ComputerWeekly.com | TechTarget

E-guide

CW+ Content

Top 10 Artificial Intelligence stories of 2018

In this e-guide:

Following the hype around Christmas 2017, 2018 became the year of the smart speaker – at least from a consumer perspective. With a growing number of Google Assistant and Alexa devices now in people's homes, industry experts see an opportunity for businesses to use these devices as a new way to connect to their customers.

Building applications for these devices requires a different approach than enterprise IT has been used to. The user interface is voice-based, so the application must be able to cope with the vagaries of human conversation. People who have used these devices have found that the way to get a smart speaker to respond with the right request or take the correct action is hardly intuitive.

Nevertheless, experts agree that it is key for organisations to understand the opportunity these devices offer going forward.

Moving on, in April, the House of Lords select committee report on artificial intelligence (AI) was published. This recommended that the UK needs to take a leading role in AI ethics. The report came out around the time that the Cambridge Analytica/Facebook data exploitation scandal came to light and the General Data Protection Regulation (GDPR) came into force.

Beyond ethics, AI and automation have found a niche within IT operations, enabling administrators to discover anomalous system behaviour by using pattern matching and machine learning. This approach is also being used outside of IT for the management of the industrial internet of things (IIoT), where digital twins simulate real-world machines.

As machine learning find more and more application areas, some experts believe there are certain problem areas that the AI approaches currently in use cannot easily solve. Problems such as modelling climate change show the limits of AI pattern matching. Instead, something is needed that can combine machine learning with physical modelling.

Here are Computer Weekly's top 10 AI stories of 2018.

Cliff Saran, managing editor

Contents

Why AI must move beyond pattern-matching

Cliff Saran, managing editor

Among the most elementary examples of machine learning is the one Google provides on identifying iris flowers via its Tensorflow machine learning framework.

Artificial intelligence (AI) practitioners are reaping the rewards of finely tuned image recognition based on the volume of images of data readily available on the internet. Without too much effort, it is possible to train the machine to identify cats or almost any new image using pattern-matching with a high degree of confidence.

Such pattern-matching has many application areas, such as in oncology, autonomous driving, chatbots, voice recognition in smart speakers and any time it is necessary to look for patterns in large datasets.

For instance, in January, Intel published an article describing how medical technologies such as computed tomography, magnetic resonance imaging (MRI) and ultrasound provide deep learning algorithms with a source of learning

data. With this data, deep learning models can be used to measure tumour growth over time in cancer patients on medication, said Intel.

But some decisions cannot simply be made by matching against known patterns. This is where physical model mathematical simulations are used.

In a blog post, Vegard Flovik, data scientist at Kongsberg Digital, wrote: “With sufficient information about the current situation, a well-made physics-based model enables us to understand complex processes and predict future events. Such models have already been applied all across our modern society for vastly different processes, such as predicting the orbits of massive space rockets or the behaviour of nano-sized objects which are at the heart of modern electronics.”

However, if there is no direct knowledge available about the behaviour of a system, said Flovik, it is not possible to formulate any mathematical model to describe it in order to make accurate predictions.

Machine learning can help

This is where machine learning can help by effectively matching an unknown problem with a pattern that has already been learnt, drawing on massive datasets.

Flovik said machine learning can be used to learn any underlying pattern between the information about the system (the input variables) and the outcome

that the AI needs to predict (the output variables). But machine learning has yet to evolve to a stage where it can confidently predict complex physics.

In a paper entitled *Deep learning for physical processes: incorporating prior scientific knowledge*, submitted to Cornell University Library in November 2017, researchers Emmanuel de Bézenac, Arthur Pajot and Patrick Gallinari described how machine learning based on deep learning methods cannot easily be applied to a problem such as predicting sea surface temperature.

In the introduction to the paper, the authors wrote: "We considered the use of deep learning methods for modelling complex phenomena like those occurring in natural physical processes. With the large amount of data gathered on these phenomena, the data-intensive paradigm could begin to challenge more traditional approaches elaborated over the years in fields like maths or physics.

"However, despite considerable successes in a variety of application domains, the machine learning field is not yet ready to handle the level of complexity required by such problems."

Commodity AI

AI platforms are being commoditised, so businesses should use the technology provided by the major AI providers, says Bill Ruh, CEO at GE Digital and chief digital officer at GE. However, he says the challenge is how to use these

commoditised AI platforms in atypical application areas where an AI cannot simply rely on the ability to pattern-match.

Instead, domain knowledge is required. According to Ruh, matching the inputs of a physical system with the desired outcome – the output variables – is something few people are actually doing in live production scenarios.

Ruh says the most prevalent use of AI is in targeted selling. “The price of being wrong is almost zero, while the value of being right is high, so the AI can be wrong a lot of the time and still do well,” he says.

But when used to assess the need for repairs on a gas or oil pipeline, the AI cannot afford to be wrong often, says Ruh.

Limitations of AI

Ruh says it is harder to find practitioners of AI than the people who build AI tools. “Commoditised toolsets can actually solve large problems, but AI people tend to be more interested in building the tools than solving application problems,” he says.

In his experience, a general- purpose AI specialist tends to lack the domain expertise to solve an application-specific problem. “When I go to conferences and the presenter asks, ‘how many people have AI projects?’, all the hands go up,” he says. “When the presenter asks ‘how many people are doing pilots?’

maybe a little more than half the audience put their hands up. But almost zero are actually in production."

The reason businesses are not yet seeing value in AI, says Ruh, is because the people building the AI systems are not sophisticated enough to engineer in domain expertise. "It is not true that with machine learning you just pump in some data and it works," he says. "A lot more work is required. The more domain knowledge built into the AI, the more valuable it becomes."

Ruh argues that although machine learning and AI basically use a pattern to achieve a goal, there are some situations where a wrong outcome is dangerous or costly. "If you put in enough data of the right level of quality, the AI will eventually become very good at spotting a pattern and can tell you about it," he says. "Now that may be good at picking salespeople, but in an oil or gas pipeline, while the AI is used to identify that corrosion has occurred, the question of what to do next is complex.

"If an expert sees corrosion, they can start to do analysis, understand the physics of the corrosion to decide whether it needs to be repaired now or can wait another year. A wrong decision is very costly, but a correct decision is highly valuable."

AI will never be able to give the user the correct answer to the question of what to do next, says Ruh. "The only way to do it is through modeling and simulation, looking at every instance of corrosion to understand the physics of what is

happening to the pipeline. AI doesn't understand this physics; it understands patterns."

Given that not every instance of corrosion can be repaired immediately, Ruh says modelling and simulation enables the analyst to assess millions of possible scenarios to decide whether the repair can be scheduled in efficiently and cost-effectively. "With machine learning, the action is not learned; it is predetermined. If you see this pattern, perform this action," he says.

Only as good as the algorithm

Ruh argues that this action is only as good as the algorithm coded by the programmers. The only way to look at all possible outcomes in real time is through modelling and simulation, says Ruh.

Over time, Ruh expects hardware will evolve to build in physical modelling. "I believe that chip technology will continue to evolve to where physical modelling will be part of the decision-making process and it will be possible to look at all possible choices to optimise the action to take in a way that presents the least risk or costs the least," he says.

At GE's pipeline inspection business, Ruh says the company produces massive datasets from miles of pipeline "It takes several years before someone can correctly identify from the data where in the pipeline corrosion is actually a problem, even when there is no leak," he says.

GE is using AI to preprocess the data, to help its analysts identify the areas of the pipeline that are most likely to have corrosion. “This is a perfect case for machine learning because domain knowledge and domain data can be coupled together to feed a machine learning tool,” says Ruh. “The more data pulled in, the more the AI can learn from the analyst identifying problems in the pipeline.”

This is effectively using AI to augment a human, where the expert trains the machine, so it can identify more complex patterns.

