IDC MarketScape

IDC MarketScape: Asia/Pacific (Excluding Japan) AI Life-Cycle Software Tools and Platforms 2022 Vendor Assessment

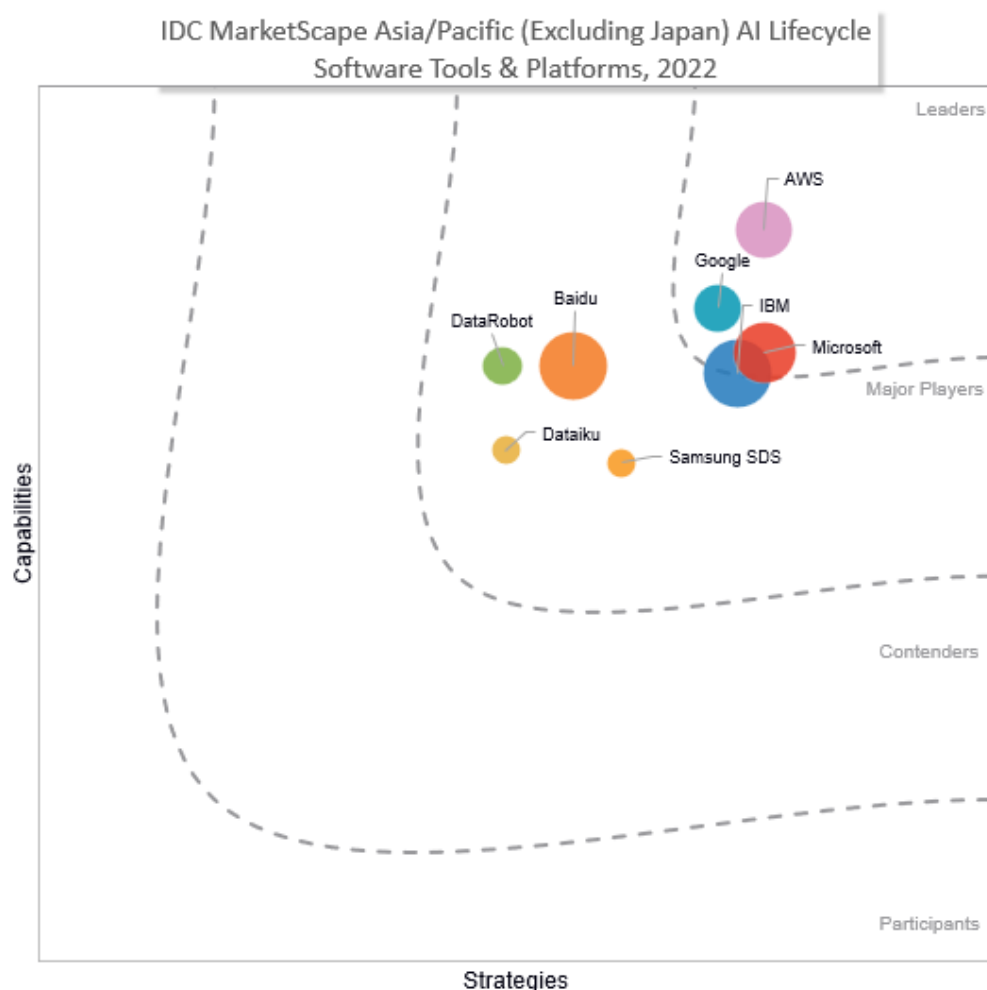Christopher Lee Marshall   Jessie Danqing Cai   Sandeep Sharma   Maxine Wong

THIS MARKETSCAPE EXCERPT FEATURES: AWS

IDC MARKETSCAPE FIGURE

FIGURE 1

Asia Pacific (Excluding Japan) AI Life-Cycle Software Tools & Platforms, 2022

Source: IDC, 2022

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

IN THIS EXCERPT

The content for this excerpt was taken directly IDC MarketScape: Asia/Pacific (Excluding Japan) AI Life-Cycle Software Tools and Platforms 2022 Vendor Assessment (Doc #AP48940522). 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 are Figure 1, Figure 2, and Figure 3.

