Hype Cycle for Artificial Intelligence, 2020

Published 27 July 2020 - ID G00448060 - 117 min read

By Analysts Svetlana Sicular, Shubhangi Vashisth

Enterprises are making tangible progress with AI initiatives, but also many mistakes. As AI grows more widespread and new solutions emerge, organizations are realizing AI's increased value, but also facing new challenges. This report will help you assess AI-specific maturity and adoption.

Analysis

What You Need to Know

If artificial intelligence (AI) as a general concept were positioned on this Hype Cycle, it would be rolling off the Peak of Inflated Expectations. On the one hand, the COVID-19 pandemic has starkly exposed the brittleness of AI solutions in fraud detection, supply chain management and recommendations that stopped working properly when incoming data abruptly changed. On the other hand, AI has come to rescue. Chatbots, for example, have helped answer the flood of pandemic-related questions, computer vision has helped maintain social distancing, and machine learning models have proved indispensable for modeling the effects of reopening economies.

It is also good news that AI projects have accelerated in the healthcare, bioscience, manufacturing, financial services and supply chain sectors — they continue unabated, despite disruptive socioeconomic issues. Gartner polls conducted in May and June 2020 found that 47% of the respondents' AI investments were unchanged since the start of the pandemic and that 30% of the respondents actually planned to increase such investments. Only 16% had temporarily suspended AI investments, and just 7% had decreased them. (For details of these polls, see the Evidence section.)

Over the coming year, enterprises will be interested in using AI primarily to increase operational efficiencies and enable digital transformations.

The Hype Cycle

Despite AI advances within enterprises, this Hype Cycle remains “trigger-heavy” — new entries keep appearing on the upward-sloping Innovation Trigger (see Figure 1). This is an indicator of continued high levels of research and development (R&D) and equity investment in AI. The impact on enterprise buyers is a plethora of first-generation, highly priced approaches that lack maturity and that will evolve significantly before achieving mainstream adoption.

This Hype Cycle gives a high-level picture of AI innovations and disciplines for CIOs, AI leaders, and data and analytics leaders. They should also consult the following closely related Hype Cycles for specific techniques and nuances:

- “Hype Cycle for Data Science and Machine Learning, 2020,” which outlines the progress of ML tools, algorithms and data science techniques.
- “Hype Cycle for Natural Language Technologies, 2020,” which is dedicated to this rapidly expanding field.

Two megatrends dominate this year’s AI landscape:

- **The democratization of AI** means that AI is no longer the exclusive preserve of subject matter experts. Instead, it is increasingly within the reach of users in various roles, of different skill levels, and especially of diverse levels of creativity and insight.
- **The industrialization of AI platforms** enables reusability, scalability and safety, which accelerate AI adoption and growth.

Both megatrends will be key influences upon whether many of the AI technologies on this Hype Cycle go through the Trough of Disillusionment quickly and achieve mainstream adoption.

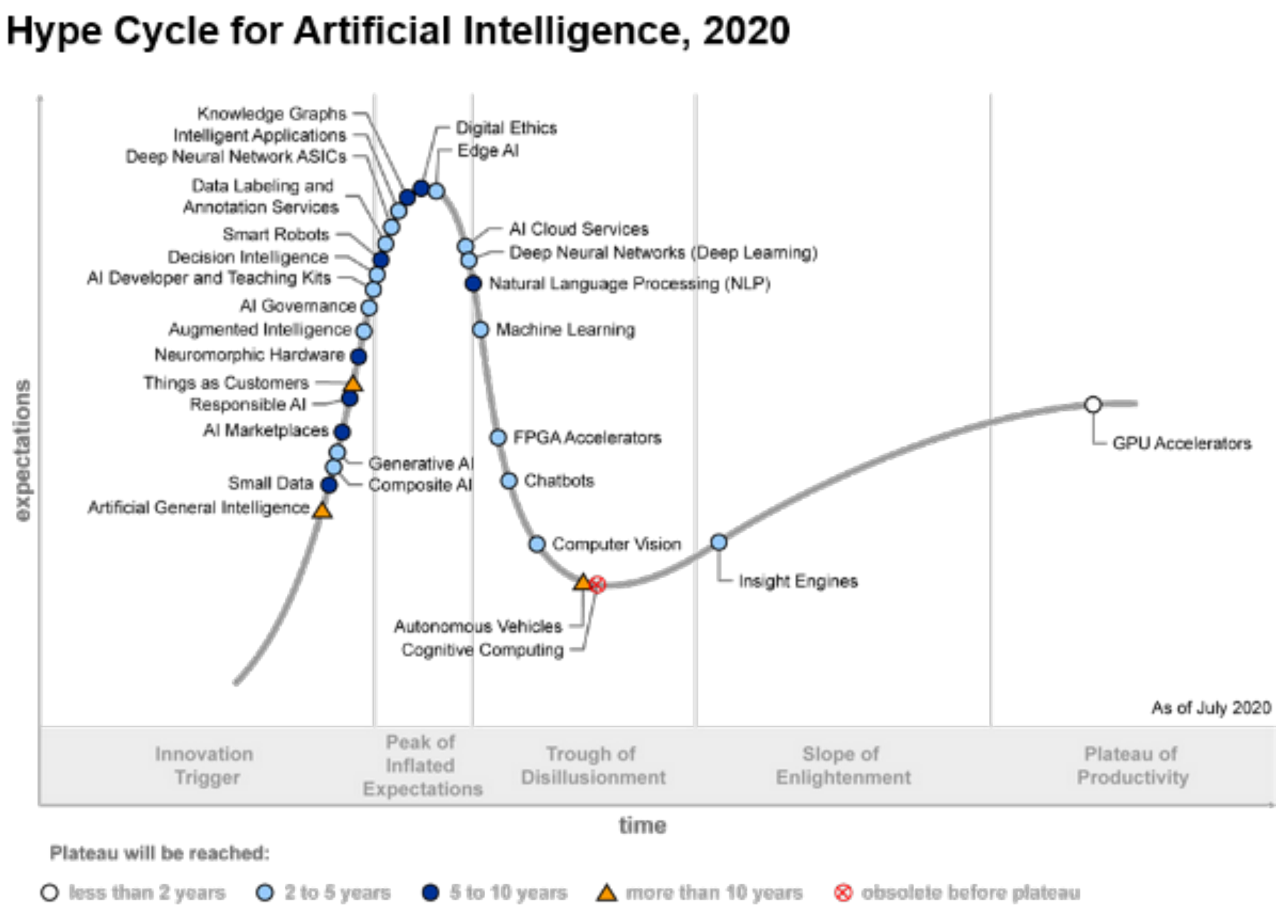
Over the past year, we have observed several developments that underpin these megatrends:

- **Data for AI is coming to the fore.** Gartner predicted three years ago that “by 2020, the focus within machine learning will shift from algorithms to high-value data.” Many enterprises are now making data strategies for AI their top priority. Small data, data labeling and annotation services, AI governance, knowledge graphs and insight engines are part of data strategies for AI.

- **Responsible AI is on the rise.** The broader AI adoption is, the more enterprises learn about their responsibility for the AI solutions and technologies they implement. For example, in June 2020, Amazon, IBM and Microsoft stopped selling AI services for facial recognition to the police. The COVID-19 pandemic has posed hard questions about trust, privacy, fairness and emergency decision making. To derive value from AI safely, we recommend adopting ideas from this Hype Cycle's profiles of responsible AI, digital ethics and AI governance.
- **Improving customer experience with AI heads corporate agendas for operational efficiency and digital transformation.** Innovations to help customers, citizens, patients, students and employees include the following: things as customers, chatbots, natural language processing (NLP), computer vision, augmented intelligence and intelligent applications (especially for CRM).
- **Compute infrastructure is being tailored to enable further AI advances.** This is a long-term trend. GPU accelerators, field-programmable gate array (FPGA) accelerators, deep neural network application-specific integrated circuits (ASICs) and neuromorphic hardware manifest different computing ideas, and more approaches are on the way.
- **Innovations continue to emerge.** They are capturing the imagination and promise new solutions to tough problems. The following innovations have gained visibility in the AI sector since 2019: generative AI, small data, composite AI, things as customers, responsible AI.

Cognitive computing is expected to be Obsolete Before Plateau because of the confusion about the differences between this concept and AI. Cognitive computing is narrower than AI, and the future of AI is wider than AI — it lies in decision intelligence.

Figure 1. Hype Cycle for Artificial Intelligence, 2020



Source: Gartner
ID: 448060

The Priority Matrix

The Priority Matrix maps the benefit rating for each technology — transformational or high for all but two — against the amount of time each requires to achieve mainstream adoption. This perspective can help when setting priorities. It's crucial to identify the use cases for the AI technologies most associated with your industry and its key business processes.

The timelines for AI technologies are accelerated in comparison with entries in other Hype Cycles due to high levels of interest and investment in AI. Most of the profiles in this Hype Cycle will reach mainstream adoption within two to five years. Gartner recommends setting priorities to ensure that business leaders have the right mindset and AI tools at their disposal within this time frame. This will prepare your organization to maximize benefits from many AI innovations when they reach the mainstream. Manage risk by implementing the emerging practices associated with responsible AI, AI governance and digital ethics. This will allow you to better evaluate innovations at the

Peak of Inflated Expectations, for example, determine your readiness for AI cloud services, and decide whether to consider AI at the edge as opposed to a centralized solution.

The COVID-19 pandemic has caused the elimination of pilot projects in favor of minimum viable products and accelerated AI development cycles. We hope this approach will become established as the best practice. We recommend employing it today to adopt and scale technologies, starting with some of those in the Trough of Disillusionment: NLP, machine learning, chatbots and computer vision. We also recommend investigating insight engines. Their progress onto the Slope of Enlightenment reflects renewed interest in, and new use cases for, search and in deriving insights from documents.

Figure 2. Priority Matrix for Artificial Intelligence, 2020

Priority Matrix for Artificial Intelligence, 2020

benefit	years to mainstream adoption			
	less than two years	two to five years	five to 10 years	more than 10 years
transformational		Augmented Intelligence Chatbots Composite AI Deep Neural Networks (Deep Learning) Edge AI Generative AI Intelligent Applications Machine Learning	AI Marketplaces Natural Language Processing (NLP) Neuromorphic Hardware	Artificial General Intelligence Autonomous Vehicles
high	GPU Accelerators	AI Cloud Services AI Developer and Teaching Kits AI Governance Computer Vision Decision Intelligence Deep Neural Network ASICs Insight Engines	Digital Ethics Knowledge Graphs Responsible AI Small Data Smart Robots	Things as Customers
moderate		Data Labeling and Annotation Services FPGA Accelerators		
low				

As of July 2020

Source: Gartner
ID: 448060

Off the Hype Cycle

- AI developer toolkits, which, however, reappear under a new name: AI developer and teaching kits.
- AI PaaS, which has been subsumed by AI cloud services.
- VPA-enabled wireless speakers, which have graduated from the Hype Cycle.

Profiles of the following now appear only in "Hype Cycle for Data Science and Machine Learning, 2020":

- AI-related C&SI services.
- AutoML.
- Explainable AI (also subsumed by "responsible AI," a more general category, in the present Hype Cycle).
- Graph analytics.
- Reinforcement learning.

Profiles of the following have been moved to the new "Hype Cycle for Natural Language Technologies, 2020":

- Conversational user interfaces.
- Speech recognition.
- Virtual assistants.

Profiles of the following now appear only in other Hype Cycles:

- Quantum computing, which appears in "Hype Cycle for Compute Infrastructure, 2020." AI accounts for just a fraction of the hype about quantum computing.
- Robotic process automation software, which in Gartner's opinion is not AI. It appears as "robotic process automation (RPA)" in numerous other Hype Cycles.

On the Rise

Artificial General Intelligence

Analysis By: Saniye Alaybeyi

Definition: Artificial general intelligence (AGI) is the hypothetical intelligence of a machine that has the capacity to understand or learn any intellectual task that a human being can. It is also referred to as strong AI.

Position and Adoption Speed Justification: Tangible progress on AI continued to be limited to narrow AI this year. On the philosophical front, no viable, agreed upon criteria to define AGI was in place. On the technical front, very small steps toward AGI were taken, such as IBM-MIT collaboration on neuro-symbolic concept learners, Microsoft and OpenAI partnership with $1B investment, and DeepMind, owned by Google, declared as its mission to build AGI. AGI safety is even less understood than AGI benefits, which makes AGI even further challenging and hypothetical. Therefore, this year, we keep AGI's position on the Hype Cycle the same. Today's AI technology cannot be proven yet to possess the equivalence of human intelligence (the lack of agreement about a test to prove such intelligence is itself a problem). It may, at some point, be possible to build a machine that approximates human cognitive capabilities, but we are likely many years away from completing the necessary research and engineering.

User Advice:

- End-user organizations should ignore AGI, however, until researchers and advocates demonstrate significant progress. Until then, ignore any suppliers' claims that their offerings have AGI or artificial human intelligence — these are generally illusions created by programmers.
- Focus on narrow AI, not on AGI. Special-purpose AI will have a huge and disruptive impact on business and personal life. Deliver business results enabled by applications that exploit special-purpose AI technologies, both leading-edge and older.
- Look for business results enabled by applications that exploit a full range of AI techniques, represented in this Hype Cycle.
- Experiment with less proven AI technologies that have precedents of success and give you competitive advantage.

Business Impact: AGI is unlikely to emerge in the next 10 years, although research will continue. When it does finally appear, it will probably be the result of a combination of many special-purpose AI technologies. Its benefits are likely to be enormous. But some of the economic, social and political implications will be disruptive — and probably not all positive.

There are currently no vendors of systems that exhibit AGI, but many companies are engaged in basic research. Examples are DeepMind (owned by Google), OpenAI, Vicarious, Numenta, Project AGI, OpenCog.

Benefit Rating: Transformational

Market Penetration: Less than 1% of target audience

Maturity: Embryonic

Recommended Reading: “Emerging Technologies and Trends Impact Radar: Artificial Intelligence”

“Maverick* Research: Being Human 2040 — The Life of the Architected Human in a More-Than-Human World”

Small Data

Analysis By: Jim Hare; Pieter den Hamer

Definition: The concept of “small data” indicates both the issue and approach on how to train AI models when small amounts of training data are enough or there is insufficient or sparse training data. There are a variety of strategies and data augmentation techniques to overcome the problem such as simulation, synthetic data, transfer learning, federated learning, self-supervised learning, few-shot learning and knowledge graphs.

Position and Adoption Speed Justification: Supervised deep learning that started the current AI hype is already fulfilling its promise, but it needs a lot of labeled data. Unlike consumer internet companies, which have data from billions of users to train AI models, collecting massive training sets in most enterprise is often not feasible. Also, most data science teams are not in a position to develop and train complex supervised models from scratch due to resource limitations. Moreover, reducing the need for big data and the ability to use small data, results in AI solutions that are more resilient and agile to handle changes. For example, the COVID-19 virus has resulted in many production AI models across different industry verticals to lose accuracy because they were trained using big data that reflected how the world worked before the pandemic hit. Retraining models using the same approach was not feasible, because more recent data of just a few weeks old are too limited to reflect the patterns of the new market circumstances. As a result, data scarcity has emerged as a major challenge, even more so with organizations becoming dependent on AI to run their businesses, also in times of disruption.

There is a growing number of data science innovations and open source projects focused on different data augmentation or other techniques. Among others, graph techniques have garnered new attention because of the ability to find patterns in small data, or to reduce dimensionality, complementing machine learning. Several new AI startups have created platforms and solutions that operate on small datasets.

User Advice: Data and analytics leaders whose teams are experiencing data scarcity issues in exploring new AI use cases, building hypotheses, or handling production models that have lost their accuracy should consider this approach first:

- Simpler models — Replacing more complex models with simpler, classical ML models such as linear regression, support vector machines, K-nearest neighbors, and naïve bayes that can be trained on small amounts of data. Proper feature engineering, the use of simpler models or ensembles thereof, should be in the toolbox of any data scientist, especially in the case of small data.