The authors of the *Deep learning for physical processes: incorporating prior scientific knowledge* paper also believe deep learning can be combined with physical modelling data. “We believe that knowledge and techniques accumulated for modelling physical processes in well-developed fields such as maths or physics could be useful as a guideline to design efficient learning systems,” wrote Bézenac, Pajot and Gallinari.

Next Article

How artificial intelligence is fuelling engine maintenance innovation at Rolls-Royce

Cliff Saran, managing editor

Engine manufacturer Rolls-Royce has been steadily building up the intelligence capabilities in its aero engines.

In April 2018, Rolls-Royce was one of the companies whose artificial intelligence (AI) initiatives were highlighted by the Department for Digital, Culture, Media and Sport (DCMS). Rolls-Royce also announced it has signed a memorandum of understanding with the Alan Turing Institute to work together on developing next-generation AI, specifically for use in industrial applications.

Artificial intelligence plays a major role in the company's strategic vision.

In February 2018, at the Singapore Airshow, the company unveiled its IntelligentEngine vision powered by AI and advanced data analytics.

To support greater numbers of airline passengers, the IntelligentEngine strategy wraps a service around the engine, supported by intelligent sensors. Rolls-Royce says this offers a wealth of opportunities to improve the way it provides power to its customers.

Smart engines

In addition to designing, testing and maintaining engines in the digital realm, the IntelligentEngine vision sets out a future where an engine will be increasingly connected, contextually aware and comprehending, helping to deliver greater reliability and efficiency.

Specifically, engines will connect with other engines and will be contextually aware, allowing them to respond to the environment around them without human intervention. Through the IntelligentEngine initiative, an aero engine will be able to learn from its own experiences and from its network of peers, and will adjust its behaviour to achieve best performance, according to Rolls-Royce.

The initiative is supported by Rolls-Royce's R2 Data Labs, an acceleration hub for data innovation launched in December 2017. The lab provides a centre for advanced data analytics, industrial AI and machine learning research, along with product development.

The lab currently provides four main products: design optimisation, integrated lifing, fleet planning and MRO insight.

The design optimisation tool uses operational and environmental data to simulate how different component designs would perform and last in service. Integrated lifing is a methodology that uses "digital twins" to predict the

remaining life of engines and their components, based on the detailed conditions they are expected to experience in service.

R2 Data also offers a fleet planning application that explores over 25 years of in-service data to understand how changes to customer fleet operations could deliver improved aircraft efficiency and availability. The MRO insight tool is said to close the learning loop by capturing the condition of used parts to inform improvements to design, manufacture and usage.

Andrew Hutson-Smith is chief business officer for R2 Data Labs, responsible for business development and strategy. His role at Rolls-Royce is to focus on using technology such as AI to enhance existing services and create new ones for customers in all of the sectors Rolls-Royce operates in.

Speaking to Computer Weekly, Hutson-Smith says: “R2 Data Labs is a central group in Rolls-Royce, which is used to accelerate the use of AI and analytics in the business and investigate ways for the company to become more digital.” It is divided into four main areas: the AI hub (*see box below*), a disruptive companies’ portfolio, a strategy and business development arm, and an operations unit for deploying technologies from R2 Data Labs out to business units.

Understanding telemetry

Rolls-Royce has a 20-year history of working with telemetry data from its engines. In the past, it operated as an internal service, but Hutson-Smith says R2 Data Labs is trying to extend its central role with self-service for business users.

“We have a highly capable central group,” he says. “We are looking at the best ways to deploy groups of people out to the business units. Everyone learns from the exercise. Everyone becomes an analytics expert.”

One of the drivers for the company is to use automation to reduce the amount of human intervention necessary to keep its engines running optimally.

“We are always looking at the performance of assets. There is often a human in the loop, but we are trying to get to a heightened level of automation,” says Hutson-Smith.

Given the sheer level of data that can be accessed, including contextual data such as the effect of weather, routing and congestion, the complexity of decision-making is growing exponentially. “You need a level of automation so you can focus on the right decision,” he says.

Machine learning helps to improve data models, according to Hutson-Smith. “We have technology on the aircraft which can be adapted.” Engine telemetry

data can be downloaded from the aircraft. “We can look at both individual assets and the whole fleet’s data,” says Hutson-Smith.

The company makes use of simulation models called digital twins – a term used by industrial companies like Rolls-Royce and GE to describe a simulation of a real machine using live operational data.

Hutson-Smith says Rolls-Royce runs many digital twins and carries out peer-to-peer analysis, through which it is able to analyse data from aircraft running in similar operating environments to understand where they differ.

The amount of data collected is not necessarily substantial. Typically, a long-haul flight generates around 500MB of data, since most of the time the telemetry data is stable. But if an event occurs, the onboard systems are able to capture far higher data rates.

This data, while not huge, can be highly informative when analysed across a whole flight, and can be used to verify that the digital twin is performing as expected.

Rolls-Royce also buys in data from external providers to enhance its data models. “There is an increasing amount of data in the market, such as weather data and environmental data. We may buy in flight aspects such as flight times and even baggage data,” says Hutson-Smith.

Computing powers data-driven services

Rolls-Royce operates cloud computing server farms where terabytes of data can be processed. This enables Rolls-Royce to run simulations to help it develop new designs. For instance, it can run a design model simulation alongside an operational model using real data downloaded onto one of its digital twins. “This requires a huge amount of computing power, which would not have been economical previously,” says Hutson-Smith.

When asked if Rolls-Royce sees an opportunity to sell data, Hutson-Smith says: “We are extremely respectful of customer data. The vast amount of our data is our customers’ data and we add value back to them. There may be aggregation opportunities, such as datasets on fleets, to make data clean, useful and meaningful.” But he says any opportunities to aggregate datasets will need to be done in partnership with Rolls-Royce customers.

As R2 Data Labs evolves, Hutson-Smith sees an opportunity to improve how the business consumes data. “As we bring in more analytics, we need to do more work to help people consume information, and how they get the AI to do its work,” he says.

The role of Rolls-Royce's AI hub

Rolls-Royce's AI hub is run by Nigel Jackson. Previously called the Capability Lab, the AI hub is used to make new technology available to the business. "We are trying to de-risk new technology," says Jackson.

Application areas being investigated at the AI hub include equipment health monitoring services and the acquisition of data in aerospace, ground-based transport and ships. Data from equipment operated by Rolls-Royce is transmitted back to base, decoded and analysed, to understand the condition of the equipment and how it is being used.

"We can take appropriate action based on the analysis of this data," says Jackson. "We can predict with accuracy what maintenance has to be done and we can optimise our workshop capacity and ensure our [parts] supply chain is ready."

Various aspects of artificial intelligence are being using at the AI hub, including machine learning. "Some of the analytics is self-learning," says Jackson. Previously, improving an engine's performance would need people to understand the impact of optimisations. But now, with machine learning, the analytics models can train themselves, but these are augmented by experts. "We use a lot of automation and use the experts to supervise, which means we get a much higher level of quality," says Jackson.

The heart of the system used at Rolls-Royce is a mathematical model of an engine, or digital twin, which is compared with the real engine. This model is updated with real operational data. “We fly the model forward and see what is going to happen. This is then used to make decisions,” says Jackson.

For instance, in R2 Data’s design optimisation tool, the digital twin model can be used as part of the process to certify engine fan blades. “The certification involves releasing a fan blade at full power to ensure the engine casing will retain debris. We do a lot of stress-test modelling and thermodynamics tests to prove a design digitally,” says Jackson. The final test is a real validation of a blade failure.

In the event of a real-world component failure, Jackson says Rolls-Royce is able to capture high-frequency data for analysis. Using the digital twin, he says it can re-run the model to develop an understanding of the problem.

Telemetry data that is collected continually updates the mathematical model of an engine, enabling Rolls-Royce to improve the accuracy of the tools to support maintenance, repair and optimisation (MRO) and the “integrated lifing” services.

Next Article

CW500: The advent of AI

Lis Evenstad, management editor

Artificial intelligence (AI) is nothing new. In fact, with more than half a century gone since Alan Turing presented the Turing Test, the technology has been around for quite a while. So why has the phrase “artificial intelligence” suddenly become such a buzzword? What exactly is AI, and what opportunities does it bring?