IDC OPINION

According to *IDC's Industry AI Path Survey* and *AI Strategies BuyerView Survey*, both conducted in 2021, 26% of AI projects in Asia/Pacific companies have reached production stage, and among them, 63% need to retrain/rebuild models at least weekly. With more models being deployed within the typical enterprise, and with shorter update cycles, the demand for integrated and end-to-end model development and life-cycle management products will mushroom. Large enterprises and digital-native businesses in this region have already started to establish or improve the end-to-end life-cycle processes associated with their artificial Intelligence (AI) models — as illustrated below, to increase governance, productivity, and reproducibility of machine learning/deep learning (ML/DL) model delivery by improving coordination between different data, analytics, and business roles within the enterprise.

FIGURE 2

The Life Cycle of Operationalized ML Models

Note: CI/CD stands for continuous integration/continuous delivery.

Source: IDC, 2022

According to IDC's software taxonomy, the various software tools needed to support end-to-end AI model development, including data preparation, model building/training, model operation, evaluation, deployment, monitoring, and so forth, are categorized as AI life-cycle software tools. This report investigates AI life-cycle software tools vendor capabilities using IDC's MarketScape research model. For detailed definitions of each submarket and the assessment methodology, please refer to the Appendix. Our findings and conclusions are drawn based on:

- request for information (RFI) survey responses, product briefings, and interviews with buyer organizations conducted between May and August 2022;
- indicative market shares maintained in IDC's Semiannual Artificial Intelligence Tracker; and
- other IDC reports and surveys covering a similar scope.

This research suggests the following:

- Generally, there are three vendor segments active in supplying AI life-cycle capabilities in Asia/Pacific (excluding Japan) (APEJ): (1) the **hyperscalers**, such as Amazon Web Services (AWS) and Google, whose wide-ranging products tend to stay closely aligned with their open-source origins, providing greater engineering flexibility, with effective support of inference at scale and price for performance; they have more customers running on single/multicloud platforms; and they offer the benefits of data gravity and economies of scale to would-be users, (2) the **solution-builders**, such as IBM, Baidu, and Samsung SDS, whose products tend to align with often consulting-led digital transformation (DX) initiatives such as smart manufacturing and digital experience; they have more customers running on private/hybrid cloud; and they offer customized and verticalized solutions through their consulting and partner networks; (3) the **capability-enablers**, such as DataRobot and Dataiku, whose products tend to be easier to learn and use and have connectors to many other solutions, and they have more customers running on hybrid/multicloud. They offer clients fast time to value.
- In Asia/Pacific, the numbers suggest **hyperscalers** have been most successful in helping companies manage their ML initiatives. They have more customers actively productizing their own ML/DL models than the others, resulting in faster feature releases, and greater investment in staffing, community engagement, and skill enablement. Still, there are many larger organizations requiring the complex solutioning capabilities of the **solution-builders** or the quick leverage of scarce skills provided by the **capability-enablers**.
- In addition, all three of the above segments face increasing competition from two emerging vendor clusters: the **hardware-providers** and the **analytics-navigators**. Hardware-providers include companies, such as Intel, Nvidia, and HPE, which are commercializing their AI/ML software platform capabilities, targeting primarily technology decision makers and organizations running on private/hybrid cloud, or having large-scale edge deployment, with demanding hardware performance requirements. Analytics-navigators are typically providers of analytic tools, now offering features and modules beyond advanced analytics. Their value propositions tend to resonate well with business decision makers but not necessarily with IT. Many hardware-providers' and analytics-navigators' products are newly available for Asia/Pacific customers. As these are established names with sizable customer bases, we expect them to significantly influence the competitive landscape in the coming years.

FIGURE 3

The Five Vendor Segments
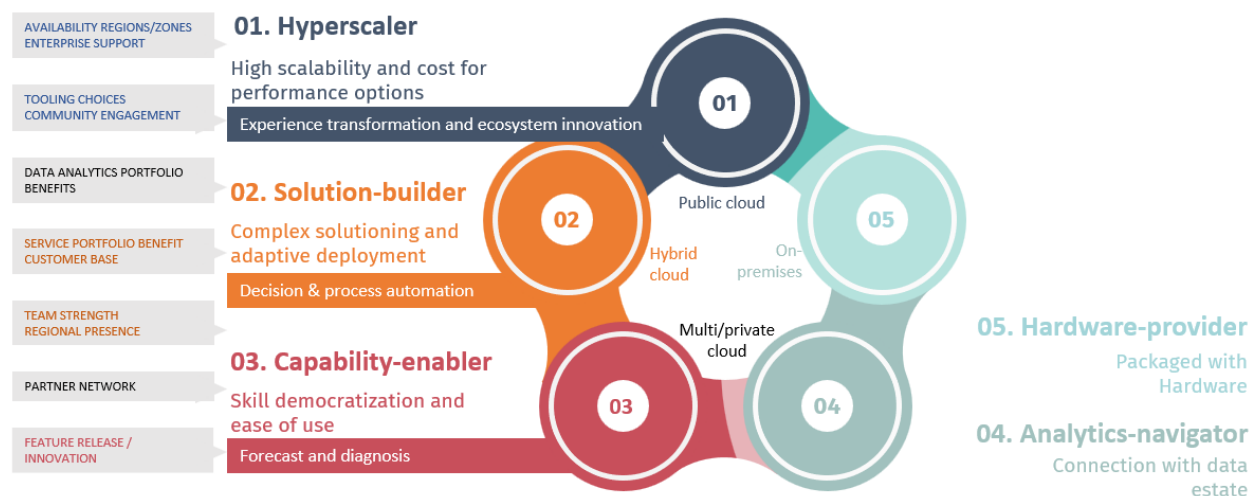
Source: IDC, 2022

