IDC MarketScape: Asia/Pacific (Excluding Japan) Vision Artificial Intelligence Software Platform 2021 Vendor Assessment

Christopher Lee Marshall   Jessie Danqing Cai
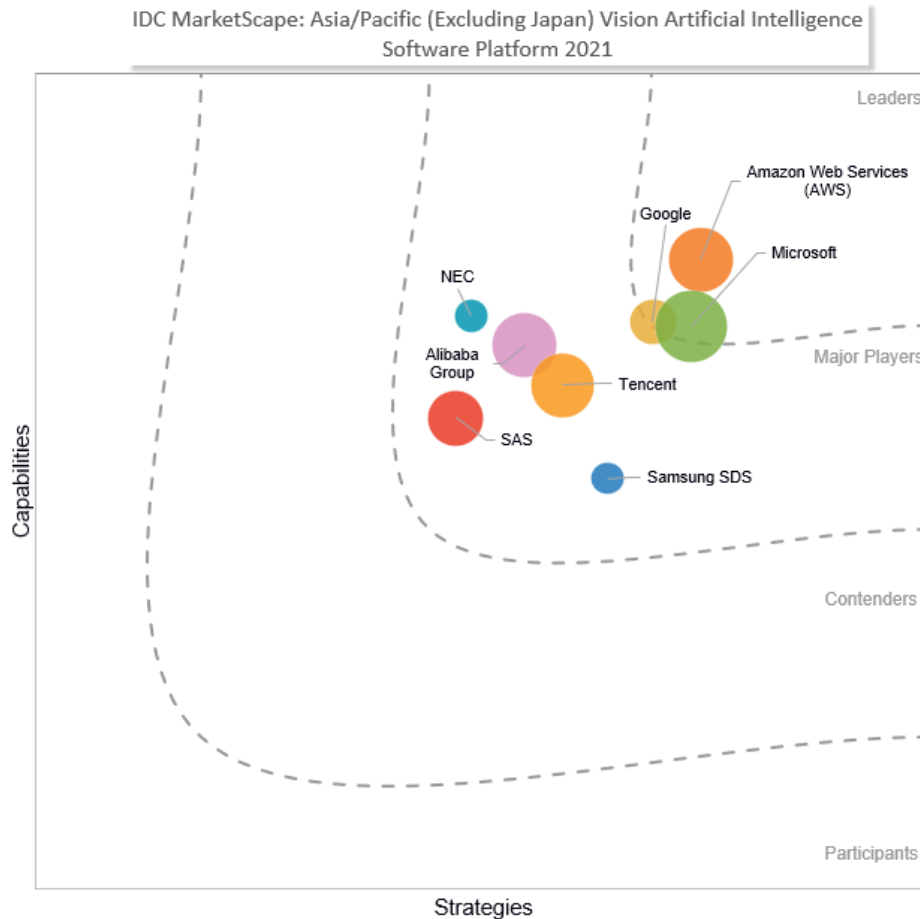Sandeep Sharma   Dylan Yap

THIS MARKETSCAPE EXCERPT FEATURES: AWS

IDC MARKETSCAPE FIGURE

FIGURE 1

Asia/Pacific (Excluding Japan) Vision Artificial Intelligence Software Platform 2021 Vendor Assessment

Source: IDC, 2021

Please refer to the Appendix for the detailed methodology, market definition, and scoring criteria.

IN THIS EXCERPT

The content for this excerpt was taken directly from IDC MarketScape Asia/Pacific (Excluding Japan) Vision Artificial Intelligence Software Platform 2021 Vendor Assessment (Doc #47490521). All or parts of the following sections are included in this excerpt: IDC Opinion, IDC MarketScape Vendor Inclusion Criteria, Advice for Technology Buyers, Featured Vendor Profile, Appendix and Learn More. Also included is Figure 1 and Table 1.

