Hype Cycle for Data, Analytics and AI in China, 2023

Published 17 July 2023 - ID G00760832 - 99 min read

By Analyst(s): Julian Sun, Ben Yan, Xingyu Gu, Fay Fei, Mike Fang, Tong Zhang

Initiatives: Digital Technology Leadership for CIOs in China; Analytics, BI and Data Science Solutions; Data and Analytics Programs and Practices; Data Management Solutions

Data and AI are crucial to China's digital economy and national strategy, with a unique, value-adding position supported by regulatory frameworks. Data and analytics leaders in China must understand the hype and reality of the regional D&A and AI ecosystem to make progress toward business outcomes.

More on This Topic
This is part of an in-depth collection of research. See the collection:

- 2023 Hype Cycles: Deglobalization, AI at the Cusp and Operational Sustainability

Strategic Planning Assumption

By 2026, more than 30% of white-collar jobs in China will be redefined, and leveraging and managing generative AI will become a sought-after skill.

Analysis

What You Need to Know

Data, analytics and AI in China exhibit many similarities with the global market, but also possess unique differentiations in organizational structure, technology focus and value proposition. In the digital economy era, data, analytics and AI serve as the foundation of every organization's strategy to drive outcome-first investment. Today, data and analytics (D&A) industry best practices tend to align with China's regulations on data, privacy and AI — even more so than those set in the U.S. or the EU.

Despite numerous inflated expectations within this space, there are very few D&A and AI innovations that have reached a plateau of technology maturity. This Hype Cycle primarily focuses on the hype surrounding emerging data, analytics and AI technologies and techniques that have varying degrees of commoditization. It also addresses the operationalization of these techniques to create systems that transcend everyday D&A and AI, as well as the impact of these innovations on people and processes within and beyond an enterprise context.

D&A leaders in China must leverage this research to understand and utilize technologies that offer high impact in the present and prepare their strategy for the future.

In addition to this Hype Cycle, D&A leaders should consult the following Hype Cycles in adjacent areas:

- Hype Cycle for ICT in China, 2023
- Hype Cycle for Smart City and Sustainability in China, 2023
- Hype Cycle for Security in China, 2022

Together, these four Hype Cycles analyze the elements required for technology CIOs in China to form a holistic view of the data and analytics ecosystem.

The Hype Cycle

Data, analytics and AI have consistently been identified as the top investment priorities for China's CIOs and are at the forefront of the national strategy supported by the Chinese government. Meanwhile, China's regulations on data, privacy, and AI are causing concerns among technology executives across the spectrum.

The most crowded part of the Hype Cycle is toward the Peak of Inflated Expectations. Innovations are often hyped as solutions to traditional bottlenecks. The expectation is that they will demonstrate clear business value by addressing hardware resource scarcity, scalability, sustainable operationalization, security risk mitigation, technology self-sufficiency and multidomain applicability of AI models — all common concerns of CIOs in China. From the end-user perspective, there is a heightened emphasis on tangible impact rather than abstract strategic concepts.

Amid the hype:

- The concept of data middle office — a highly touted practice among Chinese organizations — has fallen into the Trough of Disillusionment. Many organizations and vendors are either shying away from adopting this concept internally or removing it from their branding altogether.
- Data asset management is a transformative and distinctive practice that anchors the digital economy in Chinese organizations, enabling them to effectively articulate and deliver the value of D&A initiatives.
- The Chinese market exhibits a particular enthusiasm for generative AI and large language models (LLMs). However, end users exhibit relatively lower confidence in the model-centric approach to homegrown LLMs while favoring data-centric or application-centric approaches for adopting generative AI. This is particularly due to China's lagging ability to produce homegrown AI chips quickly, and Chinese companies remain in the early stages of designing AI chips compared with global vendors such as NVIDIA.

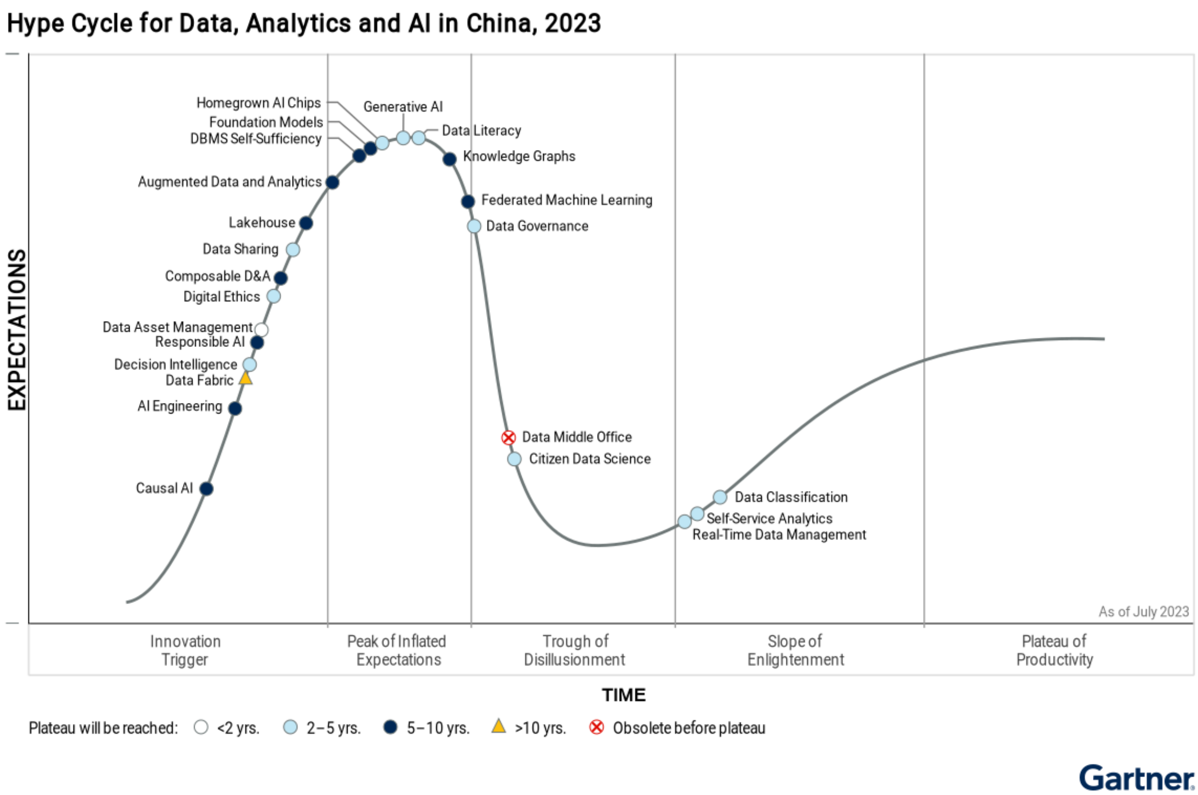
The innovations featured on the Hype Cycle demonstrate a converged and composable data, analytics and AI ecosystem in China, categorized into four main areas：

- **Analytics and data science** includes augmented data and analytics, citizen data science, composable D&A, data middle office, federated machine learning, knowledge graphs and self-service analytics.
- **Artificial Intelligence** includes AI engineering, causal AI, decision intelligence, foundational models, generative AI, homegrown AI chips and responsible AI.
- **Data management** includes data classification, data fabric, DBMS self-sufficiency, lakehouse and real-time data management.

- **Data and analytics programs and practices** include data asset management, data governance, data literacy, data sharing and digital ethics.

This year marks the inaugural release of Gartner's Hype Cycle for Data, Analytics and AI in China, where we anticipate a plethora of rapid innovations and failures with immense potential. Among them, key innovations in China include data middle office, locally developed AI chips, self-sufficient DBMS and data asset management.

Figure 1: Hype Cycle for Data, Analytics and AI in China, 2023



The Priority Matrix

Data asset management is the only innovation profiled here that has a transformational benefit rating and can achieve mainstream adoption in less than two years. It is bolstered by the Chinese government's digital economy ambitions and corresponding policies, including the establishment of the National Data Bureau. This will impact data and analytics leaders, CIOs and other technology executives in China, as data has been recognized as the fifth production material. [1] Chinese organizations thus have a unique advantage, since no other nation in the world yet aligns government regulation to what drives business outcomes.

Innovations that warrant particular attention in the next two to five years for mainstream adoption include citizen data science, decision intelligence, generative AI and real-time data management. Early adoption of these innovations will confer significant competitive advantage and mitigate issues stemming from a dearth of business technologists as well as challenges associated with capturing business context and value for outcome.

Several innovations have a five to 10-year period to mainstream adoption, but their application can be started with small-scale projects to deliver immediate impact. These include composable D&A, data sharing, data fabric, foundation models and responsible AI. The implementation of these innovations is crucial in dispelling the myth of middle office and establishing operational AI systems to move from innovation to production. Data and analytics leaders should accelerate the implementation of data classification and data literacy as a long-term strategy to establish a solid foundation for their AI ambitions given that talent or service provider shortage is the biggest obstacle hindering AI adoption, according to Gartner's 2022 China AI Use-Case Survey. [2]

Table 1: Priority Matrix for Data, Analytics and AI in China

(Enlarged table in Appendix)

Benefit	*Years to Mainstream Adoption*			
↓	***Less Than 2 Years*** ↓	***2 - 5 Years*** ↓	***5 - 10 Years*** ↓	***More Than 10 Years*** ↓
Transformational	Data Asset Management	Citizen Data Science Data Sharing Decision Intelligence Generative AI Real-Time Data Management	Composable D&A Foundation Models Responsible AI	Data Fabric
High		Data Classification Data Governance Data Literacy Digital Ethics Homegrown AI Chips	AI Engineering Causal AI DBMS Self-Sufficiency Federated Machine Learning Knowledge Graphs	
Moderate		Self-Service Analytics	Augmented Data and Analytics Lakehouse	
Low				

Source: Gartner (July 2023)

On the Rise

Causal AI

Analysis By: Ben Yan, Julian Sun

Benefit Rating: High

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Definition:

Causal artificial intelligence (AI) identifies and utilizes cause-and-effect relationships to go beyond correlation-based predictive models and toward AI systems that can prescribe actions more effectively and act more autonomously. It includes different techniques, like causal graphs and simulation, that help uncover causal relationships to improve decision making.

Why This Is Important

AI's ultimate value comes from helping us take better actions. Machine learning (ML) makes predictions based on statistical relationships (correlations), regardless of whether these are causal. This approach is fine for prediction, but predicting an outcome is not the same as understanding what causes it and how to improve it. Causal AI is crucial when it comes to determining the optimal actions to take to affect specific outcomes. Causal AI techniques help make AI more autonomous, explainable, robust and efficient.

Business Impact

Causal AI leads to:

- Greater decision augmentation and autonomy in AI systems by estimating the intervention effects.
- Greater efficiency by adding domain knowledge to bootstrap causal AI models with smaller datasets.
- Better explainability by capturing easy to interpret cause-and-effect relationships.
- More robustness and adaptabilty by leveraging causal relationships that remain valid in changing environments.

- The ability to extract causal knowledge with less costly and time-consuming experiments.

Drivers

- **The ongoing shift from predictive analytics (what is likely to happen?) toward more prescriptive analytics (what should be done?).** Making accurate predictions will remain key, but a causal understanding of how to affect predicted outcomes will be increasingly important.
- **AI systems increasingly need to act autonomously to generate business value,** particularly for time-sensitive and complex use cases where human intervention is not feasible. This will only be possible by AI understanding the effect of actions and how to make effective interventions.
- **Limited availability of data for certain use cases is pushing organizations toward more data efficient techniques like causal AI,** which leverages human domain knowledge of cause-and-effect relationships to bootstrap AI models in small data situations.
- **The growing complexity of use cases and environments where AI is applied requires more robust AI techniques.** Causal structure changes much more slowly than statistical correlations, which makes causal AI more robust and adaptable in fast-changing environments. The volatility in the last few years has exposed the brittleness of correlation-based AI models across industries that struggled to adapt since they had been trained under a very different context.
- **The need for greater trust and explainability of AI models is increasing the interest in models that are more intuitive to humans.** Causal AI techniques, like causal graphs, allow us to be explicit about causes and explain models in terms that humans understand.

Obstacles

- **Causality is not trivial.** Not every phenomenon is easy to model in terms of its causes and effects. Causality might be unknown, regardless of AI use.
- **The quality of the causal AI models depends on their causal assumptions and on the data that is used to build them,** which are susceptible to bias and imbalance. Just because a model is causal doesn't mean that it will outperform correlation-based ones.

- **Causal AI requires technical and domain expertise** to set up for properly estimating causal effects. This is often a more difficult exercise than building correlation-based predictive models and requires active collaboration between domain experts and AI experts.

- **AI experts might be unaware of causality methods** and may be overly reliant on data-driven models like ML, which could result in pushback when looking to implement causal AI.

- **A nascent vendor landscape and current low enterprise adoption could represent a challenge** when running initial causal AI pilots and identifying specific use cases where causal AI is most relevant.

User Recommendations

- Acknowledge the limitations that come from the prevalent approach to correlation-based AI and ML, which focuses on leveraging correlations and mostly ignores causality.

- Use causal AI when you need more augmentation and automation in decision intelligence, when AI is needed not only to generate predictions but to understand how to affect the predicted outcomes. Examples include customer retention programs, marketing campaign allocation and financial portfolio optimization.

- Select different causal AI techniques depending on the complexity of the specific use case. These include causal rules, causal graphs and Bayesian networks, simulation and the use of machine learning for causal learning.

