Gartner®

Hype Cycle for Data and Analytics Programs and Practices, 2023

Published 26 July 2023 - ID G00792543 - 160 min read

By Analyst(s): Donna Medeiros, Andrew White, Kevin Gabbard

Initiatives: Data and Analytics Programs and Practices; Evolve Technology and Process Capabilities to Support D&A

Data and analytics hype is at a fever pitch and inflection point with innovation and AI center stage. CDAOs are pressured to deliver on enterprise business value by driving growth, productivity and innovation. This research helps CDAOs navigate this challenging environment and plan offensive action.

Additional Perspectives

- Summary Translation: Hype Cycle for Data and Analytics Programs and Practices, 2023
  (12 October 2023)

Analysis

What You Need to Know

In this data-fueled era, chief data and analytics officers (CDAOs) must act and think differently as business executives. CEOs and boards realize data, analytics and AI can propel their business to new heights and have increased their appetite for risk accordingly (see 2022 CEO Survey — The Year Perspectives Changed and 2023 CEO Survey — The Pause and Pivot Year). Business stakeholders expect to use data and analytics ubiquitously to improve performance and uncover new opportunities. Chief data and analytics officers (CDAOs) must now propel growth and innovation across their enterprise by engaging across the C-suite, taking an offensive approach while simultaneously experiencing fast-paced change (see CDAO Agenda 2023: Presence, Persistence and Performance).

While individual innovations might be at different stages of hype, D&A overall is on a heightened peak this year, with the fuse as short as it has ever been to deliver value in stressful times. This is a make-or-break year for CDAOs; CIOs and CTOs are vying for their territory. With so many alternative technologies, design patterns and practices — there are too many options for CDAOs given the pressing need to realize value. The good news is you can ask the right questions of the opportunities and challenges in front of you and harness what is fit for purpose, bypassing the hype by focusing on business value and use cases.

A combination of compelling circumstances has converged, leading to a large array of D&A programs and practices. Confounding conditions — several perennial, others more recent — have all been increasing in intensity:

- Fast-paced generative AI has consumed attention; research is ongoing and established players are getting products into the market. Investments have not yet sufficiently transformed into reliable, repeatable and sustained business value for many enterprises — yet transformational potential exists in every industry.
- Data-driven enterprise directives are a challenge for CDAOs. Substantial time and resources to build data literacy capabilities and competencies across business lines that vary in D&A maturity.
- Serious D&A talent shortages continue to be severe, especially in AI/ML, data science, decision intelligence and data engineering. Over the long term, demographic changes will negatively impact skills availability.

- Real-time D&A demand continues to be a priority, with economic uncertainty and multiple crises impacting business and society. The pandemic demonstrated the necessity of D&A capabilities to detect and manage threats.
- In complex organizational structures, success can mean dissemination of D&A with increasingly indirect power over domain D&A.
- Data and information types, systems and organizational silos are unyielding in their complexity, both inside and outside the organization.
- The value of D&A products and ROI remains ambiguous to many.
- Working with ballooning amounts of data is complex, with myriad solutions positioned as “critical” or “must-have” to support business outcomes.

This Hype Cycle provides insight into prevalent innovations across all of D&A today. CDAOs and other D&A leaders will gain a clearer understanding of what is nascent, what is becoming more mainstream, and what has organizations facing implementation challenges where market solutions, talent and maturity have not yet caught up. It provides clarity to this chaotic time by making sense of hype and connecting the dots to pragmatically address, align and change your strategy.

The Hype Cycle

Data and analytics programs and practices operate as a set of core disciplines enabling a business ecosystem approach to thrive, which in turn enables the success of a value-based data and analytics strategy.

Over the past year, as enterprises grappled with hyped generative AI and continued uncertainty, D&A leaders made strategic gains in enabling trust for the enterprise to be data-driven. Key innovations that support this goal are data governance, data literacy, modern business intelligence and analytics, AI, ML, and advanced techniques such as data fabric and data monetization.

There is continuous innovation in and around the work of D&A. New entrants to this year’s Hype Cycle include:

- Generative AI
- AI sustainability
- Composable D&A

- Data ecosystems
- Decision engineer
- Data storytelling
- Dynamic trust-based model for governance
- D&A product management
- Data monetization

Some D&A leaders are focusing on transformative practices and emergent techniques such as generative AI. AI offers many benefits, but it is not easy to reliably generate business value from it without adopting the right best practices and foundational practices such as data governance and data literacy.

The shape of the Hype Cycle reflects that most disciplines and practices can take a long time to mature to the point where the mass market can gain repeatable value from them:

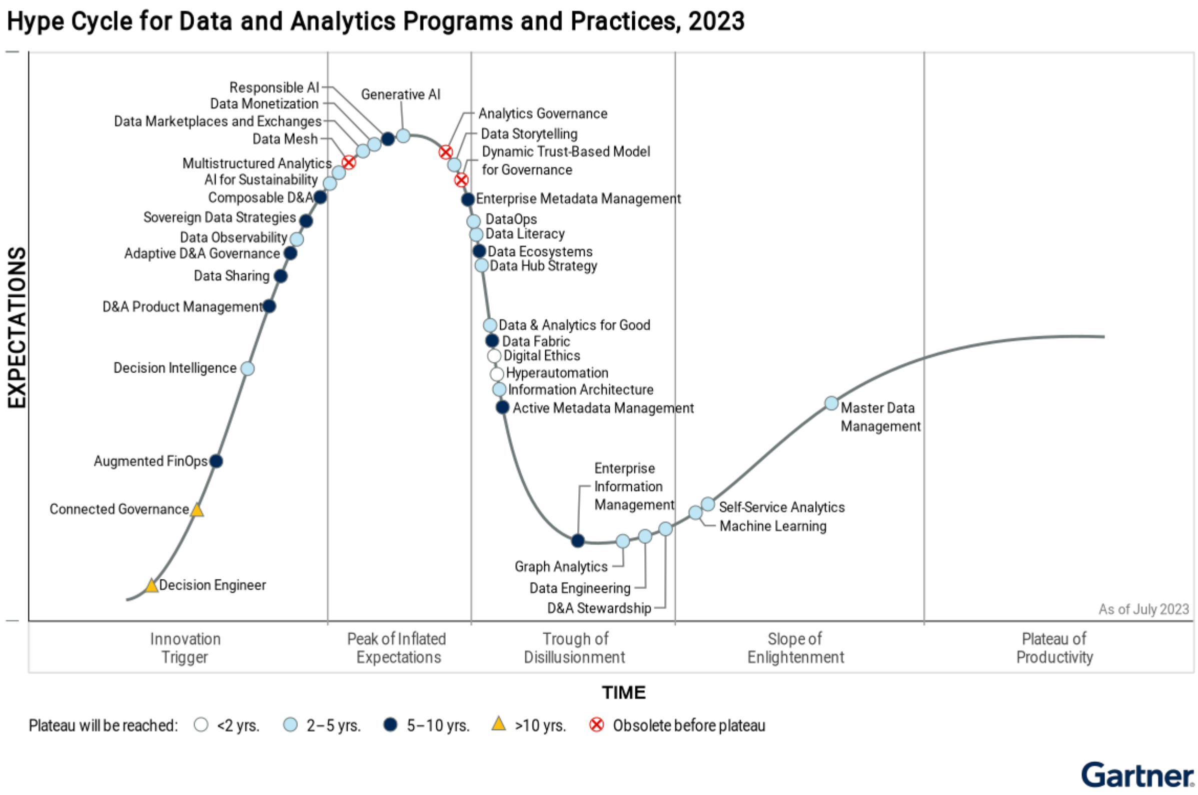
- Data literacy, at peak hype for several years, has moved into the Trough of Disillusionment as organizations struggle to deploy and scale requisite data and analytics skills across the workforce.
- D&A product management has leaped on near the peak as demands for business revenue growth for the private sector and societal value benefit for the public sector from D&A products explodes.
- Crisis and emergency management information systems reached the Plateau of Productivity within three years, demonstrating technologies can quickly reach mainstream when the business is at stake.
- The decision engineer role is hyped with hiring locally in high demand.

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

- Hype Cycle for Analytics and Business Intelligence, 2023
- Hype Cycle for Artificial Intelligence, 2023
- Hype Cycle for Data and Analytics Governance, 2023

- [Hype Cycle for Data Management, 2023]
- [Hype Cycle for Data Science and Machine Learning, 2023]
- [Hype Cycle for Data Security, 2023]
- [Hype Cycle for Privacy, 2023]

Figure 1. Hype Cycle for Data and Analytics Programs and Practices, 2023



The Priority Matrix

The data and analytics programs and practices on this Hype Cycle can be used to organize, implement and deploy any enterprise-level or business unit data and analytics initiatives. The IPs and matrix can also be used to align and link people, roles, business processes, data and technology across these initiatives. Strong programs will therefore be transformational in nature, and it will take several years before most of the practices reach the desired level of maturity.

Digital ethics is set to offer high-value benefits, given the more recent focus on ethics in particular with responsible AI by enterprise.

Hyperautomation adoption has accelerated by digital transformation efforts, with less than two years to wide-scale transformational benefit achieved.

A significant number of other initiatives and innovations will offer transformational or high-value benefits during the next two to 10 years (as shown in our Priority Matrix for Data and Analytics Programs and Practices). These include:

- Generative AI
- Adaptive data and analytics governance
- Data and analytics stewardship
- Data fabric
- Data sharing
- Decision intelligence
- Enterprise metadata management
- Responsible AI

Table 1: Priority Matrix for Data and Analytics Programs and Practices, 2023

(Enlarged table in Appendix)

Benefit	*Years to Mainstream Adoption*			
↓	*Less Than 2 Years* ↓	*2 - 5 Years* ↓	*5 - 10 Years* ↓	*More Than 10 Years* ↓
Transformational	Hyperautomation	Data Literacy Decision Intelligence Generative AI Machine Learning	Active Metadata Management Adaptive D&A Governance Augmented FinOps Composable D&A Data Fabric Data Sharing Responsible AI	
High	Digital Ethics	AI for Sustainability D&A Stewardship Data Engineering Data Hub Strategy Data Marketplaces and Exchanges Data Observability DataOps Data Storytelling Information Architecture Master Data Management Multistructured Analytics	Data Ecosystems Enterprise Information Management Enterprise Metadata Management Sovereign Data Strategies	Connected Governance Decision Engineer
Moderate		Data Monetization Graph Analytics Self-Service Analytics	D&A Product Management	
Low		Data & Analytics for Good		

Source: Gartner (July 2023)

Off the Hype Cycle

The following entries were removed for the Hype Cycle for Data and Analytics Programs and Practices, 2023:

- **Applied observability** has reached wide-scale adoption and thus moved off the Hype Cycle.
- **Chief data scientist and data science education** have achieved a level of mainstream adoption as the D&A function has continued to mature, with the data scientist considered a must-have role in D&A organizations.

- **Decision engineering.** The market focus has shifted from the practice (engineering) to the role (engineer) that exploits the practice. The role has been added to the Hype Cycle this year.
- **Data and analytics centers of excellence (COEs)** have achieved mainstream adoption.
- **Cloud D&A migration** has achieved maturity; this is now a given.
- **COVID-19 health risk mitigation** has moved off as the pandemic hype has waned and risk mitigation measures have lessened, if not completely gone away.
- **Crisis/emergency management solutions** moved off, given wide-scale adoption by enterprises as potential threats posed to business continue. Accelerated adoption of these solutions were triggered by the pandemic, and their need persists given war, conflicts, climate change, health threats and other events.
- **DaaS** has not been accepted in the market at the same level of SaaS, IaaS, and PaaS, and is also a form of SaaS and PaaS. Thus, interest has shifted toward things like data fabric and D&A ecosystem.
- **Vaccine management** has moved off the Hype Cycle as fully adopted. With the COVID-19 pandemic no longer a global public health emergency, this is mostly a program area remaining in some low and middle income countries (LMICs).

The following innovations have evolved, and the Hype Cycle has been adjusted accordingly:

- Cloud data ecosystems has evolved to “data ecosystems”
- Trust-based governance has evolved to “dynamic trust-based model for governance”

On the Rise

Decision Engineer

Analysis By: David Pidsley

Benefit Rating: High

Market Penetration: Less than 1% of target audience

Maturity: Embryonic

Definition:

Decision engineers apply analytics and engineering skills to decision intelligence platforms and practices. This practical discipline advances decision-making experiences with design thinking by engineering how decisions are modeled and made and how outcomes are evaluated and improved via feedback. Practitioners foster multidisciplinary collaboration in decision support, augmentation and automation by applying decision engineering to processes with embedded analytics, data science and AI.

Why This Is Important

Decision engineer is a decision-centric role that operationalizes embedded decision models, monitors feedback and optimizes outcomes with decision intelligence practices and platforms. The emerging role is not only focused on implementation. It is also essential in fostering multidisciplinary collaboration to bridge the gap between business domain and process experts on the one hand, and data scientists and AI experts on the other. They collaborate with the business, inventory decision models to manage and monitor, and report on the business value of decisions under management to promote reuse and deduplication of decision models.

Business Impact

Most decisions that currently use data are or soon will be at least partially automated. Decision engineers apply process, (computational) software engineering and mathematical techniques to help organizations make better decisions. Decision engineering leverages data, analytics, DSML, optimization and simulation to support decision augmentation and automation across a range of industries and contexts.