IDC OPINION

It has been estimated that more than 50% of the cortex of the human brain is dedicated to sight and vision — more than all the other senses combined. Sight and vision are different. Vision is how the mind, an aspect of the brain, interprets the images captured by the sense of sight. About 80% of what we learn about our world is because of perception and cognition through vision, and 30% of the brain's neurons are dedicated to visual processing, compared with just 8% for touch and 3% for hearing (refer to Sells, S.B. and Fixott, R. S., *Evaluation of Research on Effects of Visual Training on Visual Functions,* Elsevier, American Journal of Ophthalmology, 44(2), 1957, 230–236). Vision enables us to act on what we see and respond to it either with our muscles or through our intellect. Therefore, it is no surprise that computer vision, also known as vision artificial intelligence (AI), has long been a priority of AI research. Over the last 20 years, the advent of deep learning (DL) algorithms, the investment in systematically labeled image libraries, together with the spread of charge-coupled vision-capture devices, and the availability of massive storage and high-performance compute capabilities have revolutionized vision AI and made a host of applications practical — from fault detection to facial recognition.

In this study, we profile and evaluate the major vision AI software platform vendors according to the IDC MarketScape model. It uses surveys and interviews conducted between March and August 2021 and incorporates indicative market shares of vendors in Asia/Pacific (excluding Japan) (APEJ) as reflected in IDC's tracker products. The assessment methodology, including scoring dimensions, items, and criteria, is detailed in the Appendix section.

In this research, we have identified three basic categories of vision AI use cases from those frequently mentioned by reference customers or customer stories:

- **Vision AI for productivity.** Productivity use cases tend to support existing business-critical operations. AI is used to augment or automate tasks that are traditionally performed by people, thereby relieving or reducing their workload or lowering the level of attention required, thus making the work safer, easier, and more productive. Apart from improving productivity, vision AI in this category also improves work outcomes with better accuracy and reliability. These use cases typically have well-defined scope and outcomes (pass/fail, hit/miss, and so forth), well-established criteria for success, and restricted implementation. Common use cases include in-scene text extraction (DL-based optical character recognition [OCR]) for process automation; visual quality and conformity inspections for discrete manufacturing, logistics, and construction; crowd behavior, social distancing, missing child search, and criminal detection for public safety and law enforcement; and so on.

- **Vision AI for end-user experience**. In this category, vision AI is used to achieve less intrusive or more differentiated services and offerings for customers and end users at scale. The adoption drivers of these use cases are typically not only about productivity but also the user's broader digital experience. The ongoing COVID-19 pandemic has also accelerated the adoption of use cases in this category. It follows that buying organizations mostly look out for performance indicators, such as the scope and speed of end users adopting and sharing these experiences. The scope of adoption is less predictable, making the software service's capability in latency and scalability more important. Such experience use cases often trigger innovations within the larger ecosystem. Examples of use cases include identity authentication for customer onboarding or transaction; image and video content annotation, curation, search, and moderation for media and entertainment; video and three-dimensional (3D) image content augmentation and generation for media and entertainment; and so forth.

- **Vision AI for decision recommendation.** Decisioning use cases tend to solve more cognitively demanding problems. Traditionally, they would have been performed heuristically by experts or executives based on cumulative experience and limited data. These use cases often require multiple AI and analytics technologies in addition to vision AI, such as search, forecasting, and optimization, to produce a nontrivial recommendation regarding an action or solution. It follows that solutioning is more complex, and the contribution to business outcomes may take more time to realize. Their functionality can be high value-add and can often lead to more distributed decision making within the enterprise. Such use cases include shelf and route analysis for retail store operations, behavior detection and analysis for online education and videoconferencing, medical image analysis for healthcare, and aerial image analysis for resource surveying, among others.

These categories are not exhaustive, but they span the space of major use cases and help organizations in search of vision AI solutions make more effective like-for-like comparisons. All participating vendors produce effective vision AI solutions; however, each has its own strengths and weakness, making a particular client challenge more or less appropriate. Hence, we discourage a simplistic reading of the charts, such as looking for the "one best vendor for all use cases", because there is no such thing.