- Conduct studies on where causal AI can provide a distinct advantage over classical correlational AI by prioritizing the highest impact ones.

- Educate your data science teams on causal AI, how it differs from correlation-based AI and the range of techniques available to incorporate causality.

Sample Vendors

Causality Link; DataCanvas; Huawei; IBM; Lingxi Technology; Microsoft

Gartner Recommended Reading

Innovation Insight: Causal AI

Case Study: Causal AI to Maximize the Efficiency of Business Investments (HDFC Bank)

AI Engineering

Analysis By: Mike Fang, Tong Zhang

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Embryonic

Definition:

AI engineering is foundational for enterprise delivery of AI solutions at scale. The discipline unifies DataOps, MLOps and DevOps pipelines to create coherent enterprise development, delivery (hybrid, multicloud, edge), and operational (streaming, batch) AI-based systems.

Why This Is Important

The value of fully industrialized AI lies in the ability to rapidly develop, deploy, adapt and maintain various AI models (statistical, machine learning, deep learning, graph, linguistic and rule-based) across different environments within the enterprise. To scale AI initiatives, organizations must adopt and implement pipelines using AI engineering practices. These practices are critical given the complexity of AI engineering and the demand for faster time to market.

Business Impact

Despite hundreds of AI pilots, many organizations struggle to move AI models from pilot to production. Most organizations put only a small number of AI models into production because of significant bottlenecks in the process. AI engineering gives organizations a framework to address bottlenecks. With AI engineering methods — DataOps, ModelOps and DevOps — it is possible to readily deploy models into production in a structured, repeatable factory-model framework to realize significant value.

Drivers

- Promote the “middle office” concept. “Middle office,” a unique term used in China, refers to an organizational strategy for building reusable capabilities. The goal of this strategy is to scale up and achieve business agility. “AI middle office” has emerged from many solution sales to support the adoption of AI engineering.
- Eliminate traditional siloed approaches to data management.
- Reduce impedance mismatch across data ingestion, processing, model engineering and deployment, which can drift once the AI models are in production.
- Accelerate the productionalization of AI by standardizing, governing and automating data, model and deployment pipelines.
- Adopt DataOps, ModelOps, DevOps and other best practices for enterprise development and deployment of AI platforms.
- Enable discoverable, composable and reusable AI artifacts (data catalogs, feature stores, model stores) across the enterprise context.
- Scale AI initiatives by enabling orchestration across hybrid, multicloud, edge AI or IoT.
- Operationalize AI architectures by bringing together data engineering, data science, application development, security and platform infrastructure teams.
- Broaden the use of foundational platforms that leverage existing data, analytics and governance frameworks to scale the production of AI initiatives.

Obstacles

- Chinese enterprises often seek "unicorn" experts to productize AI platforms. Few vendors in China could offer strong AI engineering capabilities compared with global vendors. Enterprises often have to build and support these environments via open-source architecture. Although a quick win, this approach can lead to long-term scalability and maintainability issues.
- AI engineering requires simultaneous development of pipelines across domains.
- AI engineering requires integrating full-featured solutions with select tools, including open-source technologies, to address enterprise architecture gaps with minimal functional overlap. These include capability gaps around ETL stores, feature stores, model stores, model monitoring, pipeline observability and governance.
- AI engineering requires a high degree of cloud maturity and possible rearchitecting, or the ability to integrate data and AI model pipelines across deployment contexts. Potential complexity and management of analytical and AI workloads alongside costs may deter organizations that are in the initial phases of their AI initiatives.

User Recommendations

- Maximize business value from AI initiatives by adopting an "AI middle office" platform. Follow AI engineering practices to streamline the data, model and implementation pipelines.
- Simplify data and analytics pipelines by identifying the key capabilities required to operationalize end-to-end AI platforms.
- Avoid building patchwork AI engineering pipelines that integrate piecemeal functionality. Use point solutions sparingly and only to plug feature/capability gaps in fully featured DataOps, MLOps, ModelOps and PlatformOps tools.
- Develop AI model management and governance practices that align model performance, human behavior and delivery of business value. Make it easier for business users to adopt and act on AI models by incorporating stakeholder trust and speed to value as primary inputs for model design.
- Leverage cloud service provider environments to build AI engineering. At the same time, rationalize your data, analytics and AI portfolios as you migrate to the cloud.
- Upskill data engineering and platform engineering teams to adopt tools and processes that drive CI/CD for AI artifacts.

Sample Vendors

4Paradigm; Amazon Web Services; Baidu; IBM; Microsoft

Gartner Recommended Reading

Market Guide for AI Software, China

Top Strategic Technology Trends for 2022: AI Engineering

Chinese AI Survey Analysis: AI Trends Wave 3.0 — From Operational to Strategic

Go Beyond Machine Learning and Leverage Other AI Approaches

Data Fabric

Analysis By: Xingyu Gu, Fay Fei

Benefit Rating: Transformational

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Definition:

A data fabric is a design framework for attaining flexible and reusable data pipelines, services, and semantics. The fabric leverages data integration, active metadata, knowledge graphs, profiling, machine learning and data cataloging. Fabric overturns the dominant approach to data management, which is “build to suit” for data and use cases, and replaces it with “observe and leverage.”

Why This Is Important

The emergence of data, analytics and AI use cases, coupled with highly dynamic data security regulations, has led to data management complexity and uncertainty in China. Data fabric capitalizes on sunk costs while simultaneously providing prioritization and cost control guidance for new spending on the data management infrastructure. It offers infrastructure flexibility, scalability and extensibility, enabling humans or machines to ensure that data is consumable across multiple use cases and environments.

Business Impact

Benefits of data fabric include:

- Provides insights to data engineers, ultimately automating repeatable tasks in data integration, quality, delivery and access
- Enables resources to better integrate, find, govern and share data across a multicloud and hybrid infrastructure
- Adds semantic knowledge for context and meaning, and enriches data models
- Evolves into a self-learning model that recognizes similar data content regardless of form and structure, enabling broader connectivity to new assets
- Monitors data assets for optimization and cost control

Drivers

- Organizations have found that existing approaches to data acquisition and integration are insufficient. Data fabrics can deliver integrated data through a broad range of styles, including bulk/batch (ETL), data virtualization, message queues, APIs and microservices.
- Chinese organizations are struggling to boost data management efficiency across a heterogeneous environment spanning on-premises, multicloud and hybrid ecosystems. Data fabrics ensure that these organizations can integrate, govern and share data without ripping and replacing existing infrastructure.
- For most organizations in China, the accumulation of different types of data, coupled with AI technology adoption, has led to a complex data and analytics (D&A) architecture with data silos. This complexity is fueling high demand for more automated operations in D&A architecture.
- Inefficient cross-department communication is still the main obstacle to realizing persistent business value from D&A investments. The active metadata aspects of data fabric provide additional insights into data usage and patterns, thereby improving communication, insights and data sharing across data silos.
- The increasing demand for data tracking, auditing, monitoring, reporting and evaluation is inflating manual workloads, impeding both cost-effectiveness and D&A agility. A more automated data management concept is necessary.
- New technology advancements enable data fabrics to assist with graph data modeling (which is useful to preserve the context of the data along with its complex relationships). These advancements also afford business users greater access to enrich analytics models with agreed-upon semantics.

Obstacles

- Proprietary metadata restrictions hamper the data fabric, as it is wholly dependent upon acquiring and sharing metadata from a variety of applications and sources, such as log files, trace files, application logs and scripts. Even though the data fabric supports analytics and machine learning capabilities to analyze design-time and runtime metadata, such capabilities will be error-prone at first.
- China has a limited number of local D&A vendors with necessary technologies, such as active metadata and knowledge graph, embedded in their products, creating a gap between concept and successful practice of data fabric.
- The diversity of skills and platforms to build a data fabric presents technical and cultural barriers. The basis of data management must shift from analysis, requirements and design to discovery, response and recommendation.
- Misunderstandings and knowledge gaps about how to reconcile data fabric with legacy D&A governance programs will complicate implementations.

User Recommendations

- Understand and prioritize “active metadata” efforts in all data management initiatives. Active metadata is the core capability and differentiator of data fabric.
- Evaluate how vendors share and utilize internal and external metadata to automate operations among adjacent applications in terms of resource efficiency, performance, security, compliance and usability. Vendors that heavily rely on manual efforts, such as on-site technicians, for the tasks above should be deprioritized.
- Deprioritize SaaS solutions that isolate their metadata from access by adjacent PaaS/SaaS solutions that orchestrate across solutions.
- Invest in an augmented data catalog that assists with creating a flexible data model. Enrich the model through semantics and ontologies so that the business can use the catalog.
- Deploy data fabrics that populate and utilize knowledge graphs.
- Ensure that business process experts can support the fabric by enriching knowledge graph capabilities with business semantics.

Sample Vendors

Alibaba Cloud; Datablau; Denodo; ESENSOFT; IBM; JKStack

Gartner Recommended Reading

Quick Answer: What Is Data Fabric Design?

Emerging Technologies: Critical Insights on Data Fabric

Case Study: An Active Metadata Augmented Data Classification System to Boost Analytics Efficiency

Quick Answer: What Is Active Metadata?

From Logical Data Warehouse to Data Fabric

Decision Intelligence

Analysis By: Ben Yan, Mike Fang

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Definition:

Decision intelligence (DI) is a practical discipline used to improve decision making by explicitly understanding and engineering how decisions are made and how outcomes are evaluated, managed and improved via feedback.

Why This Is Important

The current hype around automated decision-making and augmented intelligence, fueled by AI techniques in decision making, is pushing DI toward the Peak of Inflated Expectations. Recent crises have revealed the brittleness of business processes; rebuilding those processes to be resilient, adaptable and flexible will require the discipline brought by DI methods and techniques. A fast-emerging market around various software disciplines is starting to provide answers for decision makers.

Business Impact

DI has the following business impacts on organizations:

- **Materialize data-driven decision making at all levels within organizations.** Leverage decision support, augmentation and automation to optimize operational/tactical/strategic decision-making outcomes.
- **Increase decision-making visibility and reduce unpredictability.** Visualize the organizations' decision models, making decisions more transparent and auditable. Capture the uncertainty factors in the business context.

Drivers

- **A dynamic and complex business environment, with an increasingly unpredictable and uncertain pace of business.** The combination of AI techniques and the confluence of several technology clusters around composite AI, smart business processes, decision management and advanced personalization platforms are creating a new market around decision systems platforms supporting the DI discipline.
- **Need to curtail unstructured, ad hoc decisions that are siloed and disjointed.** Often uncoordinated, such decisions promote local optimizations at the expense of global efficiency; this happens both from an IT and a business perspective.
- **Expanding collaboration between humans and machines,** supplemented by a lack of trust in technologies (such as AI) increasingly replacing tasks and promoting uneasiness from a human perspective. DI practices promote transparency, interpretability, fairness, reliability and accountability of decision models critical for the adoption of business-differentiating techniques.
- **Tighter regulations are making risk management more prevalent.** From privacy and ethical guidelines to new laws and government mandates, it is becoming difficult for organizations to fully understand the risk impacts of their decisions. DI enables an explicit representation of decision models, reducing this risk.
- **Uncertainty regarding decision consistency across the organization.** The lack of explicit representation of decisions prevents proper harmonization of collective decision outcomes. This is remedied by DI.
- **The emergence of software tools in the form of decision intelligence platforms (DIPs)** will enable organizations to practically implement DI projects and strategies.

Obstacles

- **Fragmentation:** Decision-making silos have created data, competencies and technology clusters that are difficult to reconcile and could slow down the implementation of decision models.
- **Operational structure:** An inadequate organizational structure around advanced techniques, such as the lack of an AI center of excellence, could impair DI progress.
- **Reputation-damaging outcomes** from autonomous decision models (from embedded analytical assets to self-contained machine agents) and the failure to understand their collective impact impede DI adoption.
- **Lack of proper coordination between business units** and inability to impartially reconsider critical decision flows within and across departments diminish the effectiveness of early DI efforts.
- **Lack of modeling in a wider context:** For organizations that have focused almost exclusively on technical skills, the other critical parts of human decision making — psychological, social, economic and organizational factors — have gone unaddressed.

User Recommendations

- **Promote the resiliency and sustainability of cross-organizational decisions** by building models using principles aimed at enhancing traceability, replicability, pertinence and trustworthiness.
- **Improve the predictability and alignment of decision agents** by simulating their collective behavior, while also estimating their global contribution versus local optimization.
- **Develop staff expertise** in traditional and emerging decision augmentation and decision automation techniques, including descriptive, diagnostic, predictive and prescriptive (optimization, business rules) analytics.
- **Tailor the choice of decision-making technique** to the particular requirements of each decision situation by collaborating with subject matter experts, AI experts and business process analysts.

Sample Vendors

4Paradigm; Alibaba Cloud; Cardinal Operations; DataCanvas; Percent; SF Technology; Trusfort; Youhualin Information Technology; Zhongke Wenge

Gartner Recommended Reading

Innovation Insight for Decision Intelligence Platforms

Reengineer Your Decision-Making Processes for More Relevant, Transparent and Resilient Outcomes

How to Choose Your Best-Fit Decision Management Suite Vendor

Data Asset Management

Analysis By: Tong Zhang, Julian Sun

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Definition:

Data asset management is the process of managing, processing and utilizing data that has already been identified as a valuable asset to business operations. It applies to multiple data models, such as digital images, videos, files, documents and transactional data in systems. This discipline spans the entire data life cycle, from acquisition to disposal. Its purpose is to treat datasets like assets and generate value from them.