IDC MARKETSCAPE VENDOR INCLUSION CRITERIA

According to IDC's software taxonomy, the AI life-cycle software market covers the tools and technologies used by data scientists and ML developers, from the experimentation to the production deployments of AI and ML solutions. It includes four named submarkets: data labeling software, AI model build (and training) software, machine learning operations (MLOps) software, and trustworthy AI software. Detailed definitions of these four submarkets can be found in *IDC's Worldwide Software Taxonomy, 2021* (IDC #US47588620, April 2021).

This report details the capabilities and strategies of major vendors that offer functions in the three named submarkets. Note that data labelling software is also an important submarket of AI model life-cycle software, but it is not included in this report as a primary subject to study, because it has greater dependency on manual work and expert supervision, aspects that are often idiosyncratic to particular end-users. In contrast, data preparation capabilities such as ingestion, integration, cleaning, and transformation, though not explicitly defined as a submarket, are included in this study.

We have tried to capture the most significant AI life-cycle players operating in the APEJ market, especially those based and operating in Asia/Pacific locations. The criteria for their inclusion are as follows:

- The vendor provides software products for various aspects of the end-to-end ML/DL life-cycle process under independent product stock-keeping units (SKUs) or as part of a general AI software platform.
- The products are based on the company's own IP.
- The products have generated software license revenue or consumption-based software revenue for at least 12 months in APEJ as of March 2022.

- For these products, the company is among the top 15 vendors by the reported revenues of 2020–2021 in the APEJ region, according to IDC's AI software tracker. Alternatively, the company can confirm that the revenue in this timeframe has exceeded US$10 million.

ADVICE FOR TECHNOLOGY BUYERS

- **Consider AI life-cycle capabilities as essential for an intelligent enterprise.** AI life-cycle capabilities will become increasingly important for all organizations aspiring to become data-driven in their decision making, operation automation, and problem solving. This is because ML/DL models are key enablers to translate company's data assets into actionable insights, and their use at scale inherently requires life-cycle management.

- **Assemble a multipersona team and define the AI life-cycle process in your own organization.** The development and management process of ML/DL models is both collaborative and iterative, as depicted in Figure 2. The process also varies in different organizations, depending on the range of targeted use cases, the platform, and the availability of internal skills. It is important to identify such a process with clearly defined roles when setting out to productize and operationalize AI, communicate widely within the enterprise, and update as needed.

- **Define your workload requirements and any dependencies.** Adding to the complexity of AI life-cycle management are the varied adoption requirements of the buyer organizations — some run all workloads in the same cloud platform, some train on-premises and then move them to cloud for inferencing, some face data sovereignty restrictions — just to name a few examples. Organizations must define these constraints explicitly and communicate clearly with the shortlisted vendors. Not all vendors can support all these requirements.

- **Maintain an elevated view on how models generate (or not) business value so as to inform improved collaboration and accelerate change management.** Change management is a common challenge associated with technology adoption, more so in established organizations and for investment in AI and ML, as the technology is inherently opaque, requiring continuous oversight from a team of specialists. Clarity of the business value, provided by the models deployed or the use cases implemented, helps prioritize efforts so that specialist teams can stay engaged and keep adding value.

- **Keep pace with updates on product and platform capabilities as they are fast evolving.** As vendors are quickly adding new features and launching new availability regions/zones, laundry lists to summarize solution features can become quickly outdated. Organizations should use this report as a start to understand vendor strengths and stay engaged with the selected vendors to get the latest updates.