Drivers

- Though they recognize the need to collaborate, executives report that too many stakeholders and unclear decision ownership cause problems and delays in taking action. Instead of supporting multiple decision types for a single business unit, decision engineers can support a specific decision type, such as cost management or product improvements, across a number of business units.
- The shift from data-driven to decision-centric enterprise accelerates the demand for emerging roles that apply analytics and engineering skills to decision intelligence platforms and practices.
- Gartner identifies decision intelligence as a top strategic technology trend that is disrupting decision-making culture, and a decision engineer describes “who” plays a key role in this.
- Data and analytics leaders are upgrading their operating models, especially for organizations and people, to ensure they can enable dynamic business outcomes amid disruptive market conditions.
- Challenges in executing high-impact reengineering of decisions will accelerate common definitions of embryonic roles to have a high benefit and mature into productivity in the coming decade.
- Decision engineers bring a deeper understanding of how effective decision-making processes work, and they provide human and social perspectives. Some top data science teams will be rebranded as cognitive science or science consultancies, increasing diversity in staff skills.
- Skills in demand for decision engineering include data science, simulations, optimization, SQL, Python, R, DAX, VizQL, process methods, software engineering techniques, design thinking and communication skills.

Obstacles

- Decision engineers may become a role, but not a job title. Similarities to other roles like citizen data scientist, data steward or D&A translator lead to confusion among candidates and hiring managers.
- Embryonic roles have less than 1% of target market adoption and are undifferentiated from adjacent D&A and decision-making roles. Despite different focuses and responsibilities, these roles all involve working with data and using quantitative methods to augment and reengineer decision making.

- Skill and staff shortages are the top roadblocks for success in D&A initiatives. Late adopters will struggle to recruit decision engineers and must instead focus on upskilling, motivating and retaining decision intelligence talent.
- Failure to operationalize decision intelligence or embed decision models into workflows and business processes makes for ineffectively integrated decision making.
- Not employing adaptive governance of decisions to ensure ongoing optimization of business outcomes by establishing clear decision-making processes, proactively identifying and addressing issues, continuously refining and optimizing decision models based on data-driven insights, and aligning decision intelligence with the organization's goals and values.
- Existing organizational structures silo decision intelligence approaches which, in fact, go across domains. Tactical, functional decisions are often compartmentalized by technology vendor or product (e.g., CRM, ERP, HCM, FP&A).
- Localized implementations may create fragmentation in organizational units, where decisions are very similar but regulated differently.

User Recommendations

D&A leaders responsible for analytics, BI and data science solutions should:

- Evolve their D&A approaches to support data-driven decisions by empowering and supporting business units to embed D&A in business processes.
- Assess the impact of the transition from data-driven to decision-centric and update your operating model for decision intelligence practices.
- Assess the impacts that demand for decision engineering will have upon existing skills shortages and how to fill the role by considering how other companies do this (see sample vendors).
- Define decision engineers' roles, responsibilities, requirements and qualifications (a bachelor's or master's degree in computer science, mathematics, statistics, operations research or a related discipline).
- Foster and develop decision intelligence talent to address staff shortages by recruiting decision engineers and data scientists, forming fusion teams with business experts and fostering communities of practice.

- Define the role's key responsibilities as collaboration with business functions outside of D&A, decision modeling using frameworks, decision model management (especially deduplication, reuse and mitigation of model drift), valuation and data storytelling, and continuous learning and trendspotting.
- Involve relevant stakeholders in the business, D&A and adjacencies in a collaborative way by applying best practices to fill the role through upskilling, attracting recruits, motivating and retaining decision engineers.

Sample Vendors

Airbnb; Amazon; Google; LinkedIn; Meta; Microsoft; Netflix; Philips; Uber

Gartner Recommended Reading

What Are the Essential Roles for Data and Analytics?

Maverick Research: Data and Analytics Roles Will No Longer Be a Priority

The Future of Data and Analytics: Create Competitive Differentiation Through Better Decision Making

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

Redefining Analysts as Decision Experts (Philips)

Connected Governance

Analysis By: Saul Judah, Malcolm Murray, Andrew White

Benefit Rating: High

Market Penetration: Less than 1% of target audience

Maturity: Embryonic

Definition:

Connected governance is a framework for establishing a virtual governance layer across organizations and business functions, or legal entities, in multiple geographies to achieve cross-enterprise business outcomes. By connecting existing governance bodies within and across enterprises, its component-based approach enables complex business challenges to be addressed without adding further layers of bureaucracy.

Why This Is Important

Governance bodies for enterprise functions such as HR, risk, and data and analytics are typically adequate for addressing their individual domain areas. However, cross-enterprise and interenterprise governance challenges are increasingly difficult to overcome. Rather than creating yet another permanent governance body, connected governance leverages existing governance bodies through a virtual framework, providing strategic oversight and accountability management across them with minimum additional overhead.

Business Impact

Senior business executives and board members spanning multiple organizations, legal entities and geographies will find value in exploring connected governance to address cross-enterprise strategic issues and opportunities. Organizations anticipating mergers and acquisitions (M&As) will find value in connected governance, enabling both value and risk management to be addressed earlier and allowing experimentation with governance bodies prior to their formal adoption.

Drivers

- The fast pace of deglobalization and digitalization is putting pressure on senior leaders across multiple business functions to respond to business and regulatory demands at greater effectiveness and speed than they are able to with their existing capabilities. Existing governance bodies are designed to address their functional areas, but understanding accountability and decision rights across these proves very difficult. This is especially relevant when some of the functional areas exist in different legal entities and different countries, and the same business asset is subject to potentially conflicting governance policies.

- A key driver for adoption of connected governance stems from the limitations of existing approaches. Traditional approaches to cross-enterprise governance challenges have been to establish another layer of governance, which adds a greater overhead cost, creates another layer of bureaucracy and is often inflexible. Furthermore, some strategic challenges (such as M&A and business model changes) require a one-off response for governance, and creation of additional governance layers in these circumstances is an excessive drain on executives’ time without accrued benefit. Consequently, adoption of connected governance becomes an attractive option.

Gartner

Obstacles

- Connected governance leverages existing governance bodies, but some of these bodies may operate poorly. As a result, the value that connected governance offers may be depleted in organizations that are not already mature in their governance.
- Siloed governance efforts might reinforce those silos and prevent the benefits of connected governance without disruptive organizational change. Either way, inertia and local success of siloed governance will slow down the adoption of connected governance.
- Once the board of directors or executive committee has approved the cross-governance initiative, an executive leader is expected to shape the cross-governance response. However, this needs support and facilitation from a strategic governance office, which requires skills that are currently in short supply.

User Recommendations

- Evaluate whether connected governance would benefit your organization. If you operate in a complex environment, across multiple legal entities and geographies, there may be challenges that are difficult to address now. In such situations, put on the agenda of your executive committee meeting to initiate a cost-benefit assessment and report its findings. If this does not apply at your organization, connected governance may not be for you.
- Connected governance needs the support of strategic, cross-enterprise governance. Analyze whether this needs a dedicated governance office or if operating as a virtual governance office will be sufficient. If your strategic challenge is a one-off situation, or if you are trialing this as a new initiative, a virtual office may be sufficient for now. However, large enterprises in diverse, complex ecosystems and expecting to address many strategic scenarios may necessitate a dedicated strategic governance office to support connected governance.

Gartner Recommended Reading

Connected Governance Orchestrates Complex Cross-Enterprise Decisions

Connected Governance Drives Adoption of Data and Analytics Governance Platforms

Quick Answer: What Kind of Governance Does Healthcare Data Interoperability Need?

Choose the Optimal Corporate Structure to Cope With Geopolitical Risks

Trends 2023: Rise and Risks From EU, U.S., China and Other Sovereign Data Strategies and Policies

Augmented FinOps

Analysis By: Adam Ronthal, Dennis Smith

Benefit Rating: Transformational

Market Penetration: Less than 1% of target audience

Maturity: Embryonic

Definition:

FinOps applies the traditional DevOps concepts of agility, continuous integration and deployment, and end-user feedback to financial governance, budgeting and cost optimization efforts. Augmented FinOps automates this process through the application of artificial intelligence (AI) and machine learning (ML) practices — predominantly in the cloud — to enable environments that automatically optimize cost based on defined business objectives expressed in natural language.

Why This Is Important

In the cloud, it is now possible to assess the cost of a specific workload or collection of workloads assigned to a project. However, price/performance — the primary measure of cloud efficiency — is difficult to assess due to the complexity and diversity of choice in underlying cloud infrastructure and service offerings and a lack of consistency in pricing models. Augmented FinOps can automate this process by applying AI/ML techniques.

Business Impact

The automation of cloud budget planning and financial operations will allow businesses to express their objectives — ideally in natural language — and allow their cloud ecosystems to automatically optimize the underlying cloud resources to meet those objectives. This will result in more efficient use of resources and, therefore, optimal spend by reducing/eliminating misaligned or poor use of cloud infrastructure and service offerings.

Drivers

- Practitioners are increasingly realizing that cloud is fundamentally a complex cost optimization exercise.
- Cloud adopters have a strong desire for transparency into cloud spending.
- Buyer inexperience is leading to either under-provisioning and associated resource contention or overprovisioning and spending more than is needed.
- Vendors are positioning cost-effectiveness as a competitive differentiator in their go-to-market strategies.
- Practitioners need to reduce the unpredictability of cloud spending when using cloud infrastructure and services for analytics, operational database management systems (DBMSs), data lakes and other applications, including custom IT infrastructure.
- Consumption-based usage remains common in earlier stages of cloud adoption, driving the need for augmented FinOps, although commit-based usage mitigates some unpredictability.
- Cost overruns are often obscured, downplayed, or dismissed by line of business implementers, requiring augmentation to achieve holistic and comprehensive cost optimization.
- Automation of financial governance controls in cloud environments provides increased predictability and cost optimization with less operational effort.
- Solid financial governance frameworks are positioning organizations to take advantage of FinOps.
- Emergence of specific roles — like FinOps practitioner or cloud economist — focused on FinOps practices and cost optimization means organizations have the expertise to address augmented FinOps.
- Owing to their complexity, cloud environments are ideally suited for the application of ML and AI methods to automate processes and track price and performance.
- Core FinOps capabilities are being delivered in three ways: Homegrown solutions, cloud service provider (CSP) instrumentation and third-party vendors. Increasingly practitioners are seeking out third-party or CSP tools to address their needs. All of these have a broad objective of adopting augmented capabilities as a means of competitive differentiation.

Obstacles

- Cloud service provider pricing models remain needlessly complex and diverse.
- Cloud ecosystems are (and will remain) open to third-party participants, which implies multiple commercial arrangements with multiple providers.
- Standards for cloud cost, usage and billing data like the FinOps Foundation's FOCUS proposal have yet to be broadly adopted. APIs for communicating performance data within the context of a broader ecosystem have yet to emerge. Both of these are required to assess the primary measure of success: price/performance.

User Recommendations

- Seek out service offerings to automate (via AI/ML) performance, consumption and pricing options. Increasingly, incorporate these capabilities into cloud data ecosystems that will learn from consumption patterns as they seek to optimize the underlying resources, and by extension, cloud spending through orchestration and optimization.
- Apply Gartner's FinOps Maturity Model to assess FinOps offerings in terms of their ability to address the following core capabilities: Observe, report, recommend, predict and optimize. The last three introduce augmented FinOps capabilities.
- Plan to use multiple tools to address the full scope of requirements. Many tools are broad in reach, but do not go deep into prescriptive recommendations. Others are tightly scoped and provide very targeted optimizations. Expect to spend time combining multiple tools to achieve broad and deep capabilities.

Sample Vendors

Acceldata; Anodot; Apptio; Capital One Software; Densify; Enteros; Finout; OtterTune; Sync Computing; Unravel Data

Gartner Recommended Reading

How to Identify Solutions for Managing Costs in Public Cloud IaaS

A Guidance Framework for Selecting Cloud Management Tools

Emerging Tech: Data Management Product Leaders Must Implement Augmented FinOps in Their Cloud Solutions

CDAOs and CFOs Must Drive Business Value in the Cloud Through Collaboration

Financial Governance Is Essential to Successful Cloud Data and Analytics

Decision Intelligence

Analysis By: Erick Brethenoux

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Definition:

Decision intelligence (DI) is a practical discipline that advances 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 (including generative AI), is pushing DI toward the Peak of Inflated Expectations. Recent crises have revealed the brittleness of business processes. Reengineering those processes to be resilient, adaptable and flexible will require the discipline brought by DI methods and techniques. A fast-emerging market (DI platforms) is starting to provide resilient solutions for decision makers.

Business Impact

Decision intelligence helps:

- **Reduce technical debt and increase visibility.** It improves the impact of business processes by materially enhancing the sustainability of organizations' decision models based on the power of their relevance and the quality of their transparency, making decisions more transparent and auditable.
- **Reduce the unpredictability of decision outcomes.** It does so by properly capturing and accounting for the uncertain factors in the business context and making decision models more resilient.

Drivers

- **A dynamic and complex business environment, with an increasingly unpredictable and uncertain pace of business.** Two forces are creating a new market around decision intelligence platforms (DIPs). The first is the combination of AI techniques such as natural language processing, knowledge graphs and machine learning. The second is the confluence of several technology clusters around composite AI, smart business processes, insight engines, decision management and advanced personalization platforms.

- **The 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 phenomenon happens from both an IT and a business perspective.

- **Expanding collaboration between humans and machines.** This collaboration, supplemented by a lack of trust in technologies (such as AI), is 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 that 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.** Lack of explicit representation of decisions prevents proper harmonization of collective decision outcomes. DI remedies this issue.