Why This Is Important

Data, as a new production factor, becomes a competitive edge for organizations. The velocity, variety, volume and veracity of data require a streamlined process to generate insight from it. Data assets add business value beyond operational excellence and decision making. New business models and direct monetization could be achieved as well. However, despite accelerating value generation, data assets have embedded risks. Organizations must carefully manage data assets to avoid regulatory violations and accidental data breaches.

Business Impact

Data is recognized as an intangible asset that can benefit business decision making, business model design, operational efficiency and cost reduction. However, becoming a data-driven enterprise is not just about encouraging the use of data in decision making. To establish data asset management, organizations must develop competencies and align work with the enterprise's goal of generating information value. Data could have a direct impact on financial statements (if recognized as an intangible asset on the balance sheet), making data asset management crucial.

Drivers

- Direct data monetization requires data asset management to lay a good foundation for value generation.
- The success of advanced analytics and artificial intelligence relies on data assets to provide a wide range of good-quality data inputs.
- Data and analytics (D&A) use cases, including customer experience, customer retention and sales conversion improvement, depend on historical transactional data, service records and interactional behavior data, which are all under the data asset management umbrella.

Obstacles

- Data assets are often siloed and protected by individual departments, making management hard to streamline.
- Under current accounting rules, the ability to recognize data as an intangible asset is constrained. There is no accounting model that a business could adopt to measure and assign a value to such an asset.
- Organizations struggle to quantify the development of data asset management. This effort includes the hours spent collecting, refining and enriching data, as well as the costs of personnel recruiting, storage, computing and other factors.
- Data is largely contextual. Raw data is alphanumeric in nature, and it increases in value as it moves through the valuation chain. Hence, depending on the usage value of the data under various circumstances, the difficulty in assigning a value to it stems from the uncertainty of future benefits. With increased regulations and data privacy requirements, unnecessary collected data may pose a compliance risk to organizations.

User Recommendations

- Create a vision of a data-driven enterprise that captures the aspirations of the executive team. Identify and prioritize information-based outcomes, such as monetizing data assets, establishing D&A governance and improving decision-making capabilities through D&A.
- Ensure that authority and accountability for D&A align with the enterprise's data-driven ambitions. Work with enterprise executives to clarify expectations and accountability for the chief data and analytics officer (CDAO) role, and to remove ambiguity concerning other executive roles.
- Change the D&A operating model to account for D&A competency gaps. Distribute competencies throughout the enterprise, and develop the data-driven culture needed for your enterprise vision to succeed.

Sample Vendors

Alibaba Cloud; Datablau; ESENSOFT; Huawei; Primeton Software

Gartner Recommended Reading

Quick Answer: How Should Chinese Enterprises Better Deliver Data Monetization Regarding "20 Data Measures"?

3 Ways to Promote Your Data Agenda at the Center of the Chinese Digital Economy

Responsible AI

Analysis By: Mike Fang, Tong Zhang, Ben Yan

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Responsible artificial intelligence (AI) is an umbrella term for aspects of making appropriate business and ethical choices when adopting AI. These include business and societal value, risk, trust, transparency, fairness, bias mitigation, explainability, sustainability, accountability, safety, privacy, and regulatory compliance. Responsible AI encompasses organizational responsibilities and practices that ensure positive, accountable, and ethical AI development and operation.

Why This Is Important

Responsible AI has emerged as the key AI topic for Gartner clients. When AI replaces human decisions, it amplifies both good and bad outcomes. Responsible AI enables the right outcomes by resolving dilemmas rooted in delivering value versus tolerating risks. This requires a set of tools and approaches, including industry-specific methods, adopted by vendors and enterprises. More jurisdictions introduce new AI regulations that challenge organizations to respond in meaningful ways.

Business Impact

Responsible AI signifies the move toward accountability for AI development and use at the individual, organizational and societal levels. If AI governance is practiced by designated groups, responsible AI applies to everyone involved in the AI process. Responsible AI helps achieve fairness, even though biases are baked into the data; gain trust, although transparency and explainability methods are evolving; and ensure regulatory compliance, despite the AI's probabilistic nature.

Drivers

- China's Ministry of Science and Technology (MOST) appointed the national New Generation AI Governance Expert Committee in March 2019, consisting of a mixed group of prominent scholars, entrepreneurs and government officials to provide advice on ethical issues related to AI.
- By June 2019, the expert committee unveiled the Governance Principles for the New Generation Artificial Intelligence — Developing Responsible Artificial Intelligence consisting of eight principles for the "safe, controllable and responsible use" of AI. They include harmony and friendliness; fairness and justice; inclusiveness and sharing; respect for privacy; security and controllability; shared responsibility; open cooperation; and agile governance.
- The Chinese government released its Position Paper of the People's Republic of China on Strengthening Ethical Governance of Artificial Intelligence (AI) in December 2022.
- On 11 April 2023, Chinese regulators at the Cyberspace Administration of China (CAC) issued draft measures to govern generative AI service provision in China.
- Enterprises that relied on the output from AI models would expect to have reliability and robustness built into their models.

Obstacles

- Legislative challenges lead to efforts for regulatory compliance while ignoring other responsible AI drivers.
- Rapidly evolving AI technologies, including tools for explainability, bias detection, privacy protection and some regulatory compliance, lull organizations into a false sense of responsibility, while mere technology is not enough. A disciplined AI ethics and governance approach that brings together multiple perspectives is necessary, in addition to technology.
- China's AI agenda represents the persistence of "top-level design" and state-centric approaches to technology development, which might make responsible AI look good on paper but ineffective in reality. This includes, for example, lacking flexibility and guidelines for enterprises to implement responsible AI from the execution level.

User Recommendations

- **Follow and align with government regulations.** Define and adjust your strategy regarding AI techniques and program adoption accordingly.
- **Combine responsible AI dimensions to publicize consistent approaches across all focus areas.** The most typical areas of responsible AI in the enterprise are fairness, bias mitigation, ethics, risk management, privacy and regulatory compliance.
- **Designate a champion accountable** for the responsible development and use of AI for each use case.
- **Define model design principles.** Address responsible AI in all phases of model development. Go for hard trade-off questions. Provide training and education on responsible AI to personnel.
- **Establish ethics principles and, optionally, an AI ethics board to resolve AI dilemmas.** Ensure diversity of participants and the ease to voice AI concerns.

Gartner Recommended Reading

3 AI Priorities to Increase Consumer Trust for Retailers in China

Quick Answer: How Should Chinese Enterprises Use Privacy-Enhancing Computation in Artificial Intelligence Initiatives?

A Comprehensive Guide to Responsible AI

Composable D&A

Analysis By: Fay Fei, Julian Sun

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Composable data and analytics (D&A) utilize container- or business-microservices-based architecture and data fabric to assemble flexible, modular and consumer-friendly D&A and AI capabilities from existing assets. This transforms monolithic data management and analytics applications into assemblies of D&A and AI or other application building blocks. This is achieved via composition technologies enabled by low- and no-code capabilities, supporting adaptive and intelligent decision making.

Why This Is Important

Organizations in China experiencing a fast-changing business environment need to speed up time-to-insight with agility. Composable D&A enables them to blend more insights and references into actions with modular D&A capabilities, instead of keeping development siloed. Organizations can accelerate the time-to-delivery flexibility in the assembly/reassembly of D&A capabilities by using composable D&A for different use cases.

Business Impact

Enterprises in China have invested in D&A projects independently:

- Composable D&A allows application/citizen/analytics developers to build analytics in applications and software, driving unprecedented business value.
- The increasing D&A silos impede the value realization of investment. Composable D&A helps to correctly identify data objects with wide reuse.
- The technology enables enterprises to embed and assemble AI-augmented features to create advanced analytics capabilities and solve more complex questions.

Drivers

- The emerging low- or no-code platform market in China opens up more possibilities for organizations to build composable D&A applications. No-/low-code solutions can be used to source data and compose more D&A capabilities with less technical effort, such as interactive visualization and predictive modeling, independently enriching more comprehensive embedded analytics.
- The growing number of analytics requirements in organizations in China drives the need for composable D&A. Organizations look for solutions to reuse and compose D&A capabilities into various scenarios.
- With the increasing popularity of AI usage in China, organizations can use composition to connect business intelligence (BI) to AI, extending BI capabilities and empowering users with a comprehensive, tailored, and even personalized, solution without having to use different applications.
- Organizations need to assemble descriptive, diagnostic, predictive and prescriptive analytics capabilities dynamically to generate insights along with the decision-making process. Use analytics to inform decision making and drive effective actions in a more connected, continuous and contextual way.
- Government-led initiatives such as “data sharing agency” and “industry cloud” are expected to be effective channels for organizations to distribute and share D&A applications. Composable D&A enables them to easily share and find the required components and contribute to D&A monetization.

Obstacles

- Composable D&A is still in the early stage in China. There are few solutions in the market that are well-developed and focused on composable capabilities.
- Organizations in China lack experience and guidance in working with business and finance stakeholders and in dealing with increasing management and governance issues.
- Cloud-based architecture drives composable D&A since many emerging capabilities are arising from cloud-first or cloud-only solutions. Leadership in Chinese enterprises may be relatively conservative in terms of cloud adoption. This may hinder the speed to benefit from composable D&A.
- In many Chinese organizations, close collaboration between development and D&A teams is not that common. There also exists a lack of standard processes and best practices to enable a working pattern to achieve composable D&A.
- Composable D&A also requires more close collaboration among vendors. The composability of the existing D&A products in China is not mature enough without deep technology partnerships.

User Recommendations

- Start working with business stakeholders, and improve decision making and business impact of D&A by incorporating and assembling modular, reusable D&A capabilities from the existing analytics portfolio.
- Leverage composable analytics to drive innovation by incorporating advanced data science and machine learning capabilities into analytics applications.
- Explore opportunities to add analytics capabilities to applications by building a joint team and ensuring ongoing collaboration between application developers and business analysts. Rethink organization, processes and skills to support agile assembly and reassembly of analytics services.
- Pilot composable analytics in the cloud, establishing an analytics marketplace to drive and support collaboration and sharing.
- Find vendors that support composable D&A both technically and in terms of commercial business models, such as API packaging, governance packaging and others.

Sample Vendors

DEEPEXI; Hengshi; Keendata; Kyligence

Gartner Recommended Reading

Use Gartner's Reference Model to Deliver Intelligent Composable Business Applications

Predicts 2023: Analytics, BI and Data Science Composability and Consolidation

Client Question Video: What Is Composable Data and Analytics?

Adopt Cloud Analytics to Drive Innovation

Digital Ethics

Analysis By: Mike Fang, Julian Sun

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Definition:

Digital ethics comprises the systems of values and moral principles for the conduct of electronic interactions among people, organizations and things.This includes areas such as AI, data & analytics and social media.

Why This Is Important

Digital ethics, especially regarding privacy and bias, one of the most discussed topics in the industry and society. The voice of society is getting louder, and the responsible use of AI is rapidly coming into focus for individuals, organizations and governments. People, increasingly aware that their information is valuable, are frustrated by lack of transparency, misuse and breaches. Organizations are acting to mitigate risks around managing and securing personal data, while governments are implementing stricter legislation.

Business Impact

Digital ethics strengthens the organization's positive influence and reputation among customers, employees, partners and society. Areas of business impact include innovation, product development, customer engagement, corporate strategy and go-to-market. Intention is key. If ethics is simply a way to achieve business performance, it comes across as disingenuous. The goal to be an ethical company serves all parties and society more broadly and leads to better business trust and performance.

Drivers

- China releases guidelines to strengthen governance over ethics in science, technology. The document clarified the ethical principles in science and technology, saying that scientific activities should serve the well-being of humanity, respect people's right to life, adhere to fairness and justice, control risks in an appropriate way, and maintain openness and transparency.
- Board members and other executives are sharing concerns about the unintended consequences of innovative technology use.
- The media is increasingly featuring high-profile stories about the impact of data and technology on business and society at large.
- With the emergence of artificial intelligence, the ethical discussion is now taking place both before and during a technology's widespread implementation. AI ethics and other responsible AI steps attempt to reverse the negative popular sentiment around AI and establish more responsible use of its powers.
- Gartner predicts that, by 2024, 30% of major organizations will use a new "voice of society" metric to act on societal issues and assess the impact on their business performance. The voice of society will put more pressure on both governments and public/private organizations to use technology ethically. "Big tech" is already a negative stereotype in societal jargon.

Obstacles

- Due to the ambiguous nature of digital ethics, organizations are struggling to operationalize it and expending significant effort to implement best practices.
- Opinions differ across people, regions and cultures on what constitutes “good” and “bad.” Even in organizations where ethics is recognized as an important issue, consensus between internal and external stakeholders (such as customers) remains sometimes difficult to achieve.
- Organizations see digital ethics as a moving target due to confusion around society’s expectations. An organization’s position and beliefs may even steer digital ethics against the majority’s opinion.
- Digital ethics is too often reactive, narrowly interpreted as compliance, confined to the technical support of privacy protection and/or viewed as explainable AI only.
- AI ethics is currently the main focus of overall digital ethics. Supporting technology needs to mature further.

User Recommendations

- Identify specific digital ethics issues and opportunities to turn awareness into action.
- Discuss ethical dilemmas from diverse points of moral reasoning.
- 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 societal and business value, rather than simply focusing on compliance and risk. Link digital ethics to concrete business performance metrics.
- Ensure that digital ethics is leading and not following digital transformation. Address digital ethics early “by design” to create methods that resolve ethical dilemmas quickly.
- Organize training in ethics, and run workshops to create awareness within all AI initiatives about the importance of an ethical mindset and clear accountability in AI design and implementation.