At the latest CW500 Club, experts examined the new era of AI, robotic process automation (RPA) and machine learning, and what it means for both organisations and citizens.

Lloyds Banking Group’s risk transformation director, Claire Calmejane, studied AI and cognitive science at university many years ago, and said that although the technology has been around for ages, with the birth of new technologies, “we are really just at the beginning of the journey” when it comes to the use of AI.

The technology is “evolving, learning and maturing”, said Calmejane, and the ever-growing amount of data being generated continues to bring new opportunities.

Some 90% of “all the data of the internet” has been generated in the last two years, she said, and technological developments have exponentially increased our capacity to produce data.

For example, the advent of cloud technology enables us to store data at a lower cost, anywhere in the world, and new software allows us to manipulate and handle the data in entirely new ways.

AI and business transformation

AI is one of those technologies that, like others before it, has, or will, lead to huge changes in business processes.

Sarah Burnett, research vice-president at the Everest Group, said the rise of digital has created a need for businesses to change the way they work.

Over the last 20 years, we have seen the rise of the arbitrage model, with shared service centres and the offshoring and outsourcing of services, she said. Then came the rise of lean methodologies and legacy tools brought together, and the rise of wrappers and connectors.

“What we are seeing today is the rise of digital and, in terms of using self-service channels, the possibility of automating the processes that organisations have to run behind the scenes to cater for this new digital world,” said Burnett.

“It’s about the elimination of manual work – really bringing the back office to the front office.”

Burnett said the emphasis now is on automation, particularly RPA, which “doesn’t inherently have any intelligence”, as well as cognitive AI-enabled tech, which “is really changing the equation completely”.

RPA is where you can, for instance, schedule a robot to work behind the scenes in the datacentre, and automate “all kinds of things, as long as the data is structured data”, she said.

RPA delivers a return on investment very quickly because it is a quick and easy deployment, said Burnett. The added benefit is that RPA goes through the user interface, and you can just switch it off if something goes wrong.

“The kind of benefits you would get is reducing error rates, with regulatory compliance embedded into the RPA robot,” she said.

“With [cognitive] AI, you have to make it a bit harder and make sure that it continues to use the regulatory instructions you’ve given it, because some AI is based on machine learning, which means they can change their behaviour, so you do need some safeguards around that.”

AI in banking

AI has been around for a while in the banking sector too, said Calmejane, but it is only recently that it has become part of the banks' identity.

"AI has become more relevant in banking now, because we have enough trusted data to enable us to maximise operations and offer data-enabled customer experience," she said.

"That's bringing a new hunger about how we want to exploit this data, because it's about how we want to make it better for the customer."

AI is being used across the board in banking, said Calmejane, including customers using speech and image recognition to log into their banking products, having their experience personalised, being offered help 24/7 through chatbots, and receiving smart alerts. AI is also used in mortgage and loan applications, with algorithms making lending decisions.

These uses of AI have led to new business models emerging, and data roles being the most sought-after in the market, said Calmejane.
But although there is a lot of AI-related innovation, Calmejane said has yet to see it cause proper disruption.

"Most of the changes we are talking about are very incremental," she said. "So they make some areas of life better, but they are not disruptive."

To take full advantage of AI, organisations need to change the way they conduct transformation programmes, said Calmejane. “You can’t think of this as a five-year project because the technology is not going to be relevant in five years,” she said.

“You need to think about how you can do it in one year, what your minimum product is and how you can release value quickly for your business, and build from there.”

AI also produces a need for new skills, which are in short supply. According to research, two-thirds of businesses lack the relevant skills to adopt AI.

The big machine takeover?

Although AI brings many benefits, some people are worried that it could have a negative impact on our future.

A recent survey by OpenText showed that 21% of British citizens are worried that their job will be replaced by AI software in the next 10 years.

Yes, AI will disrupt jobs, and many of the job roles that exist today may be taken over by robots in the future.

But Burnett said there are complexities that currently require a human in the loop, such as if you have different types of data that a machine cannot cope with.

With machine learning, the machine watches and learns “so you get this mix of human agents and cognitive robots working together”, she said, but as time goes by, “the proportion of work done by the cognitive robot could increase, and the human part could decrease”. This is where the fear of job losses comes in, said Burnett.

AI creates jobs

However, new technologies will also bring new jobs that don’t currently exist. Gartner research shows that AI will begin to create more jobs than it replaces in 2020, with two million more jobs created than are replaced by AI by 2025.

There are also concerns around machines and robots making important decisions about people. Here, transparency is key, said Burnett.

“Because AI as a machine doesn’t have an ethical framework, we have to give it that ethical framework,” she said.

“There is a lot of worry about robots being taught the wrong things, and AI going rogue on us and perhaps being used maliciously. I think it’s an evolving world and we do have to be aware of these risks.”

To tackle these risks, the House of Lords select committee on AI set out a series of principles it believes AI development should follow, including that the technology should “operate on principles of intelligibility and fairness, and not be used to weaken data rights and privacy”.

It added that “the autonomous power to hurt, destroy or deceive human beings should never be vested in artificial intelligence”, and called for a cross-sector code of ethics.

The future of AI

According to Calmejane, although AI has been around for a long time, we are only “just at the beginning of the journey”.

“There is still much more to come,” she said. “Technologies such as cloud, deep learning and speech are all fairly new, so there is a lot to explore.”

So what does the future of AI hold? With changes in technologies happening “so fast, it’s difficult to keep up”, it is impossible to say, said Burnett.

One thing is for sure, though – artificial intelligence is here to stay.

Next Article

Smart speakers open new windows into customer behaviour for CRM

Lindsay Clark, guest contributor

As consumers flit between smartphones, web pages, call centres and stores, retailers struggle to keep track of them. Eight out 10 retailers say a single view of the customer is an issue for their current or future agenda, and only 8% have successfully achieved it, a study from SAP and PwC has found.

Just as they realise the extent of what we might call the customer intelligence omni-challenge, a new source of data is about to arrive. According to Gartner, worldwide spending on virtual personal assistants, or smart speakers, will reach $3.5bn by 2021, up from $720m in 2016. By 2022, says Juniper Research, 55% of US households will have a device such as the Amazon Echo, Google Home or Sonos One.

This growth should interest retailers. Amazon is already selling products via its voice platform, while in the US, retailers Walmart and Target have formed alliances with Google. In the UK, Tesco sells through the Google Home device via an IFTTT channel.

The question for retailers and other consumer-facing organisations is what this new voice channel means for their understanding of customers and how voice data will integrate with existing investments in customer intelligence.

Kees Jacobs, Capgemini's global head of insight and data for consumer products and retail, says: "It should be the next wave of CRM [consumer relationship management]. Companies that have good omni-channel CRM can build on that. The goal is to truly understand all aspects of how the consumer thinks and acts to form meaningful, empathetic relationships with them. No company is there yet, but mature CRM may put you in a good position."

One of the initial difficulties organisations face is getting hold of the voice data, says Jacobs. "At the moment, it is early stages – a bit of a wild west," he says. "Conversations are orchestrated by Amazon and Google. If that is going to remain the case, the primary data is collected by these firms as they deliver the device and intelligence."

However, retailers and consumer goods manufacturers that have developed applications – Amazon calls them Alexa skills – of voice platforms can collect some data from the interaction.

"It is still a world where only retailers or manufacturers that have skills will collect data," says Jacobs. "That is the situation right now."

Hive, a provider of home control services and devices which is owned by UK utility British Gas, provides skills for Amazon Alexa that allow customers to control heating and lighting with voice commands. Jo Cox, Hive's commercial director for UK and Ireland, says it keeps data on interactions with its skills and uses it, along with other customer data, to develop new offers and services.

For example, after analysing data with the help of subscription application provider Zuora, it now offers a Mimic Mode, a security feature that allows users to replicate lighting patterns so it looks like they are at home when they are out.