If replacing existing models with simpler models is not feasible, consider these emerging data augmentation and modeling approaches:

- Transfer learning — Enables AI solutions to learn from a related task where there is ample data available and then uses this knowledge to help overcome the small data problem. For example, an AI solution learns to find damaged parts from 1,000 pictures collected from a variety of products and data sources. It can then transfer this knowledge to detect damaged parts in a new product using just a few pictures.
- Federated learning — Enables collaborative ML by sharing local model improvements at a central level, where the central model combines locally trained or retrained models on small data in a decentralized environment. For example, when a hospital wants to develop a model for treating a condition, but has limited data, it trains the model on its own local data. It then passes this model to the next hospital that keeps training the model on its own data and so on, combining the model improvements. It also increases data privacy as no local data needs to be shared centrally.
- Synthetic data — Used to generate data to meet very specific needs or conditions that are not available in existing authentic data. Can be useful when either privacy needs limit the availability or usage of the data or when the data needed to train a model does not exist.
- Self-supervised learning — A relatively recent ML technique where the training data is autonomously (or automatically) labelled. The datasets are labelled by finding and exploiting the correlations between different input signals. Production models can continuously be learning in production making self-supervised learning well suited for changing environments.
- Few-shot learning — Classifies new data having seen only a few training examples.

This forces the AI to learn to spot the most important patterns since it only has a small dataset. Useful when training examples are hard to find or where the cost of labelling data is high.

- Other approaches include the sharing of scarce data between organizations, together building a larger set, and the use of reinforcement learning, where data is gathered through simulations or experimentation.

Business Impact: Small data techniques enable organizations to manage production models that are more resilient and able to adapt to major world events like the pandemic or future disruptions. These techniques are ideal for AI problems where there are no big datasets available. Using smaller amounts data allows data scientists to use more classical machine learning algorithms that provide good-enough accuracy but without the need for big data training sets. It can also speed up the business exploration and model prototyping for novel solutions, as this approach reduces the time, compute power, energy and costs to collect, prepare or label large datasets.

Benefit Rating: High

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Sample Vendors: Diveplane; Google (Cloud AI); Landing AI; MyDataModels; OWKIN

Recommended Reading: "3 Types of Machine Learning for the Enterprise"

"A Guidance Framework for Operationalizing Machine Learning"

"Boost Your Training Data for Better Machine Learning"

Composite AI

Analysis By: Pieter den Hamer; Erick Brethenoux

Definition: Composite AI refers to the combined application of different AI techniques to improve the efficiency of learning, to increase the level of "common sense" and ultimately to much more efficiently solve a wider range of business problems.

Position and Adoption Speed Justification: Composite AI is currently mostly about combining "connectionist" AI approaches like deep learning, with "symbolic" and other AI approaches like rule-based reasoning, graph analysis, agent-based modeling or optimization techniques. Composite AI aims to synergize these approaches, both from a pragmatic engineering perspective (improving the effectiveness of AI) and from a more profound scientific perspective (progressing our knowledge about artificial intelligence). The ideas behind composite AI are not new, but are only recently truly materializing. The goal is to enable AI solutions that require less data and energy to learn and which embody more "common sense," thus bringing AI closer to human learning and intelligence. In addition, composite AI recognizes that neither deep learning nor graph analytics or more "classical" AI techniques are silver bullets. Each approach has its strengths and weaknesses; none is able to resolve all possible AI challenges.

User Advice: AI leaders and practitioners should:

- Identify projects in which a fully data-driven, ML-only approach is unviable, inefficient or ill-fitted. For example, this is the case when not enough data is available, when training a deep learning network requires large amounts of data, time and energy, or when the required type of intelligence is very hard to represent in current artificial neural networks.

- Leverage domain knowledge and human expertise to provide context to and complement data-driven insights, by applying decision management with business rules, knowledge graphs or physical models in conjunction with machine learning models.

- Combine the power of deep learning in data science, image recognition or natural language processing with graph analytics to add higher-level, symbolic and relational intelligence (for example, spatiotemporal, conceptual or common sense reasoning).

- Extend the skills of data scientists and machine learning experts, or recruit/upskill additional AI experts, to also cover graph analytics, optimization or other required techniques for composite AI. In the case of rules and heuristics, skills for knowledge elicitation and knowledge engineering should also be available.

- Since composite AI is still emerging, be cautious of the fact that the benefits of composite AI can only be achieved through the creative artisanship of AI experts, while avoiding the disadvantages and weaknesses of each underlying AI technique.

Business Impact: Composite AI offers two main benefits in the short term. First, it brings the power of AI to a broader group of organizations that do not have access to large amounts of historical or labeled data but do possess significant human expertise. composite AI is one of the strategies to deal with “small data.” Second, it helps to expand the scope and quality of AI applications, in the sense that more types of reasoning challenges and required intelligence can be embedded in composite AI. Other benefits, depending on the techniques applied, include better interpretability and the support of augmented intelligence. There are many possible examples:

- A heuristic or rule approach can work together with a deep learning network in AI for predictive maintenance. Rules, coming from human engineering experts, or the application of physical/engineering model analysis may specify that certain sensor readings are likely to indicate inefficient asset operations, which then can be used as a feature to train a neural network to assess and predict the asset health. Typically, such a combination is much more effective than relying only on heuristics or only on a fully data-driven approach.

- In computer vision, (deep) neural networks are used to identify or categorize people or objects in an image. This output can then be used to enrich or generate a graph, which represents the image entities and their (spatiotemporal) relationships. This enables answering questions like “which object is in front of another,” “what is the speed of an object” and so on. Using a connectionist approach only, such seemingly simple questions are extremely hard to answer.

- In supply chain management, a composite AI solution can be composed of multiple agents, with each agent representing an actor in the ecosystem, typically having its own intelligence to monitor local conditions and machine learning to make predictions. Combining these agents into a “swarm” enables the creation of a common situation awareness, more global planning optimization and more dynamic, responsive scheduling.

In the longer term, composite AI has the potential to pave the way for more generic and intelligent AI solutions with profound impact on business models, although still a far cry from the elusive artificial general intelligence.

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Sample Vendors: ACTICO; Beyond Limits; BlackSwan Technologies; Cognite; Exponential AI; FICO; IBM; Indico; Petuum; ReactiveCore

Recommended Reading: “How to Use Machine Learning, Business Rules and Optimization in Decision Management”

“Combine Predictive and Prescriptive Analytics to Drive High-Impact Decisions”

“Leverage Augmented Intelligence to Win With AI”

Generative AI

Analysis By: Svetlana Sicular; Avivah Litan; Brian Burke

Definition: Generative AI is a variety of ML methods that learn a representation of artifacts from the data, and use it to generate brand-new, completely original, realistic artifacts that preserve a likeness to the training data, but do not repeat it. Generative AI can produce novel content (images, video, music, speech, text — even in combination), improve or alter existing content and create new data elements.

Position and Adoption Speed Justification: The hype around generative AI is heating up due to its sensational successes and huge societal concerns. According to Adweek, patent filings for generative AI have grown 500% in 2019. Christie’s auction house already sells AI-generated artwork. More practical applications, like differential privacy and synthetic data, are increasingly drawing enterprises’ attention.

AI methods that directly extract numeric or categorical insights from data are relatively widespread. Generative AI, which creates original artifacts or reconstructed content and data, is the next frontier. So far, it is less ubiquitous and with fewer use cases. The hype around Generative AI is growing due to a recent notable progress of Generative Adversarial Networks (GANs), invented in 2014, and language generating models, such as Bidirectional Encoder Representations from Transformers (BERT), introduced in 2018, and Generative Pre-trained Transformer 2 (GPT-2) introduced in 2019. Other quickly progressing generative AI methods include self-supervised learning, variational autoencoders and autoregressive models.

Regrettably, generative AI technologies underpin “deep fakes,” content that is dangerous in politics, business and society. Prominent organizations, such as Partnership on AI and DARPA, are pursuing detection of “deep fakes” to counteract fraud, disinformation, instigation of social unrest and other negative impacts of generative AI. In 2020, “deep fakes” are not yet pervasive among the fake content and news spread across the web, but Gartner expects this to rapidly change in the next five years.

User Advice: Data and analytics leaders should evaluate generative AI for the following purposes:

- Creative AI, a large subcategory of generative AI to produce art and work that typically requires imagination, for example, Adobe Sensei for visual arts and OpenAI Jukebox for music.

- Content creation, such as text, images, video and sound. Content creation already penetrates marketing, for example, producing personalized copywriting. Twenty-nine percent of marketing leaders rank generative content creation among the top three, according to the 2019 Gartner Marketing Technology Survey.

- Content improvements, such as rewriting the outdated text, background noise cancelation, increasing image resolution, and modifying photos by adjusting, removing or adding artifacts.

- Data creation, often known as synthetic data, to mitigate data scarcity or privacy barriers to insight. Generative techniques create new data instances, so the generated data repeats patterns of the actual data, but is completely made up. For example, text generation for chatbots, image generation for quality analysis in manufacturing, differential privacy. Visma generated for the Norwegian Labour and Welfare Administration the entire population of Norway preserving demographic nuances.

- Industry applications in retail, healthcare, life sciences, telecommunications, media, education and HCM. For example, in healthcare, generative AI could create medical images that depict the future development of a disease. In consumer goods, it can generate catalogs. In e-commerce, it can help customers to “try-on” various makeups and outfits.

Gartner recommends that software companies that produce generative AI include methods to preclude their software from being used to generate fake content before releasing the software, delivering the antidote immediately in version 1.0.

Organizations must prepare to mitigate the impact of deep fakes, which can cause serious disinformation and reputational risk. There are several methods evolving to do this including algorithmic detection and tracing content provenance.

Business Impact: More use cases will surface and proliferate. The field of generative AI will progress rapidly, both scientific discovery and technology commercialization. Reproducibility of AI results will be challenging in the near term. Other technologies, especially those that provide trust and transparency, could become an important complement to the generative AI solutions.

Full and accurate detection of generated content will remain challenging for years and may not be completely possible. To do so will require elevating critical thinking as a discipline in the organization. Technical, institutional and political interventions combined will be necessary to fight deep fakes. We will see unusual collaborations, even among competitors, to solve the problem of deep fakes and other ethical issues rooted in generative capabilities of AI.

Benefit Rating: Transformational

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Sample Vendors: Adobe (Sensei); Bitext; Dessa; Google (DeepMind); Landing AI; LeapYear; OpenAI; Phrasee; Spectrm; Textio

Recommended Reading: "Innovation Tech Insight for Deep Learning"

"How to Benefit From Creative AI — Assisted and Generative Content Creation"

"Cool Vendors in AI Core Technologies"

"Cool Vendors in Speech and Natural Language"

"Cool Vendors in Natural Language Technology"

AI Marketplaces

Analysis By: Alexander Linden; Eric Hunter

Definition: AI marketplaces bring together buyers and sellers to support the sale of AI algorithms while supporting key infrastructure and transactional capabilities for all parties involved. AI exchanges are very similar to marketplaces, but the focus is on sharing over monetization. Some AI exchanges are used within an organization to support internal sharing of prebuilt algorithms among data scientists.

Position and Adoption Speed Justification: AI marketplaces remain a nascent technology and are only very slowly moving for now. The role of marketplaces and exchanges is to solve the “long tail” of demand by giving data science teams and citizen data scientists access to these special purpose algorithms and highly domain-specific solutions. This is done to eliminate or reduce the demands to build them from scratch. So far, the amount of traction they receive is very limited. Most likely, this concept is still too alien and maybe even complicated in its current form for most enterprise users. The concept was framed three years ago as “algorithm marketplaces” and still has a diverse mix of offerings including APIs, microservices, platform economy and possibly even blockchain for distributed use and billing. In an extremely heuristic area like AI, the exploratory nature of work can indeed be nicely supported by marketplace and exchange mechanisms, where publishing costs of AI algorithms are literally zero. This low barrier enables even smaller snippets of analytical code and data to be distributed alongside features and functions that can be commercialized. As with any marketplace or exchange, AI marketplaces and exchanges are self-regulated, and given their nascency, AI marketplaces and exchanges will be transforming and might end up as part of larger solutions.

User Advice: Market traction for AI marketplaces and exchanges remains in the early phases. Users can get ahead of the curve and learn to deploy components from major AI marketplaces and exchanges with relatively low efforts. The risks for these marketplaces and exchanges are low, while the opportunity for long-term benefits is high. While adoption is in the early phases, enterprises are encouraged to establish governance and standards for marketplaces and exchanges in their organization. Many companies failed to do this with the emergence of public cloud and they were left creating retroactive policies and remediating unplanned infrastructure, application and data sprawl in the cloud along with the unmanaged spend associated with them.

Business Impact: In the long term, AI marketplaces and exchanges will:

- Make it easier for data scientists and citizen data scientists to find and choose from the huge variety of available algorithms, experiments, datasets and solution accelerators.
- Enable organizations to build advanced analytic solutions more quickly, to solve business problems faster and accelerate time to value.
- Help address the data scientist skills shortages. There even could be potential grounds for hiring new talent.

- Provide a platform for end-user organizations and vendors to commercialize their solutions and datasets.
- Offer the flexibility and extensibility to use the best algorithms for a business's needs.
- Attract partners and developers to contribute models, to build rich model repositories.
- Enable synergies among the ecosystem's participants.

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Sample Vendors: Algorithmia; Amazon SageMaker; Google Cloud Platform; KNIME; Microsoft; RapidMiner

Recommended Reading: "Algorithm Marketplaces Are Bringing the App Economy to Analytics"

"Maximize the Value of Your Data Science Efforts by Empowering Citizen Data Scientists"

Responsible AI

Analysis By: Svetlana Sicular

Definition: Responsible AI is an umbrella term for many aspects of making the right business and ethical choices when adopting AI that organizations often address independently. These include business and societal value, risk, trust, transparency, fairness, bias mitigation, explainability, accountability, safety, privacy and regulatory compliance. Responsible AI operationalizes an organizational responsibility and practices that ensure positive and accountable AI development and exploitation.

Position and Adoption Speed Justification: Responsible AI signifies the move from declarations and principles to operationalization of AI accountability at the individual, organizational and societal levels. While AI governance is practiced by designated groups, responsible AI applies to everyone who is involved in the AI process. Organizations are increasing their AI maturity, which requires defined methods and roles that operationalize AI principles. Lately, responsible AI has been elevated to the highest organization levels by Accenture, Google, Microsoft, OpenAI, PwC, Government of Canada, Government of India, the World Economic Forum (WEF) and more. Although responsible AI is nascent in industries, pioneers include AXA, Bank of America, State Farm, Telefónica and Telus.

COVID-19 pandemic stressed the need for responsible AI, when all governments and the entire world were following AI models of pandemic projections and economies' reopening. Many AI vendors and individual data scientists immediately shifted to solving pandemic problems, where they had to balance vital deliverables and risks associated with privacy, ethics, abrupt data changes and unconfirmed facts. Using AI for virus tracking, monitoring masks distribution and social distancing are subjects of public debate regarding appropriate AI interpretation, transparent data handling and clear exit plans for such temporary measures (see "How to Use AI to Fight COVID-19 and Beyond").

User Advice: Data and analytics leaders, take responsibility — it's not AI, it's you who are liable for the results and impacts, either intended or unintended. Extend existing mechanisms, like data and analytics governance and risk management to AI to:

- Establish and refine processes for handling AI-related business decisions.
- Designate, for each use case, a champion accountable for the responsible development of AI.
- Establish processes for AI review and validation. Have everyone in the process defend their decisions in front of their peers and validators.
- Provide guidelines to assess how much risk is appropriate.
- Ensure that humans are in the loop to mitigate AI deficiencies.

Build bridges to those organizational functions that are vital to AI success, but poorly educated about AI value and dangers to:

- Open a conversation with security, legal and customer experience functions.
- Build an AI oversight committee of independent, respected people.
- Continuously raise awareness of AI differences from the familiar concepts. Provide training and education on responsible AI, first to most critical personnel, and then to your entire AI audience.

- Have an escalation procedure early on in case something goes wrong.
- Anticipate human problems with AI: Identify enthusiasts who can help establish ongoing education about responsible AI.

The biggest problem in AI adoption currently is mistrust in AI solutions and low confidence in AI's positive impact. Responsible AI helps organizations go beyond purely technical AI progress to more successfully balance risk and value. With AI maturity, you will learn a lot and will make fewer mistakes — remain humble and keep learning.

Business Impact: Societal impacts of AI are frequently depicted in a distorted way, either too optimistically or as doom and gloom, while the responsible AI approach helps get a realistic view and instills trust. AI, like no other technology, encompasses organizational and societal dangers that have to be mitigated by responsible AI development and handling.

- The way AI is developed will encompass the mandatory awareness and actions regarding all aspects of responsible AI. Gartner predicts, "By 2023, all personnel hired for AI development and training work will have to demonstrate expertise in responsible development of AI."
- New roles, from an independent AI validator to chief responsible AI officer are necessary and are already being created to operationalize responsible AI at the organizational and societal levels.
- Responsible AI paves the way for new business models for creation of products, services or channels. It forms new ways of doing business that will result in significant shifts in market or industry dynamics via confirmed responsible AI actions and protocols; for example, a cross-organizational effort to fight "deep fakes."