- **Emergence of software tools in the form of decision intelligence platforms.** DIPs will enable organizations to practically implement DI projects and strategies.

- **Generative AI.** The advent of generative AI is accelerating the research and adoption of composite AI models, which are the foundation of DIPs.

Obstacles

- **Fragmentation:** Decision-making silos have created data, competencies and technology clusters that are difficult to reconcile and that could slow down the implementation of decision models.
- **Subpar operational structure:** An inadequate organizational structure around advanced techniques, such as the lack of an AI center of excellence, could impair DI progress.
- **Lack of proper coordination between business units:** The inability to impartially reconsider critical decision flows within and across departments (also because of fragmentation) diminishes the effectiveness of early DI efforts.
- **Lack of modeling in a wider context:** In 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.
- **Lack of AI literacy:** Many organizations still suffer from a lack of understanding when it comes to AI techniques. This AI illiteracy could slow down the development of DI projects.

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 predictive and prescriptive (optimization, business rules) analytics. Upskill business analysts, and develop new roles, such as decision engineer and decision steward.
- **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.
- **Accelerate the development of DI projects** by encouraging experimentation with generative AI and expediting the deployment of composite AI solutions.

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

AI Security: How to Make AI Trustworthy

Top Strategic Technology Trends for 2023: Adaptive AI

D&A Product Management

Analysis By: David Pidsley, Sarah Turkaly

Benefit Rating: Moderate

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Data and analytics (D&A) product management involves strategizing, designing, developing, operationalizing, and maintaining data products and analytics that cater to repeatable business needs and measurable value. Distinct from, but complementary to project-based approaches, it prioritizes business outcomes and metrics that matter to consumers (beyond project metrics) through iterative, continuous improvement, multidisciplinary standing teams, managed end-to-end and driven with outcomes in mind.

Why This Is Important

Our rapidly evolving business world requires swiftly accessible, integrated and governed data. D&A product management supports this by turning some data into reusable products tailored to specific business outcomes. It enables decentralized D&A consumption through collaborative, standing teams, extending the options for delivering D&A work from just projects and services alone.

Business Impact

D&A product management:

- Improves the self-service experience, empowering consumers with curated D&A outputs that evolve with requirements.
- Brings data products to market enabling reuse in analytics, business applications, and data science projects, and unlocks potential revenue from data sharing and monetization.
- Shifts mindsets from one of service provision to co-creation of value through data unifying technical and business teams through an alternative approach for delivery.

Drivers

- **Expedited decision making:** Product management approaches in D&A enhance the speed of analytical insights through the use of consolidated, preintegrated and governance-approved data products.

- **Trust building:** Comprehensive, prepackaged data products boost confidence among domain experts in supporting business outcomes.

- **Adopting agile practices in D&A:** Enables cross-functional efficiency, efficient delivery and enhanced scalability. Product-oriented teams cater to diverse needs, optimizing organizational structures. Agile methodologies in D&A, akin to DevOps in software development, enhance adaptability. They also facilitate the rapid reuse and scaling of data products for various use cases.

- **Business ROI enhancement:** A product-focused approach strengthens the correlation between D&A activities and key metrics of stakeholders, maximizing returns.

- **Adaptable delivery models:** Amid market volatility, it is crucial for organizations to flexibly adjust delivery models, prioritize talent retention and spur cloud-era innovations.

- **Potential revenue generation:** Organizations identify and exploit the revenue potential of D&A assets through internal and strategic data sharing, as well as external data monetization.

- **Improved resilience:** D&A product management methods bolster DataOps, ModelOps and XOps, leading to more robust pipelines and faster resolution times.

Obstacles

Organization and operating model:

- D&A product management doesn't replace other delivery models. It is a method to add to others but still requires considerations in adjusting your operating model.

- D&A models often disregard cross-functional value creation. Because of this norm, multiple handovers across silos can complicate collaboration.

- True product management focuses on continuous maintenance and iteration and requires a dedicated product manager role.

- Integrating "ops" processes (DataOps, MLOps, etc.) for those who do not already work in this way requires a culture change.

Governance and value:

- Training SMEs in data product governance and achieving consensus on KPIs can be challenging.
- Thorough planning and iteration on the data product life cycle are essential to avoid creating single-use data products and technical debt.
- Reuse of products across many use cases can be challenging as the governance, performance and SLAs can be significantly different across the use cases.

User Recommendations

- Adopt a responsible, accountable, consulted and informed (RACI) matrix to identify skills gaps in transitioning to product management practices in D&A teams.
- Seek product management skills from places where it was a profession, such as retail and consumer goods. This is not a technical skill — it is a set of competencies including product planning, operationalization and introduction. Identify product managers who can articulate business needs to data engineering teams.
- Implement agile practices, cultivate product managers, develop cross-functional roles and utilize models like Gartner Decision Intelligence for decision optimization.
- Avoid mistaking any dataset for a data product. Emphasize scalability and reusability.
- Develop a minimum viable data product focusing on key use cases.
- Promote D&A product management successes to secure ongoing investment and support.
- Apply FinOps to identify novel funding approaches for D&A products in the cloud, supporting minimum viable outcomes.

Gartner Recommended Reading

Toolkit: Create a RACI Matrix for Data & Analytics Product Management

Infographic: Accelerate Your Transition to Data and Analytics Product Management

Product Management Practices Crucial for Data and Analytics Asset Monetization

How to Apply DevOps and Value Stream Mapping to Data, Analytics and AI

3 Keys to Persuading Your CFO to Use Product-Based Budgeting

Data Sharing

Analysis By: Lydia Clougherty Jones

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Early mainstream

Definition:

Data sharing is a business-facing KPI of achieving stakeholder engagement and providing 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 is an ecosystem of ecosystems enabling value through collaboration.

Why This Is Important

Data sharing is an essential business capability, enabling access to the right data to derive the right insight. It satisfies demand for more robust predictive analytics, aggregating diverse data sources and driving otherwise unknowable insights for data-driven innovation. D&A leaders at high-performing organizations who promote data sharing have more stakeholder engagement and influence than those who do not. Organizations that align data sharing with business outcomes outperform their peers.

Business Impact

Data sharing impacts the business by:

- Increasing access to more relevant data to match business goals, enabling the realization of stakeholder priorities, including conflicting ones
- Improving organizational readiness for mandated data sharing, replacing compliance centricity with digital transformation success
- Reducing duplicated efforts/redundant spend on the same data asset

- Improving deployed/newly built model accuracy
- Driving operational excellence and environmental sustainability

Drivers

- Organizations need portfolios of diverse data from different sources to feed the increasing use of D&A. These often exist in silos within the organization and, increasingly, external to it.
- Data sharing is also subject to various sovereign data strategies, such as those found in the EU, China, Japan and other countries. These highlight data sharing as a key priority for increasing government efficiency, generating public value and encouraging industry-specific market growth.
- Data sharing between organizations and public sector agencies may be regulated or even mandated.
- As part of the European data strategy, the European Commission announced the creation of data spaces in 10 strategic fields (including health, agriculture, manufacturing, energy, mobility, financial, public administration, skills, open science cloud and Green Deal). The recently proposed EU Data Act requires personal data sharing.
- D&A leaders responsible for building a data-driven enterprise require expansive data sharing, increasing cross-enterprise discoverability of more relevant data to match business outcomes.
- Data sharing assists with the demand for relevant available data for AI training while alleviating cost pressures for processing large amounts of AI training data.
- External data has become increasingly relevant 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.
- The rise in synthetic data creation illustrates the demand for personal data sharing, as “real data” can be expensive, unavailable, or deemed unusable due to privacy regulations or the perception of regulatory prohibitions.
- 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 and improved risk mitigation.

Obstacles

- Organizations know that data sharing is a key business capability for digital business, but lack the "know-how" to share data at scale and with trust.
- There is an outdated perception that alleged risks of data misuse outweigh the business benefits of sharing data, including cost savings and revenue growth.
- Organizations must overcome internal data hoarding, external data hijacking, confidentiality weaponizing and privacy shaming.
- Stakeholders often may resist based on perceived regulatory restraints, fear, outdated or inflexible data management and governance policies, and lack of tools/technologies.
- Data sharing requires key capabilities, such as integration and governance of shared data, which must be addressed to provide business value for data from multiple diverse sources.
- Fascination with technology enablers can overshadow strategic business goals and the collaboration required for data sharing success.
- Failure to make the right investments in data sharing can impede the discoverability, reuse and resharing required for efficient and value-producing digital business outcomes.

User Recommendations

- Use Gartner's "Must Share Data Unless" model to align data sharing with business outcomes (see Why Data Sharing Is Important: Introducing Gartner's 'Must Share' Model).
- Promote data sharing as a "business necessity."
- Apply the level of trust across the data sharing ecosystem commensurate with desired business outcomes (see Why Situational Trust Is Key to Data Sharing for Business Value).
- Communicate widely the risk of digital business failure (e.g., increased inefficiencies/lost opportunities for value creation) caused by not sharing data. Modernize data management, adopt trust-based D&A governance, and foster a data sharing culture, not a data "ownership" culture.
- Adopt data fabric architecture design to transform a siloed environment to a coordinated one, enabling a single architecture for data sharing across heterogeneous internal and external data sources.
- Deploy graph analytics to increase awareness of unknown relationships in combinations of diverse data, including personal data.
- Participate in or capitalize on external data sharing collaborations, even among competitors, driving organization and industry business outcomes.

Sample Vendors

Agdatahub; Brighthive; Explorium; Google; Tyler Technologies; Vendia

Gartner Recommended Reading

Data Sharing Is a Business Necessity to Accelerate Digital Business

How to Optimize Data Exchanges to Realize Business Value

Modernize Your MDM Program With External Master Data Sharing

Case Study: Data Governance to Establish Enterprisewide Data Sharing (B3)

How CDAOs Need to Prioritize Data Sharing Investments for Digital Business Success

Adaptive D&A Governance

Analysis By: Saul Judah

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Adaptive data and analytics (D&A) governance is an organizational capability that enables context-appropriate governance styles and mechanisms to be applied to different D&A scenarios to achieve desired business outcomes.

Why This Is Important

As organizations accelerate or scale out their digital business initiatives, ecosystems and platforms, their ability to deliver expected business value is limited by their current business practices — in particular, their governance of D&A assets. Despite greater diversity and complexity in business scenarios than ever before, D&A governance has typically continued to adopt a single, control-oriented approach, which is often unresponsive to business needs and causes or reinforces data silos.

Business Impact

Adaptive D&A governance has the potential to be a transformational change agent for digital business. It enables application of different governance styles (control, outcome, agility and autonomous) to different D&A scenarios, depending on business context. This allows better enterprise collaboration in D&A initiatives, allowing enterprises to respond faster to opportunities and become more competitive, resilient and risk-aware.

Drivers

- As levels of risk appetite and demands for growth have risen in organizations, so have expectations for flexibility and agility from D&A initiatives to meet these needs. As a result, chief data and analytics officers are increasingly turning to adaptive D&A governance practices that enable the greater flexibility, scale and resilience needed in D&A initiatives to deliver dynamic business outcomes.
- Both D&A and business leaders recognize that increased investment in infrastructure, such as D&A platforms, cannot yield the expected ROI without corresponding improvement in D&A governance practices.
- Organizations maturing in D&A increasingly recognize the key role that business leaders play in driving their governance initiatives. Business demand for greater flexibility, agility, responsiveness and interconnectedness of D&A requires better governance practices than currently exist. This, in turn, is leading D&A leaders to explore adaptive D&A governance.

Obstacles

- Although D&A governance practices are maturing in many organizations, maturity is still lower than in other areas, such as data management and analytics. Many organizations still take an IT-oriented, center-out, single-style approach to governance, which resembles compliance rather than governance. This is outdated and wrong and needs to change.
- Poor data literacy is prevalent in organizations. Business leaders often fail to understand or accept accountability for the information assets they create, instead expecting their data office (typically residing in IT) to sort out their data. When data offices initiate governance initiatives, business leaders fail to engage effectively, or at all.

User Recommendations

- Use the IT Score for Data & Analytics to evaluate your maturity and readiness to enhance governance capabilities. Don't establish agility and autonomous governance without foundations for control- and/or outcome-based governance.
- Create a proof-of-concept (POC) initiative to test the applicability of one of the advanced governance styles (like an autonomous governance style) in your environment; evaluate the business outcomes and value, emerging risks, technological limitations and cultural barriers to wider adoption.
- Engage senior business executive leadership to discuss the results of the POC initiative. Create a business case and strategic roadmap to establish adaptive D&A governance.
- Establish the control and outcome styles of adaptive governance first; evolve to the agile and autonomous styles subsequently. Use minimum governance, focusing on limiting the scope of data, analytics and business processes to those that deliver greatest business value and organizational outcomes.

Gartner Recommended Reading

Data and Analytics Leaders Must Use Adaptive Governance to Succeed in Digital Business

2022 Strategic Roadmap for Data and Analytics Governance

Adopt SMART Principles for Adaptive Data and Analytics Governance

Next Best Actions to Improve Your Data and Analytics Governance

Data Observability

Analysis By: Melody Chien, Ankush Jain

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Data observability is a technology that supports an organizations' ability to understand the health of an organization's data, data pipelines, data landscape, and data infrastructure by continuously monitoring, tracking, alerting and troubleshooting issues to reduce and prevent data errors or system downtime. It tells us what went wrong based on agreed upon SLAs for data quality and usage; reasons; assesses the impacts; and recommends solutions. Data observability improves reliability of data by increasing our ability to observe changes, discover unknowns and take appropriate actions.

