Ovum
TMT intelligence | informa

Tractica

AI
VIEWPOINTS

www.ovum.informa.com

Dear reader,

Welcome to AI Viewpoints, a selection of Ovum's current research on the topic of artificial intelligence (AI), published in partnership with Tractica, our sister research organization within Informa Group. Ovum and Tractica's coverage of AI represents a broad spectrum of client-driven interest from our analysts across vertical industries and horizontal technology research. Our purpose in unifying a selection of this research is the belief that while AI generates a great deal of excitement and interest, there is little coherent analysis focused on demystifying the technology while exploring its true value – the ability to improve and automate economic and social activity. AI is an undoubted work-in-progress. From its early days at the dawn of the modern computer, through today and its integration into the devices and apps, and the networks that connect them that we use every day – feeling the impact of AI has been a long time coming. For many it is still the realm of science fiction – a valid viewpoint – but with a growing economic imperative firmly supporting it, the range of practical, and value-adding uses for AI technologies is growing rapidly.

Previously I have said that, "the best technology is the one you don't realize you're using." This applies very directly to AI. Its inherent technical complexity is pushing it into layers of technology most of us may not even be aware of. Whether it is the intelligent orchestration and self-healing of your mobile carrier's network, the real time optimization of delivery route for your online order, or intelligent enhancement of the pictures on your device - most of us will feel the benefit, without being exposed to the technology.

The question remains, how do enterprises best employ these new capabilities? The answer should be subject to the usual rules of investing in and deploying new technologies – where it can either help grow revenue, reduce costs, enhance profit, or manage risk. In this regard, AI is no different to any other technology, but other, important differences do exist. Perception is often that AI will displace many humans from their current work, this is an undoubted possibility. However, for now, the majority of practical use cases for AI technologies are either about enhancing the ability of people to do their work, augmenting their activities and decisions with better, data-driven insights. Or, removing mundane, repetitive and low-value work, freeing time to focus where value may be created.

A final brief mention on the subject of responsibility. As with all technology, the opportunity for it to be misused, either by design or accidentally exists; this is keenly felt on the subject of AI where the volume of data that powers it, and the potentially sensitive nature of that data is such that the risk is amplified. Transparency and governance have a major role to play in mitigating the risk, and while legislation such as the European Union's General Data Protection Regulation are coming into force, it will be the responsibility of those enterprises that deploy it to ensure they do so in a way that grows our trust.

Whether you are of one of Ovum's enterprise end-user, service provider, technology vendor or professional services clients, we hope you find this eBook a useful and informative bird eye's view of some of the practical thought leadership that is at the heart of Ovum's research. We always welcome feedback and engagement with our clients, and the wider-market and look forward to hearing from you both at our flagship event, Digital Futures, and over the coming months.

Yours,

Tom M. Pringle
Head of
Technology Research

Clint Wheelock
Managing
Director, Tractica

Highlights

- Michael Azoff discusses the importance of data processing for machine learning applications and why this is a lifecycle issue: ML applications need to be continually monitored for changes in the data being input to ensure that they remain valid.
- Tom Pringle positions AI in terms of narrow and general AI, and how AI is being found useful for tackling problems today – not necessarily replacing people but making it possible to reassign people to have them perform higher-value tasks.
- Laurent Lioté offers five key steps to successfully implementing AI projects, addressing issues such as the role of data science in the enterprise and the importance of skills training.
- Aditya Kaul showcases explainable AI and its impact on AI adoption.
- Miriam Deasy addresses the question of how AI brings value to the organization. Enterprises must be clear about the business challenge or pain for which AI may offer a solution and support creative approaches to engineering new solutions.
- Adam Holtby asks how AI capabilities will help further empower the mobile workforce, sees increasing use of AI in mobile applications, and covers the need for AI chips at the edge to support such applications.
- Michael Azoff provides some clarity on the AI hardware accelerator market for all types of AI applications and stresses the importance of choosing the right accelerator for the different modes of AI model building: training and inference modes.
- Hansa Iyengar looks at the role of Agile and DevOps in enterprise AI application development, and the application of AI to IT operations.
- Eden Zoller offers her assessment of the AI consumer market and how vendors can improve the user experience with such applications.
- Adaora Okeleke discusses opportunities AI presents for CSPs to transform network operations.
- Alexander Harrowell examines patent applications relating to AI and questions current levels of innovation in the field.

Contents

Managing the machine intelligence application lifecycle

Making enterprise artificial intelligence succeed

Michael Azoff
Principal Analyst, Infrastructure Solutions
michael.azoff@ovum.com

Michael Azoff is a member of Ovum's IT infrastructure solutions group, leading a range of software development and lifecycle management research initiatives.

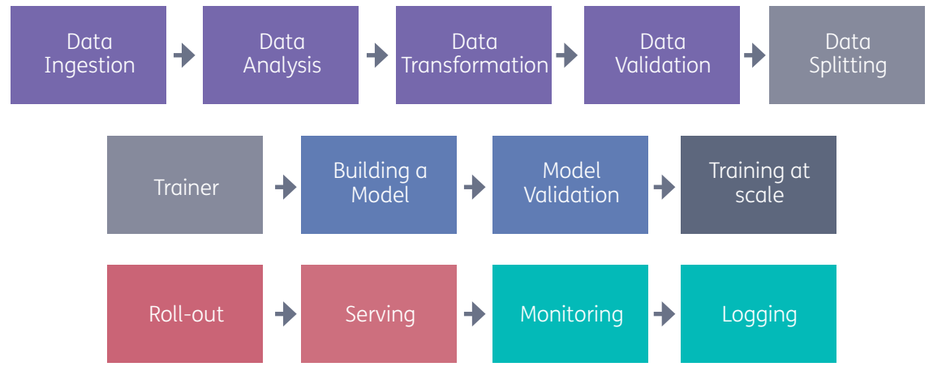
Almost every enterprise is experimenting with artificial intelligence (AI) and machine learning (ML), and as with all matters relating to technology, there are always the additional dimensions of people and processes to consider. Enterprises are rushing to test the potential benefits of AI initiatives, particularly the latest ML ideas such as deep learning, and often failing to deliver on the promise because they are failing on delivery, where people and processes are particularly important. Research by the McKinsey Global Institute in June 2017 reviewed 1,600 active AI cases and found that 88% had not progressed beyond the experimental stage.

Immense pressure from the business side of large enterprises is leading data science teams to release initial ML services to test the waters. The method by which they deliver these is more of a proof of concept than a scalable, repeatable process. However, the business needs a scalable process, often on the scale of 100 applications deployed into production with hundreds of associated models, and it is when this level of scale applies that a methodical, DevOps-style automation should be looked at.

While building ML applications requires great skills, it is a well-understood challenge, with many supporting frameworks available. However, everything else around the model building poses new challenges (especially at scale), and these can be categorized as follows:

- Composability: ML applications have their own lifecycle, separate from regular enterprise applications. Reading Figure 1 from top left to bottom right, the process starts with data processing activity, then training selection, model building, testing, scaling, and finally the process of rolling out the application, monitoring, and logging.
- Portability: Many of the steps require the experimentation of supporting technology, including UX, tooling, framework, storage, runtime, drivers, OS, AI accelerator, and hardware. Some of the steps need to be repeated because the training, dev-test, and production cloud environments are all different. If a change occurs, such as if a new driver for GPUs appears, many of the steps must be repeated.
- Scalability: Scalability of ML applications requires consideration of many factors, including AI hardware accelerators, CPUs, disk/networking, skill sets (software developers, ML experts, data scientists), teams, and experiments.

Figure 1: Machine learning lifecycle



Source: David Aronchick, Google

A lot of the work in managing the lifecycle can be separated out by using the application container paradigm, such as Docker, running on Kubernetes, and even further simplifying development management by using OpenShift or Cloud Foundry. However, even with this separation of activities and exploiting all the advantages of containers, there is still a requirement for the developer to be an expert in a host of activities, including containers, Kubernetes service endpoints, persistent volumes, scaling, immutable deployments, cloud APIs, and DevOps.

Ovum has evaluated two solutions that can help manage this challenge: the open source project Kubeflow and the startup product ParallelM MCenter. We expect this space to expand as the need to manage the ML lifecycle becomes recognized and more enterprises look for solutions that can help.

Narrow AI and its power to add value is **where the conversation should be**

Tom M. Pringle
Head of Technology Research, Ovum

tom.pringle@ovum.com

Tom Pringle is head of technology research at Ovum, focusing on data and information management, analytical technologies, and the evolution of enterprise applications. His topic expertise lies in digital transformation, automation, and customer data.