"At the moment, the services we provide are quite generic, but we will start to personalise them through our own data," says Cox. "Amazon doesn't share data with us. We will be using our own data platform to personalise Hive: 80% of that will naturally be available in Alexa."

Customer intelligence opportunities

Accenture's managing director for advanced customer strategy, Rachel Barton, says that although smart speaker and voice platform providers such as Amazon and Google currently hold the data, in the future it is likely to be available to third parties, creating opportunities for customer intelligence. Those that have access to the data will find it valuable in shaping offers to consumers, as well as future products and services, she says.

“What organisations have never really known is the behavioural pattern of consumers’ lives,” says Barton. “Voice command will provide a very different relationship. Consumer companies could know when and how consumers make decisions about what they are having for dinner. That is a huge opportunity to tailor the proposition and marketing message to a relevant template. It is much more need-driven than ever before.”

But few organisations are thinking about voice data in this way, says Barton. “There are brands that are sophisticated in the way they use and store data, but in general, there are a really small number of brands that are thinking beyond traditional datasets.”

Those that are already making progress in an omni-channel understanding of their customers will be well placed to benefit when they get greater access to voice data, says Barton.

“It is going to be a case of understanding your ecosystem and being able to respond in the most optimised way, rather than thinking of a series of channels,” she says. “It will still be about the consistency of experience across channels.”

The wake-up call for organisations will come when they realise that voice assistants could have the power over choice of brand, says Barton. Brands or retailers that are unable to serve voice content based on customer intelligence could be literally left out of the conversation, she adds.

“If you think of ordering washing powder, many consumers are less bothered by the brand,” she says. “Shopping on the internet, you can pick whether to buy brand A or B based on familiarity. If you are using voice commands, you do not get anything to prompt you to buy a brand.

“There is going to be period of transition. There will be swathes of customers buying on these platforms, so how are brands going to influence them? Through traditional means, such as advertising? Or are we going to see digital assistants helping us select the brand we want with the consumer setting the options, such as ‘I want a non-bio and eco-friendly washing powder’. If that is the case, then digital assistants have huge power to kill a brand. This subtlety has yet to be fully thought through.”

Tone of voice captured

Mike Lowndes, Gartner research director digital commerce, says one major advantage of voice data will be knowing the identity of the shopper, which is not always the case with other interactive customer technology, such as conversation bots running on a website. But organisations could also understand the mood a customer is in when he or she is making a purchase, he says.

“In terms of voice intonation and emotion, that is still in the research phase,” says Lowndes. “It is being worked on and there are papers published in *Nature*

and *Science*. It is academic now, but within five years, we could see companies using it."

Consumers using voice assistants to buy goods and services shifts shopping towards a concept that commentators are calling "conversational commerce". But enterprise data will not only have to cope with new forms of customer data, says Lowndes – consumer-facing organisations will also have to understand what their products "mean" in a conversational context.

"Many retailers have metadata for their own products, but that is not good enough," he says. "They need a service that will enrich metadata around products, that will then get a more accurate response to search queries. It is the shift from key-word to intent-based searching. It is not about the customer asking, 'can you find a radiator with capacity of 80 British thermal units?'. It is more, 'I want a radiator to work in this room'. That is something companies need to understand about their products, and we are not there yet."

SAP, as a CRM software supplier, is taking an interest in analysis of customer voice data using artificial intelligence (AI), but is not currently offering products in this area. Its chief value adviser for retail industries, Shane Finlay, says: "The way you say something to voice ordering is something that can also be collected. With voice, you can capture emotion.

"SAP is active in the discussion on AI [to analyse voice data]. We are reaching to meet customer demand with partnerships such as the ones we have with

Google, Facebook and Apple. You need a large ecosystem to work on the use-case and go-to-market strategy."

Access to customer and consumer voice data is set to be hotly debated for some time.

Amazon says audio information is not stored in CRM systems or used for targeted advertising or customer recommendations. Similarly, the company says it does not share voice and audio data with developers or third parties – it is only used to improve question answering and AI models.

An Amazon spokesman says: "There are a number of ways Alexa helps customers shop. Alexa may look to see if the item you are requesting is one you have purchased before, offer a similar item or use the Amazon's Choice algorithm to find a well-rated, well-priced item that ships with Prime.

"Amazon takes customer privacy seriously and we have taken measures to make Echo secure. These include hardware control via the 'mic off' button, rigorous security reviews, and encryption of utterances between Echo, the Echo app and Amazon servers."

The Echo device allows users to delete all voice recording history via a specific button.

However, Amazon would not comment on whether it plans to use Echo voice data and consumer queries for CRM, promotion and product development in the future.

For Capgemini's Jacobs, the dominance of a handful of players controlling consumer voice assistants and access to consumer data, whether Amazon, Google or others, should be a concern. He feels it could hinder broader industry development of the voice platform and consumer acceptance of it.

"Ultimately, there needs to be clarity and more openness in sharing voice data and insight," he says. "It relates to the decoupling of devices, the intelligence behind them and the consumer engagement and business intelligence behind that. Right now, that is in a few hands. For consumers and the industry, we need to go to a level playing field."

Whether Google, Amazon and the handful of firms popularising voice assistants agree is one question Alexa is unlikely to answer.

Next Article

Digital twins: Revolutionising product businesses

Cliff Saran, managing editor

At its annual Minds+Machines conference in San Francisco last year, GE showcased how a digital twin could be used to optimise real-world industrial assets.

Analyst Forrester defines a digital twin as an instantiation of a real, physical object in an abstracted, digital form that acts as a proxy for all communication to an actual device.

According to GE, the idea behind the digital twin is to go further than working with models. The company says the costs of maintenance versus replacing an entire asset are also considered.

“If a company orders jet engines, the revenue projections become part of the digital twin, along with the designs for the engines, specific materials used in construction and information on the factory where they were built. When the engine starts up, and when it is serviced, sensors feed that data into the twin,” according to GE.

Analyst Gartner predicts that by 2021, half of large industrial companies will use digital twins, resulting in those organisations gaining a 10% improvement in effectiveness.

“Digital twins drive the business impact of the IoT by offering a powerful way to monitor and control assets and processes,” says Alfonso Velosa, research vice-president at Gartner.

“However, to truly drive value from digital twins, CIOs will need to work with business leaders to develop economic and business models that consider the benefits in light of the development costs, as well as ongoing digital twin maintenance requirements.”

Driving adoption of emerging IoT platforms

In Forrester’s *The digital twin accelerates IoT development* report, the analyst identified the digital twin as one of the features driving adoption of emerging IoT platforms from the likes of AWS and Microsoft among others.

It says software development teams are embracing IoT platforms not only for speed and convenience, but also because such platforms have frameworks that support variations of the digital twin concept. This, in turn, offers businesses a new design pattern for managing physical assets.

Managing connectivity

The new approach involves managing connectivity between edge devices and back-end systems, and mirroring changes on a virtual model of the asset – in other words, the digital twin.

One company that is starting to look at digital twins is Dormakaba, which makes smart door locks. Since 2012, Dormakaba has been using field management software, ServiceMax, to help it monitor its installations.

Explaining how the system works, Mark Homer, vice-president of global customer transformation at ServiceMax, says: “Dormakaba has been collecting data from devices for many years. Service engineers do a ‘debrief’ on every door. They collect data on a few at a time.”

Such data helps Dormakaba and its partners to manage buildings more efficiently. As Computer Weekly has previously reported, a recent Vanson Bourne study for ServiceMax estimated that industrial companies are losing $260,000 per hour due to unplanned downtime. Predicting failure using digital twins can help overcome the problem.

One of Dormakaba’s customers is property management firm Jones Lang LaSalle, which manages Dormakaba doors and shutters at a building in South Australia. According to ServceMax, Dormakaba’s Australian service centre in Hallam, Victoria, received 338 calls and 289 support emails a day.