Other key findings in this research include:

- **The significance of data gravity.** Not surprisingly, some customers indicated that their vendor selection decision was partly because the target application in which the vision AI service is incorporated already ran on the vendor cloud/enterprise platform. The choice of a machine learning (ML) development platform also influences the vendor selection process.

- **The value of solutioning and customization.** Because not all buyer organizations invest in in-house software development, having a readily usable solution is often more important than the specific technical competencies of software products. Abilities in model customization, solution integration and implementation, and continuous management of model quality also contribute significantly to a vendor's current success. A lot of these capabilities depend on vendors' strengths in their technical sales and partner networks.

- **Performance from edge to cloud.** Solution performance, including model quality, response latency, and service availability, is key to support customer-facing and mission-critical use cases, and many of these use cases run on edge devices (this trend will continue to grow). Vendors need to architect their models for edge-to-cloud deployment, incorporating flexibility in their resource usage and optimization.

- **A virtuous cycle of learning**. Many reference customers resonate that they can no longer work without the functions enabled by vision AI, and they have seen the model accuracy improve over time. This positive feedback loop of greater use, more data, and improved model capabilities, leading to even greater use, augers well for the future of vision AI capabilities and strengthens the competitive position of early movers and adopters in this space.

- **The importance of governance.** Open source libraries and projects of vision AI can be leveraged by all. However, there is a great divide between open source resources and enterprise-grade solutions, which is largely defined by the way these solutions handle data and model governance. For example, all major vendors make efforts to ensure data privacy protection, and data encryption is a default when persistence on the cloud platform is required.

- **Tailwinds in face-based authentication**. We have noted that face detection and recognition use cases are controversial in the United States and Europe but much less so in Asia/Pacific. There are, of course, international vendors that have chosen to exclude such capabilities, leaving the field open for regional players. At the same time, the ongoing COVID-19 pandemic has provided a tailwind to accelerate face-based authentication by enabling safer, touchless digital experiences for both consumers and gig economy workforce.

IDC MARKETSCAPE VENDOR INCLUSION CRITERIA

Eight vendors are included in this report. The criteria for their inclusion include:

- The vendor provides vision AI software platform products under independent product SKUs or as part of a general AI software platform.

- The platform product provides pretrained or customized models based on computer vision technologies, including scene detection, object detection, recognition, DL, event detection, and so forth, which customers can use as a software service. The platform product supports both still image and video processing. DL-based text extraction, more commonly known as DL-based OCR, is also included as an option.

- The platform product is based on the company's own IP.

- The platform product has generated software license revenue or consumption-based software revenue for at least 12 months in APEJ as of March 2021.

We have tried to capture the most significant vision AI players operating in the APEJ market, especially those based in Asia/Pacific locations.

ADVICE FOR TECHNOLOGY BUYERS

Vision AI will be a basic component of many future software applications and is expected to enhance productivity, customer experience (CX), decision making, and knowledge discovery across industries and geographies. It is already necessary to highlight the major vendors and more mature solutions in the market because early movers in this market tend to consolidate their strengths with greater access to training data, more iterations of algorithm variations, deeper understanding of operational contexts, such as constraints and noises, and more systematic approaches to work with solution partners in the ecosystem.

We recommend the following to technology buyers:

- Refer to vision AI use cases, as mentioned at the start of this report, to gather ideas about product and service innovations that can enhance productivity, experience, and decision making. Many of these use cases are entering the mainstream.

- Have a clear understanding about the three types of competitive spaces for vision AI use cases and closely associated capabilities to look out for as summarized in Table 1 in the next section.

- Be conscious of the ongoing challenges associated with vision AI solutions, such as data residency, data privacy, model drift, model bias, model robustness, and so on, and always maintain some human oversight over more critical or controversial aspects of the technology.

- Synergize the adoption of vision AI solutions with other technology investments, such as smart edges, application modernization, and business process automation (BPA).