Discussion about AI is an intellectually engaging conversation, but often one that becomes drawn out, argumentative, and overly focused on a long run (where, according to John Maynard Keynes, we are all dead, but hopefully not as the result of human-terrorizing AI). A substantial part of the debate about AI and its use is drawn from popular science fiction concepts. Whether it is the murderous plotting of HAL 9000 in 2001: A Space Odyssey, the nuclear destruction of human civilization by Skynet from Terminator, or V.I.K.I.'s well-intentioned but destructive takeover from I, Robot, the overwhelming majority of cultural references to AI have negative outcomes for humans. More positive examples do exist – think C-3PO from Star Wars – but they do not weigh on the mind in the same way as those with a more apocalyptic result.

Add to this debate the words of caution issued by some of our most influential thinkers in science and technology, such as the late Professor Stephen Hawking or Elon Musk, and it is not hard to see why conversations about the use of AI are quite as contentious as they are. But they are often missing the point: practical, commercial applications of AI have nothing to do with the type of AI that dominates popular consciousness.

AI can be split into two types: first, general AI, which can be defined as a system that mimics human intelligence so closely that it is likely impossible that a human could tell the difference, and second, narrow AI, which can be defined as machine intelligence for specific capabilities and is applied to tightly defined tasks – for example, prioritizing sales leads based on a calculation of likelihood of success. A general AI may be developed, it may happen sooner than is predicted, or it may never be feasible. Either way, a general AI is not what enterprises with an eye to deploying technology to grow revenue, contain costs, improve profit, or manage risk should be focusing on.

The tangible, increasingly proven AI solutions enterprises can consider today are narrow AI, and of the host of technologies that can be labeled AI, it is machine learning that powers nearly all of them. Machine learning, along with many of its AI stablemates, has a long history. Machine learning was first discussed in an IBM paper in 1959 as part of a research project by Arthur Samuel that culminated in building a checkers game for IBM's first commercial computer, the IBM 701. This raises an obvious question: If this technology has been around for nearly 60 years, why is it creating so much buzz today? Two important things have changed: the volume of data and the scale of computing power available.

Machine learning feeds on data, and as the scale of the data and the complexity of the problem being tackled grow, so does the necessary processing power. We often talk about enterprises striving to be data driven, but think bigger – the world is increasingly data driven. With smart devices in the hands of billions and connected devices forming new and emerging networks, the scale and speed at which data is now created has created a problem: humans cannot keep up. This is where machines excel, sorting through the noise of data to find and surface valuable, actionable information. The machine presents information and options to users at this point, helping augment the decision process with data-derived insights and suggestions – not making the decision for them. These capabilities do raise another question: Could narrow AI replace humans for some tasks? The answer is yes, but with a substantial caveat.

Where AI technologies are used to automate work, their focus is to remove the humdrum, repetitive tasks that consume human endeavor, while adding little value. This approach is not novel in our economic history and is the method by which automation has changed the economic operating model over time – from agriculture to manufacturing, automation is now coming to services. The immediate net gain here is found in productivity: less time wasted on work that can be handled by a machine means more time focused on delivering value-adding work. These capabilities are likely to be consumed by the majority without them even realizing they are being assisted by AI – baked into their go-to apps, in either a work or personal context. This brings us to one more topic that demands our attention – the governance and ethical use of AI.

Many of the concerns associated with use of AI to help inform decisions focus on matters such as algorithmic bias and transparency of decisioning. With algorithmic bias, it may be the case that, for example, a machine learning solution either by design or by accident produces a biased result that may be related to gender or age of candidates during a job application process. This type of problem is relatively easy to tackle. As machine learning is entirely dependent on data, deny it the data that enables obvious bias risks. Other examples may be more nuanced. This is where transparency comes in to play. The explicability of the decision process operated by AI must be clear, enabling audit and therefore greater confidence and trust. Current legal frameworks and ethical standards will need to continuously evolve to tackle these challenges, and there is always room for organizations to ask themselves the question, "Just because we can, should we?"

Our use of AI today, and for the foreseeable future, is about providing people with a guiding hand to help them achieve what they need to do, faster and with greater confidence built on data, not gut instinct. The narrow AI capabilities finding their way into the applications we use at work, the smart devices we use every day, and the behind-the-scenes management of the networks of connections that join everything up are not the stuff of a dystopian future. They are a boost to our human capabilities that allow us to deal with a world where change is the only constant, and the need to adapt is continuous.

FIVE KEY STEPS TO **SUCCESSFULLY** IMPLEMENT AI PROJECTS

Laurent Lioté

Analyst, Data and Enterprise Intelligence

LaurentOlivier.Liote@ovum.com

Laurent is an analyst in Ovum's Information Management team. He focuses on self-service analytics solutions and data discovery, advising established IT vendors seeking new markets, new players providing their own spin on analytic capability, and service providers searching for new ways to effectively exploit their data.

I fear this title is a little misleading because it mentions artificial intelligence, and the expression is often associated with technologies that are far from being enterprise reality (think general-purpose AI à la J.A.R.V.I.S from Iron Man). Currently, if enterprises talk about AI projects, they refer to machine learning automation for business problem solving, and in a very limited number of cases about deep learning techniques. While we are dealing with semantics, I would like to point out that ML is often conflated with data science, and although ML projects involve data science, they also require other skills and expertise. For the successful completion of ML projects, enterprises must therefore think carefully about five issues: the project's aim, the team's composition, where the team stands in the enterprise structure, the training offered, and the feedback process.

FIVE KEY STEPS TO **SUCCESSFULLY** IMPLEMENT AI PROJECTS

How is the ML project adding value? ML is hype, but doing ML for ML's sake is pointless. The first step is to identify which areas of the organization and which business problem will benefit from ML. This should ideally be driven by the C-suite in conjunction with data scientists and line-of-business (LOB) stakeholders. This provides the right blend of technical expertise, subject-matter expertise, and financial backing to identify the business problems that are more likely to be solved using ML.

Blended expertise should be reflected in the team composition. The skills necessary and resources available will of course vary by organization and project, but the need to mix technical and business knowledge is paramount. On the technical side, key skills include statistics and programming to develop the models and data engineering to set up the right data pipeline and preparation. On the business side, subject-matter experts (SMEs) are required to ensure that the model matches the right business question and to interpret the insights generated. To truly be successful, ML should be a team sport, and project management and communication skills are essential to bridge the gap between technology (data scientists) and the business (SMEs).

Where do the data scientists sit within the enterprise This, again, depends on the size and structure of the enterprise but also what you believe should be the role of data scientists. If you think data scientists should be more involved in and understand the business, then embedding your data scientists within LOB is a good move. On the other hand, if you want to grow your data scientists' technical expertise, then having a separate unsiloed data science unit allows data scientists to exchange with each other.

Train your staff to gain the skills they do not currently have. This includes, for instance, further technical training for data scientists and high-level initiation to ML for SMEs. Not only will this help retention by keeping employees engaged, it will facilitate collaboration between more knowledgeable business and technical staff.

Have a clearly defined feedback process. Once the model has been trained, simulated, and deployed, make sure you have a feedback loop in place. This includes evaluating models for drift, peer review of successful and unsuccessful projects, and establishing ROI for deployed ML models. The final step should be to educate the wider business to ML by doing multiple, frequent presentations and writing up documentation to share the lessons learned with both technical and business stakeholders.

EXPLAINABLE AI AND ITS IMPACT ON AI ADOPTION

Aditya Kaul
Research Director
Tractica
aditya@tractica.com

Aditya Kaul is a research director at Tractica, with a primary focus on artificial intelligence and robotics. He is the lead architect of Tractica's market sizing and forecast models for the AI and robotics sectors.

Explainable AI (XAI) is not a new term. Back in the 1980s, explainability of AI expert systems was a big topic of interest, but soon after that, we went into an AI winter, and generally forgot about AI explainability, until now, when we see XAI reemerging as a major topic of interest.

XAI refers to the ability to explain the decisions, recommendations, predictions, or actions made by an AI system. Explainability is closely related to interpretability, as AI models that are not interpretable or introspected upon are hard to explain. Another term that is often used is justification, which refers to the ability to be able to justify a decision or action from a model, which does not necessarily mean that the AI model is interpretable. Explainability, interpretability, and justification are all terms that one hears very often these days in relation to modern AI, especially in relation to machine learning and deep learning models. All these terms can be put under the umbrella of XAI. AI today is largely dominated by machine learning, and particularly deep learning, which is often compared to a "black box" where it is hard to peer under the hood of models and understand why an AI model made a certain decision or gave a specific output. There is an interesting tradeoff between interpretability and accuracy, with low-accuracy models like linear regression being highly interpretable and high-accuracy models like deep learning being difficult to interpret.

