoning process

lives. The AI solutions used by the company itself and their components must also be well secured to prevent the AI from inadvertently disclosing internal information. If you fancy a game. please follow the OR code:

Despite everything:

Zero Trust is the future

or companies that are already implementing Zero Trust in their own IT systems, the next few years will be very much about integrating the technological long tail, since it does not have a simple interface to modern systems. On the other hand, it will continue to be important that everyone in the workforce is made aware of IT security and also of possible gateways.

One thing is clear: there will continue to be weaknesses and gaps. It is crucial to be prepared for this and to be clear about why the Zero Trust journey is a necessary one to take. In five years' time, analysts may no longer rate a company solely on its P/E ratio, earnings or Piotroski F-Score, but also on its cyber security resilience. Cyber security is also becoming even more relevant in due diligence processes, e.g. for company mergers. In addition to costeffectiveness, compliance with certain IT security standards could be a fundamental prerequisite for cooperation between companies. In projects with several participating companies, it is already apparent that agreeing on security requirements takes a significant amount of time before implementation even begins. In five years' time, companies without an effective cyber security strategy such as Zero Trust may no longer even exist, either because they have been the victim of an incident or as a result of non-compliance with regulatory requirements.

The author



Dr. Andreas Künsken



Page 40



Is Zero Trust only a corporate matter?

On the contrary! - says security specialist Dr. Jan Ciupka

to grips with Zero Trust? Zero Trust as a concept and paradigm has led to a certain rethinking of cyber security: we need to move away from old concepts, ideas and architectures, some of which are already ten or 15 years old and are no longer geared towards the current use of apps and services.

hy should SMEs also get

Due to their exposure, many companies have certainly dealt with the issue early on. However, SMEs are being hit just as hard. They are confronted with the same highly sophisticated phishing attacks, the same ransomware or hacking attacks as large corporations. Many SMEs are also often hidden champions in an industry, making them just as lucrative for an attack - and they even have a much harder time than large corporations.

First of all, you can't spend that much budget on cyber security purely because of your size.

Secondly, it is difficult to find and retain good IT security specialists, as they are always in competition with corporations and tech companies and the market generally provides too few employees.

Thirdly, cyber security incidents, which can also mean a loss of production, for example, are often more difficult for SMEs to handle financially.

Where, for example, can you find approaches or even solutions through Zero Trust for SMEs?

Here are three examples:

- · The Zero Trust concept also strongly embraces the topic of automation, using modern - in some cases Al-based - solutions to respond to security incidents, suspicious users or devices directly, automatically and in a personnel-friendly manner. You can't do it manually anymore anyway.
- Zero Trust does not focus solely on avoiding the incident at all costs. This is not possible anyway. Instead, the focus is also on the topic of limitation and recovery. So how do you limit an incident and eliminate damage as quickly as possible? Every reduced day of downtime can be financially life-saving.
- People and users are also part of the concept. Users are the target of phishing attacks. Even the lowest level of cyber security awareness may mean one less phishing attack that results in production downtime. A very effective measure with regard to costs and benefits.

What are the requirements before I can start And perhaps the most important thing: awarewith an implementation?

The concept of Zero Trust is not narrowly de-



Dr. Jan Ciupka, Executive Manager Consulting

fined. Every prerequisite created that contributes to improving the security architecture or contributes to certain aspects of it is already a move towards Zero Trust. In principle, the aim is to continuously improve cyber security. A few basic things to consider are:

- Support from management and the executive board
- Resources and gualified personnel or partners
- Clear view of where you stand, where you want to go and how you want to achieve this, e.g. as part of an IT strategy

ness that a cyber security incident can be at least as threatening to the company as high en-

Photo: Rifqi Ramadhar

ergy prices or supply chain problems, but can occur much more unexpectedly. This is therefore a very important component of a company's overall risk management. Zero Trust is a strategic goal that is ongoing in both large corporations and SMEs. It's not something that can be ticked off with just one project.

Which point should be tackled first?

A good identity and access management (IAM) system with solid IAM processes is the essential basis for much of Zero Trust, but also for IT security in general. This is precisely why this is a core competence of many of our consultants.