- **Engage with your solution provider as a coinnovation partner.** Much of AI life-cycle functionality is rooted in open-source projects such as MLflow (experiment-centric automation, started by Databricks) and Kubeflow (deployment-centric automation, started by Google). The limited documentation and technical support of open-source tools means that vendors can provide huge value, helping data scientists and ML engineers make the best use of all the tools available, and buyer organizations can actively influence the product development road map.

- **Leverage AutoML functionalities without losing sight of cost control.** Skill democratization tools such as automated machine learning (AutoML) and solution prebuilds, previously targeted toward ML neophytes, are now increasingly used by experienced data scientists for feature and model selection, architecture search, hyperparameter tuning, and so forth. Organizations should also take note that when such tools are in the hands of less experienced or

nontechnical users, the model life-cycle process might require more iterations than necessary, which can be expensive given the metered workloads for deep learning.

- **Consider appointing a governance steward or model risk manager**. Model trustworthiness features, such as for explainability, bias, and robustness, are increasingly required for critical customer-facing, revenue-generating, or regulated use cases. The challenge is to make this practical from the business user's perspective. For organizations using more than, say hundreds of ML/DL models for different business critical operations, it makes sense to appoint a governance steward or model risk manager (perhaps part of a larger IT risk team) to oversee this area and ensure a disciplined approach.

FEATURED VENDOR PROFILE

This section briefly explains IDC's key observations resulting in a vendor's position in the IDC MarketScape, and provides a summary of the vendor's strengths and opportunities.

AWS

AWS is positioned in the Leaders category in the 2022 IDC MarketScape for APEJ AI life-cycle software tools and platforms.

AWS' AI life-cycle software solution is branded as Amazon SageMaker. It provides end-to-end fully managed services for data preparation, model building, training, evaluation, deployment, MLOps, and trustworthiness. Amazon SageMaker consists of more than a dozen tools such as SageMaker Data Wrangler, SageMaker Studio, SageMaker Autopilot, SageMaker Canvas, and SageMaker Clarify — just to name a few. Amazon SageMaker was first released in 2017 and continues to add incremental features at a rapid pace based on customer requirements.

Amazon SageMaker's support of the life-cycle processes of organizations' ML and DL models can be summarized as follows:

- **For data preparation and model building/training.** SageMaker Data Wrangler supports data preparation as an integrated part of SageMaker Studio and can be composed into SageMaker Pipelines. Data scientists can choose the JupyterLab-based SageMaker Studio Notebooks for code-based data preparation, model building, and model training. Other options include SageMaker Autopilot powered by AutoML, the low-code/no-code approach via SageMaker Canvas, and 100+ prebuilds for use case–specific model/solution templates packaged in SageMaker JumpStart.
- **For model operations and trustworthiness**. SageMaker Model supports easy and scalable deployment into any endpoints based on Kubernetes. MLOps capabilities such as model registry, monitoring, and pipelines are provided in SageMaker Projects, SageMaker Model Monitor, and SageMaker Pipelines, respectively. SageMaker Clarify provides trustworthy capabilities such as model explainability, data bias, model bias, model drift, model robustness against adversarial attacks, and compliance report generation. It works within SageMaker Studio and is integrated with other SageMaker tooling components throughout the model life-cycle process.

Amazon SageMaker is the fastest growing product ever launched on the AWS cloud platform. It has been adopted by an array of enterprises in Asia/Pacific, including the Australian Bureau of Statistics (government organization, Australia), New South Wales Health (government organization, Australia), Pepperstone (Forex trading platform, Australia), the Bali State Government (government organization,

Indonesia), Carro SG (ecommerce, Singapore), DBS Bank (financial service, Singapore), PropertyGuru (property portal, Singapore), Hyundai (automobile manufacturing, South Korea), OPPO (manufacturing, China), Godrej Consumer Products (consumer packaged goods [CPG], India), and so forth.

Strengths

- **Functionality and offering.** Amazon SageMaker provides a rich set of tooling options for data preparation, model training, and deployment. The range of data connectors and ease of integration with other AWS offerings have received more positive mentions from customers. Amazon SageMaker offers multiple model inference options such as real time, serverless, asynchronous, and batch transform, to flexibly meet customers' scalability and cost management needs. Cost for performance options of Amazon SageMaker is supported by AWS' own silicon — AWS Inferentia for inference workloads and AWS Trainium for training workloads. SageMaker Clarify supports model explainability for real-time model inferencing jobs. Proprietary techniques 10x faster than open source, are applied to produce Shapley values for individual predictions.