Gartner Recommended Reading

Tool: Assess How You Are Doing With Your Digital Ethics

Tool: How to Build a Digital Ethics Curriculum

AI Ethics: Use 5 Common Principles as Your Starting Point

Every Executive Leader Should Challenge Their Teams on Digital Ethics

Expert Insight Video: What Is Responsible AI and Why You Should Care About It?

Data Sharing

Analysis By: Mike Fang, Xingyu Gu

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Early mainstream

Definition:

Data sharing in China is a critical business process that provides enterprise value. It drives efficiency, assisting in data reuse by diverse users. Internally, it is within commonly owned entities, such as parent/subsidiary/sister companies. Externally, it is across parties outside of commonly controlled entities, such as cross-industry initiatives or cross-public sector programs. Data sharing helps improve collaboration.

Why This Is Important

The digital economy is a strategic initiative for Chinese enterprises where data sharing is essential. Data sharing offers opportunities for enterprises to use internal and external data to support better business decisions, monetize data for an additional revenue stream, and create a collaborative ecosystem with partnerships, which is impossible using only internal data. Data and analytics (D&A) leaders at high-performing organizations remove obstacles to data access or break down data silos; they promote data sharing or increase access to data aligned to the business case.

Business Impact

Data sharing offers value to society and business by:

- Increasing access to more relevant data to match social or business goals, enabling the realization of stakeholder value while driving business outcomes
- Using both shared internal and external data to support better business decisions
- Monetizing data for the additional revenue stream
- Creating a collaborative data ecosystem with partnerships
- Improving accuracy of deployed/newly built models
- Driving operational excellence and environmental sustainability

Drivers

- The Chinese government has been treating data as the fifth production factor, after land, labor, capital and technology, since 2020 (see China Unveils Guideline on Improving Market-Based Allocation of Production Factors). The sharing of data as a new production factor is crucial for economic development and growth: it can promote optimal allocation of resources, boost productivity, and increase economic benefits.
- The government promotes data sharing to deepen public-data-resource exploitation and data-driven use cases. For example, recent launches of various government-sponsored data exchange initiatives aim to standardize data-asset trading across enterprises for better business and social benefits.
- Recent data regulations and data classification guidelines (in financial, government, healthcare, automotive, industrial and IT sectors) support the sustainable development of data sharing initiatives in a regulated, standardized manner.
- Emerging techniques developing quickly in China, such as blockchain and privacy-enhancing computation (PEC), can be used to help enhance trust and compliance during the data sharing process.
- The rise in synthetic data creation illustrates the demand for data sharing because real data can be expensive, unavailable or unusable due to privacy regulations.
- There is a lack of relevant available data for AI training, as well as sustainability and cost pressures for processing vast AI training data such as data corpus for large language models.
- As models' environments are constantly evolving, there is a parallel need to continually share near-real-time data.
- External data has an increased level of relevance in support of predictive models, as models trained exclusively with internal or first-party data have seen model drift due to phase shifts in customer behaviors.
- We observe increased demand for more robust predictive analytics generated from more diverse data sources to drive relevant, unique or otherwise unknowable insights for data-driven innovation.

Obstacles

- Lack of mature practice in data ownership, pricing, circulation and governance that block the data sharing
- Ignorance of how to share data at scale and with trust
- Stakeholder resistance based on fear, outdated data management and governance policies, and lack of tools/technologies
- The outdated perception that alleged risks of data misuse outweigh the business benefits of sharing data, including cost savings and revenue growth
- Internal data hoarding, external data hijacking, confidentiality weaponization and privacy shaming
- Fascination with technology enablers, which can overshadow data literacy demands, processes/policies and collaboration required for data sharing success
- Failure to make the right investments in data sharing, impeding the discoverability, reuse and resharing required for efficient and value-producing digital business outcomes

User Recommendations

- Modernize data management, adopt trust-based D&A governance, and foster a data sharing culture, not a data ownership culture.
- Evaluate the potential AI and analytics use cases with the possibility of achieving better/new business value by leveraging data sharing.
- Curate your D&A products from raw data to generate additional revenue streams or partnerships for your enterprise by monetizing, managing and measuring your data assets.
- Identify the value and risks associated with the data product. Bridge the delivery gap between the data provider and consumer through the appropriate data-exchange channel.
- Orchestrate the data sharing process in a trustworthy and compliant way by utilizing different techniques, such as blockchain and privacy-enhancing computation.
- Break the boundaries of the organization, fostering data literacy among external parties to expedite data sharing for better business outcomes.

Sample Vendors

BaseBit XDP; China Mobile DSSN; Huawei EDS; Transwarp Navier

Gartner Recommended Reading

Top 3 Priorities for Chinese Enterprises to Promote Effective Data Sharing

Quick Answer: How Should Chinese Enterprises Use Privacy-Enhancing Computation in Artificial Intelligence Initiatives?

Quick Answer: How Should Chinese Enterprises Better Deliver Data Monetization Regarding "20 Data Measures"?

3 Ways to Promote Your Data Agenda at the Center of the Chinese Digital Economy

Lakehouse

Analysis By: Xingyu Gu, Julian Sun

Benefit Rating: Moderate

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Definition:

A lakehouse is a converged infrastructure environment that combines the semantic flexibility of a data lake with the production optimization and delivery of a data warehouse. It supports the full progression of data from its raw, unrefined state, through the steps of refining it, to ultimately deliver optimized data for consumption.

Why This Is Important

Organizations in China continue to grapple with emerging practices for integrating data warehouses with systems like data lakes. Architecture became more complex after organizations adopted big data technologies based on Apache Hadoop. By unifying analytical databases, a lakehouse aims to simplify architecture and improve efficiency while minimizing the need to move data between them. A more efficient environment with a smaller operational footprint is the potential result.

Business Impact

Businesses will benefit from streamlined delivery, rapid access to data, and a consolidated data management platform that supports highly skilled data scientists, engineers and analysts, plus casual users who consume data via prebuilt reports or dashboards. A lakehouse provides a well-defined path from discovery-oriented analytics and model development (via the lake portion) to delivery of analytical insights and measurable value to end users (via the warehouse portion).

Drivers

- Organizations in China are struggling to operationalize data science/AI projects because of complex analytics architecture, which is amplified by the recent emergence of generative AI use cases. A lakehouse unifies different analytical use cases, helping to simplify D&A architecture complexity.
- Data lakes and data warehouses specialize in different things. Data lakes enable data science, ad hoc analytics, and management of flexible data structure, latency or containers. Warehouses excel at handling refined data that requires an audit trail, high quality and accuracy, or special data structures (dimensions, time series, hierarchies, etc.). By merging the two architectures, the lakehouse is expected to support all these data requirements and business use cases.
- Many cloud data warehouses and almost all cloud data lakes already leverage semantically flexible cloud object storage as their storage of record. It is a natural progression to unify these storage environments, thus reducing the disparate and duplicate infrastructures.
- The data lakehouse concept is maturing fast because it benefits from the market's understanding of data warehouses, data lakes, and their respective pros and cons.
- Lakehouse supports a range of user skills within an integrated platform. Data engineers can create pipelines to ingest and transfer raw data into the lake part of the lakehouse. Data scientists, power users and stewards can then refine and model data in the warehouse part. Analysts can use this data for reporting and other repeatable analytics workload demands.

Obstacles

- The vendor-built lakehouse platforms are still maturing. Some have strong data lake features, but do not support the full range of transaction consistency or workload management features expected from a data warehouse. Others have strong data warehousing features, but lack the broad data model support or data science features of a data lake.
- Some organizations in China have invested heavily in data warehouses and big data platforms based on Hadoop/Spark. From a ROI perspective, these organizations would rather improve their existing architecture than replace it with a lakehouse.
- The full scope of optimization includes data quality, security, performance and, most importantly, good metadata management and data integration. Few lakehouse platforms address all of these.
- Enterprises consistently seek rapid and unencumbered data access to overcome delivery delays associated with the data warehouse. The lakehouse is often overpositioned as a cure-all to this problem.

User Recommendations

- Employ a targeted use-case approach that solves specific problems and expands from there for long-term success. Expect your lakehouse to grow into many more use cases over time, just as lakes and warehouses do.
- Avoid “overpromising and underdelivering” by testing candidate solutions thoroughly to ensure that you can actually deliver reliable and high-performance workloads on your lakehouse.
- Carefully evaluate ROI to confirm that replacement (versus improvement) is really the right decision. Most MPP-based data warehouses and Hadoop-based data lakes can be evolved to handle a wider range of analytics use cases.
- Choose a logical data warehouse (LDW) approach when addressing a broader data and analytics scope. A lakehouse is a subset of the LDW built opportunistically. The LDW remains a mature and best practice.
- Evaluate security and governance capabilities to ensure that they meet your enterprise standards and data requirements.

Sample Vendors

Alibaba Cloud; DEEPEXI; Huawei Cloud; Oushu; SequoiaDB; SF Technology; Tencent Cloud; Transwarp

Gartner Recommended Reading

Market Guide for Analytics Query Accelerators

Does My Organization Need a Data Lakehouse?

Exploring Lakehouse Architecture and Use Cases

Cool Vendors in Data Management, China

At the Peak

Augmented Data and Analytics

Analysis By: Fay Fei, Julian Sun

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Augmented data and analytics (D&A) incorporates machine learning (ML)/artificial intelligence (AI) to enable data management and analytics capabilities. It automates workflow both in data management and analytics to drive D&A adoption and make work more efficient by applying ML/AI. It enables analytics consumers to gain more in-depth insights into the organization with less time and interpretation bias and fewer skills compared to manual analysis approaches.

Why This Is Important

Organizations are looking for solutions to better create and manage data and analytics with greater ease of use. The overall data literacy of organizations in China is still low. Adoption of augmented capabilities can reduce the manual workload for experts and can make up for the lack of people’s technical skills. In addition, with the development of ML/AI technology, augmented analytics can provide deep and contextual insights to unveil extra business findings and opportunities.

Business Impact

Augmented D&A complements and extends existing D&A platforms with more advanced analytics generated by business analysts, decision makers and operational workers:

- It reduces the time users spend managing and operating data while giving them more time to collaborate and act on the most relevant insights from data.
- It allows frontline workers to have access to context-enriched analysis and guided recommendations via automated insights to improve decision making and take the correct actions faster.

Drivers

- The hype of generative AI and its representative solution OpenAI's ChatGPT drives the attention of augmented data and analytics. Although these techniques are still in their infancy, especially when it comes to data management, it still encourages many experiments and exploration from organizations in China.
- With the rapid growth of hyperscale cloud vendors in China, the elasticity and high-performance computing workloads enabled by cloud deployment have provided a technical foundation to realize augmented D&A capabilities.
- From a data management perspective, organizations often lack the technology and experts to build efficient D&A platforms. Augmented D&A provides the capabilities to optimize the use of D&A platforms, suggest and implement new designs, schemas and queries, and infer the semantics and associations of data in order to recommend data structure improvements.
- From an analytics perspective, organizations are not satisfied with descriptive analytics but are looking for diagnostic, predictive and prescriptive analytics that are high-value-added. These kinds of analytics could be enabled by augmented D&A capabilities, such as providing data on key influencers, identifying outliers, forecasting, and autogenerating narratives and storytelling with deep analytics insights.

Obstacles

- The expectations of organizations with lower maturity and data literacy levels are too high and unrealistic when it comes to augmented technologies.
- The hype around being able to find one unified solution for both augmented data management and analytics is high, but a clear need for unification of these still does not exist in many organizations.
- The trust and transparency in AI-enabled D&A capabilities is low, and the explainability of how automated insights are generated is not yet mature, including from Chinese vendors' solutions.
- Augmented D&A doesn't remove the need for data literacy. Organizations lack training programs and need support from experts to enable users to interpret augmented results.

User Recommendations

- Embrace augmented capabilities on both data management and analytics as part of a digital transformation strategy to benefit a broader range of users. Start by identifying the business and users that will benefit most from it. Launch pilots to prove the value and build trust.
- Make augmented capabilities a “must have” selection criterion for new purchases or upgrades of D&A products. Audit the results of augmented D&A deployments to assess the risk of introducing errors and reduced performance if used more widely.
- Invest in emerging technologies, such as graph technologies and explainable AI, to harness the power of augmented D&A.
- Organize training and workshops to improve data literacy and increase internal incentives to stimulate the usage of augmented capabilities.

Sample Vendors

Alibaba Cloud; Datablau; Guangzhou Smartbi Software; Haizhi Network Technology (Beijing); Hangzhou Guanyuan Data; Kyligence; NetEase

Gartner Recommended Reading

Adopt Cloud Analytics to Drive Innovation

Market Guide for Augmented Analytics

Cool Vendors for Analytics Platforms in China

Market Guide for Analytics Platforms, China

DBMS Self-Sufficiency

Analysis By: Xingyu Gu, Julian Sun

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Adolescent

Definition:

DBMS self-sufficiency is key to the DBMSs regional market in China, in which the vendors offer DBMSs primarily to Chinese organizations and international enterprises that have operations in China. A DBMS is a product used for the storage and organization of data that typically has defined formats and structures. In China, technology self-sufficiency, which aims to make local organizations resilient in the dynamic global environment, dominates the local DBMS market.