An IAM system is undoubtedly the most important foundation of an IT security infrastructure. It authenticates users (verifies that they are who they claim to be) and controls who can access which digital asset. Without a good IAM system, for example, the comprehensive introduction of multi-factor authentication only makes partial sense.

Another example is the topic of "dynamic riskbased access". In other words, conspicuous users can be denied access to certain resources because, for example, they have logged in from a different device and a different location than is normally the case. However, this only works if you have previously followed the leastprivilege approach when assigning rights in my IAM system, i.e: as much as necessary, as little as possible.

And finally: What do you need to be aware of if you rely on the Zero Trust concept?

Both corporate groups and SMEs need to be clear: the responsibility for IT security lies with the management. Without the commitment of management, every Zero Trust project is doomed to failure.

At the same time though, all company employees should be made aware of IT security issues. Regular training and open communication about changes, progress and new risks should be the norm. A lot of this is classic change management.

Both the threat landscape and the technical possibilities of security solutions are constantly evolving. What was safe yesterday may already be a risk today. You therefore have to continually question your infrastructure, processes and measures, evaluate new technical and technological options and assess measures based on risk.

The Zero Trust checklist

Scan this QR code to access your cyber security checklist:



Gruppe



Below deck — the cabin chat with our COO Benjamin Schulte

Who is actually steering our ship through the wind and weather towards new islands? Let's take a look inside the cabin of Benjamin Schulte, who has sailed virtually all the world's oceans with Comma Soft.

Benjamin, you've been with Comma Soft for over 15 years - even though you once said you wouldn't work in consulting for more than three to four years ...

That's right! When I started at Comma Soft, I discussed precisely this with my wife. We both simply assumed the cliché that people don't stay long in consulting firms. Three or four years at the most, then it would be a corporation again, like when I started my career. At Comma Soft, however, I never experienced stagnation

or boredom and was always able to create and experience new impetus. And now it's been more than 15 years and I'm not planning on leaving any time soon.

What do you mean that it never gets boring?

We are just as dynamic as our customers' markets - routine has never been part of my job over the years. This is certainly partly due to the fact that I have changed departments and positions over time: starting as a consultant in the field of advanced analytics and data science, to managing our business intelligence product INFONEA® and my current tasks in management and on the Management Board. At the same time, the projects and customers I come into contact with are always different and very varied. The Comma Soft team members give me the most drive - dynamic, inquisitive people who are not satisfied with next best solutions.

Which development is associated with the biggest change for you?

Difficult, there are always peaks in technology and in the economy. But what I'm currently most concerned with are two innovations at Comma Soft. One is ALAN, the LLM, which we have been offering as a product for a year now. Readers have already discovered a lot about this in this travel report. And of course we have always had a product division with INFONEA. With the Comma LLM, however, it is gaining even more clout in very topical issues and is shifting our portfolio in perspective.

We used to be a consulting firm that also made software. Today, consulting, software development and product offerings are converging more and more. This is changing the way we come into contact with companies and also how we work on projects. I find it extremely enriching that we are a diverse group. It allows us to provide customers with truly holistic support and also gives Comma Soft employees even more opportunities to contribute with their individual strengths and interests.

The other big change I can see is reflected in what our readers are currently holding in their hands: with the Humboldt Group's travel reports, we have created a completely new format in which we can report on the topics we deal with and how we work together as a team behind the scenes. This allows us to engage in conversations with people in a completely different way than before, which in turn leads to new points of reference for further topics and solutions. I have to admit that I was initially a little skeptical about how it was arranged -

Page 46

Humboldt —

images of nature and travel analogies don't really seem to have much to do with deep tech. But that's exactly what gives it its special flair and when my daughter didn't want to let go of our first report because she fell in love with the horse pictures, I knew for sure: we are publishing a medium that appeals to people in a very special way and can evoke emotions. I find this a nice counterbalance to the analytical, scientific approach for which we are otherwise known.



Benjamin Schulte is COO and member of the Management Board of Comma Soft AG.

He lives with his wife and daughter near Düsseldorf and finds the perfect balance to his working day by traveling and doing things together with his family. At weekends, he can also be found at the riding stables, as his daughter is passionate about horses. Personally, this is still a project for the future: learning to ride and enjoying rides in the countryside together.