Why This Is Important

Data observability uses data profiling, AI/ML, lineage and active metadata to provide the following benefits:

- **Monitor & Detect:** Provide a holistic view to determine how components of data pipelines are operating, evaluate whether data quality meets expectations, and detect data related issues.
- **Alert & Troubleshoot:** Send right alerts to the right people at the right time and perform root cause analysis.
- **Resolve & Prevent:** Provide recommendations to fix the issues or optimize data pipelines to meet business requirements with the goal to prevent downtime or critical data issues before affecting business.

Business Impact

- Data observability allows technical teams to gain visibility of the health of data pipelines and infrastructure. They can identify possible drifts in various areas, and minimize the time to investigate and solve issues, preventing unplanned outages or critical data errors.
- Business users will also gain visibility of data quality and associated financial impacts. This will ensure appropriate use and management of data to meet governance requirements.
- Data observability allows facilitation and improvement of the data fabric with continuous observations and evaluations of the data and analytics ecosystem.

Drivers

- Data and analytics leaders face a growing number of mixed data stacks, diversity of datasets, unexpected data drifts such as change in schema or business context, high demand for data quality and near zero tolerance of downtime. All these add to the challenges in data management. They need a holistic view of the state of data quality and data pipelines within interconnected systems.

- Data pipelines move data from point to point and deliver data to consumers. This journey can be disrupted by unexpected events such as data quality issues or a lack of infrastructure resources. The data that flows through these pipelines needs to be monitored for loss of quality, performance or efficiency. Organizations need to be able to identify points of failure before they have a chance to propagate. Data observability automatically detects important events and analyzes various signals to troubleshoot the issues, and provides actionable insights of what to do next.

- Data observability goes beyond traditional monitoring. It provides a multidimensional view of data including performance, quality, usage and financial impacts to the downstream applications. Leveraging active metadata, lineage of data and AI/ML, data observability generates real-time insight by monitoring the business context and analyzing data pattern, comparing history, and developing a semantic understanding of the data. It provides an end-to-end observability to help organizations be better equipped to handle critical events and prevent business disruptions.

- This capability is essential to the data fabric design concept and becomes an important building block to further automation in data management practices.

Obstacles

- There is no standard definition of what constitutes a data observability solution. Vendors offer a range of different capabilities often branded as data observability which is causing confusion in the market and leading to issues adopting the tools.
- The current vendor landscapes are very fragmented based on coverage areas and data environments supported. Most vendors focus on observability of the data quality and data pipelines, and are less concerned about data usages and financial impacts. The full end-to-end observations are not quite there yet from individual vendors.
- Most data observability tools only support the modern data stack. This limits their application in large enterprise environments with more complex data environments in many cases using legacy data management tools.
- Most data observability tools target the data engineer persona and are positioned as IT tools. Though business users receive important insights from data observability tools, they may find them less user-friendly.
- Organizations are embracing the concept of "observability." But the actual adoption of the tools is not straightforward. The consideration of how they connect to the overall ecosystem and connecting this to data governance strategy is still a concern.

User Recommendations

- Identify the data elements or data pipelines which require high standards or SLA in quality, uptime, latency and performance. Pinpoint the gap of current monitoring capabilities vs. desired capabilities to support the requirements.
- Evaluate data observability tools available in the market that can enhance your observability based on priority of business requirements, primary users and interoperability with the enterprise data ecosystems.
- Pilot data observability program by building a monitoring mechanism as a starting point to increase visibility over the health of data. Invest in observability capabilities in a cloud environment first, as it's commonly supported by vendors and is faster and easier to demonstrate value.
- Include both business and IT perspectives when evaluating data observability tools by engaging with both personas early on in the evaluation process.
- Partner with business stakeholders to evaluate and demonstrate business value of data observation practices by tracking improvement of data quality, reduction in downtime and ability to meet SLAs to show tangible benefits.

Sample Vendors

Acceldata; Ataccama; Bigeye; Collibra; IBM; Kensu; Monte Carlo; Soda; Unravel

Gartner Recommended Reading

Data and Analytics Essentials: Data Observability

Quick Answer: What Is Data Observability?

The State of Data Quality Solutions: Augment, Automate and Simplify

Market Guide for DataOps Tools

Sovereign Data Strategies

Analysis By: Andrew White

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

Sovereign data strategies are state-controlled efforts to control data about its citizens and economy. Such strategies include regulations for privacy, security, access, use, retention and sharing, as well as processing and persistence.

Why This Is Important

Sovereign data strategies will impact every organization that does business across notable sovereign jurisdictions. They represent the coordinated approach from a sovereign entity to control and regulate how data is used in and across its economy and society. As such, these strategies and associated policies will impact data privacy, access, security, sharing, analysis, processing, storage and value.

Business Impact

Sovereign data strategies will impact where data about citizens and businesses can be stored, accessed, processed and used. Tracking such policies is hard and increasingly a challenge. The impact on organizations that trade or operate across sovereign boundaries can be as excessive as requiring you to replicate your entire operational business to exist and stay within certain sovereign boundaries. All storage and processing of such data would need to remain local and could not take advantage of the scale and synergy of global public cloud services.

Drivers

We won't analyze the politics driving sovereign motivations, but we can comment on the resulting policy or implied actions referenced by the strategies:

- Many sovereign data strategies refer to failures in public markets, public data sharing efforts, intellectual property use, and state or citizen personal data protection. The EU is a good example (see EU Data Strategy) as it seeks to create internal EU data markets to force (and reward) private firms to share data concerning the health of EU citizens. This data is to help solve health issues that otherwise would not be solved, according to the EU guidelines. Thus, the value in data will be exposed and increasingly shared.

- The U.S. (see Federal Data Strategy), China and several other states fall back on security concerns, ethics and the general well-being of the state as reasons for their sundry data strategies. Data privacy and data security are two very large drivers in this category; driving improved economic growth through better data sharing across public and private sector organizations is also referenced a lot. Together all sovereign strategies may use the same words but the motivations behind the policies and the focus tend to conflict and overlap, all at the same time.
- In 2023, there is growing recognition that geopolitical risk is elevating what started out as a set of policies and roles focused on data, to something far more visible. The shift is effectively from sovereign data strategies to sovereign digital strategies. Sovereign digital strategies describe how each seeks to govern and control access to their respective digital economy. As such, what started out as important to the CDAO and maybe the CIO is now important to the CEO and board of directors (see Choose the Optimal Corporate Structure to Cope With Geopolitical Risks).

Obstacles

- You may not always be able to order up a copy of your local sovereign data or digital strategy. Individual policies are observable even if the coordinated effort by the sovereign across all policies is not visible.
- Responding one policy at a time may well consume all your resources such that you run out of time and money should a clear sovereign strategy appear, and a coordinated response then becomes clear.
- Inability of CDAOs and other executive leaders, including CISO, CIO, and COO/CEO, along with Legal and Risk, can lead to inefficient and slow responses to the growing challenges.

User Recommendations

- Ensure your data and analytics (D&A) and digital strategy takes account of the constraints for those regions you operate in.
- Adopt a risk stance. Focus on mitigating risks as you develop global business strategies or seek to grow new businesses or services across sovereign boundaries. For example, invest tactically in privacy, security, and data sharing or processing efforts only when (and if) you have confidence in the reliability and enforcement of your targeted sovereign data strategy.
- Consider scenario planning to emulate costs and operational changes needed for extreme situations. Consider, for example, the need to air gap your entire IT landscape, or how you may advise the CEO/CFO on how data will be governed should parts of the business have to be jettisoned in worse-case scenarios.
- Don't panic or rush, as benefits may outweigh risks in a couple of years once sovereign data strategies stabilize and become more robust. At the end of the day, what was meant to be a global business environment and global cloud will become a fractured and distinct set of business environments and clouds.

Gartner Recommended Reading

Choose the Optimal Corporate Structure to Cope With Geopolitical Risks

Trends 2023: Rise and Risks From EU, U.S., China and Other Sovereign Data Strategies and Policies

Overcome the Challenges of Cross-Border Data Communication Over Internet With China

How New Privacy Laws in California and China Mirror the GDPR

Client Questions Video: How to Deal With EU-U.S. Personal Data Transfers (With Privacy Shield Gone)?

Composable D&A

Analysis By: Peter Krensky, Erick Brethenoux, Julian Sun, Carlie Idoine

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Composable data and analytics (D&A) utilizes container or microservices architectures and data fabric to assemble flexible, modular and consumer-friendly D&A capabilities from existing assets. This transforms monolithic data management and analytics applications into assemblies of D&A 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 are looking for flexibility in assembly/reassembly of D&A capabilities, enabling them to blend more insights into actions. Time to insight, reuse and agility are top requirements. Modular D&A capabilities enable faster and more proactive insight delivery.

Business Impact

The transition from monolithic D&A applications to composable D&A capabilities can be used with application development to assemble AI-augmented decision-making solutions. The focus of collaboration will transition from technology integrations to business problem solving. Organizations can create advanced analytics capabilities by composing the best capabilities from different vendors, rather than using them separately. Composability also relates to data fabric and data mesh in terms of being able to correctly identify data objects that exhibit wide reuse and separating them from those that are business-process-unique.

Drivers

- Container- or microservices-based analytics and business intelligence (ABI) and data science and machine learning (DSML) platforms with improved APIs enable the assembly of analytics applications in a more flexible way than custom code-based solutions.
- For most organizations, AI is still at the piloting stage, but ABI has been in production for years. Organizations can use composition to connect ABI to AI, extending ABI 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.
- Both D&A and software development teams will need composable data and analytics to enable emerging business technologies.
- As more data and analytics are integrated into digital platforms, traditional embedded analytics will need more modular capabilities to be assembled and reassembled for faster delivery.
- Embedded analytics are usually implemented by IT, but business users can use low- or no-code capabilities to source more data and compose more capabilities, such as interactive data visualization and predictive modeling, independently enriching more comprehensive embedded analytics.
- Cloud-based marketplaces are becoming an effective channel for organizations to distribute and share analytics applications, and composable D&A enables them to easily find the required components and add value to their applications by infusing analytics.

Obstacles

- New technologies and data have been the key drivers to evolve an analytics platform, resulting in less of a connection with business outcomes. Making data more accessible and composable often raises quality, governance and security concerns, among others.
- Software application development teams and data and analytics teams have not collaborated closely before. Composable D&A requires more involvement from the application development side, including applying XOps practices to maximize its value.
- Today's ABI and DSML markets are not zero-sum games. Many vendors of all sizes and specialties can thrive. No single vendor or tool offers all functions at the same level. It is unrealistic to implement a full D&A stack all at once, so many companies do so in stages. The composability of the existing products is not mature enough without technology partnership.

User Recommendations

- Improve decision making and business impact of data and analytics by incorporating and assembling modular, reusable D&A capabilities.
- Leverage composable analytics to drive innovation by incorporating advanced DSML capabilities into analytics applications.
- Exploit opportunities to add analytics capabilities to applications by building a joint team of application developers and business analysts with ongoing collaboration. 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.

Gartner Recommended Reading

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

Adopt Cloud Analytics to Drive Innovation

3 Steps to Build and Optimize a Portfolio of Analytics, Data Science and Machine Learning Tools

At the Peak

AI for Sustainability

Analysis By: Erick Brethenoux, Simon Mingay

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Definition:

AI for sustainability will improve business operations and optimize difficult-to-abate processes to reduce carbon and environmental footprint and mitigate material risks. AI can be made environmentally sustainable by using AI techniques that help create and run models at the lowest carbon footprint without partially or wholly compromising on accuracy. This extends to using AI to monitor, predict, mitigate and improve environmental issues.

Why This Is Important

Optimizing resource allocation and unveiling opportunities under extreme uncertainty are what combined AI techniques (composite AI) do best. AI can apply its resourcefulness to expand the social and governance aspects of ESG performance as well as deliver competitive advantage and reduced environmental impact via more efficient resource management.

Business Impact

Business impact of AI include:

- Monitoring and predicting (sense and interpret resource consumption, automate decisions and take actions)
- Optimizing resources and materials usage and enhancing circularity
- Reporting and disclosure (stakeholder-centric sustainability narratives and reporting)
- Mitigating energy costs and risks (hedge future consumption, leverage distributed energy resources and microgrids)
- Traceability and trust (record transactions and create a tamper-proof log of sensitive activity)

Drivers

AI for sustainability has both technology and use-case drivers.

Technology drivers:

- Composite AI
- Generative AI
- Adaptive systems
- Decision intelligence
- Causal AI
- Agent- and model-based reasoning methods and simulation techniques
- Advanced machine learning (ML) and optimization techniques (e.g., reinforcement learning, AutoML)
- Ambient intelligence (including computer vision)
- Verticalized AI

Use-case drivers have four categories.