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Recommended Reading: "Predicts 2020: AI and the Future of Work"

"AI Ethics: Use 5 Common Guidelines as Your Starting Point"

“Data Ethics and COVID-19: Making the Right Decisions for Data Collection, Use and Sharing”

“Top 10 Strategic Technology Trends for 2020: A Gartner Trend Insight Report”

Things as Customers

Analysis By: Don Scheibenreif; Mark Raskino

Definition: A thing (or machine) customer is a nonhuman economic actor that obtains goods or services in exchange for payment. Examples include virtual personal assistants, smart appliances, connected cars and IoT-enabled factory equipment. These thing customers act on behalf of a human customer or organization.

Position and Adoption Speed Justification: Today there are more internet-connected machines with the potential to act as customers than humans on the planet. We expect the number of machines and ambient artificial intelligence (AI), like virtual personal assistants, with this capability to rise steadily over time. They are increasingly gaining the capacity to buy, sell and request service. Things as customers start simply by alerting human counterparts that they need attention. However, things will advance beyond the role of simple informers to advisors and, ultimately, decision makers. According to Gartner research, both CEOs and CIOs agree on the potential of this emerging trend. Forty-nine percent of CIOs and 25% of CEOs we surveyed in 2019 believe demand from machine customers will become significant in their industry by 2030. These leaders believe at least 25% of all consumer purchases and business replenishment requests on average will be delegated to machines. Today, most things simply inform or make simple recommendations. We do see some examples of things as more complex customers emerging, such as smart grid technologies. HP Inc. embraced this future when it created Instant Ink — a service that already enables connected printers to automatically order their own ink when supplies run low. Some Tesla cars already order their own spare parts, and Walmart has patented grocery autoreordering based in home IoT sensing. In B2B, U.S.-based industrial supply company Fastenal uses smart vending machines that proactively place orders when stocks run low. Thinking forward, an autonomous vehicle could determine what parking garage to take its human passengers to based on criteria such as distance from destination, price, online review score, parking space dimensions, valet options, etc. In this case, it is the parking garage marketing to the car, not the humans.

This is a long-term proposition and there are major barriers, hence the early position on the Hype Cycle. The largest barrier is trust. Can the human customer trust the technology to accurately predict and execute? And, can the machine customer trust the organization that offers the service? Other barriers include: complex AI technologies, security and risk, regulatory compliance issues, and data sharing. All this will mean that things as customers across industries will not reach the Plateau of Productivity for five to 10 years.

User Advice: We recommend the following:

- Create a "tiger team" of architects, engineers, data scientists, economists, linguists, psychologists, and business decision makers to explore the business implications of machine customers. Determine whether the enterprise has the right capabilities, processes, and systems to identify, serve, communicate, and take orders from machines as customers.
- Follow examples from organizations like Tesla, Google, Amazon and Caterpillar to look for evidence of capabilities and business model impact.
- Build your organization's capabilities around artificial intelligence over the next five years. First in machine learning, then extending to other facets involved in machine customers processing information and making informed decisions.
- Identify use cases where your products and services can be extended to thing customers and pilot those ideas to understand the technologies, processes and skills required. Start with simple use cases driven by rules that can be easily configured and controlled by customers.
- Create scenarios to explore the market opportunities. Initiate collaboration with your chief digital officer, chief data officer, chief strategy officer, sales leaders, chief customer officers and others to explore the business potential of machines as your customers.
- Be mindful of the very real barriers. The complexity involved in developing a thing customer that can learn the depth and breadth of knowledge and preference trade-offs required to act on behalf of a human customer in a variety of situations is complex. Some humans may initially be uneasy about delegating purchasing functions to machines. Consider what ethical standards, legal issues and risk mitigation are needed to operate in a world of machines as customers.

Business Impact: Over time, trillions of dollars will be in the hands of nonhuman customers. This will result in new opportunities for revenue, efficiencies and managing customer relationships. Digital-savvy business leaders seeking new growth horizons will need to reimagine both their operating models and business models to take advantage of this ultimate emerging market, whose numbers will dwarf the number of human customers on (and one day perhaps off) the planet. How do you sell to a thing? What will get a thing to buy from you when its decisions are based on algorithms, not emotion? How will your human customer service agents handle requests from millions of things? What does "customer experience" even mean for a thing? Things as customers have the potential to generate new revenue opportunities, improve productivity, increase operational efficiency, improve health/well-being and enhance security of physical assets and people. They will also result in new sources of competition, fraud, legal and taxation challenges, and operational challenges (like how to provide customer service for things).

Benefit Rating: High

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Sample Vendors: Amazon; AutoCrib; Caterpillar; Google; Tesla

Recommended Reading: "Machine Customers: The Next Massive Emerging Market"

"How Customer Experience Changes When Your Customer Is a Thing"

"Why Machine Customers May Be Better Than Human Customers"

"IoT-Based Thing Commerce Requires a Differentiated Customer Experience"

"The Future of Customer Self-Service: The Digital Future Will Stall Without Customer-Led Automation"

Neuromorphic Hardware

Analysis By: Alan Priestley

Definition: Neuromorphic hardware comprises semiconductor devices inspired by neurobiological architectures. Neuromorphic processors feature non-von-Neumann architectures and implement spiking neural network execution models that are dramatically different from traditional processors. They are characterized by simple processing elements, but very high interconnectivity.

Position and Adoption Speed Justification: Neuromorphic systems continue to be at a prototype stage. IBM's TrueNorth and exploratory work on multilevel phase change memory technologies, the European Union's Human Brain Project (SpiNNaker and BrianScaleS), and BrainChip's Spiking Neuron Adaptive Processor technology are examples of neuromorphic hardware. Intel has developed a research chip, Loihi, and a range of servers and boards leveraging this chip to address a range of AI workloads: Loihi offers a higher degree of connectivity than competing implementations. Intel has also started training practitioners using its Loihi-based systems, as an early step to future adoption.

There are three major barriers to the deployment of neuromorphic hardware:

- Accessibility: Today GPUs are more accessible and easier to program than neuromorphic hardware; however, this could change when neuromorphic chips (NC) and the supporting ecosystems mature.
- Knowledge gaps: Programming neuromorphic hardware will require new programming models, tools and training methodologies.
- Scalability: The complexity of interconnection challenges the ability of semiconductor manufacturers to create viable neuromorphic devices.

At the moment, these projects are not on the mainstream path for deep neural networks (DNNs), but that could change with a surprise breakthrough in programming techniques, however, ongoing working in developing chips for neuromorphic computing continues, for this reason we have moved neuromorphic hardware closer to the peak.

User Advice: Neuromorphic computing architectures leverage spiking neural networks and have the potential to deliver extreme performance for use cases such as deep neural networks and signal analysis at very low power. Neuromorphic systems are also simpler to train than DNNs, with the potential of in-situ training. Furthermore, neuromorphic architectures can enable native support for graph analytics. Most of the neuromorphic architectures today are not ready for mainstream adoption. However, these architectures have the potential to become viable over the next five years. I&O leaders can prepare for neuromorphic computing architectures by:

- Creating a roadmap plan by identifying key applications that could benefit from neuromorphic computing.

- Partnering with key industry leaders in neuromorphic computing to develop proof of concept projects.
- Identifying new skill sets required to be nurtured for successful development of neuromorphic initiatives.

Business Impact: Rapid developments in DNN architectures may slow advances in neuromorphic hardware but NC holds the promise of enabling extremely lower power AI development. There are likely to be major leaps forward in hardware in the next decade, if not from neuromorphic hardware, then from other radically new hardware designs.

Neuromorphic systems promise of lower power operation makes them uniquely suitable for edge and endpoint devices, where their ability to support object and pattern recognition can support image and audio analytics.

We are in the midst of an extremely rapid evolution cycle, enabled by radically new hardware designs, suddenly practical DNN algorithms and huge amounts of big data used to train these systems. Neuromorphic devices have the potential to drive the reach of AI techniques further to the edge of the network, and potentially accelerate key tasks such as image and sound recognition inside the network. They will require significant advances in architecture and implementation to compete with other DNN-based architectures.

Benefit Rating: Transformational

Market Penetration: Less than 1% of target audience

Maturity: Embryonic

Sample Vendors: BrainChip; IBM; Intel

Recommended Reading: “Emerging Technology Analysis: Neuromorphic Computing”

“Forecast Database, AI Neural Network Processing Semiconductors, 1Q20”

“Forecast Analysis: AI Neural Network Processing Semiconductor Revenue, Worldwide”

“Forecast Analysis: Data Center Workload Accelerators, Worldwide”

“Product Managers Developing AI Chips Must Clearly Identify Target Markets”

"5 Questions a Product Manager Must Ask When Creating an AI-Enabled Edge Product Strategy"

Augmented Intelligence

Analysis By: Svetlana Sicular

Definition: Augmented intelligence is a human-centered partnership model of people and artificial intelligence (AI) working together to enhance cognitive performance, including learning, decision making and new experiences. Augmented intelligence is sometimes referred to as "centaur intelligence." It is different from augmented analytics: augmented intelligence is about people taking advantage of AI; augmented analytics is about data and analytics technologies enhanced with AI.

Position and Adoption Speed Justification: Increasingly, the approach to AI as a means to human augmentation outweighs the views on AI as a means to full automation. Augmented intelligence continues emerging as a design approach to get the most value from AI. It employs AI to compensate for human limitations and enables people to expand the possibilities for AI in the following key scenarios:

- Certain predetermined tasks in the process are done by AI, while the rest is done by people.
- People complete the job started with AI when AI reaches the limits of its capabilities or resources.
- Assistive AI develops and expands people's skills and talents.
- Neither AI nor people can accomplish the task without each other.

Many early-adopter enterprises had thought initially that full automation was the way to use AI. Now, they have started to realize that full automation is very expensive and complex. They are taking a more realistic view, where augmented intelligence compensates automation's limitations with people's creativity, flexibility and adaptability. Most AI vendors have also shifted their solutions and messaging from AI automation to a combination of humans and AI.

The COVID-19 pandemic further stressed the importance of the human role in AI, when an abrupt change in the data and context invalidated many models, and a human intervention was required to make those models work again. Companies that already had augmented intelligence in place fared much better in restoring AI-dependent functions.

User Advice: CIOs, data and analytics leaders and IT leaders responsible for AI should use augmented intelligence as a design approach. Implement AI to focus human attention where it is most needed, in order to accelerate organizational competencies that fulfill your vision for digital transformation. Center on what you can do for people, not what to automate — this is how you achieve human touch at scale.

Plan application and user experience design to facilitate augmented intelligence. This design could be more abstract (software, services, digital) or also in the physical space (physical robots and the like). Help people learn and improve, so the company, ecosystem and the entire society can take on more exceptional and forward-looking work. Approach augmented intelligence through three time horizons:

- In the short term, scale volume, reduce errors and automate routine tasks.
- In the medium term, scale quality, amplify human talents, further improve business efficiencies and create new products and experiences.
- In the long term, build personalized products and services at scale, reinvent your business, industry and society, and maximize customer convenience.

Add augmented intelligence to the workforce plan. Give people clarity about AI systems and ensure people's safety, for example, for AI moderation in social media. Transform from episodic to continuous, multidisciplinary learning to sustain innovation. Work with HR to upskill employees. Maximize the effects of AI-augmented roles and decisions via ongoing education, experience labs, AI-enabled just-in-time training and other methods.

Business Impact: Properly orchestrated, AI automation combined with human touch makes AI impactful. AI's mistakes are unavoidable, but people can fix them and vice versa. The use of AI in areas of "life and death" presents tremendous risk — augmented intelligence mitigates this by adding humans to AI. Conversely, when environments are dangerous for people, like working with delivering food to infected patients, robots can help. While humans see and analyze the world in hours, minutes and seconds, augmented intelligence can react much faster. With algorithms being commoditized, implementation philosophy and ethics will be the greatest differentiators of AI solutions: AI philosophy and ethics require people who know how to get them right.

Augmented intelligence already reduces mistakes and routine, positively reflecting on the customer service and transactions such as customer interactions, citizen services and patient care. The goal is to be more efficient with automation, while complementing AI with a human touch in order to keep things personal and with human common sense to manage the risks of AI automation.

Benefit Rating: Transformational

Market Penetration: 20% to 50% of target audience

Maturity: Emerging

Recommended Reading: "Leverage Augmented Intelligence to Win With AI"

"Technology Trends in Government, 2019-2020: Augmented Intelligence"

"Top 10 Strategic Technology Trends for 2020: Human Augmentation"

"Design Principles of Human-in-the-Loop Systems for Control, Performance and Transparency of AI"

"Data and Analytics Leaders: Rewire Your Culture for an AI-Augmented Future"

AI Governance

Analysis By: Svetlana Sicular

Definition: AI governance is the process of creating policies, assigning decision rights and ensuring organizational accountability for risks and investment decisions for the application and use of artificial intelligence techniques. AI governance is part of adaptive data and analytics governance. It addresses the perceptive, predictive and probabilistic nature of AI.

Position and Adoption Speed Justification: With AI having now reached the perimeter of practical enterprise application, data and analytics leaders are asking how to balance the business value promised by AI against the need for appropriate oversight, risk management and investment management. Enterprise practitioners are already making steps toward establishing AI governance. Leading organizations in the industries establish AI governance by addressing standards for AI development and operations, providing best practices, guidelines for model management and monitoring, data labeling and interpretation, AI value assurance and model risk management. The COVID-19 pandemic abruptly invalidated many patterns routinely detected by AI solutions, because pre-pandemic data stopped reflecting reality. This drew attention in organizations to advancing their AI governance to be able to restore and systemically assure stability and reliability of AI solutions.

User Advice: To develop AI governance, data and analytics leaders, CIOs and CDOs should apply the framework of trust, transparency and diversity and to data, algorithms and people to meet the new, AI-specific challenges and considerations. This framework should extend and advance existing governance mechanisms, such as risk management or data and analytics governance.

- Focus on trust in data sources and AI outcomes to ensure successful AI adoption. Develop specific testing and guidelines for "life-critical AI" that encompasses physical or moral safety.
- Identify transparency requirements for data sources and algorithms. Promote transparency and explainability of AI-enabled decision-making to minimize misinterpretations of AI results.
- Favor diversity, not just in terms of people's minds, backgrounds and cultures, but also in terms of data selection and algorithm choices. Demand new, different and even contradictory data to combine with what you already use to minimize the risk of AI biases.

Establish accountability for implementing each AI use case — all use cases differ in terms of their data, solution and outcomes requirements. Ensure ethics is considered for each use case. Develop methods for proactive regulatory compliance and outline reactive responsibilities, actions and procedures in the case of unanticipated and unintended consequences.

Plan adaptive governance to support freedom and creativity in data science teams, but also to protect the organization from reputational and regulatory risks. Little or no governance in data science teams to facilitate freedom and creativity is an acceptable approach if this is a conscious governance decision.

Business Impact: AI governance does not necessarily mean command and control; rather, it means the common ground across the entire organization when it comes to:

- Ethical and safety principles, together with mechanisms to ensure their development and adherence.
- Trust and transparency mechanisms to reach a common understanding of data and

algorithms that are used for AI via model governance and collaboration norms and capabilities.

- Diversity mechanisms to ensure the right data, algorithms and team members for each AI project.
- Nonprohibitive guidance on the standards for AI technologies, to avoid proliferation of tools in the absence of such standards.

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Recommended Reading: "AI Governance Spotlight: Early Lessons and Next Practices"

"Build AI-Specific Governance on Three Cornerstones: Trust, Transparency and Diversity"

"Governance and Best Practices for Chatbot Development"

"Healthcare Provider CIOs: Get Ahead of AI Innovation With Strong AI Governance"

"Artificial Intelligence Primer for 2020"

"Cool Vendors in Enterprise AI Governance and Ethical Response"

At the Peak

AI Developer and Teaching Kits

Analysis By: Eric Hunter

Definition: Artificial intelligence (AI) developer and teaching kits are applications and software development kits (SDKs) (some kits include hardware devices) that abstract data science platforms, frameworks, analytic libraries and devices to enable software engineers to incorporate AI into new or existing applications. Many of these kits also emphasize teaching new skills and integration best practices between software and devices for engineers.