Using ServiceMax, Homer says engineers at Dormakaba are able to do predictive maintenance. “It is very important that the work has been completed,” adds Homer.

Open my digital door

This is where the digital twin can play a role. Speaking to Computer Weekly, Andreas Häberli, chief technology officer (CTO) at Dormakaba, says: “The digital twin is something I heard about from GE Digital and I like it a lot. We can have a digital twin of a building with access control components where the physical world has a mirror in the cloud.”

This digital twin can provide the engineers at Dormakaba with an up-to-date record of every action or event the sensors on the doors have logged; it logs the installation of components and firmware updates, and can be used by Dormakaba’s service team to determine the lifetime of the product along with a security log detailing who accessed the door.

It may even be possible to provide close ties to parts and component suppliers and product lifecycle management, enabling an extremely granular level of monitoring and servicing.

Häberli says: “The digital twin represents what happens in the real world. You can get a whole history of the real product.”

The digital twin is a concept Häberli hopes will enable the company to evolve. “As a CTO what really counts is how we differentiate. Differentiation gets very hard in hardware because other people can develop products.”

Powering physical to digital transformation

He says the company looked at how it did business and how to transform it digitally. For Häberli, the digital twin offers Dormakaba a way to move beyond being a hardware company that effectively sells door locks. “Differentiation is no longer in the product,” he says. “We need to move beyond where we are as an engineering company.”

However, he says: “Buildings are real, so we can’t forget about hardware.” Rather, the company is adding sophisticated tools for planning and managing installations and the buildings themselves, which becomes part of a wider ecosystem for access control as a service.

“We have to focus on enabling the business to change from a pipeline business to a platform business,” says Häberli.

This is a digital transformation, requiring change management across the company’s entire supply chain from components to product through third-party installers to the end customer. Häberli says: “We know our customers; we know how our revenue is generated and we need partners. Digital transformation relies a lot on finding the right partners.”

This requires a mindset change in the Dormakaba sales channel. Häberli says the company's channel partners will effectively need to act as consultants and run data analytics on behalf of the customer.

Cloud-based platform

The company has developed a cloud-based platform it is now starting to share with its channel. Built on Cloud Foundry and infrastructure from Swisscom, the platform comprises a number of business applications and a back-end IoT stack.

The application uses an interactive digital floor plan of a building. "We can drag and drop templates from our products and then make sure the access control works consistently," says Häberli.

The application can also be shared with Dormakaba's distributors. "Depending on their level of security clearance everyone has access to a digital twin of the installation."

Whenever there is a problem, he said the customer can call the dealer, who then has access to the digital twin and so can determine exactly what configuration the customer has installed.

Moving to the cloud

However, like any business change, it takes a long time to move from physical product to the concept of services managed via a digital twin in the cloud. “We are rolling out across Europe and want to offer usage-based pricing,” says Häberli. “But we need to educate the dealers.”

The company is also in the process of moving from DevOps to DevSecOps – which, he said, has its own challenges. He expects it will take Dormakaba three years to complete.

But, in time, the company could support intention-based use cases. In other words, Dormakaba will be in a better position to understand the usage of its products based on context. “It is not just about technology anymore. It is also a strategic concept.”

For instance, he says an intention-based use case could be used in a cyber-physical system. “We capture intention everywhere. What is the intention of the user? We can track user actions and increase their time privileges.” This may be useful if a user requires more time to open a door.

“We are just starting out,” he says. Potentially, the tools Dormakaba is developing will enable it and its channel partners to sell its product in a pre-emptive way. For instance, he says: “We can potentially run AI-powered analytics to look at how to improve efficiency of the building.”

Greater emphasis on software

In the Gartner report, *Digital Twins Will Impact Economic and Business Models,* the analyst firm draws an analogy between the extent of data collection carried out by the likes of Google, Amazon and Netflix, and how digital twins could be deployed in industrial firms to monitor the ongoing health of internet-connected equipment.

“Developing and supporting digital twins in these environments will require the continuous updating of data collection capabilities and curating (for example, management, security, storage and staging), as well as adaptive analytics and algorithms,” Gartner analysts Alfonso Velosa and Dale Kutnick state in the report. As an example, Gartner notes that Tesla issues monthly software updates to its cars.

The analysts warn this will require even more component monitoring and software updating, and will also require that vehicle makers almost become software providers. “Asset operators will have to add software skills to their operations teams as they add smarter assets, and also add software and data terms to their contracts,” the analysts warn.

The findings reflect the conversation Computer Weekly had with Häberli, the CTO of Dormakaba, a company which is in the process of transforming itself and its channel partners to offer digital services for building access control.

Streamline infrastructure management through robotic automation

Stuart Burns, guest contributor

The IT world is ablaze with automation, as organisations strive for improved productivity and accuracy at a significantly lower cost. In line with this, a lot of companies are experimenting with the use of robotic automation tools, but adoption is still in its infancy.

And by robotic automation, we are not talking about the classic robot design that we all know and frequently see in factories, but advanced software-based robotics that can interact with existing web pages, infrastructure and systems that were never designed with robots in mind.

No robotic arms, just a virtual machine with access to system infrastructure. Simply put, software robots work around the process and complement it – they do not disrupt or break it.

What is robotic automation?

The robotic process automation (RPA) concept is well-known and is now frequently found in virtual platform management tools, running in the

background of self-provisioning and quickly moving beyond the basics of deploying and scaling virtual infrastructure.

In short, the idea is bursting out of the "faster, quicker virtual machines/physical deployment" niche and taking over day-to-day administration operations.

The concept has historically been carried out in large companies which do a lot of data collection, extraction and manipulation – think typical mail-in forms, insurance forms and the like.

Its core use case is in doing jobs that are straightforward and repetitive, such as processing forms and importing them into a digital environment, before setting aside paper copies that cannot be imported because the document is mangled or the text is difficult to read. These jobs were typically essential, but low-skilled, and have been almost completely replaced by automation.

Exploring the use cases

This could be considered "old-school" usage, however, because robotic automation is now coming for the more mundane tasks carried out by IT administrators. One of those examples is disaster recovery infrastructure testing. Again, this is quite a repetitive duty, making it ideal for automation.

While it is a necessary evil, disaster recovery infrastructure testing is very time consuming, not to mention expensive and prone to errors. Automating the failover into an isolated environment is the easy bit. The testing component,

however, requires people to set up, configure and test the disaster recovery instance. The people required to manage and implement the setup cost money. Automating the configuration, setup and testing of the application means expensive people are no longer required, but still need to be managed in the classic supervision sense.

That revelation means much bigger changes in the way tasks are tackled. Because automation essentially provides pre-programmed “free people hours”, with individuals continuously testing the disaster recovery capability of isolated environments, mimicking and – more importantly – managing and reporting the results as if a human were driving the test.

There is more to it than that, though. Robotic process automation works with existing infrastructure – and there is no requirement to retool, although that does not mean it should be discounted entirely. Unlike human input, RPA is also available for use 24/7 – the quicker, faster and more accurate selling points for the technology strike again.

Some companies have taken it as far as using automation to have a rolling test cycle for disaster recovery events, dramatically increasing the frequency and creating more trustworthy disaster recovery failover processes for when the day inevitably comes that something goes wrong.

It can also rapidly scale as required. There is no “training and upscaling period” – it’s just a case of deploying more software robots within the existing software configuration. As such, scaling work horizontally becomes a simple process.

Read more about automating IT

Computer Weekly looks at how far the suppliers of systems management tools and associated products have progressed with automation.

We look at the role of artificial intelligence and chatbots in IT service management.

But disaster recovery is just one possible use case. Technology is emerging to handle and manage tasks that aren’t a good fit for simple scripting. An example of this is disk expansion, which is quite a straightforward process to scale, but doing so to hundreds of machines introduces risk and complexity. Automation robotics can be used to modify the disks as needed, as well as carry out checks inside the virtual machine and monitor the hypervisor level for available disk space, which is not a trivial process to try to repeat.