Climate change:

- Monitoring, prediction and mitigation
- Flooding, drought and water management
- Wildfire and vegetation management
- Air quality management

Societal and environmental:

- Environment — social media monitoring (e.g., activism, citizen monitoring)
- Compliance and regulation alignment
- Society, human centricity and consumption

- Conservation and biodiversity management
- Agtech and fair trade

Sustainability and circularity:

- Equipment sustainability
- Waste management and recycling optimization
- Material productivity optimization
- Alternative resources and materials discovery

Optimization:

- Energy grid management
- Water and sanitation optimization
- Route optimization, transportation and mobility
- Sustainable buildings and facilities (including data center energy optimization)

Obstacles

- To implement AI techniques, enterprises embarking in AI engineering often seek "unicorn" experts to productize AI platforms. Few vendors provide AI engineering capabilities; as a result, such skills are hard to find.
- Lacking skills to leverage multiple AI techniques or fixating on custom-built AI when verticalized or augmented solutions are available could prevent organizations from efficiently solving specific problem types.
- Gaps exist in standards, responsible AI governance practices and data availability and/or quality (such as Scope 3 GHG emissions data).
- For organizations that have focused exclusively on technical skills, the other critical parts of human decision making — psychological, social, economic and organizational factors — have gone unaddressed.

- Rapidly evolving AI technologies — including tools for explainability, bias detection, privacy protection and regulatory compliance — lull organizations into a false sense of responsibility, while mere technology is not enough.

User Recommendations

- Develop AI model governance practices that align model performance, human behavior and delivery of business value. Make it easier for business users to adopt AI models by incorporating stakeholder trust and speed to value as primary inputs for model design.
- Extend AI experts' skills to also cover graph analytics, optimization or other required techniques for composite AI extended to generative AI. In the case of rules and heuristics, skills for knowledge engineering should also be available.
- Examine and quantify the advantages and limitations of generative AI. Use it first to improve an existing process.
- 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.
- Establish ethics principles and, optionally, an AI ethics board to resolve AI dilemmas by adopting responsible AI procedures. Ensure diversity of participants and the ease to voice AI concerns.

Sample Vendors

Amazon; Fairly AI; IBM; MathWorks; Microsoft; MoBagel; SAS; Windward

Gartner Recommended Reading

Quick Answer: How Do I Make AI Environmentally Sustainable?

Infographic: AI Use-Case Prism for Sustainability and ESG

Explore Secured, Accurate and Green AI With Federated Machine Learning

Human Controls for AI Dangers (SignatureValue Bank)

A Comprehensive Guide to Responsible AI

Multistructured Analytics

Analysis By: David Pidsley, Stephen Emmott, Tim Nelms, Anthony Mullen

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Multistructured analytics describes techniques applied to unlocking the value of wide data — the spectrum of multistructured information (structured, semi-structured, and unstructured) of any formats, including language (text/audio) and vision (image/video), sourced internally or externally. Multistructured analytics decomposes meaningful features of human-generated data for DSML modeling. It composes machine-generated data into context-enriched analysis for expert decision making.

Why This Is Important

D&A leaders must bridge the gap between the analytics of today and the context-enriched analysis for decision makers to uncover unique insights. Exponential growth in the spectrum of multistructured information sources/formats requires techniques beyond traditional, structured, transactional or relational data. To get a grip on business complexity, harness multistructured analytics to support expert decision making with richer situational awareness, augment business workflows and automate decisions.

Business Impact

Across industries/business functions, organizations accelerate their application of multistructured analytics to wide data sources/formats to reduce costs, address new uncertainties, drive growth and enable innovation in

the use cases for analytics. Multistructured analytics of audio/video streams in sales and marketing can identify behaviors and sentiments of customers and influencers across channels for new insight, experience optimization, real-time dynamic pricing and competitive intelligence.

Drivers

- Decisions are more complex, with more stakeholders and choices than two years ago. Scenarios need context-sensitive evaluation, beyond individual events, using multidimensional models of real-world uncertainties. Lacking the right variety of data stifles this.
- By 2025, 70% of organizations will shift their focus from big to wide data, providing more context for analytics. Internal (digital workplace) and external (business ecosystem) data sources continue to increase silos, leaving hidden intelligence for competitive advantage. Data marketplaces and exchanges make larger, pretrained and more diverse data assets widely available.
- Organizations are adopting multistructured analytics to move beyond storing content, to extract meaningful features and insights. By 2023, over 80% of organizations will use some form of computer vision to analyze images and videos.
- All forms of wide data can now be processed. Document topics can be tagged, speech transcribed, imaged environments annotated, emotions predicted from video, gauges digitized, opening new doors for analytics, data science and machine learning (DSML) and AI.
- Transformer models (via BERT and GPT techniques), advanced text analytics and deep learning have been a catalyst for linguistic and visual analysis. By 2025, AI for video, audio, vibration, text, emotion and other content analytics will trigger major innovations and transformations in most global enterprises.
- Improved price/performance ratio of cloud AI developer services has made experimenting accessible and scalable.
- Multistructured analytics enriches structured data with categorization and tagging. Analytics and BI and DSML platforms are adding multistructured analytics and graph capabilities so that by 2025, context-driven analytics and AI models will replace 60% of existing models built from traditional data sources, consolidating a mix of analytics solutions.

Obstacles

- Although different forms of (text) content analytics have been deployed for years, many organizations avoid leveraging multistructured information due to limited competencies, specialized tools and their perceived difficulty fueled by confusion around terminology.
- While the tools now exist to deploy multistructured analytics uniting the spectrum of multistructured information to model complex context, it's a shift in the way D&A teams undertake data modeling especially in natural language processing.
- The difficulty of combining techniques (composite AI) to handle specific formats/sources — like deep learning for videos, symbolic algorithms for text analytics, and knowledge graphs — is a challenge.
- Data sourcing, quality and privacy are common challenges that can be cost prohibitive for large datasets. Finding suitable data for a specific use case can be difficult and require governance.
- The market for multistructured analytics tools is fragmented and will likely require leveraging multiple vendors, increasing costs.

User Recommendations

- Leverage multistructured analytics for richer situation awareness and expert decision support.
- Conduct proof of value/pilots and understand the data, technical and organizational gaps.
- Apply text analytics for supply chain optimization, image analytics for diagnostic maintenance, video analytics for conferences and audio analytics for fraud prevention.
- Provide context-enriched analysis for decision makers by applying multistructured analytics to multistructured information.
- Explore multistructured analytics capabilities and roadmaps of vendors, including insight engines for text content and cloud AI developer services for image, video and audio analytics.
- Engage startups and hyperscale cloud providers for innovation.
- Estimate your compute and storage needs to train/run effective ML models that leverage multistructured information.
- Invest in taxonomy/ontology skills to accelerate the refinement and automation of information tagging/classification.
- Revise data collection, management and integration practices to take advantage of multistructured analytics.

Sample Vendors

Amazon Web Services; Databricks; Elastic; Google; IBM; Microsoft; OpenAI

Gartner Recommended Reading

Use Multistructured Analytics for Complex Business Decisions

Quick Answer: What Are the Short-Term and Midterm Implications of ChatGPT for Data and Analytics?

Buyer's Selection Spotlight: Insight Engines

Magic Quadrant for Cloud AI Developer Services

Data Mesh

Analysis By: Roxane Edjlali, Ehtisham Zaidi, Mark Beyer, Michele Launi

Benefit Rating: Moderate

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Definition:

Data mesh is a data management approach. Though not an established best practice, it supports a domain-led practice for defining, delivering, maintaining and governing data products. Data products are a packaging and delivery mechanism for data that must be easy to find and use by data consumers (business users, data analysts, data engineers or other systems). Data products must also fulfill a contract (terms of service and SLAs) between the provider and the consumer.

Why This Is Important

The definition of data mesh is evolving as the market explores the approach. Data mesh represents a potential alternative or complement to centralization-only data management strategies for analytics and other use cases. Organizations continuously seek a means to balance data requirements for repeatability, reusability, governance, authority, provenance and optimized delivery of data. Data mesh is a skills- and resource-intensive approach that shifts responsibility and authority back to subject matter experts (SMEs) in each data domain.

Business Impact

From a governance and authority perspective, data mesh relies on a federated governance approach that can delegate authority to the business and data domain SMEs. SMEs are assumed to exhibit the greatest experience in capturing and using data within their domain of expertise. They are responsible for determining guidance and processes for creating, managing and preventing unnecessary proliferation of data products. The goal of the mesh is to provide ready access to data products.

Drivers

- Data mesh provides a model that allows for decentralized data management, which aligns to organizational needs.
- Data mesh gives domains the flexibility they need to build data products that meet their required use cases. It also gives domains more control over the use of those data products across the enterprise.
- By leveraging existing assets instead of centralizing the data architecture, data mesh can reduce the time and effort required to enable data reuse throughout the enterprise. Data mesh asserts remediation for flexibility, scalability and accountability issues in approaches like centralized data warehouses, data lakes and data hubs.
- In Gartner client interactions, delays in data access and utilization are the most frequently reported issues from organizations seeking to deploy data mesh. Organizations question the success of data centralization, which can't meet all analytical use cases.
- Data mesh emerges as a compromise to respond to delivery issues, budget constraints, and misunderstandings between central teams and lines of business. Centralized approaches are often detached from the broader business domain requirements.

Obstacles

- Data management maturity and skills are required for data governance at the domain level, data completeness, application design and deployment, data quality, data provenance, systems architecture, and analytics data management.
- Data products must be able to meet the SLAs of the other groups sharing, reusing or accessing them. The associated skills may not be present in the BUs.
- Inappropriate identification of either data details or correct integrity for combining them may cause data product proliferation, thus increasing management and maintenance and necessitating reengineering to reconcile different interpretations of the data.
- Data mesh implementations and practices do not follow any specific guidelines. They cannot be vetted against standardized, or even competing, approaches. Implementations vary and may incorporate multiple approaches (e.g., marketplace experiences, virtualized views or subject-specific data marts).
- Data mesh will be obsolete before the plateau. The practice and supporting technology will evolve toward data fabric as organizations start collecting passive metadata.

User Recommendations

- Commit to building a distributed data management team, as the data mesh concept is highly dependent on the organizational model and the distribution of skills across central IT and LOBs.
- Assess data products for business domain alignment and efficiency gains upon delivery. Data strongly aligned to a single domain with broad utility across the enterprise may provide lower risk for initial data product efforts.
- Control data product proliferation by monitoring technical debt and ensuring that data products continuously evolve to meet changes in usage and scope.
- Start a metadata management program in parallel with data mesh, and collect passive metadata. This approach will allow data mesh to evolve toward metadata activation over time, guiding data product operationalization, value justification and greater transparency in the uses of data.
- Mitigate irresponsible data management by addressing management and governance contention issues in data product design within the data domains.

Gartner Recommended Reading

Quick Answer: What Is Data Mesh?

Quick Answer: Comparing Data Fabric and Data Mesh

Data and Analytics Essentials: Data Fabric and Data Mesh

Quick Answer: How Are Organizations Overcoming Issues to Start Their Data Fabric or Mesh?

2023 Planning Guide for Data Management

Data Marketplaces and Exchanges

Analysis By: Eric Hunter, Jim Hare, Robert Thanaraj, Lydia Clougherty Jones

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Definition:

Data marketplaces and exchanges provide infrastructure, transactional capabilities and services for participants as ecosystems centered around the sharing of data assets. Marketplaces prioritize data monetization via one-time or recurring subscription transactions; while exchanges prioritize sharing of assets. Internal data exchanges facilitate enterprise data sharing and remove silos to cross-organization data product provision and access.

Why This Is Important

The mass adoption of public cloud has reduced a major adoption barrier — opening the benefits of data marketplaces and exchanges (DMEs) to every organization. Although adoption remains in the early phases, DMEs provide liquidity to the data products space — enabling the sale, purchase or exchange of data products with relative ease. Participants are attracted to diverse third-party asset selection, simplified procurement, reduced operating/transaction costs and the opportunity to enrich existing data assets.

Business Impact

Easier access to third-party data assets via external DMEs benefits sellers through market liquidity and buyers through consolidated access and management of third-party data products. Organizations value internal data exchanges to simplify the sharing of data while supporting the rationalization of third-party data spend. Buyers of data products benefit through reduced provider exclusivity of specific data products from increased competition across marketplaces — driving more favorable pricing.

Drivers

- Data and analytics (D&A) products are a key enabler of digital business outcomes — increasing the demand for third-party data, the growing role of business ecosystems, and the growing awareness and need for companies about partnering.
- Vendors introducing platforms supporting DME capabilities that reduce the burden on enterprises seeking to participate in or directly operate marketplaces and exchanges.
- Enterprises seeking to rationalize their spend across third-party data projects that leverage internal data exchanges as a central point of managing and accessing external third-party data products across the organization.
- Vendors adding marketplace capabilities to their ecosystem in adjacent markets (including data management, data integration and data governance solutions).
- There is increased awareness across public and private sectors in terms of the value associated with internal and external data assets and products.
- Third-party data acquired from DMEs has become increasingly attractive as a part of feature engineering efforts for AI model features and accuracy lift.
- There is increased adoption of cloud service providers (CSPs) which has reduced on-premises data gravity limitations that slow the physical movement and integration of data across parties.
- The number of public and private data providers for DMEs continues to increase — providing both an increased level of specialization and breadth in terms of available third-party D&A product offerings.

Obstacles

- Data privacy legislation and unknown downstream data sharing risks impede the pursuit of monetizing and productizing specific data, reducing participation within DMEs.
- "Last mile" payment, integration and governance efforts for third-party D&A products acquired through DMEs are often underestimated by enterprises — leading to acquisition delays.
- Data ethics and sharing standards along with complexity in the management of data contracts/sharing agreements can prevent the adoption of third-party data.
- Enterprise procurement processes and CSP account structures provide friction for lines of business user spend within the CSP marketplaces and ecosystems.
- Terms and conditions associated with D&A products vary. This can complicate the use and derivative asset creation that often results from deep analytics.
- Funding and pricing of data products remain an obstacle — particularly for organizations that embrace data products/data monetization solely for secondary revenue streams.