Petersberger Gespräche 2023

The showroom for business, technology and science - already in its 15th round

or me, the congress begins as soon as I pass the barrier and am checked by the federal police. Only invited guests are allowed to enter. The

heavy iron gate reveals the way. The park, which never fails to impress me, appears in front as the gate closes behind me. In its cen-

ter is the classicist fountain and behind it the Villa Hammerschmidt in the first rays of the new day. As I walk through the large wooden gate, I am greeted by that special, now familiar smell of history and importance. I am delighted to be here and have high expectations for the day ahead.

The day

It is Saturday, September 23, 2023, and Comma Soft AG is hosting the 15th Petersberger Gespräche, an interdisciplinary forum for dialog between leading representatives from business, technology and science. A modern setting for collaboration at eye level, for inspiration and further development of our own perspectives.

At this very special location, the second residence and second official residence of the Federal President, it is always very important to me to bring together decision-makers from the worlds of business and technology; not to show off finished solutions, but to talk about the unfinished and to provide stimulating and inspiring food for thought.

Welcome!

While last year's Petersberger Gespräche concentrated on the importance of a change in perspective for a new way of thinking and acting in business, research and development, this Saturday's event will once again be dedicated to a technology-oriented topic. The topic is: "Cornucopia of opportunities or a field of unresolved problems? Timing and antifragility in the era of deep tech."

I open our congress at 9.00 a.m. on the dot and extend a warm welcome to all participants. My brief welcoming address once again presents the conceptual framework for the day, paying particular attention to the importance of deep tech for shaping the future of our economy. The focus is on advanced technologies such as generative AI with its large language models and quantum computing with its computing possibilities that surpass all imagination. One of the guiding principles was: "When ChatGPT was made accessible to the public at the end of last year, it hit like an asteroid. Overnight, a technology that was for most people rather cryptic became an everyday technology - and a media tsunami. No question about it: Al is not a passing fad, it is here to stay."



The Villa Hammerschmidt



Prof. Hans Uszkoreit on Generative Al



Generative AI can't do everything, **but has been able to do a lot of things well from the start**

ow generative AI has achieved this but rather stick to the truth, if I have an analyis the subject of the first speaker sis of a production chain, for goodness sake, of the day. Prof. Hans Uszkoreit please don't make anything up! ... We expect draws on an impressive 40 years this system to be God." Uszkoreit makes it clear of research in the field of computational linhere that the utilization of these models cannot guistics and AI in the USA, China and Germany. and should not be about this. The comparison The key message of his remarks: we are plachas something to it, I think, God is by definition ing too many conflicting expectations on LLMs infallible, but ChatGPT & Co. still make many (as large language models are also known). A mistakes, i.e. hallucinations. Nevertheless, the system like this "should be explanatory, truthful computational linguist emphatically underlines and creative, it should be able to think up things, the potential that lies in the individualization of it should be able to write novels and marketing these large language models and emphasizes plans. Others say it shouldn't make anything up, their many benefits for the economy.



A place for intellectual and more personal encounters

round 80 decision-makers from the worlds of business, technology and science followed his keynote speech in the main hall of the villa. As in each of the past 14 years in which our fall congress has taken place, I can see many familiar participants who I know personally as well as new faces. It was clear that many of the participants knew each other: even before the start of the congress, the atmosphere at breakfast in the atrium is very lively. This makes me very happy. After all, that was precisely my intention when I initiated the Petersberger Gespräche almost 20 years ago:

bringing people from the most diverse areas of business, technology and science together so that they can get to know each other on a personal level and engage in an inspiring exchange that broadens their own horizons.