The complexity of doing this by hand, with all the associated actions that need to be repeated precisely, makes it ideal for such things. Resizing 7,000 virtual machine system drives would be extremely error prone, not to mention expensive, if done by hand. Tasks like this can be made programmatically less difficult when done by a software robot.

Therefore, a lot of companies are using automation to permanently carry out functional testing on disaster recovery sites. Such functionality would not have been possible going back even just a few years.

The changes are coming to production infrastructure too.

Automating analytics

An associated benefit of automation is the increased scope for analytics. Each process produces a wealth of data, such as process duration – for example, how long a failover took to complete – so a complete log output for further analysis can be compiled and bundled up to create effective reports. If humans were to do this, the data would be a lot less accurate, if it existed at all, because someone would need to keep note of the very many subtle timings.

We are just starting out on the journey to software robotics. Currently, we are only taking the initial steps to automation, and the systems have to be trained to perform a task. People may think about these as akin to a macro that we all know and love. The eventual destination is that the software will truly understand what it is doing.

How far off that is, no one can say for sure, but any administrator who wants to stay employed needs to ensure they are ahead of the curve.

CIOs, prepare for the sweeping changes robotic process automation will bring

Technology leaders driving innovation should keep in mind the wide-ranging impact automation technologies will have on their businesses. These include artificial intelligence-enhanced robotic process automation (RPA) and software defined infrastructure.

Automation starts deep inside a company's infrastructure. In fact, the modern CIO's infrastructure is now largely based on software. Infrastructure-as-code, including containers, has become the new foundation of modern enterprise infrastructures.

Artificial intelligence (AI) is only one dimension of automation, which also includes many sub-AI-level algorithms and tools. But increasingly, AI suffuses and boosts many different automation technologies – or will do so in the near future.

While RPA is creating a lot of operational value today, its future lies in cognitive-AI enhancement, a technology convergence that will solve many more business problems and will do so in a more sophisticated fashion.

This content is based on a blog by JP Gownder, a principal analyst at Forrester and lead author of "The CIO's guide to automation, AI and robotics" report.

Security Think Tank: Counter application layer attacks with automation

Paddy Francis, guest contributor

Application layer attacks have been with us since the first internet-facing applications. However, a recent survey showed 89% of respondents admitted to an application attack in the past year. Historically, attackers have exploited coding vulnerabilities such as SQL injection and cross-site scripting due to poor practice and testing.

There have been improvements in coding practices and application testing, but at the same time, the Owasp top 10 application vulnerabilities still include SQL injection at number one and cross-site scripting at number seven. It is not clear what is driving this, but it may be a combination of new DevOps development environments and an increase in the sophistication of the attacker.

DevOps greatly speeds up the deployment of apps, some changing on a daily or even hourly basis, which may result in untended security vulnerabilities through a lack of time for proper security testing.

We are also seeing more sophisticated attacks. The best response to these is a combination of improved design and coding to remove vulnerabilities in

applications and websites, along with a layered approach to attack prevention and detection.

This includes denial of service (DoS) attacks that target buffer overflow, or other vulnerabilities used to crash the application or deplete application resources. Such attacks appear to be increasing faster than brute force distributed denial of service (DDoS) attacks at the network layer.

There are a number of tools that can detect coding vulnerabilities and can be integrated into the development environment. These tools will need to be configured for the deployment environment and coding policies, and there may be some push back as they start picking up problems. However, people must realise that any delays are the result of poor coding, not because of the testing.

Automated testing environments can also be used to create new rules and policies for a web application firewall (WAF) as the system evolves. This will then help to protect the applications and application programming interfaces (APIs) as they are added and detect and gather information on attacks as they occur.

A good WAF should be able to cover most of the Owasp top 10 vulnerabilities and also gather information that can be used to identify and track an attack. The other main defence is protection of third-party APIs through encryption and comprehensive access control to prevent session hijacking and ensure least

privilege principle to minimise data exposure. High-value data should, of course, be protected by encryption.

Attacks are becoming more sophisticated and more difficult to detect once the defences are breached, as illustrated by the British Airways (BA) data breach in September 2018.

An analysis by RISKIQ reveals that this was specifically tailored to the BA website, with a domain being registered specifically for the attack and exfiltration designed to blend into the regular operation of the site and so is difficult to detect.

However, on an e-commerce site with a large number of individual user transactions, it may have been possible to detect the appearance of a relatively large number of connections to a single site as something unusual.

In such circumstances, user behaviour analytics (UBA), network traffic analytics or anomaly detection can fill a gap.

UBA is typically used for insider threat detection, or fraud prevention on e-commerce sites. Typically it would flag up abnormal activity away from typical customer, or user activity.

Monitoring source and destination of the network traffic flow either through an analytic use case to detect specific types of activity, or an anomaly detection

system to detect unexpected events are probably the only way to detect such activity.

BA also identified third-party code as giving the attacker a foothold. This highlights importance of supply chain management and ensuring all code on the site is secure.

Application layer attacks are clearly still with us, with the majority aimed at e-commerce, or denial of service. The difficulty in defending against them is exacerbated by the rapid updates to such sites made possible by DevOps, meaning the response must be comprehensive and automated to remove as many vulnerabilities as possible, defend against attacks and actively look for breaches.

Next Article

AI: Black boxes and the boardroom

Tim Wright and Antony Bott, guest contributors

From Elon Musk claiming that the unchecked growth of artificial intelligence (AI) could spawn “an immortal dictator”, to the news that a major bank has started using the technology to detect money laundering, AI is seldom out of the headlines.

Whether or not AI – the science and engineering of machines that act intelligently – will herald a golden age of leisure, or spell the end of humanity, remains to be seen. What is clear, however, is that businesses globally are already embracing the technology to improve how they work.

The possibilities AI presents are vast and its current applications are similarly impressive: it can steer driverless cars; it can read and extract key information from thousands of legal contracts in minutes; and it can review MRI and PET scans and identify malignant tumours with greater accuracy than human doctors.

But AI is not without its risks, be it poor performance, misuse or reliance on bad data. Ultimately, computers can make mistakes, and decisions based on flawed, non-human advice can be costly for a business.

Even F1 driver Lewis Hamilton discovered this when, during the Australian Grand Prix, a software error and incorrect computer data prompted him not to open up a greater lead over rival Sebastian Vettel, who consequently won the race.

Lewis' experience is instructive: mistrust in AI and concerns over a reliance on bad data are two key worries for businesses, but, when planning to use AI, how can businesses avoid a pile-up in the boardroom?

Trust

AI is beset by the "black box" problem: many of its processes lack transparency and can't be easily understood by humans. For example, the developers of AlphaGo, Google DeepMind's system, could not explain why their system made certain complicated moves in beating the human world champion of board game Go.

If we can't easily comprehend AI's conclusions, how can we be sure that automated processes are playing fair with their decision-making?

Of course, this makes it difficult to accurately assess risk, but businesses still must consider the environment in which the technology is being used: systems running critical infrastructure such as nuclear power stations must set the highest bar for what is considered safe.

Before incorporating AI, businesses may need to convince a regulator, perhaps by using software to monitor the technology – algorithmic auditors that will hunt for undue bias and discrimination. This will likely impact performance, since the system will divert processing power to self-analysis, but it could mean the difference between the system getting rejected or approved for commission.

Bad data

Outcomes achieved by an AI system will only ever be as good as the quality of data on which they're based. There are many variables to determine the quality of inputted data: are the data sets "big" enough, is "real-world" data being used, is the data corrupt, biased or discriminatory?

With so much potential uncertainty, businesses should take every effort to minimise the risks. Where data is sourced from a third party, contracts should require transparency around lineage, acquisition methods, and model assumptions, (both initially and on an ongoing basis where the data set is dynamic), and there should be mandated security procedures around the data, to prevent loss, tampering and the introduction of malware – all reinforced by comprehensive rights to audit, seek injunctive relief and terminate.