User Recommendations

- Promote organizational participation in your DME of choice to accelerate the time to business value over the use of independent data asset providers or consumers.
- Adopt prebuilt DME platforms or leverage marketplace operators for the transactional infrastructure when seeking to monetize D&A products to build differentiated products and services.
- Leverage providers that operate within or are optimized for your cloud providers of choice to reduce data movement complexity and improve integration consistency.
- Include resources and support for integrating data from DMEs with established internal data as a part of your budgeting and planning process.
- Establish enterprisewide guidelines for the acquisition and management of third-party data products and analytics to rationalize and maintain budgetary oversight of third-party data spend.
- Explore the value of third-party data to increase analytic insights by either adding context as new attributes or through additional data science model features.

Sample Vendors

Amazon Web Services (AWS); Dawex Systems; DemystData; Explorium; Harbr; Narrative I/O; Nomad Data; Revelate Data Monetization; Snowflake

Gartner Recommended Reading

Tech CEO Market Foundations for Data Marketplaces and Exchanges

Top Trends in Data and Analytics, 2023

Prioritize Data Sharing Investments for Digital Business Success

How to Optimize Data Exchanges to Realize Business Value

Quick Answer: How Do I Get Started With Data Monetization?

Data Monetization

Analysis By: Lydia Clougherty Jones, Andrew White

Benefit Rating: Moderate

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

Data monetization is the practice of extracting monetary value from data, analytics and AI, quantified directly or indirectly in economic or financial terms. In this case, data, analytics and AI includes any kind of information, including model output, content, images, art and digital assets. However implemented, data monetization is central to a D&A strategy.

Why This Is Important

CEOs report that profitability is investors' top focus, while interest in revenue growth has declined sharply. Data monetization drives profitability via investments in data and analytics (D&A). At the same time, generative AI (GenAI) presents new data monetization risks and opportunities. This demands accelerating and evolving how the economic value of monetizable data is defined and used. If not, competitive advantage will be lost and the economic value of data will not be realized.

Business Impact

- Organizations can drive new revenue and cost savings. However, GenAI/ChatGPT present yet-to-be identified positive/negative economic impacts.
- Reusing/resharing data alleviates internal/external demand for diverse data sources and ML training assets.
- The increase in data monetization across finance, healthcare, life sciences and communication services means those industries will be especially impacted.
- Transactions across IoT, blockchain and metaverse create innovative monetization partnerships.

Drivers

- Sovereign data strategies promote treating data as a monetizable asset. At the same time, they are leading to a bifurcation in global trade. Together, these provide clarity and guidance, allowing organizations to monetize personal data with less risk.
- Government and business leaders/boards of directors are increasingly familiar with data monetization. As a result, public- and private-sector organizations of all sizes are realizing the power of D&A for value creation, and are pursuing external monetization opportunities.
- Digital business thrives on D&A insights. In a data-driven enterprise, D&A are no longer afterthoughts — they are fundamental to digital business transformation, and can be monetized internally and externally.
- Data marketplaces and exchanges are once again popular for data sharing between organizations and offer similar monetization opportunities. These enable publishing, curation, analysis and aggregation of data, revenue sharing, and buyer-seller matchmaking.
- Data catalogs and analytics catalogs are massively hyped in organizations as a means to increase data sharing inside organizations and improve monetization opportunities.
- Manufacturing and other industries have invested in IoT-enabled data products and analytics. Their companies amass a growing amount of data from these connections, and some have created entirely new revenue streams from the sale of data and insights to their customers.
- Privacy-enhancing computing techniques, including synthetic data and federated learning, allow for new use cases to monetize personal and sensitive data.
- The adoption of blockchain-enabled tokenization and Web3 enable new capabilities to share, extract and monetize data.

Obstacles

- Most organizations are ill-prepared to create, share, manage, measure and monetize D&A assets. Gartner's CDAO surveys over consecutive years show that 10% of CDAO respondents were effective at data monetization and 24% in the CDAO survey for 2023.
- D&A strategy and operating models have to mature enough to balance the hype related to data products and D&A product management with program/project management skills and delivery models.
- Creating data-based revenue streams requires a maturing organizational data culture and literacy.
- Developing offerings and valuation may require more analysis, market definition studies, insight and creativity than expected, due to a lack of defined markets with established guidance.
- D&A governance has to be adaptive and connected, so that ethics and data rights are considered before data monetization.
- Digital business and capital investment decision-making needs to modernize and take the economic value of data, an intangible asset, into account.

User Recommendations

- Identify data monetization opportunities by identifying stakeholders' unmet needs for economic outcomes, and map them against wide and multistructured data to conceive new revenue streams.
- Co-create data monetization programs with business/service partners by initiating collaborations that focus on developing/deploying data products, which can contribute to new revenue streams or cost savings.
- Modernize data monetization programs by establishing commercial-style D&A product management, agile delivery and DevOps practices.
- Streamline data procurement, product pricing and licensing, market definition and value, and D&A product development by hiring a D&A product manager.
- Incorporate in D&A portfolio management practices by looking across investments for the next budget cycle for maximum economic impact on outcomes.
- More-mature organizations should measure past D&A investments' "residual value" of data to look for combinatorial improvements in data monetization.

Sample Vendors

Brighthive; CARTO; DeHealth; Harbr; Infosys; Revelate; YourDataConnect

Gartner Recommended Reading

Quick Answer: How Do I Get Started With Data Monetization?

Case Study: Data Monetization Through Data Product Development (ZF Group)

Essential Product Management Practices to Monetize Data and Analytics Assets

How to Leverage Blockchain and Prepare for Data Monetization in Manufacturing

Case Study: Data and Analytics Monetization With Knowledge Graphs and AI (Turku City Data)

Responsible AI

Analysis By: Svetlana Sicular

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

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 and generates brand-new artifacts, it amplifies both good and bad outcomes. Responsible AI enables the right outcomes by ensuring business value while mitigating risks. This requires a set of tools and approaches, including industry-specific methods, adopted by vendors and enterprises. More jurisdictions introduce new regulations that challenge organizations to respond in meaningful ways.

Business Impact

Responsible AI assumes 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

- Responsible AI means a deliberate approach in many directions at once. Data science's responsibility to deliver unbiased, trusted and ethical AI is just the tip of the iceberg. Responsible AI helps AI participants develop, implement, utilize and address the various drivers they face.

- Organizational driver assumes that AI's business value versus risk in regulatory, business and ethical constraints should be balanced, including employee reskilling and intellectual property protection.

- Societal driver includes resolving AI safety for societal well-being versus limiting human freedoms. Existing and pending legal guidelines and regulations, such as the EU's Artificial Intelligence Act, make responsible AI a necessity.

- Customer/citizen driver is based on fairness and ethics and requires resolving privacy versus convenience. Customers should exhibit readiness to give their data in exchange for benefits. Consumer and citizen protection regulations provide the necessary steps, but do not relieve organizations of deliberation specific to their constituents.

- With further AI adoption, the responsible AI framework is becoming more important and is better understood by vendors, buyers, society and legislators.

- AI affects all ways of life and touches all societal strata; hence, the responsible AI challenges are multifaceted and cannot be easily generalized. New problems constantly arise with rapidly evolving technologies and their uses, such as using OpenAI's ChatGPT or detecting deepfakes. Most organizations combine some of the drivers under the umbrella of responsible AI, namely, accountability, diversity, ethics, explainability, fairness, human centricity, operational responsibility, privacy, regulatory compliance, risk management, safety, transparency and trustworthiness.

Obstacles

- Poorly defined accountability for responsible AI makes it look good on paper but is ineffective in reality.
- Unawareness of AI's unintended consequences persists. Forty percent of organizations had an AI privacy breach or security incident. Many organizations turn to responsible AI only after they experience AI's negative effects, whereas prevention is easier and less stressful.
- Legislative challenges lead to efforts for regulatory compliance, while most AI regulations are still in draft. AI products' adoption of regulations for privacy and intellectual property makes it challenging for organizations to ensure compliance and avoid all possible liability risks.
- 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 is necessary, in addition to technology.

User Recommendations

- 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, sustainability and regulatory compliance.
- Designate a champion accountable for the responsible development and use of AI for each use case.
- Define model design and exploitation principles. Address responsible AI in all phases of model development and implementation cycles. Go for hard trade-off questions. Provide responsible AI training to personnel.
- Establish operationalize responsible AI principles. Ensure diversity of participants and the ease to voice AI concerns.
- Participate in industry or societal AI groups. Learn best practices and contribute your own, because everybody will benefit from this. Ensure policies account for the needs of any internal or external stakeholders.

Sample Vendors

Amazon; Arthur; Fiddler; Google; H2O.ai; IBM; Microsoft; Responsible AI Institute; TAZI.AI; TruEra

Gartner Recommended Reading

A Comprehensive Guide to Responsible AI

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

Best Practices for the Responsible Use of Natural Language Technologies

Activate Responsible AI Principles Using Human-Centered Design Techniques

How to Ensure Your Vendors Are Accountable for Governance of Responsible AI

Generative AI

Analysis By: Svetlana Sicular, Brian Burke

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

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 Stable Diffusion, Midjourney, ChatGPT and large language models. End-user organizations in most industries aggressively experiment with generative AI. Technology vendors form generative AI groups to prioritize delivery of generative-AI-enabled applications and tools. Numerous startups have emerged in 2023 to innovate with generative AI, and we expect this to grow. Some governments are evaluating the impacts of generative AI and preparing to introduce regulations.

Business Impact

Most 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 their democratization. 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 hype around generative AI is accelerating. Currently, ChatGPT is the most hyped technology. It relies on generative foundation models, also called “transformers.”
- New foundation models and their new versions, sizes and capabilities are rapidly coming to market. Transformers keep making an impact on language, images, molecular design and computer code generation. They can combine concepts, attributes and styles, creating original images, video and art from a text description or translating audio to different voices and languages.
- Generative adversarial networks, variational autoencoders, autoregressive models and zero-/one-/few-shot learning have been rapidly improving generative modeling while reducing the need for training data.
- Machine learning (ML) and natural language processing platforms are adding generative AI capabilities for reusability of generative models, making them accessible to AI teams.
- Industry applications of generative AI are growing. In healthcare, generative AI creates medical images that depict disease development. In consumer goods, it generates catalogs. In e-commerce, it helps customers “try on” makeup and outfits. In manufacturing, quality inspection uses synthetic data. In semiconductors, generative AI accelerates chip design. Life sciences companies apply generative AI to speed up drug development. Generative AI helps innovate product development through digital twins. It helps create new materials targeting specific properties to optimize catalysts, agrochemicals, fragrances and flavors.
- Generative AI reaches creative work in marketing, design, music, architecture and content. Content creation and improvement in text, images, video and sound enable personalized copywriting, noise cancellation and visual effects in videoconferencing.
- Synthetic data draws enterprises’ attention by helping to augment scarce data, mitigate bias or preserve data privacy. It boosts the accuracy of brain tumor surgery.
- Generative AI will disrupt software coding. Combined with development automation techniques, it can automate up to 30% of the programmers’ work.

Obstacles

- Democratization of generative AI uncovers new ethical and societal concerns. Government regulations may hinder generative AI research. Governments are currently soliciting input on AI safety measures.
- Hallucinations, factual errors, bias, a black-box nature and inexperience with a full AI life cycle preclude the use of generative AI for critical use cases.
- Reproducing generative AI results and finding references for information produced by general-purpose LLMs will be challenging in the near term.
- Low awareness of generative AI among security professionals causes incidents that could undermine generative AI adoption.
- Some vendors will use generative AI terminology to sell subpar “generative AI” solutions.
- Generative AI can be used for many nefarious purposes. Full and accurate detection of generated content, such as deepfakes, will remain challenging or impossible.
- The compute resources for training large, general-purpose foundation models are heavy and not affordable to most enterprises.
- Sustainability concerns about high energy consumption for training generative models are rising.

User Recommendations

- Identify initial use cases where you can improve your solutions with generative AI by relying on purchased capabilities or partnering with specialists. Consult vendor roadmaps to avoid developing similar solutions in-house.
- Pilot ML-powered coding assistants, with an eye toward fast rollouts, to maximize developer productivity.
- Use synthetic data to accelerate the development cycle and lessen regulatory concerns.
- Quantify the advantages and limitations of generative AI. Supply generative AI guidelines, as it requires skills, funds and caution. Weigh technical capabilities with ethical factors. Beware of subpar offerings that exploit the current hype.
- Mitigate generative AI risks by working with legal, security and fraud experts. Technical, institutional and political interventions will be necessary to fight AI's adversarial impacts. Start with data security guidelines.
- Optimize the cost and efficiency of AI solutions by employing composite AI approaches to combine generative AI with other AI techniques.

Sample Vendors

Adobe; Amazon; Anthropic; Google; Grammarly; Hugging Face; Huma.AI; Microsoft; OpenAI; Schrödinger

Gartner Recommended Reading

Innovation Insight for Generative AI

Emerging Tech Roundup: ChatGPT Hype Fuels Urgency for Advancing Conversational AI and Generative AI

Emerging Tech: Venture Capital Growth Insights for Generative AI

Emerging Tech: Generative AI Needs Focus on Accuracy and Veracity to Ensure Widespread B2B Adoption

ChatGPT Research Highlights

Analytics Governance

Analysis By: Andrew White, Kurt Schlegel

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

Analytics governance is the setting and enforcement of D&A governance policy along the analytics pipeline, from data discovery through analytics deployment, and access to the analysis and insight. Though the markets use the term “governance” here, what is being delivered by technology vendors is not related to policy setting or enforcement, but execution of policy along the analytics pipeline. A more appropriate name would be analytics stewardship.

