Deloitte.



NXP gets to grip with GenAl

Patrick Attallah explains the semiconductor company's systematic approach to realizing the full potential of generative AI

In the R&D-intensive semiconductor industry, generative artificial intelligence (GenAI) could be a game-changer. By streamlining software development, among many other tasks, GenAI could significantly lower the time it takes for chip companies to start making a return on their substantial R&D investments.

That is one of the reasons GenAI represents a huge opportunity for NXP Semiconductors, according to the Eindhoven-based company's Chief Data Officer - Patrick Attallah. NXP is keen to harness GenAl enhance the to productivity of its 11,000 software lf achieve engineers. you can "productivity gains in their day-to-day work...that means we can accelerate the time to market, we can accelerate the cycle time of innovation," Attallah notes. That is particularly important to maintain NXP's leading position in the automotive sector, which is trying to accelerate its own innovation cycle.

Attallah suggests a GenAI system should be managed in a similar way to a smart intern in the sense that it should be given "bite sized" tasks and the outcome needs to be monitored carefully. The resulting productivity impact will depend on the specific task and the experience of the person in that role. "A lot of time is being spent, whenever we develop a chip, on the design and testing the specification before it goes into production," Attallah explains. "If you can get GenAI to write the script for you - how you test and potentially write the code for it - that productivity [gain] can become much, much bigger than 10%," he adds.

As well as using the technology to accelerate software coding, NXP plans to initially harness GenAI to support marketing and sales, and the development of documentation used by procurement, legal and other functions.

Putting the right processes in place

NXP's systematic approach to implementing GenAl is underpinned by three pillars. Firstly, it has created a clear governance structure across the organization. NXP requires every Al tool and use case to be approved by a new internal advisory council. Secondly, NXP is "being driven more from a use cases perspective rather than let's test whatever we want," says Attallah. The third pillar is the creation of specific teams to manage deployments "instead of having lots of people trying it," Attallah explains. "We have created mini CoEs (centers of excellence) …we have one where we have approximately 40 people, cross business line, cross function, with IT being involved, which are testing code generation and code assistant. It's concentrated and therefore, we can measure what could be the risk. We can measure what could be the value." NXP is now building a similar CoE for the sales and marketing domain.

These CoEs employs a four-phase process for each use case. In phase one, the goal is to identify the right GenAl tools. Phase two is focused on retrieval-augmented generation (RAG)¹ and fine-tuning, while phase three involves running a pilot in production, and phase four scales the solution to the rest of the organization.

This very structured approach is designed to reduce the risk of intellectual property and data "leakage", which could happen if employees were to adopt GenAI tools on an ad hoc basis. "From April (the month before I arrived last year), every GenAI tool was banned and is still banned within NXP," says Attallah, who joined

¹ RAG is a technique for enhancing the accuracy and reliability of generative AI models with facts fetched from external sources.

the semiconductor company from chemicals manufacturer DSM. "From June onwards, I started building a policy and a governance structure to make sure we don't open the Pandora's box, but at the same time, we would love those engineers to be able to go and do an assessment.... there's a need to have a balanced risks."

With GenAl evolving so quickly, the ongoing challenge for any organization is how to get the balance right between risk and reward. To that end, Attallah is employing so-called sand boxes, which allow for experimentation, while ensuring that the resulting solutions aren't put into production.

Aptitude for GenAl

Another key challenge with GenAI is training employees to use the tools effectively. "Prompting is not that easy," cautions Attallah, referring to the need to give GenAI tools carefully crafted instructions to get the results you are looking for. "Now that we're moving to phase two and phase three, the prompt engineering becomes very important," he notes. "If people don't know how to prompt, the result will be very deceiving." NXP is now building prompt engineering dictionaries/libraries for different types of engineers, so they don't have to keep reinventing the wheel.

Attallah also stresses the importance of critical thinking – people shouldn't simply assume that the output of a GenAl system will be fine: they need to review it carefully. *"How do we make sure that the engineers, the marketers have the critical thinking?"* he asks. *"The human in the loop is still very important."* When a customer service representative uses GenAl to create a response, they shouldn't simply send it to the customer without reviewing it. If a GenAl system

creates a response that would have taken a human 15 minutes to compose, they can afford to spend five minutes checking it, Attallah notes.

He also warns of the risk of being locked into a single large language model (LLM)², while noting that one solution does not fit all. In some cases, semiconductor companies might need to build their own specialist LLMs to protect their intellectual property and to ensure they have a unique proposition. "You'll potentially have LLMs that have been trained on a smaller data set, which are specific, for example, to semiconductors, which will provide potentially a much, much better response than you have with the cloud providers," Attallah adds. But for other purposes, such as human resources management, NXP could potentially use off-the-shelf models.

NXP is working with Deloitte and other consultancies to ensure that the GenAl tools and use cases it implements further strengthen its strategic goals. "The consultancy can have that benchmark or that experience because you're talking to many of our peers and therefore that can help steer where we need to focus," Attallah notes.

How to get ahead of AI

As well as identifying the right GenAl tools and use cases for NXP, one of the key priorities for Attallah's 35-strong team is to deliver the organizational and cultural changes required to successfully and responsibly scale AI-based solutions and maximize the value they can bring to the business.

This process is complicated by an evolving regulatory landscape. In January, NXP launched a responsible AI

group, which reviews various regulations around the world, such as the EU's forthcoming AI Act, and translates what they mean for NXP. The group then makes recommendations both for NXP's internal adoption of AI and the usage of AI by the products and solutions it sells externally. *"Those recommendations are basically about ethics, and about [preventing] bias,"* Attallah explains.

He also stresses the importance of robust data governance to protect intellectual property and any other sensitive information. Giving GenAl systems unrestricted access to your data can be "pretty dangerous because it's a huge magnifier," Attallah explains. "If you want to make sure that your secret data stays secret, you need to label this data and therefore, you need to have your data and information in order for you to be able to use these tools."

Although "we shouldn't look at it as the Holy Grail," the combination of GenAl and traditional AI could have a major impact, once organizations have the right building blocks and capabilities in place, he adds. "You need to have your data in order, you need to have your skillset in order. You need to have the trust in order for you to be able to move with GenAl."

² LLMs are machine learning models that can comprehend and generate human language text.

Acknowledgements

We would like to thank the following contributors who have been pivotal in the completion of this interview.

NXP semiconductors:

Patrick Attallah, Chief Data Officer, NXP semiconductors

Deloitte:

Berber Rienks, Director, AI & Data – Deloitte Netherlands

Emile Kramers, Partner, Deloitte Consulting, Deloitte Netherlands

Contacts



Berber Rienks Director brienks@deloitte.nl



Emile Kramers Partner ekramers@deloitte.nl

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited ("DTTL"), its global network of member firms and their related entities. DTTL (also referred to as "Deloitte Global") and each of its member firms are legally separate and independent entities. DTTL does not provide services to clients. Please see www. deloitte.nl/about to learn more.

Deloitte is a leading global provider of audit and assurance, consulting, financial advisory, risk advisory, tax and related services. Our network of member firms in more than 150 countries and territories serves four out of five Fortune Global 500 companies. Learn how Deloitte's approximately 286,000 people make an impact that matters at <u>www.deloitte.nl</u>.

This communication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms or their related entities (collectively, the "Deloitte network") is, by means of this communication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser. No entity in the Deloitte network shall be responsible for any loss whatsoever sustained by any person who relies on this communication.

© 2024 Deloitte The Netherlands