Finally, common sense should apply – businesses should not rely too heavily on a limited number of data points and support big data analysis with other decision-making tools.

The corollary of these concerns is to give tremendous power to those who own large repositories of accurate personal data and we therefore expect the issue to become a significant focus for regulatory and contractual protection in the coming years.

Isaac Asimov, the famous science fiction writer, once laid down a series of rules to protect humanity from AI. Perhaps it is time businesses did the same. After all, we can't know the future, but we can prepare for it. And with AI, the future is now.

Next Article

Smart speakers: How to give apps a voice

Cliff Saran, managing editor

Research from Gartner shows that about a fifth of people in mature markets are voicing a question to their smart speaker at least once a week.

However, the same study reveals that 10% admit they have stopped using the devices. “The drop-off rate is quite high,” says Annette Jump, research vice-president at Gartner. “This is very similar to wearables. Consumers try a device, and either it does not do enough for them or it does not recognise their voice. Or they can only link to a Google email account or the user does not have a smart home.”

While being able to speak to a device offers a lot of convenience, there are personal boundaries in what people will ask, she notes. This is one of the challenges businesses need to take into account when assessing how to go about developing a voice assistant app.

“Using voice control on a smartphone is location-specific, while at home people will ask the device to play music, control smart home devices and ask about the weather,” says Jump.

Amazon Echo smart speakers use in-built voice recognition software Alexa to connect to apps called “skills” that add third-party functionality.

But Jump asks: “While Alexa has 10,000 skills, how many are actually useful?”

According to analyst Rethink, voice functionality is like Apple’s AppStore, whereby it acts as a route to market for services, and those services will pay a percentage of revenue brought to market via voice.

Rethink says Amazon Alexa wants as many services (skills) as possible to build momentum.

“Later, we believe it will be in a position to charge services for a service delivery skill where it is proactive – for instance, offering a cheap taxi ride from Uber to take you to work because it senses it is raining, and asking for some of the Uber payment,” says the analyst.

The voice user interface is a growing area of interest, according to Accenture. “We are seeing an explosion in interest in voice assistants due to the launch of Amazon Alexa and Google Home. The technology is now ready for mass adoption,” says Emma Kendrew, artificial intelligence (AI) lead for Accenture Technology.

She says voice assistants provide a much more natural experience for the user. Accenture has found that the use of voice interaction for talking to a machine is

more natural among digital natives, and it is they who tend to be driving demand for new apps and smart speaker devices.

This is encouraging businesses to experiment with voice, says Kendew. “Customer-facing industries are leading on this. Financial services are also experimenting with specific use cases, such as using a voice assistant as a mortgage advisor,” she says.

“Organisations are interested in how to transform customer experience and are looking at where voice assistants should and shouldn’t be used.”

Getting started with voice

Kendew says it is very important to design voice assistants with the user in mind. As these new technologies become available, organisations need to think about what is the right way to use them to achieve their objectives.

But where should they start? Bill Kingston heads up Elixirr Creative, an agency which has built a number of voice-based apps. “With Alexa,” he says, “you can use an existing framework to turn an RSS into a skill, or take the custom skill route. The question is, is it worth developing a custom skill?”

In Kingston’s experience, it is usually better to start with an existing feed and see how well it works, then decide whether to invest in developing a custom app.

However, businesses can find that retrofitting an existing app with voice assistance may not work particularly well. Such an Alexa skill can be a bit one-way, according to Kingston, which might not fit the natural feel everyone is striving for.

Enterprise software company Unit 4 has developed an intelligent assistant called Wanda, which uses natural language queries. The company has developed a proof-of-concept voice assistant using Microsoft Cortana.

Claus Jepen, chief architect at Unit 4, has overseen the project. “Old rules-based interactive voice response applications are not beneficial to anyone,” he says. “We paid lot of attention to ensure Wanda understands complex sentences.”

He says one of the greatest technical challenges is the ability to infer meaning in a conversation – a human trait that is very hard to program.

As an example of a piece of dialogue that the voice assistant should be able to handle in a human way, Jepen says: “If you ask, ‘Please give me revenue results’, a human would automatically assume that infers the latest results. But often a voice assistant will only work programmatically and will need additional contextual information. For instance, Alexa would probably ask the user to clarify what they wanted, asking, ‘What results would you like?’.”

Since it is pretty easy to develop a simple Alexa skill based on an existing app, Jepen says a lot of people are getting involved, but they are not sure what to do with it. “Unless you put a lot of effort into designing the dialogue and handling context without having to get the user to restart the conversation, you end up with nothing,” he warns.

A good voice interaction requires specialist skills, says Jepen. “You need conversational designers, just as you have user interface designers. But we don’t know how well the voice assistant works until we get real users.”

While testing its Wanda assistant, Unit 4 collects conversation snippets that fail to provide the human tester with an appropriate conversational response. “When the Wanda assistant can’t figure out what the user wants, we pull out the dialogue and upload it back into the training,” says Jepen.

In effect, Unit 4 teaches Wanda to respond correctly. It is machine learning.

Testing is a lengthy process, according to Elixirr Creative’s Kingston, and even getting voice analytics back takes time.

“People won’t necessarily interact with an Alexa skill in the same way. After you have tested with six or seven users on the same theme, you get a good idea of what works,” he says. For instance, voice analytics can be used to identify where people pause in their interaction with Alexa, which can help developers work out how to make the interaction more like a conversation.

So creating a seamless conversation flow with a voice assistant is not easy, especially as the assistant must attempt to do something useful.

“Voice experiences will consume a significant amount of our time completing everyday duties in the coming years due to their time-saving benefit,” says Lee Mallon, managing director of Made for Voice, a company specialising in voice assistant software for the enterprise.

“These experiences must respect the user’s attention by solving their problems as quickly and in the fewest words possible. In one of our voice experiences, we found a 20-30% engagement drop-off by adding just two additional words to a three-word sentence response.”

Going forward with voice

Neither Google nor Amazon have not stopped at voice assistants. Amazon now offers a screen on its Alexa Show device.

“What is starting is combination of voice interaction with screen interaction – blending two channels into one seamless flow,” says Accenture’s Kendew.

Combining a visual user interface with voice opens up new opportunities for customer interaction, such as providing an intelligent kiosk at an airport or shopping centre.

From the conversations Computer Weekly has had, one thing is clear: there are two main voice platforms and they are very different.

“The user journey is very different on the Amazon Alexa platform in comparison with the Google Assistant platform, with Alexa requiring users to ‘subscribe’ to skills that interest them, like our own daily briefing skill,” says Rob Fricker, product manager at Time Out. “Once subscribed, you ask Alexa, ‘What’s my daily briefing?’, and it replies with Time Out’s top three things to do in the city today.”

Gartner’s Jump believes Google Assistant has the more complete voice interface, when compared with Alexa. Being linked to a Google account, Google Assistant gives users access to their calendar, so it can tell them more and prompt them to do things. But as with Alexa, she says there are lots of times it says, ‘I don’t know’.

Elixirr Creative’s Kingston agrees. “Google Assistant has so much data, it is probably a step ahead,” he says.

There is not going to be one overall winner, and as Unit 4 has shown, in business there may be an opportunity to make use of Cortana on Windows 10.

The experts Computer Weekly has spoken to recommend that businesses look at voice user interfaces. These are not quick wins or just another channel to

market. A truly compelling voice interface needs time and effort to develop and test, and teams will need dialogue experts, just as they now have user interface experts.

Time Out speaks

Time Out recently launched a conversational app on the Google Assistant, designed to provide a personal touch to keep the conversation going. Rob Fricker, product manager at Time Out, explains how the app was developed.

“One of the most interesting things about the process was the conversation design – it’s something we were all completely new to,” he says.

“We built the app on Google’s Dialogflow, which uses natural language processing to understand voice input. We spent a lot of time thinking about the different paths conversations might take and how to address questions the assistant might not know.

“This included watching people in conversation with each other when one was acting as a computer – it’s funny how polite people are when they think they’re talking to a computer. We also measured how people interacted through various iterations which was crucial to determine the kind of questions people might ask the app.