Why This Is Important

Within the enhanced digital transformation of different industries in China, DBMSs, acting as the key part of data infrastructure, are required to have high performance, elasticity, strong AI compatibility and cost-effectiveness. At the same time, the highly dynamic geopolitical environment in recent years has stimulated Chinese organizations to ensure their DBMS technologies are self-sufficient. These two factors have led to a pressing need for modern DBMSs developed in China.

Business Impact

DBMS self-sufficiency acts as a strategic component, which is aligned with local organizations' digital transformation roadmap. It helps:

- Keep organizations aligned with the "information technologies self-sufficiency" national strategy.
- Reduce the huge maintenance cost from expensive mainframe hardware.
- Address performance bottlenecks when local events with extremely high-concurrent data transactions happen (for example, the Double 11 shopping festival).
- Ensure data science and real-time analysis use cases can be further scaled and embedded into production.

Drivers

- Technology infrastructure demands in China are significant due to a large population, many internet users, and a growing desire for enhanced technology to sustain digital initiatives like smart cities and live commerce at a scale unique to China.
- Legacy DBMS solutions — largely based on implementations of non-Chinese vendors — are aging and falling short in addressing these needs.
- Rather than engaging with upgrading non-Chinese technology investments, the recent geopolitical environment has demanded that Chinese organizations become self-sufficient, driving the transition to local vendors.

Obstacles

- Immature DBMS migration tools and services are impeding organizations from migrating existing business workloads to modern DBMSs. Because unsuccessful DBMS migrations in critical business workloads have huge business impacts, organizations in China require more mature DBMS migration tools and services to help them enable a seamless migration.
- Due to the technical debt deeply embedded with legacy software and hardware, the effort and risk to migrate DBMSs from old systems to modern ones is still large. Large-scale DBMS migrations frequently take more than a year.
- Due to concerns about data security, data sovereignty and vendor lock-in on public cloud, organizations in industries like finance, government and public services prefer to modernize their mission-critical DBMSs to a hybrid cloud environment. The to-be-improved hybrid cloud capabilities of local DBMS products impede them from fully realizing the benefits of modern DBMS deployments.

User Recommendations

- Evaluate the most suitable deployment model (public cloud/on-premises/hybrid) for your use cases with respective value preferences like elasticity, flexibility, cost-effectiveness or regulation compliance.
- Observe how the cloud data ecosystem in China emerges. Choose the appropriate ecosystem best suitable for your organization from an industry, primary business region, organization scale and business portfolio variety perspective.
- Assess local DBMS vendors by evaluating not only their product capabilities, but also their openness to the local technology vendor ecosystem (such as cloud platform, hardware, business application and data security) from both product features and postsales SLA perspectives.
- Evaluate emerging DBMS capabilities (for example, distributed transactional database, lakehouse and augmented transactions) in China under consideration by leveraging Gartner research on cloud database management systems and associated RFP toolkits. Understand the pros and cons of each, and engage with business stakeholders to judge where the most suitable use cases are for them.

Sample Vendors

Alibaba Cloud; HUAWEI CLOUD; OceanBase Database; PingCAP; SequoiaDB; Tencent Cloud; Transwarp; Wuhan Dameng Database

Gartner Recommended Reading

Market Guide for DBMS, China

中国数据库管理系统市场指南

Magic Quadrant for Cloud Database Management Systems

Exploiting the Evolving Database Management System Trends in China

Foundation Models

Analysis By: Tong Zhang, Mike Fang

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

Definition:

Foundation models are large-parameter models that are trained on a broad gamut of datasets in a self-supervised manner. They are mostly based on transformer or diffusion deep neural network architectures and will potentially be multimodal in the near future. They are called foundation models because of their critical importance and applicability to a wide variety of downstream use cases. This broad applicability is due to the pretraining and versatility of the models.

Why This Is Important

Foundation models are an important step forward for AI due to their massive pretrained datasets and wide use-case applicability. They can deliver state-of-the-art capabilities with higher efficacy than their predecessors. They've become the go-to architecture for NLP, and have also been applied to computer vision, audio and video processing, software engineering, chemistry, finance, and legal use cases. Primarily text-based, large language models (LLMs) are a popular subset of foundation models. ChatGPT is based on one (GPT-3).

Business Impact

With their potential to enhance applications across a broad range of natural language use cases, foundation models will have a wide impact across vertical industries and business functions. They will accelerate digital transformation by improving workforce productivity, automating and enhancing CX, and enabling rapid, cost-effective creation of new products and services. Foundation models that process multimodel input will expand usage beyond natural language processing, and several foundation models may collaborate to generate more artifacts.

Drivers

Foundation models:

- **Require only limited model customization to deliver effective results.** Foundation models can effectively deliver value through prebuilt APIs, clever prompt engineering or further fine-tuning. While fine-tuning may deliver the best value because of customization possibilities, the other two options are less complex. Gartner clients report that they can use Google's open-source BERT implementation without training the entire network. Rather, they retrain only the top few layers to customize for their language domain.

- **Deliver superior NLP classifications.** The difference between these models and prior neural network solutions is stark. The large pretrained models can produce coherent text, code, images, speech and video at a scale and accuracy not possible before. Transformer architecture model patterns form relatively large blocks of text, as opposed to predicting the next word based on the preceding words. These improvements have materially advanced speech, language and text applications. A notable example is the improvement in Google Translate.
- **Enable low-friction experimentation.** These models can create well-formed text passages from minimal input. The past year has seen an influx of foundation models, along with smaller, pretrained domain-specific models built from them. Most of these are available as cloud APIs or as open-source projects, further reducing the time and cost to experiment. GPT-3 and GPT-4 are foundational models developed by OpenAI and licensed by Microsoft. These transformer models are designed to create paragraphs or pages of text from small excerpts. They do so by predicting the most likely next word in a sentence, based on their absorbed accumulated training.
- **Have accelerated AI innovation with massive model sizes.** For example, both OpenAI's GPT-3 and Meta AI's OPT have 175 billion parameters. Google DeepMind's Gopher has 280 billion, while Google AI's PaLM has 540 billion. Megatron-Turing NLG, a collaboration between NVIDIA and Microsoft, has 530 billion. Alibaba's M6 has 10 trillion. In addition, both Meta and Google have open-sourced certain models.

Obstacles

Foundation models:

- **Do not deliver perfect results.** Foundation models require careful training and guardrails. Because of their training methods and black-box nature, they can deliver unacceptable results or hallucinations. They also can propagate any biases or copyright issues in the datasets.
- **Require appropriate skills and talent.** The end result depends on the skills, knowledge and talent of the trainers, particularly for prompt engineering and fine-tuning.
- **Expand to impractical sizes.** Large models are up to billions or trillions of parameters. They are impractically large to train for most organizations because of the necessary compute resources, which can make them expensive and ecologically unfriendly.

- **Concentrate power.** These models have been mostly built by the largest technology companies, resulting in a concentration of power among a few deep-pocketed entities. This situation may create a significant imbalance in the future.

- **Potentially threaten humanity.** With their emergent reasoning capability, foundation models can potentially surpass human intelligence, making AI a threat.

User Recommendations

- **Introduce foundation models into existing speech, text or coding programs.** If you have older language processing systems, moving to a transformer-based model could significantly improve performance.
- **Start with established, open-source-based models** that have superior ecosystem support, have adequate enterprise guardrails around security and privacy across DSML platforms, and are more widely deployed.
- **Explore new use cases,** such as natural language inference, sentiment analysis or natural-language-based enterprise search, where the models can significantly improve both accuracy and time to market.
- **Designate an incubation team** to monitor industry developments, communicate the art of the possible, experiment with BUs and share valuable lessons learned companywide.
- **Explore new approaches,** such as lightweight fine-tuning and prompt engineering, to leverage foundation models with business knowledge.
- **Follow the value alignment with AI.** Evaluate the risks of foundation models, and seek to design methods that prevent AI systems from inadvertently acting in ways inimical to human values.

Sample Vendors

Alibaba Cloud; Baidu; Huawei; iFLYTEK; Tencent; Zhipu AI

Homegrown AI Chips

Analysis By: Roger Sheng, Mike Fang, Julian Sun

Benefit Rating: High

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Definition:

An AI chip is a type of semiconductor device optimized for processing AI algorithms, such as deep neural networks (DNNs). AI chips are typically designed to accelerate a specific element of an AI workflow, such as training and inference. AI chips are important components to both largely increase AI processing capabilities and offload CPU workload in the data centers, edge and devices. Chinese companies designing AI chips remain in the early stage compared with global vendors such as NVIDIA.

Why This Is Important

Chinese companies are investing hugely in the AI technology field in which AI accelerator chips are the key components in AI development. However, the U.S. export regulations restricted advanced AI chips to Chinese companies from 2022. This situation forced Chinese companies to seek alternative resources for AI chips. Both established IT giants and venture startups have launched the development of AI chips for the ongoing global competition in the AI field.

Business Impact

For AI-based workloads, AI chips take a more important role than traditional chip architecture. Most hyperscale cloud service providers use high-performance AI accelerator chips to increase the efficiency of massive data processing. The recently hyped generative AI requires thousands of AI accelerators for training of foundation models. These chips are typically manufactured on leading-edge semiconductor process technologies with a consequent cost impact and power budget.

Drivers

- The demand for AI innovation usages. The recently hyped generative AI is based on the foundation model that requires more powerful AI accelerator chips, especially in the training stage. Due to the U.S. government's export restrictions, Chinese companies have to seek alternative chip suppliers locally as backup. The large established companies had already built the internal team for chip development and also invested in the local emerging vendors.

- The government local-sufficient initiatives. The Chinese government is strongly supporting the development of the local supply ecosystem for the IT area, especially in the strategic high-performance computing areas by considering the nation's high-tech industry growth and cybersecurity. The Chinese government encourages more locally designed chips used in the government-guided data center projects. For example, Cambricon Technologies' AI chip business is mainly from government clients.

- The efficiency of system operation. Current GPU-architecture-based AI accelerator chips are dominated by global vendors that are priced high with limited supply. Although the GPU capabilities are superior in the general AI data training, customized application-specific integrated circuits (ASICs) that are designed for dedicated AI applications will have advantage in power efficiency. Thus, Baidu decided to use its in-house-developed AI chips in its searching business.

- The fragmented demand from AI Internet of Things (IoT) applications. The combination of AI and IoT is one of the important elements to create a digital business process. Considering the different types of AI IoT use applications and use cases, the system requirements are fragmented. It gives Chinese AI chip vendors great opportunities to develop optimized solutions for various AI IoT requirements. For example, most AI chips used in China-made surveillance systems to support AI compute vision capabilities are developed by Chinese chip vendors.

Obstacles

- The lack of general AI toolkit support. One of the key success factors of NVIDIA is its CUDA platform that is based on its GPU architecture and can support most AI development tools. Chinese AI chip companies must have more investment in software to support mainstream toolkits for AI development.
- The immaturity of the developer community. Chinese companies lack the experience to build a broad, cross-company software developer ecosystem, which is important to the development of AI applications.
- U.S. restrictions on advanced semiconductor manufacturing. The U.S. restrictions forbid Chinese companies to get the advanced chip manufacturing service (e.g., 5 nm and 3 nm) from global foundry suppliers if the product is designed for high-performance computing. The restrictions also ban the equipment for 14 nm and more advanced semiconductor chip manufacturing shipping to China.

User Recommendations

- Take a holistic solution approach in the AI infrastructure by cooperating with established Chinese semiconductor companies to both focus on the development of customized AI chips and accumulate expertise.
- Monitor and engage in research and development of new semiconductor technologies that don't require advanced process nodes, such as heterogeneous integration, RISC-V CPU and interchip optical interconnects.
- Improve the computing system performance by focusing more on the software optimization and customized chip design for dedicated applications.
- Prepare for and invest in the reasonable inventory of advanced AI chips from global vendors for AI infrastructure.
- Cooperate with industry peers to build China-specific AI ecosystems by unifying the industry standard and development platform.

Sample Vendors

Baidu; Biren Technology; Cambricon Technologies; Enflame; Huawei; Hygon

Gartner Recommended Reading

Market Trend: U.S. Restrictions Force China ICT Industry to Seek Alternative Semiconductor Solutions

Market Trends: China's New 'Secure and Trustable' Initiative for the IT System Will Accelerate the Growth of Domestic Chip Vendors

Generative AI

Analysis By: Ben Yan, Mike Fang, Tong Zhang, Tracy Tsai

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Generative AI technologies can generate new derived versions of content, strategies, designs and methods by learning from large repositories of original source content. Generative AI has profound business impacts, including on content discovery, creation, authenticity and regulations; automation of human work; and customer and employee experiences.

Why This Is Important

Generative AI exploration is accelerating, thanks to the popularity of models and generative AI applications from vendors such as Stable Diffusion, Midjourney, OpenAI (ChatGPT) and numerous others. End-user organizations in most industries aggressively experiment with it. Technology vendors in China, such as Baidu, Alibaba and SenseTime, prioritize delivery of generative-AI-enabled applications and tools. Many startups emerged in 2023 to innovate with generative AI, and we expect this to grow. Some governments, including China, are evaluating the impacts of generative AI and preparing regulations.

Business Impact

Leading technology products and services will incorporate generative AI capabilities in the next 12 months, introducing conversational ways of creating and communicating with technologies, leading to the democratization of generative AI technologies. Generative AI will progress rapidly in industry verticals, scientific discovery and technology commercialization. Sadly, it will also become a security and societal threat when used for nefarious purposes. Responsible AI, trust and security will be necessary for safe exploitation of generative AI.