- **Service delivery**. Amazon SageMaker is natively available on the AWS cloud platform — the second largest public cloud platform in the APEJ region (based on IDC Public Cloud Services Tracker, IaaS+PaaS, 2021 data), with availability regions in Japan, Australia, New Zealand, Singapore, India, Indonesia, South Korea, and Greater China. Local zones are available to serve customers in ASEAN countries — Thailand, the Philippines, and Vietnam. AWS has team presence in most, if not all Asia/Pacific countries for this product portfolio, and the quality of service provided has received many positive mentions.

- **Growth opportunities**. AWS actively contributes to open-source projects such as Gluon (deep learning framework, peers of PyTorch and TensorFlow), and engages with regional developer and student communities through many events, online courses, and the offering of SageMaker Studio Lab — a no-cost SageMaker notebook environment. Since 2018, the company has started to invest in a Shanghai-based AI lab. More importantly, AWS is the largest provider of cloud analytic data platform solutions in the Asia/Pacific region (based on IDC Big Data and Analytics Software Tracker, 2021 data). The company is known for pretrained models for vision AI and document AI in the Asia/Pacific region.

Challenges

- **Functionality and offering.** AWS has received positive mentions on the quality of its documentation; still, some customers have commented on the steep learning curve required to make the best use of all the tooling options of SageMaker, and, in addition, the need for native support of RStudio, and the need for enhanced access control/management.

- **Service delivery.** The on-demand instances of the fully managed service can have occasional performance issues. In the Asia/Pacific region, there is a rising demand to accommodate data sovereignty requirements. Thus, AWS has planned to launch new instance types to further increase availability and increase price for performance options.

Consider AWS When

IDC sees AWS as a vendor playing primarily in the space of hyperscalers. AWS has been rolling out new functions and features in its Amazon SageMaker suite quickly in recent years. Apart from being the preferred choice for customers managing data analytic workloads on the AWS cloud platform, Amazon SageMaker is also recommended to organizations having a builder/developer culture, those

working to scale ML/DL models with a relatively small team or lowered cost of operation, and those that consider the quality and timeliness of technical support to be very important.

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/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 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 mostly determined by IDC's Semiannual Artificial Intelligence Tracker 2H21, which provides an estimate of each vendor's 2021 software revenue parked under the technology category of AI life-cycle software; and subsequently put into four T-shirt size categories: very large (Baidu, IBM, Microsoft), large (AWS, Google), medium (DataRobot), and small (Dataiku, Samsung SDS).

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

According to *IDC's Worldwide Software Taxonomy* (IDC #US48990921), AI life-cycle software tools are used by data scientists, machine learning developers, and business analysts to experiment, deploy, productize, monitor, and update their own AI and ML solutions. This report covers three named submarkets: AI build software, MLOps software, and trustworthy AI software. The following are quoted from the taxonomy:

- **AI build software** facilitates the ability to build, train, and tune advanced machine learning models. It typically includes prebuilt algorithms and models that data scientists and machine learning developers could use as a starting point to customize and build their own high-quality models.

- **Machine learning operations (MLOps) software** supports model deployment and model management, including the monitoring of data and concept drift.
- **Trustworthy AI software** includes the tools used for model validation and assessment, and for ensuring safety and security of the models. They support the foundational elements of trust for AI models.

LEARN MORE

Related Research

- *Artificial Intelligence in India: Adoption So Far, Use Cases by Verticals and Key Case Studies* (IDC #AP48500722, August 2022)
- *Asia/Pacific (Excluding Japan) Artificial Intelligence Life-Cycle Software Forecast, 2020–2025* (IDC #AP48486422, May 2022)
- *IDC Worldwide Software Taxonomy, 2022* (IDC #US48990921, April 2022)
- *IDC MarketScape: Worldwide Advanced Machine Learning Software Platforms 2020 Vendor Assessment* (IDC #US45358820, October 2020)
- *IDC MarketScape: China Machine Learning Development Platform, 2019* (IDC #CHC45389019, August 2019)

Synopsis

This IDC study provides an assessment of AI life-cycle software vendors covering the Asia/Pacific (excluding Japan) (APEJ) region.

"As companies get beyond the experiment stage for artificial intelligence (AI) and data science initiatives, AI life-cycle technologies become essential parts of any intelligent enterprise," says Dr. Chris Marshall, vice president, Artificial Intelligence and Analytics Strategies at IDC Asia/Pacific.

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