“User testing throughout the development cycle also helped set expectations of the functionality the app should be able to perform, and we were able to

prioritise certain features based on this feedback and learnings that users would benefit from more."

Next Article

Automation key to unravelling mysteries of the universe at CERN

Marc Ambasna-Jones, guest contributor

In 2012, when most of the UK was gearing up for the London Olympics, strange words started to enter our vocabulary. Terms such as “Higgs boson” and “Large Hadron Collider (LHC)” were hitting the mainstream headlines and CERN, a European particle physics research centre based in a north-west suburb of Geneva in Switzerland, was suddenly all the rage.

On 4 July that year, two experiments had led to the discovery of the Higgs boson, or “God Particle” as it was dubbed – a discovery that saw theoretical physicists Peter Higgs and Francois Englert receive the Nobel Prize for physics a year later.

It was CERN’s gold medal moment, but as is the case with the athletes that were about to embark on their own voyages of discovery, a lot of the hard work that goes into such high-profile successes is unseen. It had taken decades of research and experimentation to reach this point, and it wasn’t all down to the work of the physicists.

There are, in fact, two CERNs: the one we all hear about that is trying to solve the mysteries of the universe, and the other one, which is somewhat less glamorous.

There are, according to David Widegren, head of asset and maintenance management at CERN, 13,000 people working at the complex at any point in time, including up to 10,000 visiting particle physicists. The rest are looking after the nuts and bolts and the daily running of the place. This is a challenge, with 700 buildings, roads, car parks, an electricity grid, complex research equipment and, of course, the accelerator complex with its 100 million components.

In 2008, CERN had an accident when a faulty electrical connection between two magnets led to an explosion. CERN released a full explanation, and Widegren claims that if they'd had machine learning and an automated asset management system in place back then, it could potentially have been avoided. That's the theory anyway.

Automation required

The point Widegren is making is that CERN has grown so big it needs automation. It already uses Infor's enterprise asset management software EAM to help keep track of around 2.1 billion assets, including the 100 million components that make up the collider. CERN is, in fact, one of Infor's oldest customers. It's been using EAM for over 20 years, and although the earlier iterations were, in Widegren's words, "quite basic," today he says EAM has to

be powerful and scalable enough to cope with the increasing demands of CERN.

“The two million assets we manage through EAM generate 800GB of data every day,” says Widegren. “If we are to minimise unplanned downtime at CERN, and given that we get a billion Swiss Francs a year to research physics, we need to behave like an enterprise and use this data to maximise the visibility of our systems and assets.”

CERN now has lifecycle tracing of all its assets, from manufacturing through to waste management – important given that some of the components become radioactive. Not everything has sensors, but Widegren talks about the site’s industrial internet of things (IIoT) network, the need to use sensors more on machines and components to improve management, and how automation will eventually help reduce downtime by enabling alerts to potential issues.

“The next phase is to use the data to drive automation and predictive technology,” he says.

CERN has been in discussion with Infor to trial its machine learning engine Coleman AI, so Widegren and his team of 12 can look for correlations and pattern recognition to see how they can better understand how the colliders behave and predict potential failures in the future.

It's the latest step in a CERN-wide initiative, that started seven years ago, to modernise its IT and asset management. While Widegren has been focused on the rapid escalation in assets and services – he says they have tripled since 2011 – managing the IT function for a number of users that jumped from a few hundred to 2,000 has also been a challenge.

Automatic ignition

According to Tim Bell, compute and monitoring group leader at CERN, automation has been an ongoing development and something CERN has been trying to increase in all its processes, when it comes to providing IT services for the community.

In 2011, the IT team was using a self-built tool for IT configuration management, which by Bell's own admission, was "very limited". Something had to change, especially given the scaling of the whole community and the LHC preparing for its second, historical run.

The team adopted Puppet, an open source configuration management tool, with a specific aim of making the deployment of its IT infrastructure more palatable. The idea that the IT team would have to configure and manage thousands, rather than hundreds, of machines, documentation and development was a big enough driver to get the Puppet software in place as soon as possible.

“We wanted to remove the limitations of our old solution – mainly that there was not much expertise outside of CERN and we felt we could profit from a larger skills pool by using a more popular solution used elsewhere,” says Bell. “That was also making it difficult for us to hire engineers with the right experience. We also had to take care of documenting and evolving our configuration system, basically on our own.”

Reducing deployment time, while essential to the ongoing viability of IT systems at CERN, has also had a knock-on effect in terms of the IT team. Automation is already changing the way things are done and the roles of key staff.

“Currently, new services can be deployed in a matter of hours and, more importantly, resources dedicated to each service can also be enlarged or reduced dynamically. This is of great help when coping with service load-related problems, as well as to make transparent to users hardware interventions,” says Bell.

“Adding a new node to a given service is a matter of executing a command using the tools that have been developed at CERN, which integrate Puppet with our OpenStack-based service provisioning infrastructure. As a result of automation, we are reducing the size of the team of engineers that had access to our computer centre, as the number of calls they get has been reduced dramatically.”

Keeping the physicists happy is, of course, one of the priorities, but that can only really be achieved by making sure they don't have to be hamstrung by the IT. Reducing the need for support calls was essential, something which Bell believes they have already achieved.

The team adopted a DevOps approach to enable the continuous introduction of service changes to minimise service disruption. It was a new way of working but, says Bell, it fitted with the pattern of a large and constantly changing team. It gave the team structure, and support tickets went from 60,000 for the system administration team in 2011 to a few hundred today.

So, has Puppet and increased automation led to loss of jobs in the IT function, or has it meant the redeployment of roles?

"Definitively, the rate of managed services per headcount has increased significantly, as has the total amount of physics-compute resources we run," says Bell. "At the same time, the number of members in the IT department at CERN has stayed stable, as the number and size of services has increased with time. We have also been able to enhance IT functionality, such as improving service monitoring or working on new software developments for the physics community, in which we can employ more resources."

While Bell contemplates his next challenge – to be able to provision services using public cloud resources (he claims the team has already done some proof of concepts provisioning batch worker-nodes in external clouds using Puppet for

managing their configuration) – he says there are some lessons for all IT teams to learn from his experience.

“We believe that reusability and specialisation have been key for the success of our Puppet deployment. We’re making use of plenty of upstream Puppet modules, which has contributed significantly to reducing engineering time spent on writing configuration,” he says.

“As well, we have lots of domain-specific experts in the department who are in charge of providing and maintaining configuration for other service managers to build their services on top. For example, if you’re responsible for a content delivery service, you can focus on integrating components, delegating, for instance, the configuration of your back-end storage and the monitoring by simply including centrally maintained Puppet code that you can customise if needed using Hiera.”

Both Bell and Widegren are front and central to CERN’s infrastructure modernisation. It’s the hidden work, the hard yards that are needed to make the physics possible. What’s incredible is that it has taken CERN so long to get here. You would think that the organisation that gave us the world wide web would always be playing with the latest toys. It did, after all, have some of the first touchscreens back in 1971.

Infrastructure, though, is a little different. It's now big and can be unwieldy, which is why CERN has turned to powerful management tools such as Infor and Puppet, and is looking to increase automation and its use of machine learning.

As Widegren says, "it's a constant simplification of complex machines and systems". While it might not be the answer to the mysteries of the universe, it's making the physics possible. As the LHC goes into a period of hibernation for repairs and upgrades, it's a maxim by which, for the next 18 months at least, everyone at CERN will be living.

Getting more CW+ exclusive content

As a CW+ member, you have access to TechTarget's entire portfolio of 140+ websites. CW+ access directs you to previously unavailable "platinum members-only resources" that are guaranteed to save you the time and effort of having to track such premium content down on your own, ultimately helping you to solve your toughest IT challenges more effectively—and faster—than ever before.

Take full advantage of your membership by visiting www.computerweekly.com/eproducts

Images; stock.adobe.com