TABLE 1

Core Capabilities Associated with Use Case Categories

Use Case Category	Characteristics	Closely Associated Capabilities
Productivity	- This targets frontline workers. - This includes scenario-specific objects to be detected, classified, or recognized. - The next-step actions are clearly given according to the original business process. - These use cases typically have well-defined outcomes and established criteria for success, and they face restrictions in implementation. - System availability and real-time performance are more important.	- Provision of model training using customers' own image data - Architecture of a model for optimal edge deployment - Model drift and feedback management - Image data annotation - Synthetic data generation - Data encryption and protection
Experience	- This targets mostly consumers. - This includes common objects to be detected, classified, or recognized, such as faces, product types, and ambience types. - The next-step actions are embedded in a CX application or interface. - Indicators of success are the scope and speed of end users adopting and sharing these experiences. - Apart from system availability and latency, the robustness and scalability of such services are also important.	- Provision and deployment of pretrained models - Fine-tuning of pretrained models for enhanced accuracy and robustness against noises and constraints - Data encryption and protection - Model fairness management - Model drift and feedback management - Image/Video data repository

TABLE 1

Core Capabilities Associated with Use Case Categories

Use Case Category	Characteristics	Closely Associated Capabilities
Decisioning	▪ This targets business users and, often, team leads, managers, and executives. ▪ The next-step actions require a separate engine to predict and recommend. ▪ The contribution to business outcomes may take more time to realize. ▪ These use cases often lead to more distributed decision making within the enterprise. ▪ System availability and real-time performance are less important.	▪ Provision of model training using customers' own image data ▪ Complex solutioning that involves multiple data science building blocks ▪ Data integration and management ▪ Image data annotation ▪ Model drift and feedback management ▪ Data encryption and protection

Source: IDC, 2021

FEATURED VENDOR PROFILE

This section briefly explains IDC's key observations resulting in a vendor's position in the IDC MarketScape. Although every vendor is evaluated against each of the criteria outlined in the Appendix, the description here provides a summary of the vendor's strengths and opportunities.

AWS

AWS is positioned as a Leader in this IDC MarketScape: APEJ Vision AI Software Platform 2021 Vendor Assessment study.

AWS provides vision AI capabilities through Amazon Rekognition and industry solutions, such as Amazon Rekognition custom labels, Amazon Textract, AWS Panorama, and Amazon Lookout for Vision. They are referred to as "AWS products" for the rest of this document. AWS vision AI products were first released in 2016.

AWS products have been adopted by regional customers, including CoinJar (financial technology [fintech], Australia), Tape Ark (data management services, Australia), REA Asia (property management, Australia), Mindtickle (sales enablement platform, India), GE Healthcare Japan (healthcare, Japan), and Hyundai Motor Company (manufacturing, South Korea), among others. AWS also has customers in other global locations, such as Dafgards (consumer goods, Sweden).

Strengths

AWS' strengths include:

- **Core capability categories.** AWS plays in all three competitive spaces. Deployed use cases include quality and personal protective equipment (PPE) inspection for manufacturing, search for missing children, explicit content moderation for media and gaming, scene/activity detection for video editing, video content quality control, OCR for process automation, anomaly detection for field maintenance, defect classification for insurance, face-based authentication for customer onboarding, check-in and gig economy staff verification, retail goods recognition for inventory management, brand and logo recognition for campaign management, video and content sentiment analysis, among others.

- **Closely associated capabilities for productivity**. AWS offers Amazon Lookout for Vision as its quality inspection solution, in which feedback loops are designed to enable continued learning and model enhancement. Transfer learning–based few-shot learning allows companies to start their model training with fewer than 30 images. AWS Panorama is the company's video analytics platform for performing vision AI on IP cameras at the edge. AWS provides Amazon Rekognition Custom Labels to allow companies to use their own image data to train vision AI models. It is a specialized automated ML (AutoML) service for image contents, supported by automated image labeling, multilabeling, and object detection. Amazon Textract extracts typed and handwritten text in seven languages from structured text documents containing defined fields and tables.