Drivers

- The popularity of ChatGPT is accelerating the hype around generative AI, pushing it to the peak of the Hype Cycle. It is garnering substantial attention from executives, driving enthusiasm for broader generative AI adoption.
- Foundation models largely improve both the quality of AI-generated content and the speed of generation across a broad array of artifacts, including language, code, images, voice and multimodal data.
- Prompt engineering with zero-/few-shot learning has been rapidly improving generative modeling while reducing the need for training data and fine-tuning.
- Machine learning (ML) and natural language processing (NLP) platforms are adding generative AI capabilities, along with transfer learning, for reusability of generative models, making generative models customized and accessible to AI teams.
- When integrated with other models or applications — such as visual generation models, scientific computing models or web applications — generative AI models unlock more potential business use cases.
- Generative AI offers the industry a much-needed opportunity to drive productivity, which is a required component of any growing economy. Few other technologies offer such an opportunity to reinvent how work gets done.
- Synthetic data draws enterprises' attention by helping augment scarce data, mitigate bias or preserve data privacy. It boosts the accuracy of brain tumor surgery.

Obstacles

- The foundation models underlying generative AI are not mature enough. Fully relying on foundation model outcomes is risky as it may introduce incorrect results. A human in the loop (HITL) is required.
- Generative AI can be used for nefarious purposes. Generated deepfakes are dangerous in politics, business and society. Full and accurate detection of generated content will remain challenging for years and may not be completely possible.
- Fragmented and specialized technology offerings (such as generating only images or only text) currently lead to a combination of tools rather than a single solution.
- Compute resources for training large models are heavy and not affordable to most enterprises. Most enterprises can exploit existing models, but can't develop their own.
- Generative AI vendors will need to adjust their approaches as regulations on the technology are introduced.
- Local generative AI models in China with similar performance to ChatGPT/GPT-4 are not available yet. Organizations that are impressed by ChatGPT/GTP-4 and looking for local alternatives have to wait or leverage other specialized solutions instead.

User Recommendations

- Investigate how generative AI techniques can benefit your industry or sector. Identify initial use cases where you can rely on purchased capabilities or partner with researchers.
- Explore how synthetic data could accelerate the development cycle, lessen regulatory concerns, mitigate data bias, facilitate data monetization and lower the cost of data acquisition, especially if you lack data for rare events.
- Examine and quantify the advantages and limitations of generative AI. Use it first to improve an existing process. Provide guidelines where generative AI could bring breakthroughs, as it requires skills, funds and caution. Weigh technical capabilities with ethical factors.
- Prepare to mitigate the impact of deepfakes, which can cause serious risk. Mitigation methods, such as applying algorithmic detection and authenticating content provenance, are still evolving. Technical, institutional and political intervention will be necessary to fight deepfakes.
- Pay close attention to the generative AI techniques, as we expect their adoption to be rapid.

Sample Vendors

aiXcoder; Alibaba Cloud; Amazon Web Services; Baidu; Huawei; Microsoft; SenseTime; Tencent; Zhipu AI

Gartner Recommended Reading

Innovation Insight for Generative AI

Top Strategic Technology Trends for 2022: Generative AI

Predicts 2022: Generative AI Is Poised to Revolutionize Digital Product Development

Quick Answer: China Perspective — Frequently Asked Questions on ChatGPT and Large Language Models

Data Literacy

Analysis By: Fay Fei, Mike Fang

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Definition:

Data literacy is the ability to read, write and communicate data in context, with an understanding of the data sources and constructs, analytical methods, and techniques applied. Data-literate individuals have the ability to identify, understand, interpret and act upon data within business context and influence the resulting business value or outcomes.

Why This Is Important

Digital experiences pervade both our personal and professional lives. At both the country level and the industry level, digital transformation is becoming increasingly critical. The ability to understand, interpret and act upon data — data literacy — is foundational to the digital economy and society. It helps stakeholders:

- Unlock the business acumen that knowledge workers possess
- Draw insights from data relevant to specific business use cases
- Establish trust around the data to support confident decisions

Business Impact

To both leverage data for competitive advantage and echo the top-level digital strategy of the Chinese government, enterprises must become data-driven. This transformation requires explicit and persistent organizational changes. Enterprises need to promote and orchestrate cultural changes at critical points affecting business success. Central to success will be the ability to guide the workforce by addressing both data literacy and data-driven culture.

Drivers

- The Chinese government has revealed a strong will to drive the digital economy and digital business. In recent years, it has published numerous policies and initiatives to encourage digital transformation, which increases society's overall data awareness.
- There is growing recognition of the role that employees' data literacy plays within an organization's overall digital dexterity. Industries with high data literacy levels have achieved promising business growth in China.
- The role of the data and analytics (D&A) function has evolved. It is now at the core of an organization's business model and digital platforms, making everyone an information worker. Thus, business use of data and analytics is broader than ever before.
- The value of data and analytics has been limited to visualization for years. Organizations have started to realize that enhanced data literacy will contribute to more value-driven D&A use cases.
- Many individuals in China, especially the younger generation, naturally welcome digitalization and are passionate about building their data literacy.

Obstacles

- Industries lack common models, frameworks, standards and terminology for data literacy.
- The correlation between "data literacy level" and "business optimization" is not clear. It is still hard to explicitly quantify the business value coming from a higher data literacy level.
- Organizations lack guidance and practices to foster data literacy. Approaches to training and certification are sporadic and inconsistent.
- Talent and skills shortages in the data and analytics domain lead to poor data literacy within the current workforce and inhibit any ability to even get started.
- Data literacy does not get sufficient attention from high-level management. Thus, organizations still lack initiatives to address cultural and data literacy challenges within strategies and programs.
- Overall adoption will still take years, due to the complexity of upskilling entire workforces.

User Recommendations

- Digest China's digital policies and align with internal data literacy initiatives to gain competitive advantage in the changing environment.
- Engage stakeholders, raise awareness and create a strong shared vision of desired business outcomes for data literacy, particularly with respect to data-driven innovation.
- Partner with HR and business leaders to incorporate data literacy into the employee value proposition. Call out examples of "good" and "bad" data literacy to reinforce desired behaviors.
- Assess the data literacy levels, learning goals and outcomes for various job roles and personas, and work data literacy into career paths and advancement.
- Go beyond vendor product training by tailoring the curriculum to your organization's industry, business domains and roles. Use a mix of training methods (classroom, online, community, on-the-job) to improve overall learning.
- Start with D&A quick wins to build momentum, but set reasonable expectations for different stages of data literacy improvement. Lasting change takes time because it requires people to learn new skills and behave in new ways.

Sample Vendors

Coursera; Heywhale; iCourse; IMOOC; Udacity; Udemy; XuetangX

Gartner Recommended Reading

How CDAOs Must Lead Data Literacy and Data-Driven Culture

Address Both 'Skill' and 'Will' to Deliver Data-Driven Business Change

Drive Business Outcomes by Measuring the Value of Data Literacy

Tackle Data Literacy Head-On to Avoid Data and Analytics Program Failure

Knowledge Graphs

Analysis By: Tong Zhang, Ben Yan, Mike Fang

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

Knowledge graphs are machine-readable data structures, representing knowledge of the physical and digital worlds including entities (people, companies, digital assets) and their relationships, which adhere to a graph data model — a network of nodes (vertices) and links (edges/arcs).

Why This Is Important

Knowledge graphs capture information about the world in an intuitive way that is often easier to understand, manipulate and use than other types of data models. Amazon, Facebook, Google and other tech companies use graphs as the backbone of a number of products and services due to their ability to encode and interrelate disparate data at source. Knowledge graphs support collaboration and sharing, search and discovery, and the extraction of insights through analysis.

Business Impact

Knowledge graphs can drive business impact in a variety of settings including:

- Investigative analysis (e.g., law enforcement, cybersecurity or financial transactions)
- Digital commerce (e.g., product information management and recommendations)
- Data management (e.g., metadata management, data cataloging and data fabric)
- Presentation of human knowledge and business acumen, composing a good component to be leveraged by multiple ML and AI models, including large language models

Drivers

- Ongoing digitization and globalization initiatives with complex dynamics creating a need for more adaptive and integral approaches, as offered by knowledge graphs, replacing more static and siloed approaches.
- Increasing awareness of the use of knowledge graphs in consumer products and services such as smart devices and voice assistants, chatbots, search engines, recommendation engines and route planning.
- The emerging landscape of Web3 applications and the need for data access across trust networks, leading to the creation of decentralized knowledge graphs to create immutable and queryable data structures.
- Improvements in graph DBMS technology that can handle the storage and manipulation of graph data structures at scale. This includes PaaS offerings that take away the complexity of provisioning and optimizing hardware and infrastructure.
- The desire to make better use of unstructured data held in documents, images and videos, using standardized metadata that can be related and managed.
- The need to manage the increasing number of data silos where data is often duplicated, and meaning, usage and consumption patterns are not well-defined.
- The use of graph algorithms and machine learning to identify influencers, customer segments, fraudulent activity and critical bottlenecks in complex networks.
- Prompt engineering and fine-tuning of large language models need collaboration with the knowledge graph.

Obstacles

- Awareness of knowledge graph use cases is increasing but business value is difficult to capture in the early stages of implementation.
- Moving knowledge graph models from prototypes to production requires engineering and system integration expertise. Methods to maintain knowledge graphs as their size increases — to ensure reliable performance, handle duplication and preserve data quality — remain immature.
- Fragmentation of the graph DBMS market across the types of knowledge graph data models (resource description framework [RDF] or property), implementation architectures (native or multimodal) and differences in optimal workloads continue to cause confusion and hesitation among adopters.
- The challenge of enabling data within organizations to be interoperable with external knowledge graphs to enable the ingestion, validation and sharing of ontologies and data relating to entities (such as geography, people and events).

User Recommendations

- Identify use cases where there is a need for custom-made knowledge graphs through the use of a pilot project that delivers not only tangible value for the business but also learning and development for data and analytics staff.
- Take an agile approach to knowledge graph development to decrease time to value. Assess the data needed to feed a knowledge graph, both structured and unstructured, creating a minimum viable subset that can be used to capture the information of a business domain.
- Utilize vendor and service provider expertise to validate use cases, educate stakeholders and provide an initial knowledge graph implementation.
- Include knowledge graphs within the scope of data and analytics governance and management. To ward against perpetuating data silos, investigate and establish ways for multiple knowledge graphs to interoperate and extend toward a data fabric.

Sample Vendors

Alibaba Cloud; Haizhi Stargraph Technology; Tencent Cloud; Transwarp; Vesoft; Zhejiang Chuanglin Technology

Gartner Recommended Reading

Top Trends in Data and Analytics, 2022

Graph Steps Onto the Main Stage of Data and Analytics: A Gartner Trend Insight Report

Working With Graph Data Stores

Quick Answer: What Is Data Fabric Design?

AI Design Patterns for Knowledge Graphs and Generative AI

Federated Machine Learning

Analysis By: Ben Yan, Mike Fang

Benefit Rating: High

Market Penetration: Less than 1% of target audience

Maturity: Emerging

Definition:

Federated machine learning aims to train a machine learning (ML) algorithm on multiple local datasets contained in local nodes without the explicit sharing of data samples. Federated ML helps to protect privacy, enables ML and specifically DNNs to use more data, resolves data transfer bottlenecks, and empowers collaborative learning for better accuracy.

Why This Is Important

China issued the Data Security Law and Personal Information Protection Law in 2021, and public awareness regarding data and privacy protection is rising. Federated ML (FedML) highlights an important innovation in training ML algorithms in a decentralized environment without disclosing sensitive business information. FedML enables more personalized experiences with local learning in smartphones, softbots, autonomous vehicles or IoT edge devices and enables organizations to build collaborative learning models across data silos.

Business Impact

FedML enables collaborative ML by sharing local model improvements at a central level while storing the data locally. It especially benefits the IoT, cybersecurity, privacy, data monetization and data sharing in regulated industries. Utilizing data across individuals, organizations and industries to benefit society is an attractive concept for both enterprises and public sectors in China.

Drivers

- The proliferation of privacy regulations requiring protection of local data during machine learning.
- With the increasing hype around edge AI, the data becomes distributed across multiple, heterogeneous edge devices and clouds. FedML allows organizations to keep the data in place.
- Data volumes are still growing rapidly, making it harder to collect and store big data centrally. This is especially pronounced in the IoT scenarios, where sensor data is collected on the devices; often, there is no time or reason to pass it centrally. Due to scalability issues, excessive power consumption, connectivity and latency, we see a move toward edge infrastructure in the form of federated computing architectures.
- Upstream and downstream partners in the supply chain need FedML to leverage data together across different data silos.
- As the large language model (LLM) evolves, research on federated LLM emerges so that a group of organizations could collaborate to train LLM together.
- Swarm (federated) learning, enabled by blockchain, is emerging as a promising approach in decentralized ML, uniting edge computing, peer-to-peer networking and coordination.
- FedML is often combined with other privacy enhancing computation techniques as complete secured computing solutions.

Obstacles

- Building trust between organizations for collaborative learning models takes time.
- The incentive mechanism needs to be defined and agreed on, with all parties engaged to keep participants motivated and keep the FedML group in the long run.
- System and data heterogeneity requires much coordination and standardization among systems to be fully functional.
- Enabling FedML requires a complete end-to-end infrastructure stack that integrates capabilities across DataOps, ModelOps, deployment, and continuous tracking/retraining, necessitating a high degree of implementation maturity.
- Creating a new, more accurate and unbiased central model from local model improvements can be difficult, as the diversity or overlap between local learners and their data may be hard to assess and may vary greatly.
- Security and privacy validation concerns require additional security-related techniques.