Why do we need explainable AI?

One can argue that it is hard to really nail down the specific reasons why humans make certain decisions, which could range from personal preferences, social factors, and environmental factors or be subconscious factors, which are even harder to interpret. While it is hard to provide clear and precise explanations around human decisions, we can perform some level of introspection, interpretability, or justification to explore the reasons why and how a decision was made. This is what XAI is trying to achieve, for the most part. Also, XAI allows humans to have more confidence in an AI system, and have faith and trust in the results. We can hold AI systems accountable and liable for their decisions, ensure that they are compliant to regulations, help verify and debug that the AI system is making the right decisions, and eventually help improve AI systems.

Tools for AI Explainability

There are several tools and frameworks that are working toward delivering explainability in AI today. Below is a selection of some of the key tools and techniques in XAI:

- **Introduction to Local Interpretable Model-Agnostic Explanations (LIME) Framework (University of Washington):** This is a tool and framework that is aimed at classifiers for both text and image domains. LIME is model agnostic, breaking down images and text into subcategories that can then be flagged as key reasons why a classification decision was made.

- **Generating Visual Explanations (UC Berkeley):** This is tool that is used to provide explanations around visual classifiers like a deep neural network (DNN) or a convolutional neural network (CNN), providing reasons why a specific classification decision was made. This is dependent on the human labeling and supervised features, providing some reasoning behind why a bird was labeled in a certain class.

- **Building Blocks of Interpretability (Google):** This is a powerful and rich user interface (UI)/user experience (UX) and visualization tool, going one step beyond the LIME framework, which breaks down the process of how a neural network works, especially across its hidden layers, not just input and output. Essentially, every neuron activation is paired with a visualization, but the tool also provides a way to see which groups of neurons are being activated for a spatial location in an image.

- Most enterprise AI platforms have been slow to adopt explainability, but some are ahead of the curve. For example, ai has embedded the LIME framework into its AI development platform. Bonsai is another AI platform company that is incorporating XAI techniques into its platform. DataRobot is another AI platform that provides feature impact analysis, which is a step toward explainability. Kyndi is marketing itself as one of the world's first AI platforms that is focused on explainability, pushing machine reasoning techniques, rather than deep learning techniques as the way to drive XAI.

- Some startup companies have also emerged that are pushing for explainability in specific domains. Factmata is applying explainability to media and news analysis tackling the issue of fake news.

Verticals that will see the most impact

XAI is not a mantra that is necessarily needed for all AI applications and use cases. For example, if Google Maps is recommending a specific restaurant or Spotify is playing a certain song from my recommended playlist, it does not really add much value in terms of seeing the detailed guts of the AI model that is behind those recommendations. But if Facebook's AI is tagging a specific news story as fake, then it would be useful to see why it has made that decision. Similarly, if an AI doctor makes a diagnosis of a medical image scan, it is necessary to have explainability as a core part of the AI model. The same would apply to decisions about loans, insurance claims, or legal contract analysis where explainability becomes almost mandatory from a regulatory and ethical business practice standpoint. Autonomous cars would also fall under the XAI category when accidents caused by autonomous cars will be under heavy scrutiny from regulators, as has been recently seen in the Uber crash in Arizona. Heavily regulated sectors like healthcare, finance (including insurance and investment), legal, and automotive will all need to adopt XAI soon. The military and defense sector is also likely to see some impact of XAI, especially in drone footage and surveillance.

Overall, the largest impact of XAI will be in vision and text-based use cases, with vision likely to be one of the hardest to implement because of the complex nature of algorithms and techniques that are used. Based on Tractica's AI Market Forecasts, vision-based AI is largely dominated by verticals such as the public sector or government, driven by video surveillance, and in the automotive sector because of self-driving cars. For text, or specifically language-based AI, legal and insurance are the two biggest sectors in terms of revenue potential.

Currently, there seems to be very little adoption within these verticals of XAI techniques or understanding of the risks associated with not using explainability, justification, or interpretability as core principles. The success of these sectors and the use cases will be very much dependent on how well explainability is adopted.

TO AI OR NOT AI, WHAT IS THE QUESTION?

Miriam Deasy

Principal Analyst,
Advanced Digital Services
Miriam.Deasy@ovum.com

Miriam Deasy is a principal analyst in Ovum's Advanced Digital Services team. She focuses on application services for large enterprises, AI, automation, and business process outsourcing. Her key areas of expertise include IT services contracts evaluation and industry analysis.

For enterprises entering the AI journey and wondering where to start, the most important overarching question should be how to put AI to good use for the longer-term benefit of the organization. Further questions then rapidly become more difficult. This could be likened to the point when children who have been learning addition, subtraction, multiplication, and division for some time gain sufficient mastery of the basics of mathematics to start applying it to problem solving. Remember problems such as this: "A train leaves the station at <time> and travels at an average speed of <x> kph for <y> hours…" The new skill here is the application of the calculating skills to the problem at hand and first figuring out what the sum is, then performing the calculations, then expressing the answer in a useful way.

In order to translate the information they have been given, they need to understand what a train is and how it travels (i.e., across terrain, as opposed to how a rocket travels), and they need to understand what kph means. They need to figure out which mathematical functions are necessary – multiply OR divide? They need to understand what answer is sought, how to calculate the answer, and what unit to express the answer in – in this case, speed x hours = distance in km. It gets easier the more problems they tackle as similarities emerge and they discover that what was useful to solve one problem can be applied to another similar scenario.

The application of AI has similar challenges. The mathematics of machine learning's algorithms are complex, and there is greater depth and breadth of process and industry-specific detail that is pertinent to determining what the question is. The easy questions are the short-term ones, while the harder questions are the longer-term ones. And it will be difficult to find the people who have the breadth and depth of education and experience to understand all the mathematics along with the industry and process specifics. It is more realistic that the communication and collaboration skills of teams of disparate specialists will be the differentiator. Think of any heist movie with a hotchpotch of gloriously diverse characters, each with unique specialisms, who manage to work together effectively as a team.

Most people are limited in imagining the art of the possible. The institutions that educate us and the large organizations that we work in are increasingly seeing the need for creativity and imaginative thinking, but therein lies a paradox. Large organizations and institutions demand acceptance of the status quo to keep daily operations moving – the vast majority of people have to be followers as opposed to leaders for the functioning of large institutions, organizations, and society at large. This is why design thinking workshops matter, pulling people out of their day to day, into an environment more conducive to creativity to imagine a better and different way of doing things given the technology available today.

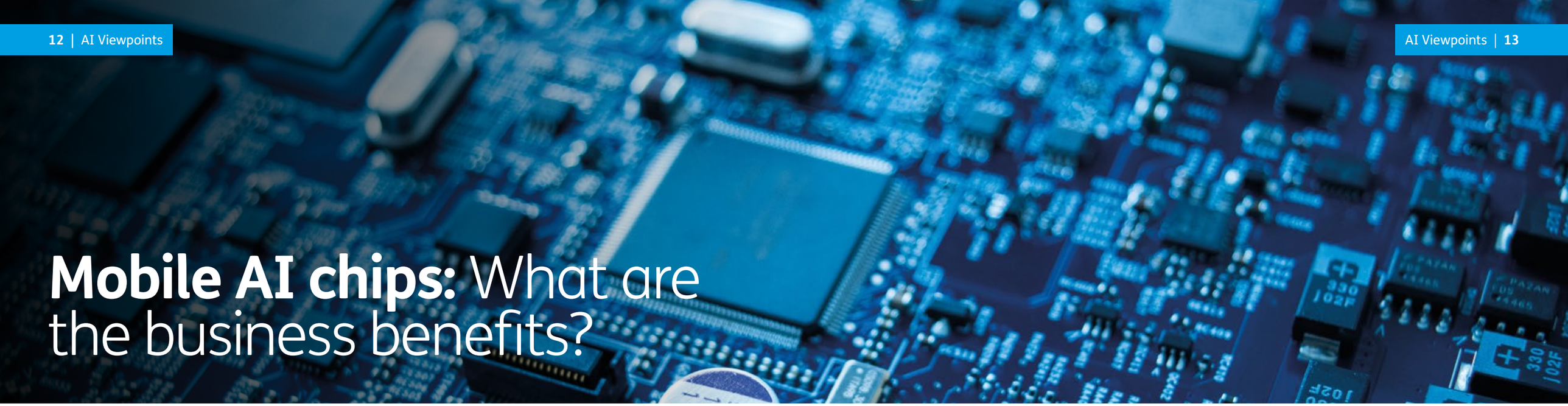
Henry Ford once said that if he had asked people what they wanted to improve their journeys, most would have said "faster horses," thereby missing the point that a series of new technological developments had enabled a totally different solution, initially termed "horseless carriage" and eventually "car."