- **Closely associated capabilities for experience.** AWS already offers hundreds of pretrained models to enable end-user experience–related use cases. The range grows as the company works backward with their existing customers to roll in requirements continuously on new objects to be detected. In-scene text detection, for example, already covers many of the regional languages of Asia/Pacific. AWS products are supported by a huge body of curated image and video assets to increase both the range and quality of models they create. With regard to data security and privacy protection, AWS products have been certified by SOC, Payment Card Industry (PCI), the Federal Risk and Authorization Management Program (FedRAMP), and Health Insurance Portability and Accountability Act (HIPAA).

- **Closely associated capabilities for decisioning.** Amazon SageMaker provides the necessary model-related adjacency services for decision-related use cases. It includes fully managed data labeling, model builder's studio, pretrained feature sets for jump-start training, and model life-cycle management as well as data and model parallelism to reduce the training time on the AWS cloud platform.

- **Regional availability and direct sales and service.** AWS products are provisioned on public cloud, private cloud, and edge devices, with minimal 99.9% availability. AWS has direct sales in most, if not all, APEJ locations. Regional field services are delivered across all key markets, including India, Japan, South Korea, Southeast Asia, China, Australia, and New Zealand. AWS has dedicated computer vision experts not only in the actual solutioning and implementation but also in early-stage engagement. AWS practices sales and service interlocks and tunes best practices around the world for regional realities.

- **Partner network.** Apart from direct sales, AWS products also go to the market with a wide variety of partners, such as 1CloudHub, Accenture, Aspire Systems, China Mobile, Capgemini, Cognizant, Deloitte, Dynatrace, Indus.AI, National Computer Systems (NCS), NTT, Pegasystems, Searce Technologies, Talend, and Tech Mahindra, among others. AWS plans to expand their partner network, especially in the category of OEM edge device and gateway providers.

- **Product strategy.** AWS' product development strategy is more aligned with requirements in the competitive spaces of **experience** and **productivity.** The company is working toward building solutions focused on completing face authentication workflows, expanding from moderation to

compliance, additional object and language types, multimodel analysis, latency reduction, and higher scalability.

- **Growth.** Data-wise, AWS, as the largest cloud platform provider, has the benefit of data gravity — the notion that large data sets and applications operating on these data sets tend to attract smaller data sets and relevant services and applications. AWS further leverages its Data Flywheel initiative, encouraging opted-in data to propel algorithm improvement. Engagement-wise, AWS runs immersion events, workshops, trainings, and certifications in many locations in Asia/Pacific to remain close to the developer community. Location-wise, AWS has planned additional regional launches of Amazon Rekognition across key Asian markets and has announced new regions in Melbourne (Australia), Hyderabad (India), New Zealand, and Indonesia to further bolster regional presence.

- **Research and innovation.** AWS has research and development (R&D) labs in APEJ locations, including China and India, with researchers focusing on computer vision. It has ongoing collaborations with institutions of higher learning such as Tsinghua University. AWS research papers addressing technical and ethical topics around vision AI appear regularly in conferences, such as the CVPR, International Conference on Computer Vision (ICCV), European Conference on Computer Vision (ECCV), Conference on Neural Information Processing Systems (NeurIPS), and International Conference on Learning Representations (ICLR).

Challenges

AWS products cover a wide variety of practical use cases. The company also grows regionalized capabilities, such as language coverages for Bahasa Indonesia and Vietnamese.

Considerations

Consider AWS when:

- There's a need to centrally plan vision AI capabilities in a large-scope initiative, such as digital transformation (DX).
- Best practices and knowledge sharing–related services are considered important.
- Composite capabilities for AI, data management, and cloud compute, with more availability of cloud regions are required.
- Flexible ways to control costs are required.

APPENDIX

Reading an IDC MarketScape Graph

For the purposes of this analysis, IDC divided potential key measures for success into two primary categories: capabilities and strategies.

Positioning on the y-axis reflects the vendor's current capabilities and menu of services and how well aligned the vendor is to customer needs. The capabilities category focuses on the capabilities of the company and product today, here and now. Under this category, IDC analysts will look at how well a vendor is building or delivering capabilities that enable it to execute its chosen strategy in the market.