Position and Adoption Speed Justification: AI developer and teaching kits have moved slightly higher along the Hype Cycle this year. Vendors have increased the number of offerings for developer-oriented AI developer and teaching kits and SDKs during the past year. The developer and teaching kits cover three maturing categories: kits for virtual assistants, AI design kits, and AI mobile serving SDKs. Representative offerings include:

- Kits for virtual assistants: Alibaba Group's AliGenie, Amazon Alexa Skills Kit, Apple SiriKit, Baidu DuerOS Open Platform and Google Dialogflow
- AI design kits: AWS DeepLens, DeepRacer and DeepComposer, Google AIY Vision and Voice Kits along with Google Coral, Intel Neural Compute Stick and Microsoft Vision AI DevKit
- AI mobile serving SDKs: Apple's Core ML, Clarifai Android and Apple SDK, Google ML Kit

Across all categories, vendor offerings require distinct deployment considerations and have varied feature coverage differences, but we expect greater consistency in the future.

AI developer and training kits support a limited set of native use cases, such as computer vision, image recognition, image labeling, natural language processing and text analytics. Developers can also deploy prebuilt models and optionally update those models from cloud services at model runtime.

AI design kits leverage custom hardware devices (such as cameras, musical instruments, speakers or vehicles) with developer-friendly APIs and SDKs to encourage platform developer adoption. These kits have also driven new vendor innovations targeting mainstream enterprise use cases — the Google Coral initiative is a good example of this.

AI mobile serving SDKs such as Core ML (Apple iOS) and ML Kit (Google Android) simplify mobile device model deployments for developers. A lack of standardization on a single format requires conversion utilities, which continue to be released for models from numerous formats, including Open Neural Network Exchange (ONNX) and MXNet. Commercial vendors have also introduced services (IBM's Watson Services for Core ML, Fritz AI) to extend AI serving SDK support.

User Advice: Vendor offerings are being released at a rapid pace in the market with a desire to attract new development communities. Application development leaders adopting these offerings to incorporate AI capabilities and features into applications should:

- Consistently evaluate internal project and platform roadmaps against the continuously evolving capabilities of AI developer kits along with traditional platforms (SAP, Oracle, SFDC) to avoid duplication of features or capability efforts.

- Leverage these kits to upskill developer knowledge and skills, which can translate to present and future enterprise needs that may directly or indirectly relate to kit-specific use cases.

- Carefully evaluate, and stress-test employed offerings, along with fully understanding the going concern support for each specific offering.

- Abstract adopted vendor offerings where possible to minimize portability constraints and lock-in.

- Avoid directing disproportional investments or effort in migrating established applications to a new platform for a small set of differentiating features.

- Ensure deployed capabilities are aligned to direct end-user benefits that cannot be easily achieved without AI.

- Adopt offerings in alignment with larger organizational cloud and mobile development standards and strategies.

Business Impact: The demand for AI is significant and is increasing at a rate beyond which experienced data scientists can meet alone. Adoption of AI developer and teaching kits will continue to increase. These offerings will equip software developers to become a key contingent for AI development and implementation. As these offerings continue to mature, Gartner expects offerings to:

- Expand support for edge and device-centric AI models through lightweight runtime frameworks

- Mature into distinct categories (AIPaaS) and influence other categories (Google's Coral) in future Hype Cycles

- Increase support for higher-level, focused AI use cases across specific business verticals and consumer demands

- Continue to reduce adoption barriers in the deployment of AI capabilities for software engineers and citizen data scientists

- Increase user gravity and stickiness to broader, vendor-based cloud and platform offerings, including platform as a service (PaaS)

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Sample Vendors: Alibaba Group; Amazon; Apple; Baidu; Clarifai; Google; IBM; Intel; Microsoft; Tencent

Recommended Reading: "How to Fast-Track Your Product Roadmap With Cloud Vendors' AI Development Accelerators"

"Democratization of Computer Vision Presents New Opportunities for Differentiating Personal Devices"

"A Framework for Applying AI in the Enterprise"

"How to Move Beyond AI Trials, to AI in Production"

"Technology Insight for Cloud AI Developer Services"

Decision Intelligence

Analysis By: Erick Brethenoux

Definition: Decision intelligence (DI) is a practical domain framing a wide range of decision-making techniques. DI provides a framework that brings multiple traditional and advanced disciplines together to design, model, align, execute, monitor and tune decision models and processes. Those disciplines include decision management (including advanced nondeterministic techniques such as agent-based systems), decision support, continuous intelligence and process management; and techniques such as descriptive, diagnostic, predictive and prescriptive analytics.

Position and Adoption Speed Justification: In a dynamic and increasingly complex business environment where business processes are siloed and disjointed, preventing the proper harmonization of collective decision outcomes, decisions are often ineffective. Five key reasons contribute to this:

- The pace of business is increasing.

- Unstructured, ad hoc decisions are becoming more frequent.
- Collaboration between man and machines is expanding.
- Tighter regulations are making risk management more prevalent.
- Consistency of decisions across the organization is questionable.

The current hype around automated decision making and augmented intelligence fueled by the integration of artificial intelligence (AI) techniques in decision making, is pushing DI toward the Peak of Inflated Expectations. Fueled by the COVID-19 crisis, the decision management space and by extension the DI space is bound to see an acceleration in its adoption. What the crisis revealed is the opacity and brittleness of decision models; rebuilding those models into adaptable and flexible models will require the discipline brought about by DI methods and techniques. A fast-emerging market around various software disciplines is starting to provide sensible answers for decision makers, but it will take between two to five years for DI to reach the Plateau of Productivity.

User Advice: Many of the disciplines encompassed within the DI domain are already leveraged by a large number of enterprises, unfortunately inconsistently across (and even sometimes within) processes, resulting in inconsistent decisions and sometimes contradictory outcomes.

Data and analytics leaders should:

- Improve the outcome of decision models and accommodate uncertainty factors by evaluating the contributing decision-modeling techniques.
- Promote the sustainability of cross-organizational decisions by building models using principles aimed at enhancing their traceability, replicability, pertinence and trustworthiness.
- Improve the predictability and alignment of cooperating decision agents, by mapping and simulating their collective behavior while also estimating their global contribution versus their local optimization.
- Develop staff expertise in the full range of traditional and emerging decision-augmentation and decision-automation techniques, including descriptive (dashboards and reports), diagnostic (interactive data exploration tools), predictive (machine learning) and prescriptive analytics (optimization, business rule processing and

simulation with rules).

- Tailor the choice of decision-making technique to the particular requirements of each decision situation by collaborating with business unit managers, subject matter experts and business process analysts.

Business Impact: The harmonization of decision-making disciplines through practically implementing Gartner decision intelligence model is applicable to a wide range of decisions within any industry or organization.

Decision intelligence helps organizations:

- Reduce the unpredictability of the outcomes of today's decision models that stems from the inability to properly capture and account for the uncertainty factors linked to their "behavior" in the business context.
- Improve the impact of business processes by materially enhancing the sustainability of organizations' decision models which is based on the power of their relevance, the quality of their transparency and the strength of their resilience.
- Increase the scrutiny of autonomous decision models (from embedded analytical assets to self-contained machine agents) at design time, so that their collective impact can be better understood, and disastrous outcomes can be avoided.
- Automate their fully structured decisions, provide information to improve the accuracy and effectiveness of semi-structured, augmented decisions and make decisions more transparent and auditable.

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Sample Vendors: ACTICO; Decision Management Solutions; Enova Decisions; Exponential AI; FICO; Noodle.ai; Petuum; Quantexa; r4 Technologies; ReactiveCore

Recommended Reading: "Decision Intelligence Is the Near Future of Decision Making: A Gartner Trend Insight Report"

“When to Automate or Augment Decision Making”

“Develop Good Decision Models to Succeed at Decision Management”

“How to Manage the Risks of Decision Automation”

Smart Robots

Analysis By: Annette Jump

Definition: Smart robots are electromechanical form factors that work autonomously in the physical world, learning in short-term intervals from human-supervised training and demonstrations or by their supervised experiences on the job. They sense environmental conditions, recognize and solve problems. Some can interact with humans using voice language, while some have a specialized function, like delivery or warehouse robots. Due to advanced sensory capabilities, smart robots may work alongside humans.

Position and Adoption Speed Justification: Smart robots have had significantly less adoption to date compared with their industrial counterparts (predefined, unchanged task) — but they received great hype in the marketplace, which is why smart robots are positioned climbing the Peak of Inflated Expectations. In the last 12 months, many of the established robot providers expanding their product line and new companies entering the market. Here are few examples:

- Whiz robot from SoftBank Robotics that will be sold under robot-as-a service (RaaS) model and originally be available only in Japan.
- Furhat Robot from a Swedish startup (Furhat Robotics) developing social robots.
- Smart Robotics that has introduced a robot valet for parking cars in France (Lyon).
- Temi robot from temi that will target home assistance for elderly and will incorporate Amazon’s Alexa.

The market is becoming more dynamic though the cost of entry and user tech sophistication are still high. Also, the time lag between product announcements and launch dates remain quite long at six to 12 months. Some products are killed before they reach broad availability. Recent market examples of slow adoption and withdrawals are Rethink Robotics, very low rate on renewal contracts for SoftBank Robotics’ Pepper three-year contracts and decision of Henn na Hotel, a Japanese hotel, the first hotel chain to replace smart robots with humans. Specialization also is very important to success, as no smart robot can address all industry specific use cases. Despite some advancements in AI, product and material experimentation in 2019 and early 2020, the progress is still slow, as companies are still trying to identify business valuable use cases. Therefore, the position of “smart robots” has not changed versus 2019 and still remains on the Innovation Trigger curve. Hype and expectations will continue to build around smart robots during the next few years, as providers execute on their plans to expand their offerings and explore new technologies, like reinforcement learning to drive continuous loop of learning for robots.

User Advice: Users in light manufacturing, distribution, retail, hospitality and healthcare facilities should consider smart robots as both substitutes and complements to their human workforce. Begin pilots designed to assess product capability, and quantify benefits. Examine current business- and material-handling processes into which smart robots can be deployed; also, consider redesigning processes to take advantage of the benefits of smart robots with three- to five-year roadmaps for large-scale deployment. Smart robots could also be a quality control (QC) check at the end of the process, rejecting product with faults and collecting data for analysis.

Business Impact: Smart robots will make their initial business impact across a wide spectrum of asset-centric, product-centric and service-centric industries. Their ability to do physical work, with greater reliability, lower costs, increased safety and higher productivity, is common across these industries. The ability for organizations to assist, replace or redeploy their human workers in more value-adding activities creates potentially high — and occasionally transformational — business benefits. Typical and potential use cases include:

- Logistics and warehousing: Product picking and packing, e-commerce order fulfillment, locating and moving goods
- Medical/healthcare: Patient care, medical materials handling, prescription filling
- Customer care
- Goods delivery due to social distancing and quarantine with COVID-19
- Manufacturing: Product assembly, stock replenishment, support of remote operations
- Delivery of packages and food
- Reception/concierge in hospitality, retail, hospitals, airports, etc.

- Other: Disposal of hazardous wastes

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Sample Vendors: Aethon; Amazon Robotics; Google; iRobot; Panasonic; Rethink Robotics; Savioke; SoftBank Robotics; Symbotic; temi

Recommended Reading: "Top 10 Strategic Technology Trends for 2020: Autonomous Things"

"Forecast: IoT Enterprise Robots by Use Case, Worldwide, 2018-2028"

"Preparing for a Future When Your Next Manufacturing Employee Will Be a Robot"

"Top 10 AI and Sensing Technology Capabilities for Personal Assistant Robots in 2020"

Data Labeling and Annotation Services

Analysis By: Anthony Mullen

Definition: Data labeling and annotation services support enterprises in labeling/annotating data for artificial intelligence (AI) projects. These services and associated platforms route and allocate this work to both internal staff and external third-party knowledge workers.

Position and Adoption Speed Justification: The proto provider in this space was Amazon Mechanical Turk (MTurk) launched in 2005 designed to coordinate labor to perform micro jobs that computers were unable to perform. As AI adoption has picked up among enterprises, the need for labeled data has dramatically increased in order to remove the bottleneck in developing AI solutions. As a result, offerings in this space have grown to help companies turn their unstructured data into structured data.

Baseline offerings in this space are access to pools of prequalified knowledge workers (either internal or external) who can label and annotate training data such as street scenes, speech, music, photos, documents and other assets. Many providers are now beginning to adopt a combination of machine learning techniques and human workers to accelerate the classification and annotation of training data. Semantic support in the form of taxonomies further speeds the labelling process. Some vendors have begun to differentiate with fair pricing and ethical approaches.

Increasingly these solutions are enlarging from preproduction focus to real-time human-in-the-loop solutions designed in real time to call upon a pool of workers (internal and external) to handle automation exceptions where model confidence is low, e.g., classifying and answering customer support questions. Further, annotations, classifications and content provided by third-party knowledge workers can be synchronized back to enterprise platforms such as content management systems, CRM, conversational platforms and knowledge management systems. Challenges remain around the skills that third-party knowledge workers have to annotate the data but are ameliorated somewhat by the development of reputation systems and prequalification tests.

While tech heavyweights like Facebook, Amazon, Google and Microsoft have used these providers for a while, many end users are quite unaware that such services exist.

User Advice: We suggest:

- Ensure the provider you choose has methods to test their pool of knowledge workers for domain expertise and measures around accuracy and quality.
- Model costs to avoid surprises by exploring and estimating the spend across the variety of business models which range from label volumes and project based to per annotator/seat costs.
- Allow data scientists to focus their time on more valuable tasks and lighten their load in classifying and annotating data by using these services.
- Use providers with real-time human-in-the-loop solutions for production systems like chatbots and recommenders to handle low confidence thresholds, spikes in demand or access to real-time knowledge not present in the enterprise.
- Design development and production workflows to leverage a mixture of knowledge workers — both internal and external staff.

Business Impact: While the supervised learning approach is predominant in AI, these services will continue to grow in usage. Scenarios that do not require deep domain knowledge of a field or datatype can expand annotation more quickly by using external knowledge workers. While there are many applications for this capability in preproduction environments, the real-time human in the loop solutions where models are continually trained and calibrated, such as chatbots or recommendation engines, will provide ongoing benefit. Business users need to join the human-in-the-loop workflows to route and train handover and moderation tasks to subject matter experts.

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Sample Vendors: Alegion; Amazon SageMaker Ground Truth; Appen; CloudFactory; Hive; Labelbox; Mapillary; Playment; Prolific

Recommended Reading: "Individuals, Groups and Society in the Loop of Artificial Intelligence Design and Development"

"Maverick* Research: Use Simulations to Give Machines Imagination"

"Clarify Strategy and Tactics for Artificial Intelligence by Separating Training and Machine Learning"

"Boost Your Training Data for Better Machine Learning"

Deep Neural Network ASICs

Analysis By: Alan Priestley

Definition: A deep neural network (DNN) application-specific integrated circuit (ASIC) is a purpose-specific processor designed to execute the DNN computations utilized in a wide range of artificial intelligence applications.

Position and Adoption Speed Justification: DNN ASICS are being used in a divers set of data center, edge and endpoint applications, some examples include object detection and classification in images and video streams, natural language processing, social media recommendation engines, autonomous vehicles and pharmaceutical analytics.

There are two major phases of DNN-based AI application development:

- Training: Large volumes of known data are processed by the DNN ASIC. These operations are data throughput intensive and typically require the use of floating point math.

- Inferencing: New or unknown data is analyzed by the DNN. These tasks are latency dependent and can utilize integer math.

A majority of training and inferencing tasks currently use GPUs, but the use of DNN ASICs can deliver significantly higher performance and lower power consumption than CPUs or GPUs when executing DNNs.

While it is possible for the same DNN ASIC to be used for both training and inference tasks, devices are being developed that are optimized for a specific task and often for a specific class of DNN. The training phase typically takes place in a data center and leverage large scale designs, typically high power chips, optimized for data throughput. Having developed and trained a DNN-based AI application it is typically deployed on a DNN ASIC optimized for inference operations. These chips may be used in data center deployments but often will be utilized in edge or endpoint systems where there may be constrained formfactors and power resources, requiring highly optimized chip designs.

Many companies have announced plans to DNN ASICs for both training and inference ranging from traditional semiconductor vendors to startups. The large hyperscale cloud service providers are also developing ASICs optimized for their specific DNN-based workloads, examples include Google's tensor processing units (TPUs) optimized for its TensorFlow-based applications.

User Advice: The benefits of DNN ASICs in processing the highly parallel operations required for today's AI-based applications are significant. However, widespread use of DNN ASICs will require the standardization of neural network architectures and support across diverse DNN software development frameworks. Plan an effective long-term DNN strategy comprising DNN ASICs by choosing ASICs and vendors that offer or support the broadest set of DNN frameworks to deliver business value faster. General purpose CPU vendors are also adding new instructions to their CPUs to support DNN-based workloads and these should also be evaluated when assessing the use of ASICs to accelerate these DNN-based applications.