The easy and obvious route to both automation and AI centers on cost reduction and achieving a version of events that already happens today but better, faster, cheaper. Right now, the low-hanging fruit is in areas where technologies such as robotic process automation can perform large volumes of routine, repeatable, rules-based tasks. It is possible to add AI components that are sufficiently advanced, like natural language processing and image recognition. Focusing on short-term cost reduction is a short-term way forward; it is easy to imitate, and once all competitors are there, no one has competitive advantage. It will be those who dig deeper into their long-term future possibilities, truly exploring beyond what is familiar into the great unknown, who will find a new way of doing things with today's cloud-based, digitized, agile, and flexible technologies and processes.

So we may think we are tired of hearing people talk about searching for the "Blockbusters to Netflix" moment or Uber-ization. But the supply chains of today are ripe for disintermediation and disaggregation following digitalization. New alternative supply chains building gradually over the coming years will render current supply chains irrelevant no matter how much better, faster, or cheaper they become.

Large enterprises and organizations need to imagine how they would and could operate if limitations or friction points were removed, with customer, employee, and supplier journeys at the heart of their thinking, and how data can inform decisions in the processes.

Then once the questions are formulated, AI and other technologies will need to be woven together to form the workings of tomorrow's business model.



Mobile AI chips: What are the business benefits?

Adam Holtby
Senior Analyst, Workspace Services

adam.holtby@ovum.com

Adam Holtby is a senior analyst at Ovum, specializing in enterprise mobility, workplace productivity, and IT service management.

It is hard to have a conversation about enterprise technology without the subject of artificial intelligence being a focal point. Recent Ovum research shows that 71% of enterprises are either considering or planning to use AI, bots, and machine learning over the next 12 months (Ovum ICT Enterprise Insights, 2017/18), and 65% of organizations believe artificial intelligence will have an impact on their workplace over the next three to five years (Ovum Workspace Services Survey, 2017). As mobility is such a vital digital workplace element, it is important to understand how AI capabilities will help further empower the mobile workforce going forward.

AI chips in mobile devices will help power new ways of working

AI can be very processor-intensive, so to overcome this issue, some device manufacturers are introducing dedicated AI processors into their mobile devices, with Huawei being one example. The company introduced a dedicated AI chip called a Kirin 970 into some of its recently released smartphones. Having a dedicated chip to handle AI processes means that device power can be reserved for other tasks that a user may wish to undertake on the device, improving the user experience as a result. Having this dedicated hardware capability means AI processing can take place on the device itself, with no reliance on the cloud or app, as is often the case with AI applications.

Dedicated AI chips open up new possibilities as to how mobile devices can be used within organizations. A good example is how these chips can help speed up image recognition and processing. When powered by AI and machine learning capabilities, smartphone cameras can add another layer of business security and authentication through functionality such as face ID and behavioral analysis. Cameras can also be used to profile an environment via visual scanning, providing people with a greater level of real-time information not only on the environments they are working in, but on things and objects within that environment. Samsung Bixby, for example, enables people to learn about items that can be scanned via a smartphone camera. These scanned items could potentially be anything, and they need not be connected to the internet. This camera use case also highlights shows how AI and machine learning, in combination with other capabilities, can help businesses improve user experiences and the processes and practices that support them.

An important benefit often touted by hardware manufacturers relating to on-device AI chips is that they are not dependent on any network capability – all tasks and processing can be handled on the device itself. However, it is important to note that many enterprise application providers, especially those that have platforms built in the cloud, are developing their AI capabilities to be network dependent, as this is the way such applications are commonly used anyway. Since many enterprise apps are now hosted and utilized via the cloud, this lack of dependency on the network may not in fact be a feature that many employees will see great benefit from day to day, as many employees work online and with apps that offer AI capabilities that have been developed to always be online. Developers do benefit, however, as they have flexibility in whether they create apps leveraging on-device AI capabilities, those in the cloud, or a combination of both.

Choosing hardware acceleration for your **machine learning applications**

Michael Azoff

Principal Analyst, Infrastructure Solutions
michael.azoff@ovum.com

Michael Azoff is a member of Ovum's IT infrastructure solutions group, leading a range of software development and lifecycle management research initiatives.

Artificial intelligence and its highly active branch of machine learning are being applied across every major industry, bringing automated intelligent decision-making and processing. To run these AI systems requires hardware acceleration for both training the ML systems and in inference mode. The choice of which accelerator to use needs to weigh many factors: performance, power constraints, edge or server based, power availability, frequency of change, type of model being implemented, and especially whether the model is being trained or used in production. The choice will widen as accelerators in incubation reach the market.

In working with AI/ML applications, the starting point is the central processing unit (CPU). The addition of hardware accelerators is necessary to speed up training and running in production, known as inference mode. The porting of deep learning algorithms onto Nvidia's early-generation, general-purpose, programming-enabled graphics processing units (GPUs) in 2010–11 was crucial to the emergence and rise of these AI systems, accelerating AI research progress. Since then, the AI hardware acceleration space has grown considerably, with Nvidia transforming itself into a specialist for deep learning computing, new players entering this market, and VCs funding novel architectures.

Nvidia has established itself as the leader in deep learning training. Each year it introduces a superior GPU architecture. For example, its latest-generation GPU, the Volta, has a 4x performance improvement over the previous-generation Pascal GPU. AMD entered this market in 2017 with its high-end GPU for AI, the Radeon Vega.

Another contender in this space is the field-programmable gate array (FPGA), an ASIC with programmable cores (essentially lookup tables) that allows a developer to configure (program) it to implement any desired algorithm. FPGAs are typically used where algorithms are likely to change, but having a hardware implementation for its high performance is desired and burning an ASIC, which then cannot be changed, is not suitable.

FPGAs have a long history in multiple markets. The most well known is microprocessor design prototyping and emulation, with a 1,000x speedup over CPUs. Applications where machine learning will benefit from running on FPGAs include the following:

- **Data center compute acceleration:** This is deployed by Amazon EC2 F1, Baidu, Tencent, and Microsoft Azure for flexible compute acceleration (e.g., AI inference, database search, and genomics sequencing with 40x speedup over CPUs).
- **Signal processing in 5G wireless infrastructure and radar:** Adaptive mobile networks are one of many examples where the integration of machine learning inference at the edge has huge benefits for FPGA companies.
- **Programmable memory hierarchy:** The ability to flexibly segment memory bandwidth at the hardware level maximizes its efficiency and creates dramatically (10x) higher performance systems. Because this is dictated by the algorithm, this feature will shine most during periods of highest algorithmic change.

Finally, several novel processor architectures are already in production or about to reach the market; novel designs such as Graphcore and Intel Nervana are eagerly awaited.

Final conclusions and takeaways for how to choose an AI hardware accelerator:

- There is no single best AI hardware accelerator. The type of AI algorithm used and the constraints of the task will determine the most suitable processor.
- Training deep learning neural network systems is best performed on GPUs. The case for inference mode is more open, with FPGAs and other dataflow architectures proving to be key contenders.
- The most superior AI algorithm of the day is likely to change over time as research continues. ASIC implementations designed for today's algorithms will therefore be short-lived as the preferred algorithm for AI acceleration continues to change.
- Venture capital is highly active in the AI hardware accelerator space.
- The whole AI industry, from vendors to users, will benefit from standards being introduced. Benchmarking AI accelerators is an important step, and the MLPerf benchmark initiative is most welcome.

Intelligent automation is transforming IT Operations

Hansa Iyengar
Senior Analyst, Advanced Digital Services
Hansa.Iyengar@ovum.com

Hansa Iyengar is a senior analyst in Ovum's Large Enterprise Services practice and specializes in the application-modernization and digital-transformation spaces.

The cornerstone of a digital enterprise is its ability to deliver better user experiences using real-time data and insights into customer behavior. This requires the ability to build, test, and release products in quick intervals, and collect and analyze data at incredibly fast rates. DevOps has become the go-to solution for enabling speed and managing the conflicts between development and operations teams. However, DevOps itself is in an evolutionary phase that will only pick up pace over the next few years. One of the key factors influencing this evolution is the rapid growth of intelligent automation that combines automation technologies with artificial intelligence to enable humans and machines to work in tandem to optimize productivity and efficiency across the operations function.