Positioning on the x-axis or strategies axis indicates how well the vendor's future strategy aligns with what customers will require in three to five years. The strategies category focuses on high-level decisions and underlying assumptions about offerings, customer segments, and business and go-to-market (GTM) plans for the next three to five years.

The size of the individual vendor markers in the IDC MarketScape represent the market share of each individual vendor within the specific market segment being assessed. For this IDC MarketScape, vendor size was partially determined by IDC's 2020 Software Tracker, which represents an estimate of each vendor's 2020 software revenue. Because not all vendors have vision AI as an independent SKU to report revenue, the revenue from the general AI software platform is used as a proxy and subsequently put into four t-shirt size categories: very large, large, medium, and small.

IDC MarketScape Methodology

IDC MarketScape criteria selection, weightings, and vendor scores represent well-researched IDC judgment about the market and specific vendors. IDC analysts tailor the range of standard characteristics by which vendors are measured through structured discussions, surveys, and interviews with market leaders, participants, and end users. Market weightings are based on user interviews, buyer surveys, and the input of IDC experts in each market. IDC analysts base individual vendor scores and, ultimately, vendor positions on the IDC MarketScape on detailed surveys and interviews with the vendors, publicly available information, and end-user experiences in an effort to provide an accurate and consistent assessment of each vendor's characteristics, behavior, and capability.

Market Definition

The AI software services market is divided into three submarkets that handle different aspects of unstructured information analysis and processing: text and audio AI, image and video AI, and document understanding AI. An extended reading about the market definition can be found in *IDC's Worldwide Software Taxonomy, 2021* (IDC #US47588620, April 2021). This study looks primarily at the image and video AI submarket and, supplementally, the document understanding AI submarket using the term of vision AI. It includes computer vision technologies that use DL techniques to process, understand, and extract information from digital images or videos.

The elements of this submarket include scene reconstruction, object detection and recognition, pose and gesture recognition, image annotation, text extraction, video tracking, and so forth using DL neural network algorithms.

To cater to the actual delivery of vision AI software services, we have included the delivery of both pretrained vision models and customized models based on the user's own image sets. These models can then be deployed or integrated in other software solutions to process images and videos in the targeted workflows to detect, classify, or recognize content, object, face, or activity in the context of front-end user interaction, middle-desk analytics, or back-end metadata management. Vision AI services are usually provided as part of a general-purpose AI software platform. The output is primarily used by software developers in the form of APIs or system development kits (SDKs), allowing various types of integration and deployment.

Related Research

- *Tackling Artificial Intelligence Solutions' Trust Issues* (IDC #AP47386521, June 2021)
- *Asia/Pacific (Excluding Japan) Artificial Intelligence Applications Spending Forecast, 2019-2024* (IDC #AP46313020, May 2021)
- *IDC's Worldwide Software Taxonomy, 2021* (IDC #US47588620, April 2021)
- *IDC MaturityScape Benchmark: Artificial Intelligence in Asia/Pacific (Excluding Japan), 2021* (IDC #AP47369521, March 2021)

Synopsis

This IDC MarketScape provides an assessment of vision artificial intelligence (AI) software platform service vendors, covering multiple Asia/Pacific (excluding Japan) (APEJ) locations and discussing the important success factors for companies looking to leverage vision AI for improved productivity, user experience (UX), and decision recommendation.

"Vision AI is one of the emerging technology markets," says Christopher Lee Marshall, associate VP, AI and analytics strategies, IDC Asia/Pacific. "It is critical to watch major vendors and more mature market solutions because the early movers tend to consolidate their strengths with greater access to training data, more iterations of algorithm variations, deeper understanding of the operation contexts that are inclusive of constraints and noises, and more systematic approaches to work with solution partners in the ecosystem," he adds.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC Asia/Pacific Headquarters (Singapore)

83 Clemenceau Ave
17-01 UE Square, West Wing
Singapore 239920
65.6226.0330
Twitter: @IDC
idc-community.com
www.idc.com