Business Impact: Hardware acceleration will enable DNN-based workloads to address more opportunities in a business through improved cost and performance. Use cases that can benefit from DNNs include video analytics and object detection, image recognition, natural language processing and recommendation systems.

IT leaders deploying deep neural network applications should include DNN ASICs in their planning portfolio. We expect this market to mature quickly, possibly within the three-year depreciation horizon of new systems.

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Sample Vendors: Amazon; Baidu; Google; Graphcore; Hailo; Huawei; Intel; SambaNova Systems; Syntiant

Recommended Reading: "Forecast Analysis: Data Center Workload Accelerators, Worldwide"

"Forecast Database, AI Neural Network Processing Semiconductors, 1Q20"

"Forecast Analysis: AI Neural Network Processing Semiconductor Revenue, Worldwide"

Intelligent Applications

Analysis By: Alys Woodward; Helen Poitevin; Jim Hare

Definition: Intelligent applications are enterprise applications with embedded or integrated AI technologies such as intelligent automation, data-driven insights, and guided recommendations in order to deliver a more personalized interface, improve productivity and decision making.

Position and Adoption Speed Justification: AI is the near-term next major battleground for enterprise application providers with every tech provider now incorporating some type of AI capability into their product or service offering. Enterprise application vendors are embedding AI technologies within their offerings as well as introducing AI platform capabilities — from ERP to CRM to HCM to workforce productivity applications. This drives greater customer loyalty and dependence on the applications rather than driving specific additional revenue lines for vendors — the AI enhances the usefulness of the whole application.

Intelligent applications will use AI in the following ways:

- Data capture and response: AI technologies such as NLP, text analytics, deep neural networks (DNNs) and image recognition can be used for intelligent invoice matching, extracting terms and conditions or clauses from contracts, or analysis of images for photographic recognition.

- Process augmentation: AI technologies like machine learning, decision intelligence, knowledge graphs and explainable AI can provide more intelligent actions for an application. In the future, this can be extended further to identify patterns of work, from which process models can be built and executed. When processes or recommendations change due to AI, the business user responsible for the process and decision being taken needs to understand the reason for the changes, hence the use of explainable AI.

- User experience: Conversational UI platforms are used to develop language-based interfaces that use text or speech to interact with the user. Natural language processing used to create virtual assistants is one application of AI to the user experience. Further examples include facial recognition and other AI applications for understanding user emotions, context or intent, and predicting user needs.

- Analytics: AI technologies like augmented analytics can be used to create more predictive and prescriptive analytics that can then be presented to users for further evaluation, or plugged into a process to drive autonomous action.

Although intelligent applications will have a widespread transformational effect, the hype around them has not particularly advanced since 2019. It may be that the message has been obscured by the move to the composable enterprise, or that application leaders tend to consider intelligent components in the same way as other components. It is important to recognize the different nature of smart components that require updating from machine learning models, because the need to update models causes deployment challenges.

User Advice: Enterprise application leaders should:

- Explore how AI can improve your organization's processes and operations by adding more diverse ways to capture a variety of information, more intelligent automation, conversational user interfaces and better decision making.

- Challenge your packaged software providers to outline in their product roadmaps how they are incorporating AI to add business value in the form of a range of AI technologies.

- Be aware of "AI washing" as more and more startups, and even aging solutions, claim AI as part of their solution. Ask them how they use AI to deliver improvements on the original promise of the application. Some technologies under the AI banner are fairly mature and well-established.

- Prioritize investments in highly specialized and domain-specific intelligent applications delivered as individual point solutions which help solve problem areas such as customer engagement and service, talent acquisition, collaboration, engagement and more.
- Understand how the vendor is mitigating bias in the models used in its application. Verify that the vendor is safeguarding your data if used as part of a benchmarking service.
- Bring AI components into your composable enterprise thinking to innovate faster and safer, to reduce costs, and to lay the foundation for business-IT partnerships. Remain aware of what makes AI different, particularly how to refresh and rebuild machine learning models, as this can cause implementation and usage challenges.

Business Impact: Intelligent enterprise applications that leverage AI can offer the following benefits:

- Reduce or eliminate human-based manual tasks allowing workers to focus on more value-adding activities through the use of intelligent automation and insights — via bots, sensors and machine learning.
- Improve business efficiency via packaged AI technologies embedded in enterprise business processes.
- Make business operations more agile with business processes that adapt and reshape themselves as they run.
- Enhance the user interface with conversational capabilities that can ingest and interact through text or speech.

For example, in the area of human capital management (HCM), AI is increasingly being added to HCM applications to match talent supply and demand, predict recruitment success, or optimize recruitment marketing. Candidate-facing chatbots are becoming increasingly common in enabling further automation of this process, such as recommending which jobs to apply for and answering questions or conducting initial candidate screening.

In procurement, applications can ingest contract terms and conditions, and provide analytics on the level of risk introduced by contracts from a newly merged or acquired organization.

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Sample Vendors: Google Docs; Microsoft Office 365; Oracle Applications; Salesforce Einstein; SAP Leonardo; ServiceNow; Sievo; Workday

Recommended Reading: "Application Leaders: Master Composable Enterprise Thinking for Your Post-COVID-19 Reset"

"Technology Options for Talent Analytics"

"Drive Better Digital Workplace Employee Collaboration Using AI, Chatbots and Advanced Analytics"

"Predicts 2020: AI for CRM Sales Technology Must be Balanced With Analytics, Training and Change Management Considerations"

"Artificial Intelligence Maturity Model"

"How to Apply 'Intelligent' to Your ERP Strategy"

Knowledge Graphs

Analysis By: Stephen Emmott

Definition: Knowledge graphs are data structures within which disparate data about entities (people, companies, digital assets, etc.) is codified as a graph — a network of nodes (vertices) and links (edges/arcs). This enables information ("knowledge") to be located (knowledge graph as an index) or synthesized (knowledge graph as a data source).

Position and Adoption Speed Justification: Google's Knowledge Graph, Facebook Social Graph, LinkedIn Graph and Microsoft Graph are evidence of the growing popularity of knowledge graphs due to their ability to encode and interrelate disparate data, whether structured or unstructured, at source. All use knowledge graphs to enhance the relevance of search, and provide information “cards” that are synthesized directly from their graphs. This supports collaboration and sharing, search and discovery, and the extraction of insights through analysis.

Specialist vendors offer graph-based capabilities that support the creation and management of knowledge graphs that can serve a range of use cases within the enterprise. Products in a variety of markets — such as text analytics, insight engines, and metadata management solutions — also utilize or are based upon knowledge graphs.

Knowledge graphs are entity-centric applications of graph technology. They are modelled using an ontology which must be defined and instantiated by encoding data sourced either within or external to the enterprise. Embedded within cloud offices services, and increasingly other application categories, most enterprises will have — in some form or other — knowledge graphs in operation. However, few enterprises are building their own, or taking ownership of what they have. While awareness is growing, purposeful utilization and a strategy for attaining this is lagging, hence the ascent toward the Peak of Inflated Expectations coupled with a long time to the Plateau of Productivity.

User Advice: IT leaders should approach knowledge graphs as databases for storing data about entities and their relationships. This is especially so where the data has many different sources and forms, such as documents in a content services platform, updates in a data feed, audio from a video, or tables in a database. Where data is unstructured (for machines), use AI to extract and structure data.

Application leaders should collaborate with data and analytics leaders to identify the knowledge graphs already in operation within their applications portfolio. IT leaders responsible for data and analytics must include knowledge graphs within the scope of their data and analytics governance and management. To ward against perpetuating data silos, investigate and establish ways for multiple knowledge graphs to interoperate. This is likely to extend to third party data knowledge graphs.

Ensure data and analytics staff are familiar with knowledge graphs and the technology used to support them including graph query and analysis. Developing competence requires embracing ontologies, to define and describe how knowledge graphs should be structured and constrained. Venture forward through the use of a suitable pilot project that delivers tangible value for the business, but also learning and development for data and analytics staff: an opportunity to distinguish graph- from RDBMS-based approaches.

IT leaders should gain access to their knowledge graph(s) so as to be accessible for inspection and management. In particular, facilitate access for subject-matter experts within the business so they can moderate the data and contribute to its modelling. Beyond “received” knowledge graphs, you should identify use cases where there is a need for custom-made knowledge graphs and evaluate products that facilitate this.

Business Impact: Organizations can expect significant value from knowledge graphs. They have the potential to impact all business function and industry domains. However, as an entity-centric databases, they will primarily be the concern for IT, especially data and analytics.

Business units will be impacted through the applications that benefit from knowledge graphs — such as insight engines to help discover and retrieve insights. They will also work directly with knowledge graphs through the involvement of their subject matter experts responsible for authoritative data.

Key use cases for knowledge graphs have emerged in:

- Digital workplace (e.g., collaboration, sharing and insight);
- Automation (e.g., ingestion of data from content to RPA);
- Supporting machine learning (e.g., augmenting training data);
- Data analysis (e.g., augmented analytics especially in the context of business intelligence reporting and cybersecurity);
- Digital commerce (e.g., product information management and recommendations); and
- Data management (e.g., metadata management, data cataloging, and data fabric).

Key to the long-term success of knowledge graphs is enabling data within organizations to be interoperable with external knowledge graphs so as to enable the ingestion, validation and sharing of ontologies and data relating to entities e.g., geography, public institutions and events, etc.

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Sample Vendors: data.world; Diffbot; Franz (AllegroGraph); Mininglamp; Neo4j; Ontotext; Semantic Web Company (PoolParty); Sisense; Stardog; TopQuadrant

Recommended Reading: "Top 10 Trends in Data and Analytics, 2020"

"Augmented Data Catalogs: Now an Enterprise Must-Have for Data and Analytics Leaders"

"An Introduction to and Emerging Use Cases for Natural Language Processing"

"An Introduction to Graph Data Stores and Applicable Use Cases"

"Financial Data Strategy and Knowledge Graphs"

"COVID-19 Demands Urgent Use of Graph Data Management and Analytics"

"Magic Quadrant for Metadata Management Solutions"

"Magic Quadrant for Insight Engines"

Digital Ethics

Analysis By: Jim Hare; Frank Buytendijk; Lydia Clougherty Jones

Definition: Digital ethics comprise the systems of values and moral principles for the conduct of electronic interactions among people, organizations and things.

Position and Adoption Speed Justification: Digital ethics remains at the Peak of Inflated Expectations. Digital ethics and privacy remain growing concerns for individuals, organizations and governments. Consumers are increasingly aware that their personal information is valuable, and they're frustrated by lack of transparency and continuing misuses and breaches. Organizations increasingly recognize the risks involved in securing and managing personal data, and governments are implementing strict legislation in this area.

The coronavirus outbreak has demonstrated the important role of digital ethics in how governments and healthcare organizations are using technology and personal data to address the pandemic. However, no matter how urgent the response to the crisis is, decisions about how technology and data are used could result in more harm than good if those decisions are not grounded in digital ethics. The pandemic has shown that regardless of the hype around digital ethics, many organizations are still not applying them. And, as a result, the innovation hasn't yet passed the Peak of Inflated

Expectations.

Board members and other executives are sharing their concerns about the unintended consequences that the innovative use of technology can have. There is frequent, high-profile press coverage of stories that concern the impact of data and technology on business and society more broadly. More universities across the globe are adding digital ethics courses including the University of Oxford and the University Melbourne that recently launched programs and centers to address ethical, policy and legal challenges posed by new technologies. Government commissions and industry consortiums are actively developing guidelines for ethical use of AI. See "How Forthcoming EU Legal Framework Will Affect Your AI Initiatives."

User Advice: Business value and digital ethics need not be in conflict. Intention is key. If the only goal is business performance, and ethics is seen only as a way of achieving this goal, this may lead to window dressing. However, if the goal is to be an ethical company, and this leads to better business performance, then this serves all parties, and society more broadly. It will only strengthen the organization, helping you to have an even greater positive influence in the future.

Business and IT leaders responsible for digital transformation in their organizations should:

- Identify specific digital ethics issues, and opportunities to turn awareness into action throughout the various business domains.
- Discuss ethical dilemmas from different points of moral reasoning, such as outcome determinative versus empathy-focused. Ensure that the ethical consequences have been accounted for and that you are comfortable defending the use of that technology, including unintended negative outcomes.
- Elevate the conversation by focusing on digital ethics as a source of business value, rather than simply focusing on compliance and risk. Link digital ethics to concrete business performance metrics.

Business Impact: There are ethical consequences that arise through the use of digital technology in every business domain. Digital ethics should be treated as a tangible business practice discipline rather than an academic discussion. It does not have to be at odds with optimizing business performance. In fact, ethical behavior can have business value in itself.

Areas of business impact include influencing innovation ideas, product development, customer engagement, corporate strategy and go-to-market.

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: Avanade; Hypergiant; IBM; Microsoft; Salesforce; SAP; SAS

Recommended Reading: "Data Ethics and COVID-19: Making the Right Decisions for Data Collection, Use and Sharing"

"Digital Ethics: What Every Executive Leader Should Know"

"Digital Ethics by Design: A Framework for Better Digital Business"

"Top 10 Strategic Technology Trends for 2020"

"The CIO's Guide to Digital Ethics: Leading Your Enterprise in a Digital Society"

"Data Ethics Enables Business Value"

"Use Privacy to Build Trust and Personalize Customer Experiences"

Edge AI

Analysis By: Alan Priestley; Erick Brethenoux; Eric Goodness

Definition: Edge AI refers to the use of AI techniques embedded in IoT endpoints, gateways and edge servers, in applications ranging from autonomous vehicles to streaming analytics. While predominantly focused on AI Inference, more sophisticated system may include a local training capability to provide in-situ optimization of the AI models.

Position and Adoption Speed Justification: Edge AI will be implemented in a range of different ways, depending on the application and design constraints of the equipment being deployed — form factor, power budget (i.e., battery powered vs. mains powered), data volume, decision latency, location, security requirements etc.:

- Data captured at an IoT endpoint and transferred to an AI system hosted within an edge computer, gateway or aggregation point: In this architecture, the IoT endpoint is a peripheral to the AI system. The endpoint acts as a data gatherer that feeds this data to the AI system. An example of this is environmental sensors deployed for a smart agriculture application.

- AI embedded in the IoT endpoint: In this architecture, the IoT endpoint is capable of running AI models to interpret data captured by the endpoint and drives some of the endpoints' functions. In this case, the AI model (e.g., a machine learning model) is trained (and updated) on a central system and deployed to the IoT endpoint. An example is a medical wearable that leverages sensor data and AI to help visually impaired people navigate the world in their daily lives.

The applications that are starting to see increasing adoption of edge AI include those that are latency sensitive (e.g., autonomous navigation), data intensive (e.g., video analytics), and require an increasing amount of autonomy for local decision making. While many of these applications are still in R&D or trial phases, and widespread adoption is at least a few years away, other such as video analytics (leveraging deep learning methods and deployed as deep learning models at the endpoints or in edge servers) are starting to see adoption — driven by the rapid growth in deployment of surveillance cameras and the need for real-time interpretation of captured video streams.

User Advice: Enterprise architecture and technology innovation leaders should:

- Determine whether the new AI developments in machine learning (ML) are applicable to their IoT deployments, or whether traditional centralized data analytics and AI methodologies are adequate.
- Evaluate when to consider AI at the edge vs. a centralized solution. Applications that have high-communications costs are sensitive to latency, or ingest high volumes of data at the edge are good candidates for AI.
- Assess the different technologies available to support edge AI and the viability of the vendors offering them — many potential vendors are startups, which may have interesting products but limited support capabilities.
- Estimate carefully and pragmatically the appropriate level of autonomy and trustworthiness for AI systems deployed on edge systems.
- Assess the risk associated with the nondeterministic nature of many AI techniques

where it may not be possible to control or replicate the analysis results.

- Use edge gateways and servers as the aggregation and filtering point to perform most of the edge analytics functions. Make an exception for compute-intensive endpoints, where AI-based analytics can be performed on the devices themselves.

Business Impact: By incorporating AI techniques at the edge, enterprises may be positively impacted as follows:

- Improved operational efficiency, such as enhanced visual inspection systems in a manufacturing setting.
- Enhanced customer experience, with faster execution time, performed at the edge.
- Reduced latency in decision-making, with the use of streaming analytics and migration to an event-based architecture.
- Communication cost reduction, with less data traffic between the edge and the cloud.
- Increased availability even when the edge is disconnected from the network.

Enhanced local decision autonomy for edge systems.

Reduced storage demand through a more reactive exploitation of the data and a better estimate of its potential obsolescence.