Development (Dev) has been witnessing increasing levels of automation over the years, and Ops (operations) is evolving to a stage where self-service infrastructure, continuous integration, and deployment solutions completely automate routine operational tasks such as deployment, maintenance, and bug fixes, freeing engineers to attend to new challenges in the operational environment. The major advantages of automation of IT operations are:

- increased business agility as it enables IT to keep pace with the speed at which business models evolve as enterprises transform digitally
- better service availability of IT as intelligent systems identify and resolve outages automatically
  - In some cases, the automated system is also able to predict when an outage might occur and take remedial action to prevent it.
- improved accuracy and efficiency of service delivery as automated systems can reliably take over repeated tasks such as configuration and provisioning, which would also provide cost savings
- improved sourcing, governance, and compliance as automated systems offer a real-time view of systems, applications, and infrastructure, allowing for better visibility across the IT landscape
- support for innovation by freeing up resources from the drudgery of routine tasks and acting as a catalyst for speedy implementation of innovative business models that keep pace with changing technology and customer preferences.

Despite intelligent automation offering obvious benefits, enterprise opinions around the adoption of automation remain polarized. While most enterprises believe it would boost their existing functions, a sizeable number of businesses also believe it would replace their existing workforce, and these enterprises fear the backlash of such an eventuality. The dearth of skilled labor, difficulty in training and reskilling, cybersecurity concerns, and the challenges of integrating intelligent systems with existing applications and infrastructure are some of the other concerns that need to be addressed before intelligent automation of IT operations finds mainstream adoption.

Automation of the IT operations function combined with the rapid growth and increased adoption of mobile technologies, cloud computing, and analytics provides a fertile ground for business innovations to take shape and bear fruit. Investments by vendors and systems integrators such as IBM, AWS, TCS, and Infosys emphasize the importance that is being placed on automation and the role it is expected to play in revolutionizing IT operations in the coming years.

The way forward for consumer AI assistants

An increasingly crowded market where it is hard to stand out

Eden Zoller
Principal Analyst, Consumer Services

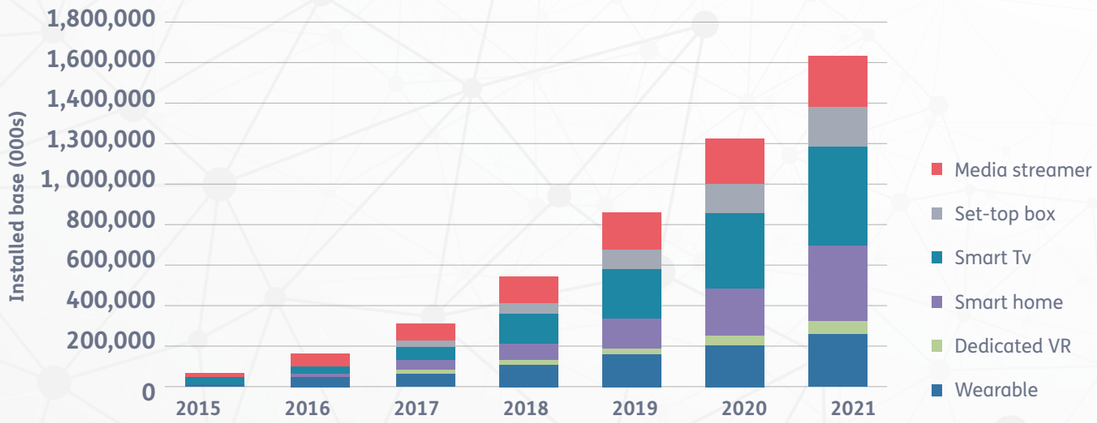
Eden.Zoller@ovum.com

Eden Zoller is responsible for leading Ovum's consumer digital commerce research, focusing on consumer dynamics, service provider strategies, market outlooks, and competitor analysis.

Consumer tech and OTT and e-commerce players are putting enormous effort and investment into AI-powered assistants. The front runners are Amazon Alexa, Apple's Siri, Google Assistant, Microsoft Cortana, Samsung Bixby, Alibaba's AliGenie, and Duer from Baidu. They are being joined by telcos launching their own proprietary AI assistants, notably Telefónica Aura, Orange Djingo, Deutsche Telekom Magenta, and, in South Korea, Nugu from SK Telecom and KT's GiGA Genie.

The heat around AI assistants is understandable, and if executed well, they are compelling propositions. AI assistants enable new forms of engagement and interactivity with consumers. This is partly a function of the deep data insights that AI assistants can generate but also because of their increasingly sophisticated voice interface along with advances in visual computing and cognitive abilities. AI assistants with these attributes coupled with telco-specific domain knowledge have the potential to pull customers more deeply into an operator's service ecosystem. AI assistants can give consumers the ability to manage and interact with telco services in a unified framework and via the convenience and fun of a conversational or visual interface. AI assistants can support semi- or fully automated customer care functions, enhance upselling and cross-selling opportunities, and drive automated personalization at scale. The bottom line is that if executed well, AI assistants can pull customers more deeply into a service provider's ecosystem.

Figure 1: Global installed base of voice AI–capable devices by subsegment, 2015–21*

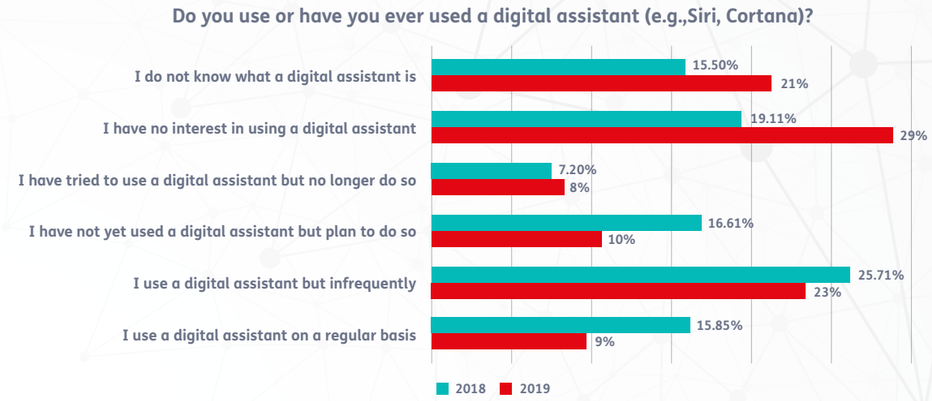


Source: Ovum *excludes smartphones and tablets

The installed device base for AI assistants is expanding – and moving cross-platform

The installed base of active devices is a better representation of the addressable market for AI assistants than unit sales. Smartphones clearly lead the voice AI-capable device market today, with 3.3 billion active devices of this kind by the end of 2017 that will increase to 4.65 billion by 2021. But AI assistants are rapidly moving into other connected platforms and devices beyond smartphones, as shown in Figure 1, which is opening new opportunities for service providers of all kinds. Ovum expects an exponential uptake of voice AI capabilities among these new devices, with a combined installed base of 1.63 billion active devices in 2021, a tenfold increase over 2016.

Figure 2: Consumer usage of digital assistants



Source: Ovum Digital Consumer Insights: Comms, Commerce & Media 2018
N=5,122

Converged experiences will be a point of differentiation

Service providers that can successfully deploy their AI assistant across multiple device segments and connected platforms stand to gain a significant competitive edge. Cross-platform AI assistants will be able to realize highly converged services and experiences. The ultimate scenario, which we have not come close to reaching today, is when AI assistants are able to process and understand multiple, connected environments and accompany and support human needs across them. AI assistants in this context will provide access to personal data across platforms, offering information, advice, and other applications across different domains with an integrated, seamless user experience.

Figure 3: Reasons for rating an AI assistant poor or average

What are your reasons for your rating of your digital assistant?