User Recommendations

- Apply FedML to create and maintain decentralized smart services or products while protecting the users' privacy and preventing the need to centrally collect massive amounts of data.
- Explore FedML use cases with upstream and downstream partners and look for opportunities to improve overall operation efficiency.
- Give a head start to decentral ML applications by deploying a common, centrally pretrained model while still providing personalization and contextualization by locally retraining the model based on local data and feedback.
- Continually improve decentralized ML applications with collaborative learning by repeatedly collecting local model improvements to create a new, superior central model and then redeploying it for decentral usage and fine-tuning.

Sample Vendors

4Paradigm; Ant Group; Baidu; BaseBit; Nuowei Tech; Tencent; WeBank

Gartner Recommended Reading

Innovation Insight for Federated Machine Learning

Quick Answer: Why Is Federated Learning Prominent in China?

Top 3 Priorities for Chinese Enterprises to Promote Effective Data Sharing

Sliding into the Trough

Data Governance

Analysis By: Tong Zhang, Julian Sun

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Adolescent

Definition:

Data governance is the specification of decision rights and an accountability framework to ensure the appropriate behavior in the valuation, creation, consumption and control of data and analytics. It encompasses the people, processes and technologies required to manage, protect and utilize data assets.

Why This Is Important

Data governance defines the guardrail of data assets. It provides a holistic approach to collecting, managing, securing and storing data and ensures value generated from data through the proper process with the right people using relevant technologies. The rapid business-value-focused D&A initiatives from Chinese organizations have left data governance risks unaddressed, which impedes the scale-up of D&A and organizational data-driven strategy.

Business Impact

- The accuracy, timeliness and reliability of data affect the quality of business decision making and data-driven insights. Data governance helps organizations to avoid the costs associated with poor data quality, such as rework, data cleaning and data integration.
- Effective data governance builds trust and confidence among stakeholders, including customers, investors, and employees, that the organization is managing data responsibly.
- Data is defined as one production factor in China. The economy of data requires a good governance guardrail and framework to be applied to all organizations.

Drivers

- Compliance with regulations such as China Cybersecurity Law (CSL), Data Security Law (DSL), Personal Information Protection Law (PIPL) are key drivers for implementing data governance to ensure that data is managed, secured, audited and utilized properly.
- As a new form of product, data becomes a key differentiator and could provide a competitive edge for businesses. Extracting business value from data relies on good data quality and various data elements enabled by data governance.
- Businesses rely on data governance to avoid risks such as data breaches, data loss, unauthorized access and the costs associated with insufficient management of data.

Obstacles

- **Inability to persuade business partners that data is a business concern, not an IT concern.** While business partners conceptually agree on the importance of data quality, consistency and access, they push back on taking direct ownership of data decisions, perceiving that data is an “IT thing.”
- **Difficulty to define and adopt consistent data governance processes and policies.** Without a formal governing body with both business and IT representation and a targeted scope, governance processes will fail to be adopted and benefit will be diluted. The comparatively unstreamlined and vague business processes in China have amplified this problem.
- **Failure to come to consensus on common enterprise data definitions.** Business unit heads tend to view their data needs as unique rather than consistent with other areas of the enterprise.
- **Inconsistent approach to data across projects.** An inconsistent approach across the enterprise puts project solution outputs at risk of being at odds with each other and fails to achieve scalability, slowing down projects as tasks and documentation are reinvented for each project. The lack of mature data governance vendors has amplified this problem.
- **Difficulty to define and sustain a path toward a target state of data competency.** Achieving data quality, accuracy and consistency is a vast and amorphous objective. Organizations struggle to identify a realistic target state of data competency and the interim milestones to reach in the near term and midterm.

- **Uncertainty of regulation changes makes data governance more compliance-driven rather than value-driven.** Data governance is expected to deliver value to business rather than to avoid risk. Business needs to manage risk to generate value.

User Recommendations

- **Clearly scope the data governance mandate and objectives.** Clarity of purpose will provide the focus to apply scarce resources to the most important governance activities.
- **Apply varying degrees of governance depending on the data asset.** Prioritize the highest level of rigor on the data assets with the highest business value and broadest enterprise use.
- **Involve the business in data stewardship roles and tasks.** Data is a business asset and needs to be managed as such by business owners. Ensure business representation on stewardship boards.
- **Take an opportunistic approach to data standardization.** Pursue opportunities for data standardization and integration that are highlighted by events in the business environment, such as a merger or new senior executive, that would trigger new information requirements and needs.
- **Prioritize data quality improvement based on the data's quality gap relative to its importance to the organization.** Highlight data quality concerns from multiple user perspectives and in varying contexts to ensure the right data is targeted for quality improvement.
- **Study China's National Data Bureau regulation and China data policy updates.** Exploring new opportunities is key to success in data governance. The National Data Bureau will set the direction of new use cases of data.

Sample Vendors

Alibaba; BONC; Datablau; ESENSOFT; Huawei; JKStack

Gartner Recommended Reading

Playbook: Building a Modern Data Governance Program

Distilling Data Governance Essentials for FP&A Leaders

Market Guide for Data and Analytics Governance Platforms

Data and Analytics Governance as a Business Capability: A Gartner Trend Insight Report

3 Ways to Promote Your Data Agenda at the Center of the Chinese Digital Economy

Data Middle Office

Analysis By: Fay Fei, Xingyu Gu

Benefit Rating: Moderate

Market Penetration: 5% to 20% of target audience

Maturity: Early mainstream

Definition:

Data middle office (DMO) is an organizational and technical practice that empowers users from different lines of business to efficiently use enterprise data to make decisions through a single source of truth. Think of building a data middle office as the way a company curates the composable and reusable data and analytics (D&A) capabilities that deliver distinct digital operations and connect them throughout the value chain via its technology stack.

Why This Is Important

Many companies in China adopt data middle office practices to reduce technology redundancy in their data and analytics architecture, bridge data silos across different systems, and drive reusable data and analytics capabilities.

Business Impact

Data middle offices optimize data utilization and eliminate D&A silos by:

- Triggering CIOs/D&A leaders to rethink how to build business-driven D&A with a composable architecture. Business users could access and utilize the enterprise's information resources with flexibility and agility.
- Redirecting organizations to leverage and complement the existing assets instead of rebuilding everything.
- Transforming the organization's project-oriented D&A initiatives into platform-oriented initiatives by continuously building up reusable data analysis capabilities.

Drivers

- The well-known success stories of the data middle office in some industries and organizations has made it a key project for digital transformation across a wide range of industries.
- The growing capabilities from local vendors, especially on metadata management, data integration and data preparation based on open-source technologies, will improve the practical implementation of the data middle office.
- It has been promoted by vendors and the market as the most advanced and holistic D&A platform and, thus, vital to all organizations.

Obstacles

- Data middle office has been diluted in the market since it fails to deliver its commitment to composable and agile D&A capabilities in many cases. Organizations that are not digital-native are not prepared to build reusable D&A capabilities within their own business context.
- The goal and the value proposition of DMO is often not clear, which leads to vague project scope. Business users underestimate the challenges involved in — and the resources required to — implement a data middle office. They subsequently face overwhelming obstacles in the process of cooperating with the implementation, making DMO seem like an impossible goal.
- DMO uses a “collecting” data management practice that moves the data into one place and lacks a “connecting” practice that links diverse data from different systems. This results in a long time to value and fails to react to the fast pace of changes.
- Instead of focusing on constant business outcomes, the traditional project delivery method focuses on the go-live time stamp, which cannot retain users and create sustainable value.
- Most of the organizations in China lack the data literacy to engage broader citizen data users with tools.

User Recommendations

- Switch the focus from the "data middle office" terminology itself to nurturing core data and analytics capabilities behind it.
- Establish an ongoing data literacy program, which is essential for the data middle office to act as a "force multiplier."
- Use composable D&A (e.g., analytics packaged business capabilities and data fabric) as a guiding principle for the vision and architecture of data middle office.
- Define the reusability of data and analytics capabilities via business exploration by creating a hybrid organizational model that features a centralized team working with decentralized lines of business teams through the organization.

Sample Vendors

DEEPEXI; Kangaroo Cloud; Kejie; Mininglamp Technology; Shulan Technology; StartDT

Gartner Recommended Reading

Demystify Data Middle Office by Nurturing Core D&A Capabilities

Video: Demystifying the Data Middle Office

From Logical Data Warehouse to Data Fabric

Data and Analytics Essentials: Architect an Analytics Platform

Market Guide for Analytics Platforms, China

Citizen Data Science

Analysis By: Fay Fei, Julian Sun

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

Citizen data science is the collective set of capabilities applied to deliver analytic insights where the personnel are not the experts and their role or job function may not be within the data and analytics (D&A) discipline. Citizen data scientist is a persona rather than a title or role within an organization.

Why This Is Important

- Citizen data scientists deliver business-oriented insights and add to the impact of the D&A discipline on the organization through the creation and delivery of analytics models.

- Unlike the academic or expert data scientist, the functional knowledge of citizen data scientists adds a dimension of efficiency, efficacy and depth to the solutions and experience. With emerging technologies, citizen data scientists often unlock new insights beyond descriptive and diagnostic capabilities.
- The cost of professional data scientists is high due to the shortage in the China market.

Business Impact

With proper data science and machine learning (DSML) tools and training, citizen data scientists could infuse their industrial expertise into analytics life cycle (such as feature generation and selection, and algorithm selection). The business-oriented analytics model or insights found by citizen data scientists are often aligned to making business decisions with more expected outcomes.

Drivers

The most significant drivers of citizen data science include:

- **The increasing need of data science talents from business teams:** The sheer volume of personnel needed continues to outstrip demand in China. Citizen data scientists help fill a portion of the talent gap. Historically, building data science and machine learning (DSML) models required expert data scientists, who are difficult and expensive to hire and retain. Citizen data science helps overcome such limitations.

- **Data science requires in-depth functional knowledge:** Citizen data scientists' primary knowledge base is an in-depth understanding of the business domain. This combination of functional knowledge, data science skills and technology drives results.
- **Product maturity of vendors' offerings is improving:** Vendors have recognized this additional population as a target-rich environment for their offerings. As such, many vendor offerings now commonly include tools and features designed specifically for usage by citizen data scientists.
- **Augmented analytics blurs the boundary between BI and data science:** These include automated, streamlined data access and data engineering; augmented user insight through automated data visualization and exploration; modeling and pattern detection, including feature engineering, model selection and validation; automated deployment and operationalization; and capabilities to support collaboration and sharing.

Obstacles

- Upskilling in advanced DSML techniques and approaches is important to derive value from citizen data science. Considering the industrial knowledge needed, citizen data scientist is still not a common role in Chinese organizations.
- In many scenarios, especially for new projects, there is still a need to (statistically) validate the results of citizen data science by expert data scientists.
- Citizen data science is deemed to be just a preliminary, elementary step and not a fully functional DSML approach. Leaders in Chinese organizations often underestimate this discipline because the role is hard to be matched, and the enablement and education cost is high.
- There is a lack of standards for citizen data science working models and principles in the China market. Citizen data scientists often work in silos with no oversight or collaboration among experts and others with a vested interest in DSML.

User Recommendations

- **Success starts with leadership:** Educate business leaders about the potential impact of a broader range and larger pool of delivery capability. Work with leadership to scan opportunities for citizen data science to complement existing analytics and expert data science initiatives.
- **Inviting and inclusive environment:** Create communities of practice, and provide training and tools to make an inviting and supportive environment for all to explore the value of the citizen data scientist persona.
- **Expert data scientist value:** Acknowledge that you may still need specialist data scientists to validate and operationalize models, findings and applications.
- **Tools and technologies:** Provision augmented analytics tools (including but not limited to augmented data science and machine learning tools), platforms and processes to support and encourage collaboration between business users, application developers and data science teams. Track the capabilities (technology) and roadmaps of existing business intelligence (BI) and data science platforms and emerging startups for support of augmented features.

Sample Vendors

4Paradigm; Baidu AI Cloud; DataCanvas; Heywhale; Transwarp; Profet AI; Sensors Data

Gartner Recommended Reading

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

Build a Comprehensive Ecosystem for Citizen Data Scientists to Drive Impactful Analytics

The State of Data Science and Machine Learning

Market Guide for Multipersona Data Science and Machine Learning Platforms

Best Practices to Enable Effective Citizen Data Science

Climbing the Slope

Real-Time Data Management

Analysis By: Xingyu Gu

Benefit Rating: Transformational

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Definition:

Real-time data management is a set of technologies needed to enable real-time automated situation awareness and near-real-time responses to emerging threats or opportunities. It consists of database management systems with real-time capability, and event stream processing (ESP) and change data capture (CDC) based data integration, to sustain the modern real-time digital use cases.

Why This Is Important

Real-time data management is the set of data management technologies needed to sustain modern digital use cases that require real-time or near-real-time reactions to external change. There are two types of use cases: augmented transactions embedded in business processes, and near-real-time human reactions. Both of them have reached early mainstream adoption because of maturing DBMS and data integration (DI) technologies, and will eventually be adopted by multiple departments within every large company.

Business Impact

Real-time data management plays an essential role in most transformational digital use cases and innovations in China, including metaverse, smart manufacturing, smart city, autopilot, quantitative calculation, and others. Many of these use cases are well-established concepts, and have been under practical exploration for many years. With the maturing of real-time data management, these use cases will further evolve, and finally become the new norms in Chinese society in the next 10 years.