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Sample Vendors: Amazon; Baidu; Google; Huawei; Intel; Matroid; Microsoft; Neurala; NVIDIA; Qualcomm

Recommended Reading: "Exploring the Edge: 12 Frontiers of Edge Computing"

"5 Questions a Product Manager Must Ask When Creating an AI-Enabled Edge Product Strategy"

"Use Edge AI to Drive Revenue Growth, Forecasting, Customer Engagement and

Workforce Planning"

"Cool Vendors in AI Semiconductors"

AI Cloud Services

Analysis By: Van Baker; Bern Elliot

Definition: Artificial intelligence cloud services provide AI model building tools, APIs and associated middleware that enable the building/training, deployment and consumption of machine learning models running on prebuilt infrastructure as cloud services. These services include automated machine learning, vision, and language services.

Position and Adoption Speed Justification: The use and sophistication of AI cloud services continues to increase, with the leading cloud service providers, including Alibaba, Amazon Web Services (AWS), Baidu, Google, IBM and Microsoft, competing to become the platform of choice. Over the past several years, AI applications utilizing cloud services have continued to gain traction and acceptance in the market both by data scientists and developers alike. The promise of using cloud services to more quickly and easily build and deploy AI solutions has pushed this technology to the Peak of Inflated Expectations. However, this will be followed by some level of disillusionment as organizations experience and understand the limitations of the offerings.

The AI cloud approach is continuing to disrupt the on-premises data science and machine learning platform market, especially as organizations experiment and build AI prototypes. The availability of specialized hardware instances with AI-optimized chips and large amounts of data storage makes the cloud an ideal environment for organizations to build and deploy AI applications without the risks, costs and delays of conventional on-premises procurement. Cloud service providers are also offering packaged APIs and tools that make it easier for developers to integrate AI capabilities into existing applications.

User Advice: IT leaders responsible for building and deploying AI-enabled applications should take these steps:

- Consider AI cloud services over on-premises options to reduce the overhead of developing and for easier deployment and elastic scalability.
- Improve the chances of success of your AI strategy by experimenting with different AI techniques and AI cloud services providers, using the exact same dataset, and then selecting one that best addresses your requirements. Consider using an A/B testing

approach.

- Increase your organization's AI project success by selecting AI cloud services that addresses your data science, developer and infrastructure requirements and skill limitations. Pretrained AI cloud services often require no (or limited) data science expertise.

- Use features like automated algorithm selection and training-set creation to offload some of the complexity of the project and leverage existing expertise on operating cloud services. This will assist technical professional teams with little to no data science expertise.

Business Impact: AI cloud services offerings will become ubiquitous in the three key AI services of automated machine learning (ML), natural language processing and computer vision.

- **Automated machine learning:** Packaged autoML services offered by the AI cloud service providers to unify the end-to-end ML workflow. Advanced solutions providing integrated access to all phases of the project — from data preparation to deployment in a managed training and execution environment accessible through APIs. All providers offer automated model building but many fail to deliver data preparation and augmentation capabilities.

- **Natural language processing:** Organizations can use pretrained NLP systems to create cloud-based language solutions for a variety of use cases. Major AI cloud services vendors provide a language processing catalog as part of their portfolio. This includes tools for developing and maintaining chatbot solutions or more sophisticated conversational virtual assistants. Additional language services include transcription, translation, speech-to-text, text-to-speech and text analytics. Developers can also use autoML to customize language models.

- **Computer vision:** This enables organizations to use pretrained visual models for generic images, though not for custom images. This may enable more rapid development of applications that process visual information. Major AI cloud services vendors provide a catalog of services for both images and video that can categorize elements of the images or video. Additional visual services include optical character recognition (OCR), handwriting recognition (HWR) label extraction. Pretrained systems often require no data science expertise and allow developers to gain unique and new insight by invoking an API. Developers can also use autoML to customize vision models.

The combination of the above as cloud services will accelerate digital business technology platform viability in the short term.

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Sample Vendors: Alibaba Group; Amazon Web Services; Baidu Cloud; Google (Cloud AI); IBM (IBM Cloud); Microsoft (Azure AI Platform)

Recommended Reading: "Magic Quadrant for Cloud AI Developer Services"

"Critical Capabilities for Cloud AI Developer Services"

Deep Neural Networks (Deep Learning)

Analysis By: Farhan Choudhary; Svetlana Sicular; Alexander Linden

Definition: Deep learning is a variant of machine learning algorithms that use multiple layers to solve problems through extraction of knowledge from raw data and transforming it at every level. These layers incrementally obtain higher-level features from the raw data allowing the solution of complex problems with higher accuracy, less features and less manual tuning.

Position and Adoption Speed Justification: Deep learning techniques have successfully made an impact in industries that include healthcare, transportation, national security, military, criminal justice, cities, finance, social media, among many others. However, convolutional neural networks and recurrent neural networks remain the most widely adopted and acclaimed approaches to deep learning due to ease of applicability with a variety of frameworks, prepackaged solutions, platforms and APIs: However, the end users are mostly unaware of the deep neural network's (DNN) role in these solutions and their performance remains brittle from a lack of transfer of learning capabilities.

The applicability of DNNs have vastly been in cognitive domains — voice, vision and text. DNNs are, however, tricky to build and train. To achieve consistently good results, you need large quantities of labeled data, data science expertise and special-purpose hardware — which are difficult to obtain and may even require a great deal of capital expenditure.

The level of hype about DNNs has worn out a bit — the commercial traction is falling slightly behind expectations due to the overall challenges with AI projects in terms of data and skills required, upkeep and operational costs.

User Advice: Data and analytics leaders of modernization initiatives should:

- **Examine and select** business areas where deep learning could provide best value from the available wide, heterogenous and quality data.
- **Explore** prepackaged capabilities first from various vendors and then proceed with platforms provided by cloud providers: Wherever possible, begin by using tools and platforms currently available from the major cloud providers. They have dedicated years of research on state-of-the-art systems, and in some narrow use cases, their systems will likely outperform almost anything you build and deploy yourself.
- **Develop and acquire skills:** Improve your machine learning experts' skills through trainings and allowing them to participate in competitions.
- **Acquire quality data for deep learning:** Working with DNNs is a long-term investment and adopters will not find value overnight. Utilize data annotation services or synthetic data generators to curate datasets for deep learning.

Business Impact: Of the many factors that created the current AI hype, DNNs were vastly responsible for it. They have transformational potential for all industries. Hyperscalers such as Amazon, Netflix, Baidu, Google, Tesla and so forth have set standards toward the use of deep learning in business applications. DNNs will be created by a handful of "creators" which will have the largest impact on the industry. Moving further along, there will be some "innovators" responsible for expanding DNN applications created by the "creators." Ultimately there will be "adopters" and "beginners" that will adopt and experiment with the available solutions from DNNs.

The basis of a DNN's potential is its ability to produce more nuanced representations of highly dimensional and complex data. A DNN can, for example, give promising results when interpreting medical images to diagnose cancer early; help improve the sight of visually impaired people; enable self-driving vehicles; colorize black-and-white photographs; and the possibilities are unlimited.

Systems relying on DNNs are most prominent now in military, law enforcement and healthcare.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: Amazon; Baidu; Facebook; Google; H2O.ai; Landing AI; Microsoft; NVIDIA; Vyasa Analytics

Recommended Reading: "Innovation Tech Insight for Deep Learning"

"4 Ways to Obtain AI Solutions Using Deep Neural Nets"

Natural Language Processing (NLP)

Analysis By: Bern Elliot; Erick Brethenoux

Definition: Natural language processing (NLP) enables an intuitive form of communication between humans and systems, i.e., NLP includes computational linguistic techniques aimed at parsing and interpreting (and sometimes generating) human languages. NLP techniques deal with the pragmatics (contextual), semantics (meanings), grammatical (syntax) and lexical (words) aspects of natural languages. The phonetic part is often left to speech-processing technologies that are essentially signal-processing systems.

Position and Adoption Speed Justification: Enterprise NLP usage is increasing as capabilities improve, along with new use cases based on conversational agents and automatic machine translation among others. Existing syntactic- and semantic-based methods are increasingly augmented and displaced with deep neural networks (DNNs) approaches, which are also referred to as sub-symbolic techniques.

Visible accomplishments include technologies that:

- Improve natural language parsing (via Google's SyntaxNet, an open-source, DNN-based, natural language parsing framework for TensorFlow)
- Translate in real time from one spoken language to another (as in Microsoft's Skype Translator)
- Build large-scale knowledge graphs (illustrated by the work of Google, IBM and

Microsoft)

- Offer answers instead of a list of page links (as in Google's information cards)
- Use of transfer learning to bootstrap training of new languages (research report by Amazon)

However, human language is complex and deeply influenced by cultural and other idiosyncratic conditions. So while NLP solutions have made progress, there are many subtleties and nuances that require human intervention to enable proper interpretation. These limitations are slowing adoption. For instance, dialogue capabilities are weak, DNNs are experimental and fragile, and understanding, inferences, context and synthesis pose significant challenges. Additionally, many NLP solutions require specialists in order to ensure continued accuracy of the grammars and models.

User Advice: NLP offers enterprises significant opportunities to improve operations and services. For many enterprises, the strongest and most immediate use cases for NLP are related to improved customer service (impacting cost, service levels, customer satisfaction and upselling), employee support (including making them smarter and more effective in their work) and automation of legal tasks (such as contract analysis, compliance enforcement, etc.).

Initial projects should start with modest goals in order to demonstrate success. As experience is obtained, projects should iterate, and scope can increase. More accessible use cases include translation of blogs and other casual documents, or mining text from customer interactions for insights on sentiment or issues is one of the more accessible use case.

Additional current NLP opportunities exist for enterprises but are not as mature or will require effort before they provide consistent returns on investment. Translation or transcription services, for instance for meetings or documents, offer opportunities to improve operations and lower costs. However, these NLP-based solutions are less accurate than similar human-based options and may benefit in some cases from human involvement.

As enterprises enhance their NLP implementations, new skills should be explored. Computational linguists, for example, are versed in the manipulation of various linguistic techniques and the impact of natural communications on users. Upskilling of data scientist talents might also be necessary given the increasing use of data science techniques in NLP applications.

Finally, the quality of NLP solutions offering knowledge-based consolidation, content mapping, search enhancements and text summarization will vary. As a result, enterprise planners should test and verify the effectiveness of these solutions before making significant commitments. If enterprises invest in specialized grammars, care should be taken that these be compatible across vendor solutions.

Business Impact: To obtain clear near-term ROI and to build enterprise knowledge and skills in the area of NLP, planners should leverage NLP applications such as:

- Virtual assistants and chatbots to improve interactions, including employee and customer services in select environments.
- Text mining to extract and summarize the focus of textual reports and preview what questions are most common before building chatbots.
- Basic transcription and translation services.
- Language-generation applications that produce natural language descriptions of tabular data, making it easier for many to understand.
- Keyword tagging in documents, making it easier to determine relevant sections or to extract other information such as intent and entities.
- Content moderation services that examine user-generated content (text or images), to flag potentially offensive content or to identify fake news in social media.
- Sentiment analysis to identify the feeling, opinion expressed in statements — from negative to neutral, to positive.
- Search improvements by better understanding the intent of a search query as well as by summarizing content that is retrieved.
- Text analytics to quickly process large numbers of organizations' documents and determine their compliance or legal validity.
- Advancement in insight engine text capabilities combined with more advanced NLP functionality.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Sample Vendors: Bitext; Clarabridge; CognitiveScale; Digital Reasoning; Google; IBM Watson; Microsoft; NLTK; SAS

Recommended Reading: "Cool Vendors in Speech and Natural Language"

"Cool Vendors in AI Core Technologies"

"A Framework for Applying AI in the Enterprise"

Sliding Into the Trough

Machine Learning

Analysis By: Pieter den Hamer; Carlie Idoine; Shubhangi Vashisth

Definition: Machine learning is an AI discipline that solves business problems by utilizing mathematical models to extract knowledge and patterns from data. There are three major approaches that relate to the types of observation provided: supervised learning, where observations contain input/output pairs (also known as "labeled data"); unsupervised learning (where labels are omitted); and reinforcement learning (where evaluations are given of how good or bad a situation is).

Position and Adoption Speed Justification: Machine learning is still a popular concept, given its extensive range of impacts on business. The triggers of its massive growth and adoption have been growing volumes of data, advancements in compute infrastructure and the complexities that conventional engineering approaches are unable to handle. As organizations continue to adopt these technologies, we recently see focus on aspects that relate to ML explainability and operationalization. Augmentation and automation (of parts) of the ML development process improve productivity of data scientists and enable citizen data scientists in making ML pervasive across the enterprise. In addition, pretrained ML models are increasingly available through cloud service APIs, often focused on specific domains or industries. New frontiers are being explored in synthetic data, new algorithms (e.g., deep learning variations) and new types of learning. These include federated/collaborative, generative adversarial, transfer, adaptive and self-supervised learning, all aiming to broaden ML adoption.

User Advice: For data and analytics leaders:

- Focus on the business problem. Start with simple business problems for which there is consensus about the expected outcomes, and gradually move toward complex business scenarios.
- Assemble a (virtual) team that prioritizes machine learning use cases, and establish a governance process to progress the most valuable use cases through to production.
- Utilize packaged applications, if you find one that suits your use case requirements. These often can provide superb cost-time-risk trade-offs and significantly lower the skills barrier.
- Nurture the required talent for machine learning. Partner with universities and thought leaders to keep up to date with the rapid pace of advances in data science. Create an environment conducive to continuous education, and set explicit expectations that this is a learning process and mistakes will be made.
- Provide guidelines and monitor compliance with respect to security, privacy, bias and explainability.
- Leverage the augmentation and automation of ML activities, avoiding unnecessary low level coding and alleviating labor intensive tasks for expert data scientists, while making ML accessible for citizen data scientists.
- Explicitly manage “MLops” for deploying, integrating and monitoring ML models, not underestimating time and complexity. To be successful, early involvement is required of both business stakeholders and IT for integration and operations.
- Focus on data as the fuel for machine learning by adjusting your data management and information governance strategies to enable your ML team. Data is your unique competitive differentiator and adequate data quality, such as the representativeness of historical data for current market conditions, is critical for the success of ML.

Business Impact: Machine learning drives improvements and new solutions to business problems across a vast array of business, consumer and social scenarios:

- Automation
- Drug research
- Customer engagement

- Supply chain optimization
- Predictive maintenance
- Operational effectiveness
- Workforce effectiveness
- Fraud detection
- Resource optimization

Machine learning impacts can be explicit or implicit. Explicit impacts result from machine learning initiatives. Implicit impacts result from products and solutions that you use without realizing they contain machine learning.

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: Alteryx; Databricks; Dataiku; DataRobot; H2O.ai; IBM; MathWorks; Microsoft; SAS; TIBCO Software

Recommended Reading: "Magic Quadrant for Data Science and Machine Learning Platforms"

"Critical Capabilities for Data Science and Machine Learning Platforms"

"Toolkit: RFP for Data Science and Machine Learning Platforms"

"3 Types of Machine Learning for the Enterprise"

"Top Organizational Pitfalls of Machine Learning Initiatives"

FPGA Accelerators

Analysis By: Alan Priestley

Definition: Field-programmable gate array (FPGA) accelerators are server-based, reconfigurable computing accelerators that deliver extremely high performance by enabling programmable hardware-level application acceleration.

Position and Adoption Speed Justification: FPGA accelerators feature a large array of programmable logic blocks, reconfigurable interconnects and memory subsystems that can be configured to accelerate specific algorithmic functions. This allows FPGAs to offload tasks from the main system processor. In the data center, FPGAs can be used in a range of use cases that require applying consistent processing operations to large volumes of data, such as high-frequency trading (HFT), hyperscale search, video analytics and DNA sequencing. For example, Microsoft is leveraging FPGAs for search analytics and networks, and Illumina's FPGA-based DRAGEN Bio-IT Platform enables high-performance genome-sequencing workflows.

FPGAs are typically configured using hardware programming languages, such as register transfer level (RTL) and VHSIC Hardware Description Language (VHDL), that are very complex to use; this has held back widespread adoption. However, major FPGA vendors (Intel and Xilinx), along with a number of startups such as Mipsology and Swarm64, are working to address this with libraries and toolsets that enable FPGAs to be configured using software-centric programming models.

Adoption is also becoming easier, helped by frameworks such as OpenCL that lower the time and skills required to use FPGAs. Workloads like deep learning (inference) and easier access to development platforms, exemplified by Amazon Web Services' FPGA-enabled instance types, are also driving adoption of FPGAs within the data center.