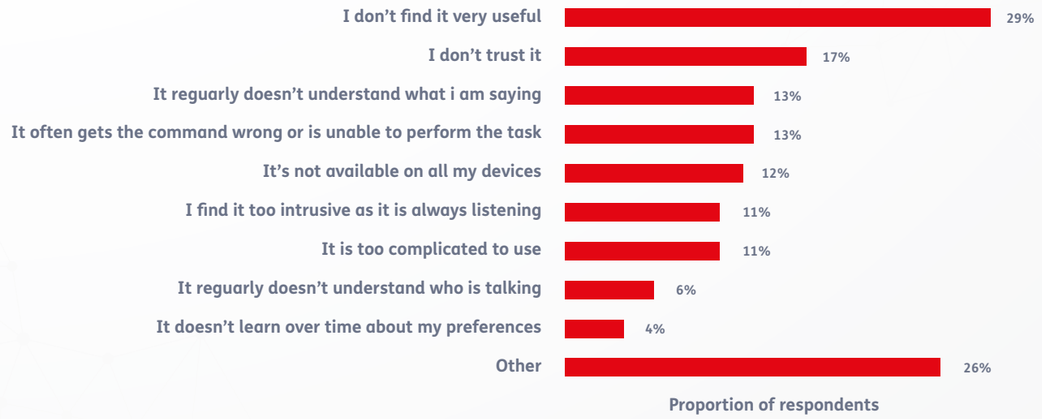
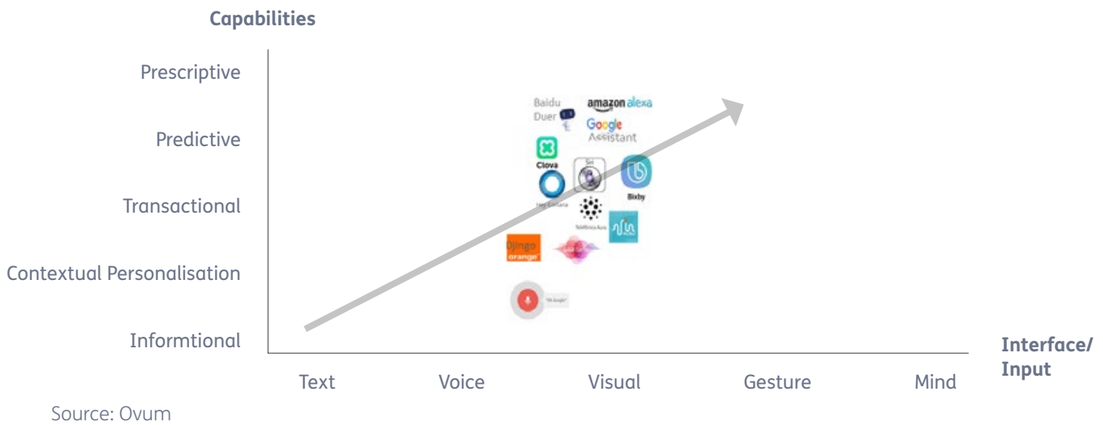


Figure 4: AI assistant evolution



Source: Ovum

The user experience needs to improve

Although AI assistants are becoming more popular, they are far from perfect, and the elements that most frustrate people largely relate to a poor user experience. This came out strongly in Ovum's 2018 Digital Consumer Insights survey when we asked respondents why they gave their AI assistants an average or poor rating, as shown in Figure 3. Consumers are frustrated and disappointed because their AI assistants frequently do not understand questions or who is talking to them, often get commands wrong, and are intrusive because they are always in listening mode. Lack of trust is also an issue, which is partly linked to AI assistants getting things wrong, but also to broader issues related to lack of trust and agency (i.e., people prefer to do things of this kind for themselves). Most of the user experience issues cited by consumers in our survey relate to the shortcomings with the AI assistant voice interface. This is obviously in need of improvement and an area where telcos and indeed all AI providers need to increase their efforts.

What to expect next: key trends and developments

A richer interface, deeper predictive and prescriptive functions

The way that AI assistants interact with people is going to become much more sophisticated over the next five years and beyond. Figure 4 illustrates key developments along two parameters: AI assistant capabilities and the AI assistant interface. We expect to see big advances in speech recognition, natural language processing, and emotion detection. The visual interface for AI assistants will become richer, driven by advances in visual computing and support for gesture control, augmented reality applications, and holograms. A longer-term scenario will see advances in the brain-machine interface reach the point where the human mind can trigger certain actions. There will also be a shift from the predominantly informational remit and basic recommendations that dominate AI assistants today to scenarios where advanced predictive analytics will drive prescriptive next best actions tailored to a user's specific needs and context.

Voice interactions will become increasingly sophisticated

One of the biggest frustrations consumers have with AI assistants is a poor experience due to the voice interface. AI assistant providers are aware of the issue and are working to improve this by leveraging the advances that machine learning is bringing to speech recognition, natural language processing, and emotion detection. The recently announced practical improvements and technology concepts for the Google Assistant voice interface are a useful barometer of how things are shaping up today and will develop going forward. During the Google I/O 2018 developer conference, the company introduced a continued conversation feature that removes the typical stop-start nature of dialogue with Google Assistant, bringing a new level of convenience and flow. A new setting in the Google Home keeps the conversation open for up to 8 seconds after a question has been asked or a command given. This means that if users have a follow-up action, they have a window to introduce it without the need to repeat wake words.

The conversation feature is not a game changer, but it is the kind of practical, simple feature that brings immediate benefits for users. A more far-reaching and potentially game-changing development is Google's experimental Duplex technology. Duplex is designed to enable Google Assistant to conduct natural, human-sounding conversations and carry out specific tasks over the phone. The demo highlighted two scenarios where a Duplex-powered AI persona interacted with a human – the first to book a restaurant reservation and the second to organize a hair dresser appointment. The Duplex persona used varied intonation, latency, pauses, and dysfluency (e.g., "um," "uh") and was able to elaborate on and modify questions and responses during the recorded conversation. The result was an almost natural-sounding conversation – it certainly came close to crossing "uncanny valley" territory. This is a substantial achievement, as natural conversations are incredibly hard to model. There is no doubt that in the longer term, Duplex could revolutionize voice interactions and what they can do, but at this point it the technology is experimental and has limitations. The conversation in the demo was recorded, and the technology is directed toward completing specific tasks in narrow domains, after being deeply trained in such domains.

Advances in visual computing for richer interactions

The visual interface for AI assistants will assume greater importance and become more sophisticated as developments in AI-powered computer vision accelerate, with increasingly accuracy and nuance in objection detection and image recognition. This will enable AI assistants to support more accurate visual search and, related to this, help users better understand what they are looking at and take actions based on that information.

IBM, Microsoft, Google, Facebook, and others are all working to push the boundaries of computer vision. At Facebook's 2018 F8 developer conference, the company announced a milestone in its progress with image recognition – helped by Instagram hashtags. Facebook said that after training the computer vision system with 1 billion images and 1,500 hashtags, it achieved a top-scoring 85.4% image recognition accuracy on ImageNet, an influential AI benchmarking competition in the image recognition field. The score is impressive, but what is highly significant is the use of hashtags as a data label and how this moves closer to the goal of unsupervised training. Training deep learning models to recognize images and/or objects typically requires people to label images, which is laborious and hard to scale. Facebook has used public images that are in essence already labeled via hashtags. This is the unsupervised element, and in theory it opens the way for deep learning models to access large data sets.

Personalization and authentication

AI-powered computer vision and cognitive computing are also driving advances in facial recognition and analysis, which could enable AI assistants to support a range of compelling use cases. Facial recognition is already being used to authenticate identities and transactions on smartphones and could be applied to AI assistants embedded in home hub speakers with a visual display. The ability to detect a user's distinct features can help personalize services and provide safeguards by, for example, recognizing a child and only providing her with access to certain programs for services on a AI assistant. The ultimate scenario could see an AI assistant use facial recognition to surface specific, tailored content.

Gesture control

Advances in computer vision are also driving developments in gesture detection and control. These technologies are already being used with VR headsets and could have applications for AI assistants. Gesture controls could act as a complement for voice commands in the connected home when it may be easier, faster, or simply more fun to execute a command via gesture rather than the spoken word – for example, using a gesture to scroll the menu on a AI assistant display or a wave of the hand to turn down music played by the AI assistant. Gesture does not necessarily have to be via a person's hand or fingers – developments in progress are working on eye control.

Augmented reality and holograms

More augmented reality applications and holograms will make their way onto AI assistants. Holograms in the AI assistant context are very new but are manifest in a few smart home hub speaker devices. At Mobile World Congress 2018, SK Telecom showcased HoloBox, a proof-of-concept hologram device based on its Nugu AI assistant platform. HoloBox contains a virtual 3D character called Wendy that users can interact with via voice and that can respond with gestures and facial expressions. Japanese technology firm Gatebox has launched a home hub-based AI assistant of the same name that contains a hologram character called Azuma Hikari, which, besides controlling other connected devices/applications in the home, is also designed as a personalized companion that will interact with users throughout the day via smartphone messaging, just as a real-life friend might.