Why This Is Important

With increasing interest in analytics and AI, the shift to cloud, and the increased risk from failure and exposure to regulatory controls, the interest in governance along the entire analytics pipelines has ballooned. Unfortunately, there is a huge misunderstanding in the market. What needs to take place in the analytics pipeline is the same as what takes place upstream in operational systems. As few people understand this, hype and confusion reign, and redundant investments exist.

Business Impact

Organizations that balance investments between the latest analytics, BI and data science technology and the needed support for governance will get a greater return on both investments. With the right business outcome and adaptive governance focus, the least amount of mission-critical data and analytics will be governed. This would assist with trusted and reliable analysis and insight leverage.

Drivers

- 2023 is marked with wide data, “lakehouses,” data and analytics producers, data fabrics, and data mesh. At the same time, the analytics pipeline is stretched across organizations and clouds. Complexity and confusion is a big driver of analytics governance.
- A shift in focus from truth to trust in governing data and analytics assets, as the vastness of data can no longer be managed with current truth-based (i.e., yes/no) approaches.
- Increase in demand and deployment of self-service analytics and BI, and more rapid prototyping of analytics outputs by users closer to the point of decision.
- The interest and need to govern data inbound to a data warehouse or lake, to have access to that data in the analytics development, and to model, share and create new analytics.
- More complex organizational structures lead to increased demand, whereby various works of data, analytics or governance are widely distributed across business units, business functions, fusion teams, and IT.
- Third-party and regulatory compliance with data privacy, security, access, quality, and ethics drive increased hype to fever pitch.
- Preservation of privacy that may even conflict when operating across multiple jurisdictions.

Obstacles

- Many organizations think that “analytics governance” is different from data and analytics governance, instead of observing the same patterns and solutions that emerge in both.
- Over time, analytics governance will be recognized as part of D&A governance. Hence, stand-alone capabilities will become obsolete before achieving mass-market adoption.
- Vendors who offer analytics, BI, data science and AI solutions are not naturally familiar with or capable of meeting needs of analytic governance. But vendors will often want to meet requests from clients in the positive, so they are trying to develop analytics governance solutions.
- Some solutions that are more capable upstream of data governance in operational use cases are not actually effective at policy execution in analytics use cases. For example, implementing operational MDM in a data lake would create more problems than it might actually solve.

User Recommendations

- Validate your governance charter with the work of policy setting (i.e., governance), policy enforcement (i.e., stewardship) and policy execution (i.e., management) along your analytic pipeline. This will reduce redundancy and save money, and lead to improved outcomes.
- Don’t assume your analytics solutions support your requirements for analytics stewardship (or governance). At most, they might execute a technical rule in their application. You may need to build your own stewardship capability until the vendors meet your needs.
- Don’t assume you need to start with a data or analytics catalog. In case you don’t know your organization’s data and analytics, simply ask your business leaders.

Sample Vendors

Alation; ALTR Solutions; Collibra; ZenOptics

Gartner Recommended Reading

Data and Analytics Leaders Must Use Adaptive Governance to Succeed in Digital Business

Next Best Actions to Improve Your Data and Analytics Governance

Infographic: Data and Analytics Governance Survey: IT Says Mission Accomplished; Business Disagrees

Use Enterprise Metadata Management to Extend Information Governance to Analytics

Data Storytelling

Analysis By: Aura Popa, Peter Krensky

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Early mainstream

Definition:

Data storytelling combines interactive data visualization with narrative techniques to deliver insights in compelling, easily assimilated forms. Analytic data stories aim to prompt discussion and collaborative decision making. Data stories can take several forms (data-connected slideshows/storyboards, annotated dashboards and graphic-design-style infographics). In contrast to dashboards used for monitoring, data storytelling techniques are used to inform and educate decision makers.

Why This Is Important

Too many decision makers still overlook the data insights delivered to them, going with their gut feeling rather than decision making based on data facts. This can be a cultural issue; however, there is also a simpler factor at play — how data insights are delivered. In many cases, even where the insight does spark interest, it may lack the context required to drive a decision. Data storytelling can help break down managerial inertia and apathy toward data by adding context and making it more accessible.

Business Impact

The impact of data delivered as a story can be much higher than dashboards and reports, as a story form is familiar to everyone. From an ROI perspective, data storytelling can help drive adoption of analytics tools by repositioning them from just data visualization tools to becoming the key medium for effective communication of data insights. This is important, as our research shows that adoption of ABI platforms is still less than what it should be to be most impactful.

Drivers

- KPI-centered dashboards are not the only or even the most effective way of delivering data. A data storytelling approach can transform how analytics and data science teams work by getting them to focus on their audience and the business decisions they need to take based on insights and the data provided. The audience, made out of nontechnical decision makers, often needs data to be presented to them in narrative format to be most compelling and actionable.
- The functional capabilities to create data stories are now widely available. Most ABI platforms now include a basic functionality to create and share data stories. These stories can take several forms — most frequently these are data-connected slideshows or storyboards, annotated dashboards, and graphic design-style infographics, but they can also be simple alerts in a chat.
- A considerable proportion of vendors have already implemented ChatGPT within their platform, enabling news-style headlines and narratives generated automatically and specifically personalized for individuals.
- As self-service analytics matures, users are beginning to use data storytelling tools and techniques to better communicate to decision makers.

Obstacles

- The use of data storytelling draws on an evolving set of skills, practices and behaviors around how data is socialized and used in organizations. Many organizations do not have these skills in place.
- Data storytelling is a part of a broader movement oriented around data literacy, and explaining and expressing data and analytics in a consumable, engaging and relevant way. Poor data literacy is a chronic inhibitor to effective data storytelling.
- There should be no compromise on quality from content governance in the world of data storytelling. Data storytelling may allow bias or overfitting to create false narratives both from the human creation form or the machine-generated route.
- Machine-generated data stories may not gain traction if they are not relevant, understandable or explainable to the intended recipients. Metadata and context around data stories may need to be automated and utilized to prevent this.

User Recommendations

- Evaluate and experiment with the data storytelling capabilities of ABI platforms. Examine how their incumbent portfolio of technologies supports the creation of storyboard-style presentations with embedded analytical content.
- Leverage the power of machine-generated data stories, collecting as much metadata and context information as possible for enrichment, but making sure human peer reviews and content governance rules are put in place for a continuous sustainable quality.
- Task members of your analytics team with investigating data storytelling as an extension to their use of interactive visual exploration and analytic dashboarding. This will provide a richer delivery of information by adding narrative and context.
- Prepare programs to develop and instill the mix of data visualization design, narration and presentation skills needed to support effective data storytelling. Identify a team of business analysts and citizen data scientists to act as a virtual team of data storytellers.

Sample Vendors

Domo; Oracle; Pyramid Analytics; Qlik; Salesforce (Tableau); ThoughtSpot; TIBCO Software; Toucan; Yellowfin

Gartner Recommended Reading

Communicate Insights Effectively With Augmented Data Visualization and Storytelling

Market Guide for Augmented Analytics

Data Storytelling: Analytics Beyond Data Visualizations and Slideshows

Engage and Influence Business Stakeholders Using Data-Enabled Storytelling

Data-Centric Translators Are Crucial for Facilitating Data Literacy

Dynamic Trust-Based Model for Governance

Analysis By: Andrew White, Saul Judah

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

A dynamic trust-based model for D&A governance leverages ML to dynamically discover, inform, and then set desired trust, reliability and efficacy of data, analytics, systems, partners and organizations, to ensure appropriate asset usage and risk mitigation. It offers a graduated approach to governance compared to traditional dimensions common in data quality and MDM that focus on a single dimension (e.g., yes/no compliance or single version of the truth) for the policy or definitional rules.

Why This Is Important

Common and long-standing governance models use yes/no conclusions, yet this binary approach no longer serves business needs adequately. A dynamic trust-based model uses a graduated capability such that an attribute might be useful in some cases, but perhaps not all. It also introduces multiple layers of “yes” or “may be.” Trust-based governance models are becoming dynamic as they are being enhanced to use machine learning (ML) to discover and profile data that might well demonstrate trust as evidenced by frequent usage.

Business Impact

Traditional data quality and entity resolution definitions tend to focus on yes/no or single responses, and these limit the flexibility of data and analytics governance. A trust-based approach alleviates the challenge and will reduce costs, mitigate risks and help businesses achieve their expected goals more effectively. Enhanced with ML, dynamic trust models achieve their goals much faster.

Drivers

- Hype and expectations around trust and trust-based governance models is being reinvigorated by augmentation with ML to help speed up implementation. ML-augmented knowledge graphs, entity resolution and data quality are helping discover relationships in data to help infer or inform additional insights on data use, which can help reinforce trust.
- The use of data to drive decisions and outcomes is still at a fever pitch, and the idea that trust over truth is an appropriate concept is now common.
- Trust is at the center of the effective use of analytics and artificial intelligence (AI) since it aligns with the vagaries and contexts of complex situations and models. Absolute forms of data quality and definitions do not work well in these new environments, which are predicated on openness, shareability and exploration.
- Trust cannot be assumed; it needs to be evaluated and calibrated to the business outcome, and is hard to earn when people or relationships are involved.
- Master data management (MDM), data quality, active metadata techniques, and various elements of data and analytics (D&A) governance are adding trust-supporting capabilities widely.
- This model will be adopted into D&A governance platforms over time as a common capability and will become obsolete before the plateau.

Obstacles

- Lack of maturity in governance is the biggest obstacle to trust-based governance. D&A teams often spend much of their time firefighting operational issues (e.g., data quality issues that prevent a business transaction or a report with “bad” data). They don’t have the opportunity to step back and assess their landscape and understand the lineage, curation and usage of their ecosystem by their organizational users. As a result, they are unable to put in place a framework that will help them reduce the issues they face on a daily basis.

- Widely deployed technologies, such as data dictionaries, glossaries, catalogs, data quality and entity resolution, and business rule engines, are also being augmented with ML. Thus, they are also converging and evolving, which makes it harder to work out where and when to innovate with a trust-based approach.

- Use and adoption of ML in many different use cases continues apace but reliably producing high-value insights remains elusive.

User Recommendations

- Pilot a dynamic trust model to some critical data and its source where governance policy efforts to date appear costly or onerous. Explore how it can help users of the data align their risk mitigation efforts to the value and use of the (trusted) data.

- Use a simple three-tier framework (untrusted, unknown, trusted) when data quality has only a single dimension (e.g., yes/no), and test the boundaries and the savings in time and effort to govern such data.

- Augment your data catalog, data dictionary or glossary program with dynamic trust model capability to help speed up implementation and leverage of such data.

- Align trust-based governance to your data use case and enterprise goals. Many use cases will still work very effectively with traditional rules and policies, such as those defined by yes/no qualifications. But for departmental use, that data might be treated differently, especially if enriched with third-party data for analytics, such as a customer data platform (CDP). It is in the CDP that trust may be more useful than a truth-based approach.

Sample Vendors

Alation; Informatica; Precisely

Gartner Recommended Reading

7 Must-Have Foundations for Modern Data and Analytics Governance

Reset Your Information Governance Approach by Moving From Truth to Trust

Why Situational Trust Is Key to Data Sharing for Business Value

Predicts 2023: Distributed Business Decisions Need Balanced Governance Approaches

Enterprise Metadata Management

Analysis By: Guido De Simoni

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Enterprise metadata management (EMM) is a business discipline for governing shared metadata assets (active and passive) between and across analytics and operational projects and programs, such as master data management (MDM), business intelligence (BI) and records management. The aim is to achieve the benefits of enterprise information management (EIM). EMM differs from project-specific metadata management, which manages metadata for specific uses within that single project.

Why This Is Important

The need to align semantics across data and analytics silos is driving demand for modern data management. Without metadata management, organizations struggle to identify data that delivers value and to manage new data sources.

Active metadata, a core part of data fabric, promises to expose metadata in more scalable ways. While patterns like data fabric may use all (active and passive) metadata to develop inference engines and insight, the most important metadata must still be stewarded by business roles. Without this link, “person and machine” can’t connect.

Business Impact

EMM can offer the following advantages:

- EMM enables coordinated efforts in data and analytics use cases of all kinds. As a result, it provides an immediate benefit to business teams by allowing them to search, request and reuse governed metadata across siloed projects for their self-service use.
- EMM extends the benefits of individual programs — such as those for analytics, MDM, data quality, data integration, business process management and service-oriented architecture — by supporting reconciled semantics in and across the information sources they use.

Drivers

- Most organizations manage metadata within individual initiatives — that is, within the confines and needs of each data and analytics program, business initiative or system. For example, an MDM program, a BI initiative and a data warehouse implementation will include a specific metadata management focus. EMM supports the discipline of aligning and governing shared and common metadata among all such programs.
- Technological innovations, such as active metadata management and data fabric, are generating significant hype. This hype is sparking new interest in linking information silos to discover data relationships. These innovations enable faster governance of information assets across multiple information management investments, which, in turn, creates fresh demand for EMM and EMM-enabled processes.