Machine teaching and benefits for companies as data-based expert systems

According to computational linguist Uszkoreit, the development of generative AI has currently entered a phase in which it is no longer as much about the algorithms as it is about machine education and machine teaching: "The progress we are currently seeing is all machine

teaching. We're giving the system different cur-According to Fischer, the consistency, relevance ricula, different materials, that's the biggest and topicality of the texts generated by ChatGPT difference at the moment, and it's a whole new & Co. can be improved by so-called retrieval ball game." Uszkoreit passionately alludes here augmented generation (RAG). He explains that to the increasing importance of large language this method, which Comma Soft often uses models, trained and fine-tuned with company successfully, is an extended use of context data. data, for companies. Their big moment comes The accuracy and reliability of language models mainly when they take a company's corporate are greatly increased by searching in previously knowledge to a whole new level. prepared document libraries filled with company data and through the targeted use of spe-ChatGPT im praxisorientierten cial plug-ins. In conjunction with the maximum Turbomodus degree of transparency and trustworthiness of the results, there is also full control. "This way," Our AI specialist and member of our Executive Management team, Dr. Andrej Fischer, will Fischer continues, "I can choose which libraries focus on this pragmatic side of large language to search in, I can determine what the user is models in the last presentation of the mornauthorized to see with their role, with their rights in the companies, what they are allowed to see ing. He classifies their performance in an arena and what not, and then I add that to it." Fischer that he calls "armchair criticism": an approach enthusiastically praises these diverse functionfrom which criticizing and judging the solutions generated accelerates the human ability to solve alities, which are not found in the basic models.

problems tremendously. Conclusion from the Comma Soft expert:



Dr. Andrej Fischer: Generate & critique

: Gregory



Space & time for encounters

"What ChatGPT can already do very well straight "out of the box" is ... to use the world's accumulated knowledge from its training material. And something that is already possible today is the use of agents that solve complex software problems very independently. The increase in productivity in the area of software development, preparation for customer meetings, solving problems of all kinds, research work ... is, well, phenomenal."

Naturally, I am very familiar with the method used by our AI experts. What's also very gratifying for me - and this includes the feedback I have received from many participants - is to see how much interest this individualized approach has generated. Fischer provided the guests at the Petersberger Gespräche with an impressive demonstration of the new era of technology we are entering with GenAl and how companies can use this innovation in areas such as corporate knowledge. Completely candid about strengths, weaknesses, potential and risks.

And it is precisely this kind of openness that is an essential element of our forum's philosophy: A discursive, open space in which it is also possible to discuss abstract issues and perhaps also the limits of one's own imagination.

Quantum computing - the new dimension of computing

The physicist, IBM Fellow and renowned guantum computing expert Dr. Heike Riel puts our imaginations to the test and pushes it to its limits with the next advanced technology. It involves a highly complex calculation that cannot be carried out even with the most powerful of conventional computers. Calculating the molecular structure of the hexabenzocoronene molecule, for example, would require a computer with a capacity of 10 to the power of 99 bits - which corresponds to more atoms than there are on Earth.

What on the one hand represents physical limits and a dead end for classical computers that are based on binary processes, represents



Dr. Heike Riel on quantum computing



Impressions during the keynote speeches



on the other hand the beginning of what is the first time in history that a completely new known in research as the "quantum advancomputer technology has been built that works in an entirely different way. No longer based on tage". According to Riel, a quantum computer would (only) require a full 500 quantum bits, or digital processes, but on quantum physics." gubits for short, for the above calculation. The quantum computers currently being developed by IBM have an output of around 1,000 gubits. Petersberger Gespräche The company expects to achieve a total processor performance of 4,000 qubits by 2025 and **2023 video playlist** around 100,000 qubits by 2033. I am also very impressed by the following comparison. Based The day in fast-forward, interviews with the speakers, voices on the exponential formula 2 to the power of of the participants: All 275, we are entering realms that until recently were reserved for science fiction at best. Riel videos from the Petersberger adds: "If you were to assume 275 qubits, you would have more base states than there are in Gespräche (available the observable universe." in German)

The physicist draws the conclusion: "This is





What the quantum advantage is good for

If you consider that every single additional qubit doubles the performance of the quantum computer, you get a rough idea of the kind of complex tasks that can finally be tackled by this new technology, in business and research for example.

Dr. Riel lists a whole range of possible applications. Among many others, these include materials research, faster drug development in pharmaceutical research thanks to more powerful simulation options, new financial technology including the optimizing of financial instruments and the simulation of other highly complex systems, which our current computer architectures have to capitulate to in purely physical terms. In Dr. Riel's estimation, traditional computer systems will continue to perform their current tasks. There will also be increasingly powerful systems specifically for AI applications - which brings us full circle to the other speakers. This will be accompanied by a new generation of quantum computers whose architecture will stand out due to their built-in modularity, new cooling systems, miniaturization and significant energy reduction.