Today, the biggest growth opportunity for FPGAs in the data center is in the inference portion of deep-learning workloads. Given the evolving nature of this new use case and the maturing of the surrounding software ecosystem, FPGA accelerators have been positioned pretrough.

User Advice: FPGA accelerators can enable dramatic performance improvements within significantly smaller energy consumption footprints than comparable commodity technologies. I&O leaders need to evaluate applicability of FPGA accelerators by:

- Identifying application subsets that can be meaningfully impacted using FPGAs.
- Evaluating the availability of FPGA-based hardware for use in data center server deployments, such as FPGA-based PCIe add-in cards.
- Outlining costs associated with necessary skill set and programming challenges of FPGAs and the maturity of the software-centric programming toolsets.
- Leveraging cloud-based FPGA services to accelerate development.

I&O leaders should use FPGA accelerators when:

- Preconfigured solutions exist that can help dramatically transform key workloads (e.g., financial trading analytics, genome sequencing, etc.).
- Algorithms will evolve requiring frequent updates at the silicon level that can be utilized by broader applications.

Business Impact: FPGAs can deliver extreme performance and power efficiency for a growing number of workloads. They are well-suited for AI inference workloads as they excel in low-precision (8 bit and 16 bit) processing capabilities in energy-efficient footprints. While programmability continues to be a major challenge, limiting broader adoption of FPGAs, I&O leaders should evaluate FPGA-based solutions for genome sequencing, real-time trading, video processing and deep learning (inference). I&O leaders can further insulate against risks by utilizing cloud-based infrastructures for provisioning FPGAs (e.g., Amazon EC2 F1 instances, Microsoft Azure, Baidu cloud).

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Sample Vendors: Amazon Web Services; Baidu; Intel; Microsoft Azure; Xilinx

Recommended Reading: "Forecast Database, AI Neural Network Processing Semiconductors, 1Q20"

"Forecast Analysis: AI Neural Network Processing Semiconductor Revenue, Worldwide"

"Forecast Analysis: Data Center Workload Accelerators, Worldwide"

"Top 10 Technologies That Will Drive the Future of Infrastructure and Operations"

Chatbots

Analysis By: Magnus Revang

Definition: A chatbot is a domain-specific conversational interface that uses an app, messaging platform, social network or chat solution for its conversations. Chatbots range in sophistication from simple, decision-tree-based, to implementations built on feature-rich platforms. They are always narrow in scope. A chatbot can be text- or voice-based, or a combination of both.

Position and Adoption Speed Justification: Chatbots represent the No. 1 use of artificial intelligence (AI) in enterprises. Primary use cases are in customer service, human resources, IT help desk, self-service, scheduling, enterprise software front ends, employee productivity and advisory. There are also a variety of offerings in the market, such as developer self-service platforms, managed products, middleware offerings, integrated offerings and best-of-breed approaches.

Chatbots in social media, service desk, HR or commerce, as enterprise software front ends and for self-service, are all growing rapidly. Still, the vast majority of chatbots are simple, relying on scripted responses in a decision tree and relatively few intents. Similar to chatbots are virtual agents, which are broader in scope and sophistication, require more infrastructure and staffing to maintain, and are designed for an extended relationship with their users outside of single interactions. Users will interact with hundreds of chatbots, but few virtual agents.

The majority of implemented chatbots are unsophisticated and rule-based — failing to live up to expectations of stakeholders. The number of proofs of concept (POCs) is high, as is the failure rate to bring even unsophisticated chatbots into production. Gartner is seeing a backlash against chatbots, primarily focused on unsophisticated implementations.

User Advice:

- Start POCs for chatbots today, because most enterprises experience trouble scaling from the initial POC to production. The focus should be on uncovering the hindrances that will stand in your way.
- Treat vendors as tactical, not strategic — acknowledge that you'll most likely want to switch vendors in the future.
- Focus on vendors offering platforms that can support multiple chatbots.

Business Impact: Chatbots are the face of artificial intelligence and will impact all areas where there is communication between humans today. Customer service is a huge area where chatbots are already influential. Indeed, this will have a great impact on the number of service agents employed by an enterprise and how customer service itself is conducted. For chatbots as application interfaces, the change from "the user learns the interface" to "the chatbot is learning what the user wants" has significant implications for onboarding, training, productivity and efficiency inside the workplace. To summarize, chatbots will have a transformational impact on how we interact with technology.

Chatbots have played a strategic role in several companies' response to COVID-19. This might have an acceleration effect on the technology.

Benefit Rating: Transformational

Market Penetration: 20% to 50% of target audience

Maturity: Adolescent

Sample Vendors: Amazon; Cognigy; Google; IBM; Microsoft; NTT DOCOMO; Oracle; Rasa; Rulai

Recommended Reading: "Architecture of Conversational Platforms"

"Market Guide for Conversational Platforms"

"Market Guide for Virtual Customer Assistants"

Computer Vision

Analysis By: Nick Ingelbrecht

Definition: Computer vision is a process that involves capturing, processing and analyzing real-world images and videos to allow machines to extract meaningful, contextual information from the physical world.

Position and Adoption Speed Justification: Computer vision capabilities have advanced through the Hype Cycle as a result of improvements in the application of machine learning methods including deep neural networks, the availability of tooling and services as well as greater processing efficiencies. Enterprises everywhere face the challenge of how to exploit their visual information assets and automate the analysis of exponential volumes of image data. However, they face difficulties activating computer vision models in business processes, along with security and privacy concerns that impact their ability to realize business value. Gartner anticipates early mainstream adoption in the 2023-2025 time frame. Computer vision has progressed through the Hype Cycle in line with the growing maturity of machine learning solutions, including advances in optical character recognition products and object/behavior recognition models. Computer vision has broad applicability across numerous domains including automotive, retail, robotics, security, healthcare, manufacturing and many IoT applications, both in the visible and nonvisible spectrum including thermographic systems for remote fever and vital signs detection and facial recognition.

User Advice: Use computer vision to augment your organization's workforce capabilities by automating the processing of image and video data. Audit your organization's image/video assets and engage with business stakeholders to discover how computer vision applications can alleviate operational pain points, improve productivity and create new business opportunities. Ensure business stakeholders clearly articulate the tangible business benefits they are expecting from the computer vision assets to be developed.

In addition, we recommend:

- Focus initially on a few small projects, use fail-fast approaches and scale the most promising systems into production using cross-disciplinary teams. Do this by ensuring that sufficient software engineering resources are available to activate AI models in business processes and that governance and maintenance costs of image-based machine learning models are properly accounted for in ROI estimates.

- Critically assess change management impacts of implementing advanced analytics on the organization and its people as this has high potential to derail computer vision projects.

- Test production systems early in the real-world environment since lighting, color, object disposition and movement can break computer vision solutions that worked well in the development cycle.

- Build internal computer vision competencies and processes for exploiting image and video assets. This will enable the organization to make better procurement choices and lay the groundwork for more advanced innovation and product development opportunities.

- Exploit third-party computer vision tooling and services to accelerate data preparation and time to value by deploying early production systems.

- Evaluate legal, commercial and reputational risks associated with deploying computer vision solutions at the outset of customer/employee experience improvement projects.

- Be warned that the fast-evolving regulatory environment may derail computer vision projects due to privacy concerns.

Business Impact: The ability of organizations to capture value and generate insight from their own video/image data assets will become a question of competitiveness and ultimately survival. Key impacts of computer vision include:

- Greater levels of automation by reducing the demands on human monitoring staff and resulting in improved quality, speed and reliability of monitoring camera surveillance feeds.
- Improved decision support via event correlation, alarm management/prioritization and policy and rule engines for predetermined workflows.
- Enhanced customer experience in features such as queue monitoring and management, enhanced customer service and technical support.
- Reduced costs due to the ability to scale video systems without requiring greater human monitoring resources or manual processes.

Data is viewed as potentially one of the most important and unique strategic business assets that organizations control. Gartner estimates around 80% of these dark data assets — including uneventful surveillance video, video meetings and unsearchable text and graphics — are composed of image or video data which typically gets discarded because it has no apparent value. Key use cases today include the use of advanced analytics for video surveillance automation, health and safety compliance (PPE detection, COVID-19 mitigation, etc.), visual search, shopper and shelf analysis, automotive applications, OCR and quality assurance/production line automation in manufacturing. Increasingly, in the future, organizations that are unable to value and leverage their computer vision assets strategically will become uncompetitive.

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Sample Vendors: Amazon Web Services; Another Brain; Baidu; Clarifai; Deepomatic; Google; Matroid; Microsoft; Nyris; Tencent

Recommended Reading: "Venture Capital Growth Insights: Computer Vision"

"Competitive Landscape: Computer Vision Platform Service Providers"

"Survey Analysis: Computer Vision Drives Enterprise Adoption of Artificial Intelligence"

"Market Trends: Facial Recognition for Enhanced Physical Security — Differentiating the Good, the Bad and the Ugly"

"Market Trends: Machine Vision Will Be the Game Changer Across Markets"

"Innovation Insight for Video/Image Analytics"

"Emerging Technologies: Top Advanced Computer Vision Use Cases for Retail"

"Critical Steps to Cash In on the Computer Vision Gold Rush"

Autonomous Vehicles

Analysis By: Jonathan Davenport

Definition: Autonomous vehicles use various onboard sensing and localization technologies, such as lidar, radar, cameras, GPS and map data, in combination with AI-based decision making, to drive without human intervention. This Innovation Profile does not cover ADAS features that require humans to supervise vehicle operations. While self-driving cars are getting most of the attention at present, the technology can also be applied to nonpassenger vehicles for transportation of goods.

Position and Adoption Speed Justification: There have been a number of signs of autonomous driving moving into the Trough of Disillusionment during the past year. Drive.ai and Starsky Robotics failed, Cruise cut 8% of its workforce, and Zoox is looking for a buyer, Continental has delayed AV investments after Q120 earnings plummeted and Audi has abandoned its plans to introduce the Level 3 traffic jam pilot feature into its A8 vehicles, which it had originally announced back in 2017. Likewise, Ford Motor Company made the decision to shift the launch of its self-driving services to 2022 to evaluate the long-term impact of COVID-19 on customer behaviors.

But there has been increased investment too. For example, Intel's Mobileye has acquired Moovit and is developing an autonomous mobility as a service (MaaS) solution for the emerging robotaxi market. This plan shifts Intel from being a supplier of chips and self-driving systems for the automotive industry and places it in direct competition with automakers' own mobility ambitions and the likes of Waymo, Baidu and Yandex. Likewise, autonomous vehicle pilots and trials have continued to be undertaken, though most continue to be supported by safety drivers. To overcome regulatory issues, many autonomous shuttle buses have been demonstrated on private road networks, such as at airports.

The efforts of automobile manufacturers and technology companies to develop autonomous vehicles have been prominently featured by mainstream media, leading to unrealistic and inflated expectations for the technology. Artificial intelligence (AI) is a critical technology for enabling autonomous vehicles, and development of machine learning algorithms for autonomous vehicles has accelerated. Key challenges for the realization of autonomous vehicles continue to be centered on cost reductions for the technology and industrialization. However, the challenges increasingly include regulatory, legal and societal considerations, such as permits for operation, liability, insurance and the effects of human interaction.

Continued advancements in sensing, positioning, imaging, guidance, mapping and communications technologies, combined with AI algorithms and high-performance computing capabilities, are converging to bring the autonomous vehicle closer to reality. However, in 2020, complexity and cost challenges remain high, which is impacting reliability and affordability requirements, as well as hindering the ability for companies to get regulatory approval.

User Advice: The adoption of autonomous vehicle technology will require increasing levels of technical sophistication and reliability that rely less and less on human driving intervention. Automotive companies, service providers, governments and technology vendors (for example, software, hardware, sensor, map data and network providers) should collaborate on joint research and investments to advance the required technologies, as well as work on legislative frameworks for self-driving cars.

Furthermore, consumer education is critical to ensure that demand meets expectations once autonomous vehicle technology is ready for broad deployment. Specific focus must be applied to the transitional phase, where autonomous or semiautonomous vehicles will coexist with an older fleet of nonautonomous vehicles.

Look for use cases, such as mining, agriculture or airports, where autonomous vehicles can operate in restricted areas safely without regulatory restrictions. Use these implementations to drive early revenue and gather data and insights to improve the performance of self-driving systems.

Autonomous vehicles will have a disruptive impact on some jobs, such as bus, taxi and truck drivers. Develop policies and programs to train and migrate employees who will be affected by automation to other roles.

Business Impact: The main implications of self-driving vehicles will be in the economic, business and societal dimensions. Automotive and technology companies will be able to market autonomous vehicles as having innovative driver assistance, safety and convenience features, as well as being an option to reduce vehicle fuel consumption and improve traffic management. The interest of nonmobility companies (such as Intel, Waymo, Apple and Baidu) highlights the opportunity to turn self-driving cars into mobile computing systems. These systems offer an ideal platform for the consumption and creation of digital content, including location-based services, vehicle-centric information and communications technologies.

Autonomous vehicles are also a part of mobility innovations and new transportation services that have the potential to disrupt established business models. For example, autonomous vehicles will eventually lead to new offerings that highlight mobility-on-demand access over vehicle ownership by having driverless vehicles pick up occupants when needed. Autonomous vehicles will deliver significant societal benefits, including reduced accidents, injuries and fatalities, as well as improved traffic management, which could impact other socioeconomic trends.

When autonomous driving enters the Trough of Disillusionment, it might be a good opportunity for new market entrants.

Benefit Rating: Transformational

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Sample Vendors: Audi; AutoX; Daimler Group; General Motors; Mobileye; Pony.ai; Tesla; Uber; Waymo

Recommended Reading: "Market Trends: Monetizing Connected and Autonomous Vehicle Data"

"Forecast Analysis: Autonomous Vehicle Net Additions, Internet of Things, Worldwide"

"Utilize Partnerships to Secure a Winning Position in the Autonomous Driving Ecosystem"

"Market Insight: Use Situationally Aware Platforms to Enable Safe Autonomous Vehicle Handovers"

"Maverick* Research: Autonomous Mobile Structures Will Fuel the Sharing Economy"

Cognitive Computing

Analysis By: David Pidsley

Definition: Cognitive computing is a class of computing systems that are designed by analogy to nervous systems and mental processes, particularly the human brain. Cognitive computing mimics and improves the sensing and reasoning of a human. Cognitive systems are typically adaptive, interactive, stateful and aware of context. We recognize "cognitive computing" is used as marketing term by vendors in the marketplace; we do not believe these systems are truly capable of cognition by themselves.

Position and Adoption Speed Justification: The term "cognitive computing" rapidly climbed to the Peak of Inflated Expectations due to marketing by major vendors and startups seeking differentiation in the AI marketplace. The analogy of computer vision was self-evident while computer mental processes (e.g., reasoning) was not. Marketing of "cognitive" became of secondary importance as user focus shifted to pragmatic industry/functional use cases. It is sliding rapidly into the Trough of Disillusionment and market penetration remains low for the target audience who struggle to train components to work effectively using generalized, code-free interfaces. Consequently, organizational/cultural acceptance is hampered by models which are imprecise, inaccurate, with poor recall. These are obstacles to cognitive computing significantly increasing revenue or cost savings in an enterprise.

"Cognitive" focuses on a biological approach to solving general classes of computing problem, rather than being tailored to industry, enterprise or business unit needs. Such assortments of loosely related AI components make it hard for users to assemble services for the data and processes that they have. This analogy is not a new way of doing business within/across industries that will result in major shifts in industry dynamics, and we expect the term to be obsolete before plateau.

"Cognitive computing" as a term should be superseded by the practical mechanisms and AI techniques that solve business problems by simulating cognitive processes, such as those available within AI Developer and Teaching Toolkits and AI cloud vendor offerings.

User Advice: Users should avoid the ambiguity created by the marketing term “cognitive computing” and instead focus on practical and pragmatic problem solving using AI techniques.

CIOs should:

- Recognize high-profile retreats from large-scale “cognitive” programs (MD Anderson Cancer Center, Nordnet) and emphasize AI.
- Evaluate RPA with AI in the background to create intelligent applications for valuable use cases.
- Experiment with several vendors to develop a long-range business case.
- Resist the temptation to select “winners,” at this stage.

IT leaders should utilize AI methods for practical business goals, test a hypothesis in short iterations toward a minimum viable product (MVP) for use cases where:

- Skilled people are detecting, labelling, categorizing or predicting entities — regularly and in mature business processes, e.g., form processing.
- High speed and process consistency are important to the outcome.
- Data formats are text, audio, image, video, Internet of Things (IoT).