The dawn of the mind-computer interface

A longer-term scenario will see AI solutions that enable the human mind to trigger certain actions, and it is conceivable that AI assistants could one day leverage a brain-computer interface. MIT Media Lab is working on a project called AlterEgo, which it describes as a closed-loop, noninvasive, wearable system that allows people to converse in high-bandwidth natural language with machines, AI-powered assistants, services, and other people without using their voice (opening their mouth or any other externally visible facial movement). The solution is based on a wearable device placed on the face. Electrodes in the device intercept neuromuscular signals sent by the brain to the face, and an advanced machine learning application and neural network matches specific neuromuscular signals with specific words. In the video demo from MIT, responses that no one else can hear are relayed verbally to the user.

Current-generation AI-powered brain-computer interfaces are cumbersome, slow, and prone to errors and inaccuracies, but this will inevitably improve going forward, as evidenced by the MIT AlterEgo project. It is not hard to envisage how this kind of AI technology could be used to enable Cortana to control HoloLens, the Microsoft mixed reality headset with which Cortana is already integrated.

Leveraging AI to improve network operations

Adaora Okeleke

Senior Analyst, Telecoms Operations and IT

adaora.okeleke@ovum.com

Adaora Okeleke is a senior analyst in Ovum's Telecoms Operations and IT team. She specializes in big data and analytics, customer experience management, customer relationship management, and vendor services.

AI presents opportunities for CSPs to transform network operations. While most of the activities we see are still in trial phases, a few CSPs have started implementing the technology in their network operations. They are recording impressive results with respect to operational efficiencies; for example, NTT DoCoMo is using an AI-assisted system to detect silent failures that previously were handled by 8,000 employees. To retain momentum and achieve the objective of autonomous operations, several factors must be fulfilled, and these include providing access to high-quality data sets and standardization of data formats and AI models. A few CSPs have started addressing these challenges; however, further progress will require stronger collaboration between industry players.

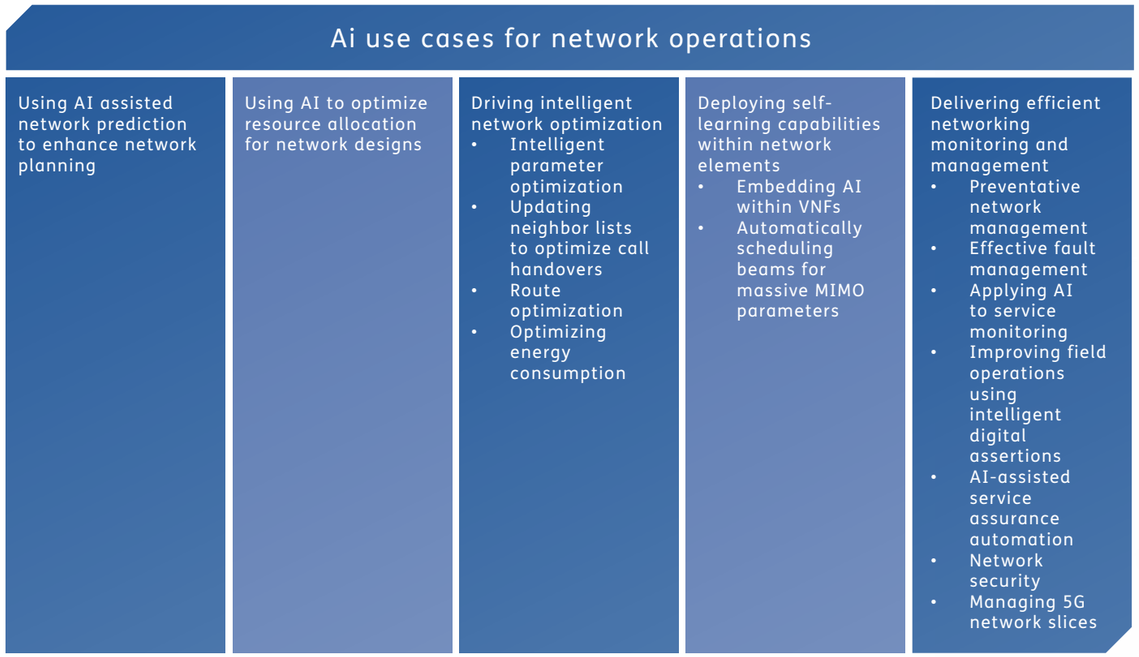
AI has applications in CSP network operations

CSP networks are undergoing dramatic changes, resulting in more complexity. It is therefore getting increasingly difficult to manage networks using human effort alone. Consequently, CSPs are looking toward automation to cope with this complexity. The rules-based approach that drives automation cannot scale well enough to support all possible use cases – preexisting as well as unknown scenarios.

AI presents an alternative approach to the existing rules-based automation method. AI systems, especially those based on machine learning techniques, use a data-driven approach to automatically build a model based on patterns associated with the workflows that are to be automated. Therefore, the benefit that AI systems bring to the automation challenge is that minimal human effort is required to generate the rules that drive automation. AI systems also enable the setup of more appropriate rules over time based on learning from data associated with the operations.

An AI-based approach is well suited to several CSP use cases, given how dynamically the CSP network is evolving, the vast amounts of data, the speed at which CSP data assets are being generated, and the increasing need to perform operations in real time. The use cases are summarized in Figure 1.

Figure 1: AI use cases for network operations



Source: Ovum

The decision on which use case to deploy must be driven by the company's objective. Vendors with managed services contracts are best positioned to support their CSP customers to identify suitable AI use cases.

CSPs are deploying AI to support their customer experience strategies

CSPs are actively trialing and implementing AI technologies in their networks and operations. Several top-tier CSPs, such as NTT DoCoMo, SK Telecom, T-Mobile US, and Vodafone, have deployed AI to enhance network infrastructure, OSS platforms, and associated operations, the ultimate objective of which is to improve customer experience by delivering high quality of service (QoS).

High QoS requires a completely different approach to monitoring network performance, one that is service-centric and not network-centric. Service-centric approaches require the correlation of service performance with the health of the network infrastructure. Tying service performance to network performance is not trivial as more data sets (due to increasing number of services and traffic) need to be monitored to guarantee service quality.

Silent failures, for example, degrade customer experience and are difficult to detect as they do not trigger any network alarms. For NTT DoCoMo, it took about 8,000 people to detect and resolve these issues. However, the CSP utilized AI techniques to classify the traffic data in different cells into normal state traffic and abnormal state (fault) traffic. Silent failures are then detected as deviation of current traffic from the normal state. Consequently, the amount of time and labor required to detect silent failures reduced, improving workforce productivity and costs.

For CSPs to achieve fully automated network operations, some housekeeping rules must be satisfied

The future goal is to utilize AI to achieve autonomous control of operations. However, before CSPs can consider autonomous use cases, several conditions must be satisfied to facilitate implementation of AI in CSPs' environments. These include access to high-quality data sets, standardizing data formats and AI models for use cases, and identifying ways to deal with CSPs' complex legacy systems and ways to deploy AI technologies efficiently in terms of time and cost. Given that most use cases are still in their trial phases, existing models are not fully mature, impacting the confidence levels CSPs have in deploying these use cases.

To make progress toward autonomous operations, the AI systems that CSPs adopt (developed in-house or purchased from a vendor) must function in an explainable way (providing insights on how decisions are made and factors driving these decisions). Having explainable AI systems allows CSP engineers to validate that the decisions taken by an AI system align with organizational policies.

Some CSPs have started addressing some of these requirements that would facilitate autonomous operations. SK Telecom and Telefónica, for example, have prioritized their big data strategies to ensure that high-quality data is available to drive AI use cases. SK Telecom converged its 150 OSS platforms into a single unified OSS platform, addressing the challenge associated with unifying data formats. The industry would, however, need to take a more unified approach to addressing the standardization of data formats and models for the AI use cases required to drive autonomous operations. There is therefore the need for more collaboration and interworking between CSPs, telecom vendors, and standard organizations such as ETSI and ITU.