Everything relating to the Petersberger Gespräche

If you would like to find out more about the Petersberg Talks, please visit the event website:

www.petersberger-gespraeche.de



Panel discussion – **lively and, as so often is the case, far too short**

he three excellent and highly stimulating presentations clearly triggered a lot of new ideas and left many questions unanswered. The

subsequent panel discussion with the speakers will be driven primarily by questions from the audience. The focus is, among other things, on the power of generative systems over us humans and their influence on our judgment and decision-making. The experts recommend the increased use of the systems in companies and simultaneously a constant critical attitude towards the results.

The audience also still has unanswered questions about quantum computing. The panel discussion is by no means the end of the exchange and the beginning of the discussion. The discussion continues into the lunch break,



Exchange with the audience

Humboldt —

Mark Ba

only briefly interrupted by the serving of the three-course meal.

China and the topic of the middle

Where does the topic "China" fit into this thematic concert of the technological music of the future? Prof. Heinz-Otto Peitgen, the long-standing moderator of the Petersberger Gespräche and renowned mathematician, provides the answer in his introduction of the next speaker. In his introduction to the interview with the China expert, journalist and author Frank Sieren, Peitgen makes the main thrust of his interview clear: where is China as a player, partner and competitor at the same time, perhaps even as a (prime) example in this rapid technological development? How should the West actually deal with this awakened, highly innovative giant?



Prof. Heinz-Otto Peitgen (I.) & Frank Sieren (r.)



Daring more realism...

Even if it is necessary and right to respect one's own values, according to Frank Sieren, the contrast "values or economy" is a false alternative "that no longer has anything to do with reality." His main thesis, which the interviewee repeatedly puts forward: "We will only have a seat at the table if we are economically strong" - and unlike in the past, we are no longer the only ones deciding who sits at the table.

The biggest mistake in politics, as in business, is to underestimate the competition. This is particularly true for China.

"How is it the case that a country that has ideologically trapped itself then becomes the innovation center of the world in just a few decades ... and what kind of internal change is that, what qualifications were needed for it, what does that say about the elites who rule China today?" - In response to this question posed by Prof. Peitgen, the China expert cites several reasons:

- Every Chinese knows where China once stood and wants to get back to the 27% world market share.
- An unprecedented opening and privatization of the economy under the leadership of Xi Jinping, including the establishment of ideal conditions for talented and excellently educated returnees from abroad, primarily the USA.
- 3. Pragmatism: The Chinese "said quite matter-of-factly that they couldn't match the quality of Audi's diesel engine, so let's leave that and go straight into electric cars, and they then became such a leader relatively quickly that they can now actually set the rules of the game for one of the West's central industries for the first time in its 1000year history."

Understanding China - a lesson in ambivalence and a change of perspective

"We have to learn to understand this country in all its ambivalence. If we don't do that, we will be increasingly on the defensive." Hence Sieren's recommendation to adopt a change of perspective in our multipolar world order: it is important for us to learn why and how people in these countries tick - and this applies not only to China, but also to other Asian and African countries. Furthermore, we need to understand how they see the world and what their main interests are. Sieren concludes with a warning: "We are already in the minority anyway, and our influence is getting weaker every day."







Exchange and get-together in the villa

Page 59 🕨

"We are God - and then what?"

he next lecture was also about the role of us humans, but the last speaker of the day confronted us with a challenging change of perspective. Thea Dorn is, as she says herself, "from a completely different planet" within the framework of the Petersberg Talks as a philosopher, author and presenter.

While Prof. Uszkoreit spoke at the beginning of the day about the fact that we sometimes have such high expectations of the major language models as if they were God, Thea Dorn picked up on this metaphor at the end of the day, as if by chance, in the title of her lecture: "We are God - and then what?" It takes its starting point in relation to today's dominant homo technologicus. From here, the philosopher builds a bridge in the course of her lecture to today's technological age and the influence of technology such as generative AI on us humans, our self-image and our society.