Drivers

Five factors are driving real-time data management growth:

- Fast-evolving social infrastructures and digital lifestyle penetration in China have led to ever-increasing amounts of data from sensors, cameras, digital control systems, corporate websites, transactional applications, social computing platforms, news and weather feeds, data brokers, government agencies, and business partners.
- Real-time data management technologies, including time series database (TSDB), in-memory database, event stream processing (ESP) and change data capture (CDC), are rapidly maturing, which makes organizations more capable to handle the real-time data mentioned above in an efficient architecture.
- The highly competitive business environment in China makes local business leaders demand more real-time, continuous intelligence for better situational awareness and faster, more-precise and nuanced decisions.
- Open-source technologies have made it less expensive for more vendors to offer real-time data management capabilities. At the same time, cloud offerings of these capabilities have lowered the hurdle of deployment.
- Chinese vendors are offering easy-to-use development interfaces, predefined business functions, and industrial services that enable faster implementation. Power users can build real-time applications using low-code techniques and off-the-shelf templates.

Obstacles

- Less-mature local data ecosystems lead to huge integration efforts to embed real-time data management technologies into existing DBMSs and data middle offices. Complicated use cases requiring both real-time and historical data still need significant implementation efforts to ensure performance and data consistency.
- A wide variety of programming languages in real-time data management technologies has led to technology hurdles and high study curves. Some local vendors have started building their integration platform as a service (iPaaS) or in-DBMS machine learning (ML) functions to address this problem, but the solutions are still emerging.
- Increasing performance of architectures supporting near real time provides a wider offering to users. However, these “good enough” products also add complexity for clients evaluating whether they are capable of handling real-time use cases.

- Many architects and software engineers are still unfamiliar with design techniques that utilize real-time data. They are more familiar with processing data at rest in databases and other data stores, so they use those techniques by default unless business requirements force them to use real-time data.

User Recommendations

- Educate business leaders about real-time data-management-enabled use cases and their importance. Brainstorm with them to identify concrete opportunities to rethink business processes and create applications that could not be implemented using traditional architectures.
- Use vendor-supported closed-source platforms or open-core products with value-added closed-source extensions for mainstream applications that require enterprise-level support and a full set of features. Use free, community-supported, open-source technologies only if developers are familiar with open-source software, and if license fees are more important than staff costs.
- Select products optimized for stream data integration with existing mainstream DBMSs and data middle offices.

Sample Vendors

DataPipeline; DolphinDB; KaiwuDB; RapidsDB; SelectDB; SF Technology; TDengine

Gartner Recommended Reading

Market Guide for Event Stream Processing

5 Essential Practices for Real-Time Analytics

Time Series Database Architectures and Use Cases

Exploiting the Evolving Database Management System Trends in China

Self-Service Analytics

Analysis By: Fay Fei, Julian Sun

Benefit Rating: Moderate

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Definition:

Self-service analytics (SSA) refers to technology and processes in which line-of-business professionals are enabled to autonomously prepare and visualize data, perform queries, and generate reports, with nominal IT support or involvement. SSA is often characterized by low-code/no-code tools that are increasingly augmented via AI. These tools provide increasingly sophisticated data preparation and analytics capabilities, but are simplified for ease of understanding and frictionless data access.

Why This Is Important

Self-service analytics can be a way for business users to easily create analytics prototypes and pilot them faster than relying solely on IT for D&A projects. Organizations in China are looking for the fast delivery of analytics content and to speed up time to insight. Unlike the traditional centralized method, self-service analytics offers a decentralized way to create data and analytics content, with autonomy and flexibility to quickly adapt to a changing environment.

Business Impact

Self-service data and analytics is critical for scaling the benefits of data-driven decision making. Many centralized D&A functions struggle to keep up with requests for data and insights coming from decentralized teams. Emerging citizen analysts or citizen data scientist personas who understand the business context of the data are able to use powerful no-code/low-code data preparation and analytics platforms to quickly discover insights.

Drivers

- Chinese analytics vendors are building self-service capabilities into their products. With the slowing momentum of enterprise reporting, vendors are investing and broadcasting their self-service capabilities to their clients.

- AI, ML and other augmented technologies lower the threshold of operating analytics. After fundamental training, users with no technical skills are able to create analytics content by themselves.

- The recent hype of generative AI with unprecedented capabilities has been quickly captured by the analytics market and will change the way users interact with and analyze data. Everyone can be an analytics producer.

- As business users advance in terms of information requirements, they expect to be able to extend the use of self-service into data preparation and data wrangling for analytics purposes. It enables more business users to conduct end-to-end analytics by themselves.
- Budgets for purchasing D&A tools are increasingly coming from business units and not just from central IT/data teams.

Obstacles

- Chinese organizations often lack official strategies and rules for managing the process and promoting the content from self-service data and analytics. The unclarified rights and obligations increase the concern and hinder people's interest to be self-served.
- Governance of self-service analytics is a common concern. With the self-service analytics model, the exponential proliferation of analytics content will create a huge burden for IT governance. Organizations need to achieve a balance of agility and control.
- Self-service analytics require business users to have a certain level of data literacy to understand and interpret data and to manipulate the data preparation process. The overall data literacy level in most Chinese organizations impedes the spread of the self-service analytics practice.
- High-quality data is still a struggle for many organizations. Despite having powerful tools, poor data quality can lead to greater potential for misunderstanding, misuse of the data and poor decision making.

User Recommendations

- Segment your users by their ability and inclination to become self-servicing, and deliver to the most prepared users first. Success often compounds and drives further successes, and it improves D&A maturity over time.
- Evaluate data catalog and self-service data management capabilities to allow business users to add curated or external sources to their data landscapes.
- Form communities of self-service D&A creators and consumers alike. Self-service should not be self-serving; communities that include sharing, collaboration, education, project overviews and success evangelism are critical as analytics audiences grow.
- Build data literacy and certification programs to ensure users are prepared to add value from self-service without mistakenly delivering bad or siloed information.

Sample Vendors

FanRuan Software; Hangzhou Guanyuan Data; Hengshi Technology; NetEase; SmartBI; Yonghong Technology

Gartner Recommended Reading

Data and Analytics Essentials: 3 Steps to Implement Self-Service Analytics

Quick Answer: 4 Easy Ways to Promote Trust in Self-Service Analytics

Magic Quadrant for Analytics and Business Intelligence Platforms

Market Guide for Analytics Platforms, China

Critical Capabilities for Analytics and Business Intelligence Platforms

Data Classification

Analysis By: Anson Chen

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Early mainstream

Definition:

Data classification is the process of organizing information assets using an agreed-upon categorization, taxonomy or ontology. This can include applying a tag or label to a data object to facilitate its use and governance, employing controls during its life cycle or activating metadata using data fabric. Typically, data classification results in a large repository of useful metadata for making informed decisions.

Why This Is Important

Data classification facilitates effective and efficient data prioritization of data within data governance and data security programs concerned with value, access, privacy, storage, ethics, quality and retention. China's data security regulatory requirements make data classification a vital step for security, data governance and compliance programs. Data classification helps organizations distinguish the sensitivity of the data and improves the effectiveness of data protection controls.

Business Impact

Data classification augments analytics on a dataset, structures data within repositories and allows immediate control over the use of data assets. Security controls such as data loss prevention (DLP) and data access governance (DAG) benefit enormously from data classification or labeling. Data classification enables organizations to meet regulatory compliance obligations cost-effectively by making data easier to find and validate while avoiding overprotection and retention.

Drivers

- The current legal and geopolitical situation has increased concerns regarding data residency and sovereignty, particularly for important data and personal information. However, the inefficiency of current firefighting data security governance practices boosted the desire to streamline and automate these processes, starting from data classification.

- The maturing data classification approaches, which include classification by type, owner, regulation, sensitivity and retention requirement, enable organizations to focus their security, privacy and analytics efforts on important datasets and their classification.

- The emergence of automated data classification tools featured with predefined industry-specific categories — for example, for finance, telecommunications, healthcare and government — lowers the amount of business and security knowledge required to initiate data classification programs.

Obstacles

- Legacy data classification initiatives have often failed because of insufficient training and dependence on user-driven classification.
- Data classification efforts mainly reflect a security-centric mindset. This means their purposes are not explained to users using business language and context, which results in low levels of engagement.
- Although many vendors offer automated data classification tools that can classify data more accurately while minimizing user effort, the accuracy of results does not meet expectations. This applies especially to machine learning or artificial intelligence algorithms for which models require ongoing training.
- From a compliance perspective, organizations not operating in heavily regulated industries and without classification standards published by the industry regulators might find it difficult to measure and justify the effectiveness of data classification results.

User Recommendations

- Determine organizationwide data classification use cases and efforts by conducting a thorough assessment of the types and sensitivity of data present within the organization and collaborating with business departments and data analytics teams to identify specific use cases where data classification is crucial.
- Implement user training and a combination of user-driven and automated data classification as part of a data security governance program.
- Analyze data classification guidance and standards released by industry regulators or national standard committees to develop a data classification scheme that aligns with regulatory requirements in China.
- Prioritize data classification tools that can be better integrated and interoperable with other data security technologies — such as anonymization, encryption, DLP and data security platforms (DSPs). Also, other aspects include richer built-in categorization templates and flexible self-defined tagging and labeling.

Sample Vendors

DAS Security; Guan An Info.; Meichuang Technology; NSFOCUS; Quanzhi Technology; Topsec Technologies; Wondersoft

Gartner Recommended Reading

Still a Moving Target — What to Do With the Chinese Data Security Law

Building Effective Data Classification and Handling Documents

Case Study: An Active Metadata Augmented Data Classification System to Boost Analytics Efficiency

Improving Unstructured Data Security With Classification

How to Succeed With Data Classification Using Modern Approaches

Appendixes

Hype Cycle Phases, Benefit Ratings and Maturity Levels

Table 2: Hype Cycle Phases

(Enlarged table in Appendix)

Phase ↓	*Definition* ↓
Innovation Trigger	A breakthrough, public demonstration, product launch or other event generates significant media 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 innovation is pushed to its limits. The only enterprises making money are conference organizers and content publishers.
Trough of Disillusionment	Because the innovation does not live up to its overinflated expectations, 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 innovation'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 innovation 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 innovation to reach the Plateau of Productivity.

Source: Gartner (July 2023)

Table 3: 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 2023)

Table 4: Maturity Levels

(Enlarged table in Appendix)

Maturity Levels	*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-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 2023)

Evidence

[1] See China Unveils Guideline on Improving Market-Based Allocation of Production Factors.

[2] 2022 Gartner China AI Use-Case Survey: This survey was conducted to understand AI implementations in China, and to understand where organizations have been most successful in deploying AI use cases. The research was conducted online from 14 November through 16 December 2022 among 300 respondents from organizations in China. Quotas were established for company sizes (in terms of annual revenue) and industries to ensure a good representation across the sample. Quotas included 45 small businesses (less than $50 million), 105 midsize enterprises ($50 million to less than $500 million), 120 large enterprises ($500 million to less than $10 billion), and 30 global enterprises (over $10 billion). Organizations were required to have developed AI to participate. Respondents were required to be in a manager role or above and have a high level of involvement with at least one stage of the life cycle from ideating to measuring AI use cases. *Disclaimer: The results of this survey do not represent global findings or the market as a whole, but reflect the sentiments of the respondents and companies surveyed.*

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

Understanding Gartner's Hype Cycles

Tool: Create Your Own Hype Cycle With Gartner's Hype Cycle Builder

Predicts 2023: Strategic Growth for China in Times of Uncertainty

Forecast Analysis: IT Spending, China

Chinese AI Survey Analysis: AI Trends Wave 3.0 — From Operational to Strategic

Exploiting the Evolving Database Management System Trends in China

3 Reasons to Invest in a Metadata-Driven Data Fabric Design in China

Demystify Data Middle Office by Nurturing Core D&A Capabilities

Quick Answer: China Perspective — Frequently Asked Questions on ChatGPT and Large Language Models

3 Ways to Promote Your Data Agenda at the Center of the Chinese Digital Economy

---

Table 1: Priority Matrix for Data, Analytics and AI in China

Benefit	*Years to Mainstream Adoption*			
↓	***Less Than 2 Years*** ↓	***2 - 5 Years*** ↓	***5 - 10 Years*** ↓	***More Than 10 Years*** ↓
Transformational	Data Asset Management	Citizen Data Science Data Sharing Decision Intelligence Generative AI Real-Time Data Management	Composable D&A Foundation Models Responsible AI	Data Fabric
High		Data Classification Data Governance Data Literacy Digital Ethics Homegrown AI Chips	AI Engineering Causal AI DBMS Self-Sufficiency Federated Machine Learning Knowledge Graphs	
Moderate		Self-Service Analytics	Augmented Data and Analytics Lakehouse	
Low				

Source: Gartner (July 2023)

Table 2: Hype Cycle Phases

Phase ↓	*Definition* ↓
Innovation Trigger	A breakthrough, public demonstration, product launch or other event generates significant media 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 innovation is pushed to its limits. The only enterprises making money are conference organizers and content publishers.
Trough of Disillusionment	Because the innovation does not live up to its overinflated expectations, 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 innovation'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 innovation 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 innovation to reach the Plateau of Productivity.

Phase	*Definition*

Source: Gartner (July 2023)

Table 3: 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 2023)

Table 4: Maturity Levels

Maturity Levels ↓	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-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 2023)