WINTER IS COMING? THE AI **HYPE** IS RUNNING ON AN EVER THINNER **BASIS OF RESEARCH**

Alexander Harrowell
Senior Analyst, SoHo & SME Services
alexander.harrowell@ovum.com

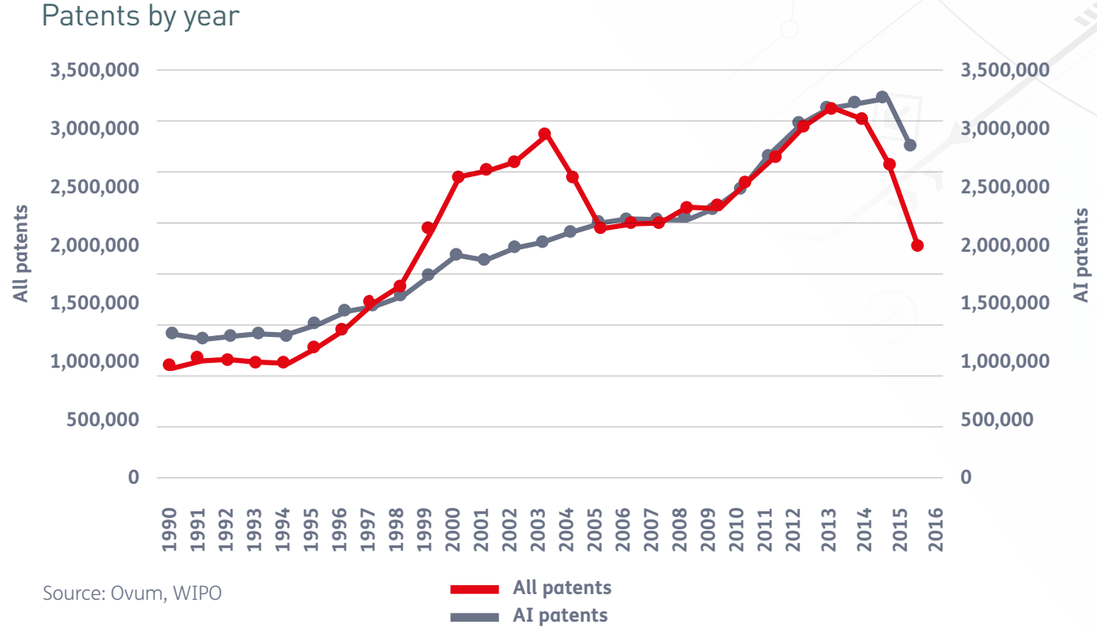
Alexander Harrowell is a senior analyst in Ovum's Enterprise Services team. He focuses on all forms of IT and telecoms services for SMEs, and especially on the intersection of the cloud, 5G, and the Internet of Things.

AI patent filings have peaked – some time ago

You cannot get away from the buzz about "AI" or at least machine learning. The surge of investment into the sector has led data center providers to pour extra-thick concrete floors so their facilities can take the weight of extremely power-dense, GPU-optimized server racks. There is open speculation about which bit of the 1950s AI research program might finally be realized first.

Yet, in some important ways, there are reasons to doubt how much real innovation is happening. For example, what about patents? Even with the growing importance of open source software and hardware, the creation of intellectual property remains a classic indicator of innovation. New patents registered in the World Intellectual Property Organization's PATSTAT database under classifications G06F (Machine Learning), G06N (AI), and G06T (Neural Networks & Image Processing) surged after 2010, peaked in 2013 with 327,366 new filings, and have declined steeply since then.

Figure 1: Annual patent filings, total and classifications G06F/N/T

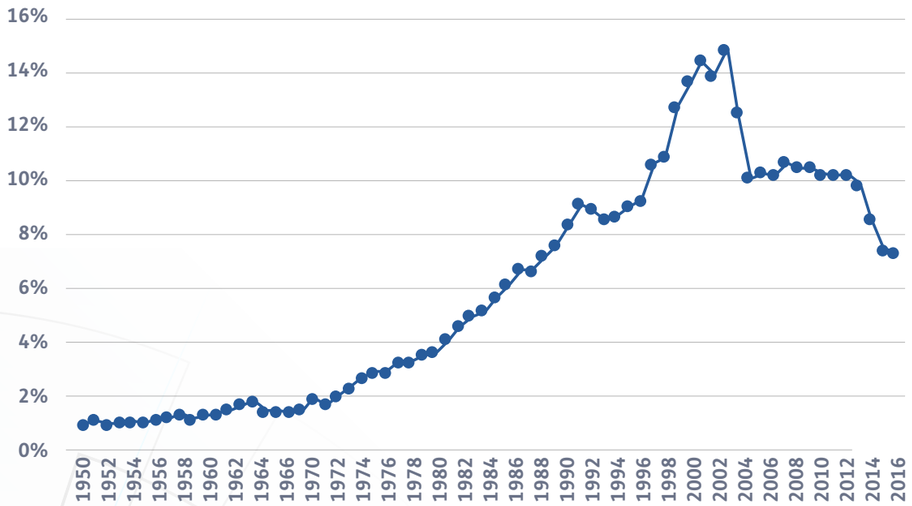


Source: Ovum, WIPO

Patents take a while to work through the system, so one possible explanation is just that recent innovations are still in the pipeline. Total filings certainly show this effect, but the drop-off begins in 2016, implying about a two-year pipeline delay. This cannot explain a peak in 2013. Also, AI-related filings were flat through the 2000s and 2010s and dropped sharply in 2013 as a percentage of the total, and this cannot be explained by the time taken to publish. Subsets of the data show a similar pattern for both China and the United States.

Figure 2: AI-related patents peaked as a percentage of total patents in 2003-4

AI patents, % of total filings

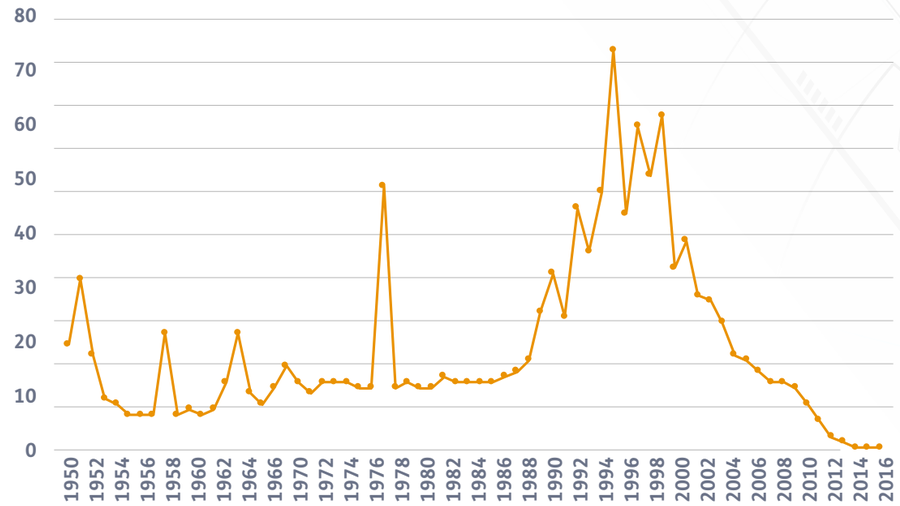


Source: Ovum, WIPO

Something similar seems to have happened with the rate at which new patents are cited by others, a metric of their influence – it rose steeply in the 1990s and has since declined. This surge seems to be accounted for by the massive Japanese R&D programs of the ‘80s and ‘90s.

Figure 3: AI-related patents were most influential in the 1990s

Average forward citations for AI-related patents 1950-2017



Source: Ovum, WIPO

Historically, AI research has been subject to a steep boom-and-bust cycle, the so-called “AI winter” phenomenon. Developments like the discovery of trivial ways to deceive all known image recognizers, the recent paper showing that most neural networks are actually solving polynomial regressions, or backpropagation inventor Geoffrey Hinton’s remark that we might need to “throw it all away and start again” may be early warnings that another round might be in prospect. Alternatively, the slowdown in patents may suggest that we have moved into a phase of commercialization and deployment rather than innovation as such. On one hand, this is positive – it is time to reap the economic harvest of past innovation – but on the other hand, AI enthusiasm is all about the idea of continuous, accelerating innovation, and the WIPO data shows nothing like that.

About Ovum

We provide authoritative data & forecasts, market research and analysis, bespoke consulting and end-to-end marketing services to help companies thrive in the connected digital economy.

Ovum helps service providers and their technology partners create business advantage by providing actionable insight to support their business planning, product development and go-to-market initiatives.

Visit **ovum.informa.com** to learn more.

Ovum Consulting

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Ovum's consulting team may be able to help you. For more information about Ovum's consulting capabilities, please contact us directly at **consulting@ovum.com**.

Contact Us

www.ovum.informa.com

marketingdepartment@ovum.com

International Offices

Beijing
Dubai
Hong Kong
Hyderabad
Johannesburg
London
Melbourne
New York
San Francisco
Sao Paulo
Singapore
Sydney
Tokyo

Ovum
TMT intelligence | informa

Tractica