Thea Dorn with a philosophical perspective

She directs the attention of the highly interested audience to the question of why human existence has changed over the course of human history and still seems to be adaptable.

The human with the capital "H" and AI

Thea Dorn considers the Enlightenment to be a paradoxical intermediate step towards today's human self-image, in which man, although no longer at the center of the universe, becomes the ultimate greatness: He is "the individual to whom we attribute reason, dignity, autonomy and free will. [...] Suddenly humans were born with the capital 'H'." And this world view has endured for centuries and withstood dissonance between the religious and scientific world views.

With a view to technological developments, including in the bio- and neurosciences, the speaker believes that we are in the process of saying goodbye to this world view. "Instead, we are told that we living beings are ultimately just algorithms" and that we can recreate and simulate our brains. Ultimately, the brain is "viewed as a neural network".

How this is done ultimately depends on the legitimate question: "How do we look at people, how do we look at ourselves, when we elevate the functioning of artificial intelligence, when we elevate neural networks to a kind of ideal, to the new God?" Possibly, the philosopher continues, we have created a new god-like entity with the AI we have developed. According to Dorn, we fail to recognize the fact that these are just tools.

clergy is emerging

Dorn advocates in this context the reinvigoration of our human power of judgment regarding the reception of information that surrounds us, as otherwise we would also run the risk of losing our sense of plausibility and truth. "Is it really in the logic of this technology that we have to give up the human being with the capital 'H' [...] because we have to recognize that we are deficient, neural networks that unfortunately still have a few problems ..." - "or do we want to continue on the path that the West took a good 200 years ago in the form of the Enlightenment?"

This "firework of thoughts", as moderator Peitgen introduces the panel discussion with Thea Dorn, seems to catch the imagination of our participants. At any rate, the audience discussed the theses very intensively at the end of the main lecture; apparently Thea Dorn opened a few windows with her thoughts. For me, the best indication of how fruitful the last lecture was is the question posed by one participant, based on the philosopher Immanuel Kant, as to whether an excess of automation is not the "entrance (of man) into self-mutilating comfort and immaturity" - as the precise opposite of Kant's famous definition "Enlightenment is the departure of man from his self-inflicted immaturity".

Gruppe



Judgment versus ChatGPT or: How a new This ends the main program, but only so that the participants can retreat into discussion groups or dialogues, in which I also take part. Afterwards, as every year, we will set off for the cultural program of our congress. This refers to both the joint dinner and the concert by world-famous jazz pianist Michael Wollny as part of the Bonn Jazz Festival program.

> Here I conclude my observation of the exciting and entertaining congress day, already in planning thoughts for the 16th congress. Petersberg Talks in 2024.



Stephan Huthmacher



In Humboldt's words

Excerpt from the "Cosmos: A Sketch of a Physical Description of the Universe"

lexander von Humboldt's "Cosmos" is his life's work cast in language. In the five volumes, he shares with readers what he has experienced, learned and analyzed on his expeditions and investigations. In this excerpt from volume 1, which was published in 1845, he sums up his thoughts on experiencing the world and nature by means of human, meaningful physical research. These thoughts will bring our travel report #02 to a close. Foto: Rachel Clair

[...] Those who consider the results of natural research not in their relation to individual stages of education or to the individual needs of social life, but in their great relationship to humanity as a whole, will find the most gratifying fruit of this research to be the benefit of seeing the enjoyment of nature increased and refined through insight into the connection of phenomena. Such refinement, however, is the work of observation, intelligence and time, in which all directions of the spiritual powers are reflected. The history of mankind teaches those who know how to trace the ancient trunk of our knowledge through the deep layers of prehistory to its roots how, for thousands of years, mankind has worked to find the persistence of the law in the eternally recurring changes of the world's formations and thus gradually conquer the wide world through the power of intelligence. To question this prehistory means to trace the mysterious course of ideas, in which the same picture, which early appeared to the inner sense as a harmoniously ordered whole, cosmos, finally presents itself as the result of long, painstakingly collected experiences. [...]