Data and analytics leaders should focus on valuable business applications of AI in functional business units, supported by data scientists using the AI cloud:

- HR: AI for the workforce: cognitive diversity, work stress, mental workload, decision-making, skilled performance, cognitive ergonomics, human reliability and training.
- Field-working: Computer vision, augmented reality (AR) and wearables, e.g., data capture, field communications, health and safety.
- Supply Chain: To leverage virtual assistants to classify interactions, label user provided data, make recommendations make decisions in dynamic environments.

Data scientists and developers should apply AI developer toolkits using big datasets to solve problems and optimize for areas of uncertainty (Bimodal IT — Mode 2).

Citizen data scientists, citizen developers, RPA developers should make incremental but significant improvements to established, narrow processes to increase revenue and optimize costs.

Business Impact: The term “cognitive computing” has recognition in multiple industries and business functions, recently being subsumed by the broader term “Artificial Intelligence” (AI), and so will be obsolete before Plateau.

Related technologies include:

- Virtual assistants
- Chatbots
- Conversational user interfaces
- Natural language generation
- Mixed reality environments

AI increasingly impacts business functions that adopted cognitive computing, including:

- Supply chain
- Procurement
- Human resources
- Customer services
- IT/Data
- Security

Industries impacted by AI include:

- Banking
- Insurance
- Healthcare

- Telecommunications
- Retail
- Utilities
- Mining
- Construction
- Manufacturing
- Maintenance
- Knowledge workers

Business benefits from AI include improved:

- Cycle times
- Output per dollar of company overheads
- Worker productivity
- R&D return on investment
- Risk mitigation and avoiding opportunity costs of poor/late decisions
- Employee safety and satisfaction

Data and Analytics Leaders can apply AI to:

- Augmented data management, augmented analytics, data labeling, annotation services
- Intelligent applications
- Robotic process automation (RPA)
- Predictive analytics, prescriptive analytics, social analytics
- Continuous intelligence, augmented intelligence

- Decision intelligence, decision management

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Sample Vendors: Aera Technology; Amazon SageMaker; Cognitive Scale; Digital Reasoning; IBM Watson; IPsoft; Microsoft Azure Machine Learning; SparkCognition; Tata Consultancy Services (Digitate); Wipro HOLMES

Recommended Reading: "Magic Quadrant for Cloud AI Developer Services"

"Cognitive Computing Trends in Shared Services"

Climbing the Slope

Insight Engines

Analysis By: Stephen Emmott

Definition: Insight engines apply relevancy methods to discover, organize, describe and analyze data. This enables existing or synthesized information to be delivered interactively or proactively in the context of digital workers, customers or constituents at timely business moments.

Position and Adoption Speed Justification: The hype behind insight engines stems from the use of AI to reinvent enterprise search, enabling enterprises to shift from keyword- to entity-centric discovery and unlock patterns inside unstructured and structured data, sourced both internally and externally. This shift enables insight — accurate and deep understanding — needed for purposeful action by placing data in context to inform. Data must be extracted from myriad sources, enriched and indexed; user queries must be analyzed and interpreted; and the touchpoint used must align with the task at hand. This comes packaged at a foundational level but must be developed by vendors, partners, and/or clients at the domain and situational levels where vendors do not offer prebuilt applications tailored to select domains and situations, e.g., CRM. Vendors have extended their use of AI (especially machine learning and knowledge graphs), new products are entering the market, and both Google G Suite and Microsoft Office 365 now include insight engines in their cloud office. Yet, the majority of enterprises have yet to shift from enabling search to delivering insight, and application of insight engines to the many and varied use cases they have the potential to serve. As such, insight engines have moved through the Trough of Disillusionment and ascend the Slope of Enlightenment.

User Advice: Focus the purpose of insight engines on informing employees to deliver insight rather than searching for information. At the highest level of maturity, insight engines retrieve and synthesize facts, deliver these through other tools, and do so proactively. For instance, a chat with a bot through Microsoft Teams or Slack can be powered by an insight engine that delivers answers as snippets from documents. More typical is a traditional search page with enhancements to guide the user using (1) autosuggest or autocomplete, (2) structured results with relevant facets to allow refinement, and (3) recommendations. Moving from the latter toward the former requires clarity of purpose and discrete application of the underlying insight engine. The beneficiaries of insight — people — must be placed at the center of the initiative: personify them, identify their use cases, the applications they use to conduct work, and the sources of content and data they need to draw information from. Then, relate these back to specific business outcomes and their measures.

With most enterprises using or contemplating cloud office, many application leaders will find their cloud office includes an insight engine — Microsoft Search (in the case of Office 365) or Google Cloud Search (in the case of G Suite). These products are deeply embedded and demonstrate what is possible with a focus on collaboration and sharing. Breadth and depth of capabilities can be obtained by looking at other insight engine vendors, with customer use cases and case studies exemplifying what is possible. See “Magic Quadrant for Insight Engines” for more information.

Enterprises have one or more cloud offices, multiple search engines operating, and search and insight capabilities within CRM, ITSM, and other categories of their application portfolio. An essential step therefore is reviewing how these various search and insight engines perform and interrelate. Deciding the right portfolio of insight engines, configuring these, and enabling users to profit from them is key to ensuring insight can be facilitated across the foundational, domain and situational levels of the enterprise.

Business Impact: The principal impact of insight engines is on an organization’s digital workplace and its capability to elevate the digital dexterity of its employees. They impact all functional domains across all industries, but are most impactful when utilized as a platform upon which to develop applications aligned with specific domains and situations. For example, proactively informing customer support agents in the context of a CRM. Such localization improves digital dexterity by enabling employees to better orient themselves to decide/act, acquire knowledge and collaborate.

A lesser-known but significant impact of insight engines is in terms of supporting automation. Insight engines can be integrated with other software such as RPA, to support the automation of various workflows relating to content, e.g., claims processing in insurance. Insight engines also have a role to play in support of digital experiences provided to external constituents, such as customer and suppliers, in the form of self-help knowledge bases, decision support, retrieval of content assets, etc.

Given these impacts, the semantic models and knowledge representation underlying insight engines and other applications will increasingly be a foundation for enterprises' natural language ambitions.

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Sample Vendors: Coveo; Funnelback; Google; IBM; IntraFind; Lucidworks; Micro Focus; Microsoft; Mindbreeze; Sinequa

Recommended Reading: "Magic Quadrant for Insight Engines"

"Critical Capabilities for Insight Engines"

Entering the Plateau

GPU Accelerators

Analysis By: Alan Priestley; Martin Reynolds; Chirag Dekate

Definition: GPU-accelerated computing is the use of a graphics processing unit (GPU) to accelerate highly parallel compute-intensive portions of the workloads in conjunction with a CPU.

Position and Adoption Speed Justification: GPUs are highly parallel floating-point processors designed for graphics and visualization workloads. Over the last decade, NVIDIA and others have added programmable capability to GPUs, enabling software applications to access deep, fast-floating-point resources. A number of GPU designs also host very high-bandwidth memory subsystems. For many highly parallel, repetitive, compute-intensive applications, GPUs deliver dramatic performance improvements over traditional CPUs. GPU subsystems are actively deployed in two key markets: high-performance computing (HPC) and AI.

In HPC, compute-intensive applications including molecular dynamics, computational fluid dynamics, financial modeling and geospatial applications can, in many cases, can be dramatically accelerated using GPUs. Programming GPUs can be challenging because execution order and code optimization are critical. However, toolkits like NVIDIA's CUDA can dramatically lower the programming challenges.

In AI, DNN technologies are maturing quickly, supported by open-source software frameworks from the large cloud providers. Today most of the DNN frameworks, including TensorFlow, Torch, Caffe, Apache MXNet and Microsoft Cognitive Toolkit, support GPU acceleration. Although many ASICs are emerging, few offer broad ecosystem support. Ease of programming GPUs, using tools such as CuDNN, and broad ecosystem support continue to be distinct differentiators and advantages.

GPU computing has moved forward on the Slope of Enlightenment primarily due to maturity of system stacks resulting in easier adoption.

User Advice: GPU-accelerated computing can deliver extreme performance for highly parallel compute-intensive workloads in HPC, DNN training and inferencing. GPU computing is also available as a cloud service and may be economical for applications where utilization is low but urgency of completion is high.

Leverage GPU-based solutions to accelerate compatible applications by:

- Selecting GPU compute platforms that offer the most mature software stack.
- Optimizing infrastructure costs by evaluating cloud-hosted GPU environments for proof of concept (POC) and prototype phases.

Use GPU accelerators when applications require extreme performance and have high degrees of compute parallelism (example: many high-performance computing and deep learning applications).

Business Impact: HPC and deep learning are essential to many digital business strategies. For this fast-growing workload, traditional enterprise ecosystems based on CPU-only approaches will not suffice. Leverage mature GPU technologies for select HPC applications and deep learning infrastructures. Programmability challenges have been largely solved in GPU-accelerated computing by toolsets such as CUDA. I&O leaders can minimize risk by using cloud-hosted GPU environments for testing and evaluation.

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Mature mainstream

Sample Vendors: AMD; Intel; NVIDIA

Recommended Reading: "Top 10 Technologies That Will Drive the Future of Infrastructure and Operations"

"Forecast Database, AI Neural Network Processing Semiconductors, 1Q20"

"Forecast Analysis: AI Neural Network Processing Semiconductor Revenue, Worldwide"

"Forecast Analysis: Data Center Workload Accelerators, Worldwide"

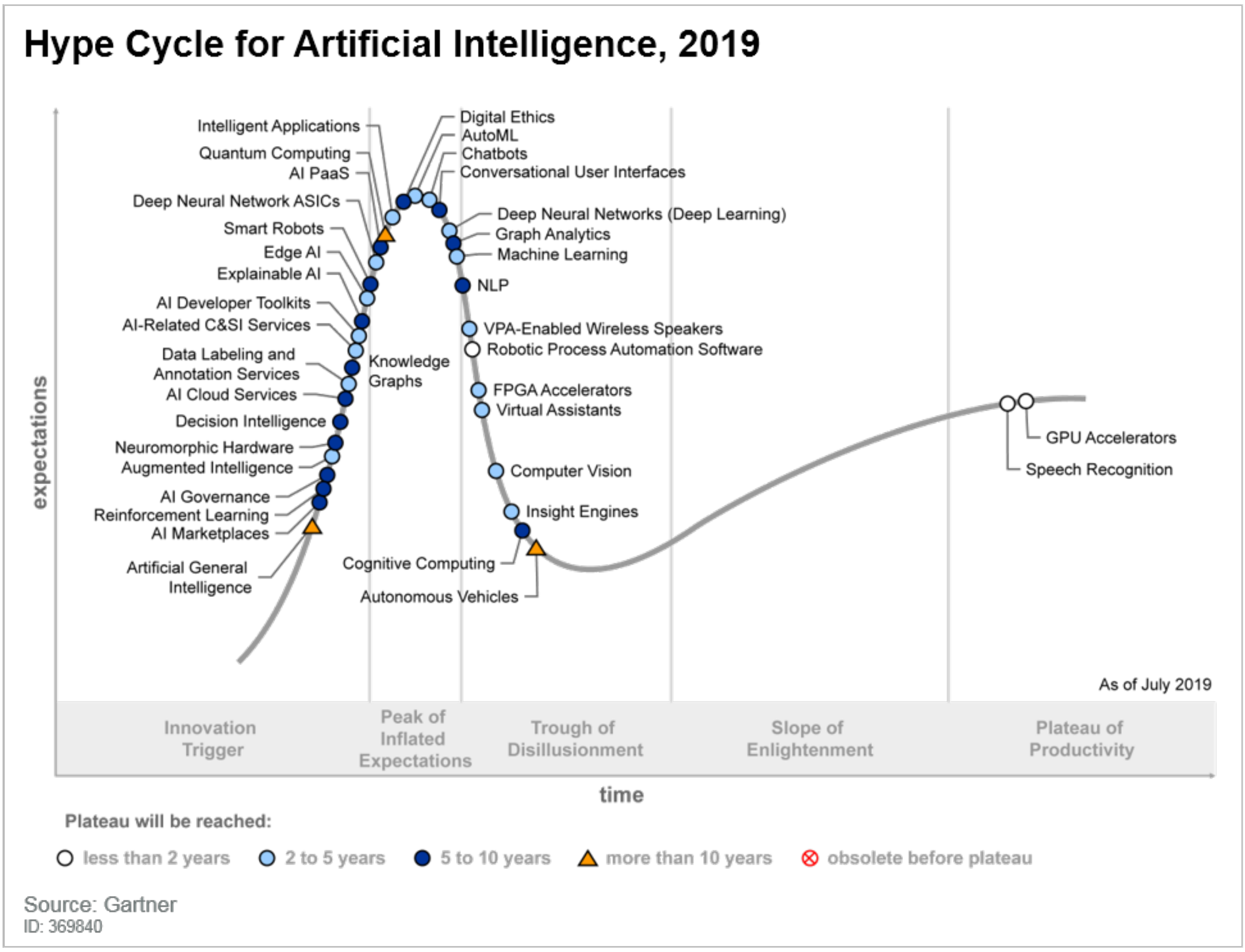
"Forecast Analysis: Discrete GPUs, Worldwide"

"An Action Plan for Growing AI-Accelerator-Enabled Server Revenue"

"Predicts 2019: Artificial Intelligence Core Technologies"

Appendixes

Figure 3. Hype Cycle for Artificial Intelligence, 2019



Hype Cycle Phases, Benefit Ratings and Maturity Levels

Table 1: Hype Cycle Phases

Phase	Definition
Innovation Trigger	A breakthrough, public demonstration, product launch or other event generates significant press and industry interest.
Peak of Inflated Expectations	During this phase of overenthusiasm and unrealistic projections, a flurry of well-publicized activity by technology leaders results in some successes, but more failures, as the technology is pushed to its limits. The only enterprises making money are conference organizers and magazine publishers.
Trough of	Because the technology does not live up to its overinflated expectations,

Disillusionment	it rapidly becomes unfashionable. Media interest wanes, except for a few cautionary tales.
Slope of Enlightenment	Focused experimentation and solid hard work by an increasingly diverse range of organizations lead to a true understanding of the technology's applicability, risks and benefits. Commercial off-the-shelf methodologies and tools ease the development process.
Plateau of Productivity	The real-world benefits of the technology are demonstrated and accepted. Tools and methodologies are increasingly stable as they enter their second and third generations. Growing numbers of organizations feel comfortable with the reduced level of risk; the rapid growth phase of adoption begins. Approximately 20% of the technology's target audience has adopted or is adopting the technology as it enters this phase.
Years to Mainstream Adoption	The time required for the technology to reach the Plateau of Productivity.

Source: Gartner (July 2020)

Table 2: Benefit Ratings

Benefit Rating ↓	**Definition** ↓
Transformational	Enables new ways of doing business across industries that will result in major shifts in industry dynamics
High	Enables new ways of performing horizontal or vertical processes that will result in significantly increased revenue or cost savings for an enterprise
Moderate	Provides incremental improvements to established processes that will result in increased revenue or cost savings for an enterprise
Low	Slightly improves processes (for example, improved user experience) that will be difficult to translate into increased revenue or cost savings

Source: Gartner (July 2020)

Table 3: Maturity Levels

Maturity Level	Status	Products/Vendors
Embryonic	■ In labs	■ None
Emerging	■ Commercialization by vendors ■ Pilots and deployments by industry leaders	■ First generation ■ High price ■ Much customization
Adolescent	■ Maturing technology capabilities and process understanding ■ Uptake beyond early adopters	■ Second generation ■ Less customization
Early mainstream	■ Proven technology ■ Vendors, technology and adoption rapidly evolving	■ Third generation ■ More out-of-the-box methodologies
Mature mainstream	■ Robust technology ■ Not much evolution in vendors or technology	■ Several dominant vendors
Legacy	■ Not appropriate for new developments ■ Cost of migration constrains replacement	■ Maintenance revenue focus
Obsolete	■ Rarely used	■ Used/resale market

only

Source: Gartner (July 2020)

Evidence

This Hype Cycle draws on data from Gartner polls conducted with a view to understanding how AI investment strategies have changed since the onset of the COVID-19 crisis. Data was collected in May and June 2020 from 202 respondents who participated in a webinar entitled “The Reset: Reexamining AI Investment Strategies” on 26 May 2020 and local briefings, entitled “The Future of Data Science and Machine Learning: Critical Trends You Can’t Ignore,” held in Russia on 2 June 2020 and South Korea on 4 June 2020. The results are not representative of the whole world or the whole market. They are simple averages of results from the respondent sample.