For the thinking observer, nature is unity in multiplicity, a combination of the manifold in form and mixture, the epitome of natural things and natural forces, as a living whole. The most vital result of meaningful physical research is therefore this: to recognize unity in diversity; to embrace from the individual all that the discoveries of the latter ages offer us; to scrutinize the details and yet not to succumb to their mass; to be mindful of man's sublime destiny to grasp the spirit of nature, which lies concealed under the cover of appearances. This way, our endeavors reach beyond the narrow boundaries of the sensory world; and we can succeed in comprehending nature, in mastering the raw material of empirical perception through ideas, as it were. [...]

Source: Alexander von Humboldt, Cosmos. A Sketch of a Physical Description of the Universe. Stuttgart and Augsburg: J.G. Cotta, 1845-58, vol. 1, pp. 4-7

Humboldt -----

AI tools to try out

After this extensive reading section, you, dear reader, may already be itching to try out (Gen) Al for yourself. If you are interested, please contact us to arrange a demo of the Comma LLM, ALAN. We have also compiled a hit list of exciting AI applications for you to try out. Have fun!

GenAI for the ears

Podcast.ai

"Podcast.ai is maintained by a speech synthesis startup as a lighthouse project. The quality of the voice output and dialog is really good. A nice glimpse into the future, but also shows how hard it will be to recognize real interviews."

suno.ai

"Creating music with AI is really exciting and somehow also a bit scary. The application is another exciting and playful AI use case. It demonstrates very well what is possible."

AI would like to wish you a pleasant experience

David Attenborough is now narrating my life

"Al-generated, humorous satire of a nature documentary with self-deprecation and the magnificent voice of David Attenborough. Note: you need your own developer environment for testing. The code is available in its own Github repository. https://github.com/cbh123/

narrator"

Emoji Scavenger Hunt

"A particularly interesting application to help children get to grips with AI for the first time. But it's also a great way to liven up on-site meetings. Interactive and with motion. You need a smartphone to be able to run it."

BirdNET

For amateur ornithologists or the curious: with Birdnet you can identify bird calls, whether you are hiking, on your balcony at home or on vacation. Good user interface and everyday operation."

Flora Incognita

"Similar to Birdnet, except that it concerns itself with recognizing plants. Can be used from anywhere where there is internet!"

t3n-Quiz

almost indistinguishable."

Spurious Correlation

"The website lists examples of spurious correlations between unusual facts from all areas of life and provides Al-generated explanations. This makes fans of 'fun facts' in particular smile and illustrates in an amusing way why correlation and causality are not synonymous and why LLM outputs should be treated with healthy skepticism. Tip: Simply click on 'random' for random examples."

TAAFT: There is an AI for that

"TAAFT is a collection of many other AI gimmicks, both entertaining and applicable in a business context. Just have a browse and try it out!"

Gruppe

AI tools for nature lovers

AI facts

"Anyone who wants to can check here whether they are being duped by Al. In some cases, the differences between real photographs and Al-generated images are

And there's much more:

Imprint

Issue 02 / Spring 2024

Comma Soft AG Pützchens Chaussee 202-204a 53229 Bonn +49 (0)228 – 97 700 info@comma-soft.com

Responsible within the meaning of press law

Comma Soft AG Benjamin Schulte, Member of the Management Board, Chief Operating Officer

Responsible editors

Julia Pedak, Hanna Wangler

Comma Soft AG authors

Dr. Jan Ciupka Executive Manager Consulting

Dr. Carsten Felix Draschner Data Science Consultant

Dr. Andrej Fischer Executive Manager Consulting

Dr. Hua-Jing Han Senior Data Business Consultant

Stephan Huthmacher Chairman of the Management Board, Chief Executive Officer

Dr. Andreas Künsken Head of Consulting Identity & Access Management

> Michael Tannenbaum Lead Consultant Generative Al Products

Idea and concept Giso Weyand, www.gisoweyand.de

> Design Tanja Schell

Graphic concept and image research Haus am Meer, www.hausammeer.org

> Proofreading logonia GmbH

Print inpuncto:asmuth druck + medien GmbH www.inpuncto-asmuth.de



Gruppe



Everywhere an early intuition precedes later knowledge.

Humbeldt.