Obstacles

- Poorly planned EMM makes it prohibitively costly to implement technologies that can manage the enterprisewide variety, volume, velocity and complexity of metadata about vital information assets.
- EMM is often poorly planned because organizations assume EMM is a wall-to-wall program, which it is not meant to be. As with other informed and modern EIM efforts, not all metadata is equal. Failure to adopt this view leads to bloated and costly programs that don't add business value.
- Hype about automated technologies powering data fabric with active metadata will appeal to those clients who assume that acquisition of technology solves the problem for which EMM is needed. This mindset will just slow down the success and adoption of EMM.
- As a result of the previous obstacles, EMM is past the Peak of Inflated Expectations, but most organizations' adoption remains at an early phase. Concurrently, various technology innovations, while trying to fill EMM gaps, are disrupting the discipline's maturation. Thus, EMM movement is not noticeable on the Hype Cycle.

User Recommendations

- Explore EMM when you have common corporate goals yet disparate information management programs (each with its own metadata) that are neither aligned nor sharing consistent information.
- Use EMM to govern the most important metadata and information assets between these discrete programs. EMM is valuable when your organization needs to incorporate its information management programs into a more mature EIM framework.
- Grow the "connections" between the programs and datasets as needed, over time, if your goal is to align information across these metadata elements and use EMM to govern the shared metadata.
- Adopt an EMM strategy to improve the situation by drawing on other planned initiatives, which may involve the participation of individuals from different organizational units.
- Account for people and process issues, as well as technological issues and choices, to create and sustain an EMM program.

Gartner Recommended Reading

Leverage Semantics to Drive Business Value From Data

Data and Analytics Essentials: Metadata Management

Sliding into the Trough

Data Literacy

Analysis By: Alan D. Duncan, Donna Medeiros, Sally Parker

Benefit Rating: Transformational

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

Data and analytics (D&A) are pervasive in all aspects of businesses, communities and our personal lives. Thus, data literacy is foundational to the digital economy and society. It helps stakeholders:

- Draw a directlink between D&A and desired outcomes
- Unlock knowledge workers' business acumen
- Explain how to identify, access, integrate and manage datasets
- Draw insights relevant to specific use cases
- Describe advanced analytics techniques and enable AI
- Reduce risk through improved decision making

Business Impact

To become data-driven and equipped to use data and analytics to their competitive advantage, enterprises require explicit and lasting organizational change. Chief data and analytics officers (CDAOs) need to promote and orchestrate "leadership moments" where they act as role models, exemplifying new cultural traits at critical points. To be successful, they will need to guide the workforce by addressing both data literacy and data-driven culture.

Drivers

- The continued growth in digital transformation is amplifying a focus on D&A best practices. Employee data literacy is becoming increasingly recognized as an important factor in an organization's overall digital dexterity.
- The role of the D&A function has evolved. It is now at the core of an organization's business model and digital platforms, and with everyone being an information worker, the footprint of business use of data and analytics is broader than ever before.
- Effective D&A strategies require an increased focus on change management. Higher-performing CDAOs prioritize their emphasis, energy and effort on change management requirements, including data literacy.
- Defining what data-driven behaviors are expected — using a "from/to/because" approach — is central to employee development plans. It ensures that creators, consumers and intermediaries have the necessary D&A skills, knowledge and competencies.
- Data literacy is not a one-off project. CDAOs need to take immediate action to create and sustain data literacy through assessment of maturity, awareness, and education. Quick wins build momentum, but lasting and meaningful change takes time because it requires people to learn new skills and behave in new ways. (For example, there is a hunger for this type of skills development within Gen Z, especially in order to future-proof their careers.)

Obstacles

- Lack of common data literacy models/frameworks/standards and terminology.
- Varying interpretations of the term "data literacy" in terms of training, curriculum and understanding, ranging from enhanced data visualization skills to fostering business curiosity about data.
- Failure to measure contribution of data to business outcomes.

- A sporadic and inconsistent approach to training and certification.
- Not recognizing that data use is a behavioral change or change management initiative.
- Lack of talent and poor data literacy within the current workforce.
- Lack of 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.
- Data literacy is treated as a checkbox activity, especially when delegated to more junior (and unempowered) resources.
- Lack of a designated leader accountable for the development and execution of the program, roadmap and communication plan.

User Recommendations

- Make the business case for data literacy by identifying stakeholder outcomes and linking these to underlying learning needs.
- Designate a leader who will be accountable for developing and executing the roadmap.
- Foster data literacy during D&A requirements gathering by bringing data and business experts together around the problem to be solved.
- Call out examples of “good” and “bad” data literacy to promote desired behaviors.
- Nurture data literacy by rewarding stakeholders who recognize this as a factor for success and sharing their stories.
- Partner with HR and business leaders to incorporate data literacy learning outcomes into job descriptions, career paths and employee value proposition.
- Use data literacy assessments to evaluate current skill levels and desire to participate.
- Go beyond vendor product training to focus on people’s role- and industry-related D&A skills. Improve learning effectiveness by using a mix of training delivery methods (classroom, online, community, on the job).

Sample Vendors

Avado; The Center of Applied Data Science (CADS); Coursera; The Data Lodge; Data To The People; Pluralsight; Skillsoft; Udacity; Udemy

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

Partner With Data Literacy Providers to Accelerate the Time to Value for Data-Driven Enterprises

DataOps

Analysis By: Robert Thanaraj, Ehtisham Zaidi, Sharat Menon, Aaron Rosenbaum

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Emerging

Definition:

DataOps is an agile and collaborative data management practice focused on improving the communication, integration, automation, observability and operations of data flows between data engineers and data consumers. The goal is to assist data and analytics (D&A) leaders in driving operational excellence in data delivery in support of their data management solutions.

Why This Is Important

DataOps eliminates various inefficiencies and misalignments between data management and consumption use cases by streamlining data delivery processes and operationalizing data workloads. DataOps practices:

- Improve organizational speed and trust in delivering data

- Manage interdependencies across business processes
- Increase reusability of data engineering work product
- Provide reliable data delivery service levels
- Govern data, leading to trust and use among consumers

Business Impact

- Data engineers benefit from increased productivity and robust change management, ensuring data delivery service levels like quality, lineage and security.
- Data consumers benefit from reduced cycle time of accessing ready-to-use data and improved data trust.
- Organizations thrive on data literacy, productivity gains, self-service enablement and a collaborative culture.
- Eliminate unwanted data delivery efforts by focusing on value flows tied directly to business impact.

Drivers

- Organizations strive to improve speed and efficiency of producing trusted and usable data. DataOps practices reveal bottlenecks in the current D&A delivery process and guide toward improving the lead time (process efficiencies) and cycle time (technical efficiencies).
- DataOps improves the shareability and reusability of the data across the organization. It involves formal processes around data architecture, quality and modeling, and ensures that the data governance requirements are being applied as part of the operational processes. Otherwise, the initial data pipeline would only be designed for a narrow use.
- DataOps practices enable reorganization of teams, which helps overcome the challenges caused by fragmented teams/processes and delays in delivering data in consumable forms.
- DataOps tools eliminate the various inefficiencies and misalignments across data management technologies by streamlining data delivery processes and operationalizing data workloads. It is an emerging technology market.

Obstacles

- Setting up DataOps is a challenge as it needs efforts on justifying operating model shifts, focus on metadata management practices (which are currently nascent) and continuing to provide effort or cost over value justifications which require aligning business outcomes to DataOps activities.
- Organizations have substantial domain expertise and siloed functional capabilities. It is challenging to retain the advantages of the legacy approaches to data management and the people holding those skills, while also aggressively pursuing DataOps.
- Organizations lack a holistic view of various stand-alone technologies that are often managed by multiple teams with varying levels of operational maturity.
- Many strive for end-to-end automation of pipelines driven by code and integrating diverse technologies to make them work together is complex as it involves diverse skill sets.

User Recommendations

- When introducing DataOps, target projects that are struggling due to lack of collaboration, overburdened by the pace of change, or where service tickets from data consumers are piling up.
- Apply the core DevOps approaches to data management such as automating deployment to test environments continuously and managing schema drifts in pipelines. Reach out to your application leaders who have successfully applied DevOps practices to application development.
- Track metrics such as time to deploy changes, degree of automation, developer productivity, code quality, failure rates in production, cost-efficiencies and business impact in dollar amount.
- Plan for data pipeline operations ownership and service-level management. DataOps teams generally own the full development life cycle from inception to production. They must be cross-functional teams that combine data management, software engineering and I&O expertise — some organizations even hire for new roles such as data product manager.

Sample Vendors

Astronomer; BMC; DataKitchen; DataOps.live; GitHub; Kensu; StreamSets; Tengu; Torana (iCEDQ); Unravel

Gartner Recommended Reading

Data and Analytics Essentials: DataOps

Market Guide for DataOps Tools

5 Ways to Enhance Your Data Engineering Practices

How to Apply DevOps and Value Stream Mapping to Data, Analytics and AI

Toolkit: Delivery Metrics for DataOps, Self-Service Analytics, ModelOps and MLOps

Data Ecosystems

Analysis By: Adam Ronthal, Robert Thanaraj, Aaron Rosenbaum

Benefit Rating: High

Market Penetration: 1% to 5% of target audience

Maturity: Emerging

Definition:

Data ecosystems provide a cohesive data management environment that supports all data and analytics workloads. They have a common governance and metadata management framework and unified access management. They integrate augmented data management capabilities with a set of services accessible by the business user. They provide streamlined delivery and comprehensive functionality that is straightforward to deploy, optimize and maintain. Third-party vendors also participate in data ecosystems.

Why This Is Important

Data and analytics leaders report that the cloud experience today requires a significant integration effort to ensure that components work well together. Cloud service providers (CSPs) and independent software vendors (ISVs) are responding with more refined data ecosystems as the market moves from “some assembly required” to a “packaged platform experience.”

Business Impact

Data ecosystems unify data management and associated use cases with streamlined delivery and easy integration within a holistic management framework. They address key data management disciplines, such as data integration, data quality, data sharing, governance, metadata and observability — via augmentation — and provide the basis for operational and analytics capabilities. They are delivered both by a combination of native CSP offerings, as well as ISV components that provide enhanced capabilities.

Drivers

- Data and analytics architectures are under significant stress on two fronts: hybrid and multicloud deployment environments, and the diversity of data persistence models required to meet the increasing demands of data and analytics.
- Cloud practitioners need to rationalize data silos, which span multiple deployment environments and frequently require different and potentially conflicting operating models.
- Enterprises are looking to unify the way they engage with different data models, platforms and use cases to improve efficiency and time to value for data-driven initiatives.
- Data ecosystems serve as a unifying approach to resolve these pressure points. Built on a common foundation of governance, metadata and emerging data fabric design, they enable new practices like DataOps, FinOps and PlatformOps. They will become self-optimizing and self-tuning, and support financial governance efforts through cost optimization.
- Data ecosystems promise improved productivity and ROI based on the value of not having to do explicit data and application integration as they are based on a common set of services.

Obstacles

- While data ecosystems have a vision of unifying data management with common governance, security and metadata, significant work is still needed to make this a reality. Gaps exist in data integration, data quality, metadata and governance, which need to be addressed either through native CSP offerings or partnerships with ISVs to fully realize the vision of the cloud data ecosystem.
- When combining native CSP offerings with third-party ISV offerings, end users may find that additional effort is required to integrate these components. This undermines the core concept of a unified, holistic data ecosystem, though the end result leveraging CSP and ISV offerings is likely to be more capable.
- While CSPs are working with third-party ISVs to provide open ecosystems, their initial focus remains on ensuring that their own components are working well within their own cloud and addressing the basic needs of their customers.
- For a cloud ecosystem (that encompasses a CSP and ISVs) to function well, CSP and ISV components must have standard interfaces that exchange metadata bidirectionally. It is essential that participating ISVs and CSP agree on common metadata sharing standards. This remains slow to emerge in the market today.

User Recommendations

- Assess the maturity of these ecosystems and the degree to which they deliver on the promise of a unified environment.
- Assess points of integration between various components (data persistence, use cases, data integration, observability, governance and metadata capabilities) to determine how cohesive the resulting ecosystem is. A less cohesive ecosystem will require significantly more integration time and effort.
- Ensure that the data ecosystem has a well-articulated path to production for a full data life cycle (from discovery to production-optimized delivery).
- Define what CSPs need to deliver as part of the solution and what capabilities to obtain from third-party ISVs; expect to spend more time on integration efforts when combining CSP and ISV offerings.

Sample Vendors

Amazon Web Services (AWS); Cloudera; Databricks; Google Cloud Platform; IBM; Microsoft; Oracle; SAP

Gartner Recommended Reading

The Impacts of Data Ecosystems: A Cloud Architectural Perspective

Innovation Insight: Data Ecosystems Will Reshape the Data Management Market

Strategic Roadmap for Migrating Data Management to the Cloud

Data Hub Strategy

Analysis By: Andrew White, Thornton Craig

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Definition:

A data hub strategy effectively determines where, when and how data needs to be mediated, governed and shared in an enterprise. It layers data and analytics governance requirements atop sharing demands to establish the patterns for data flow. The strategy drives the implementation of one or more data hubs — architectures that enable data sharing and access by connecting data producers (applications, processes and teams) with data consumers (other applications, processes and teams).

Why This Is Important

Digital business cannot scale by continuing to piecemeal all the programs and practices that have evolved in the last 10 years. A connected, aligned approach is needed. A data hub strategy provides that connected approach to streamline and simplify how all programs related to D&A governance operate: data quality, MDM, ADM, metadata management, data catalogs and so on. You end up getting more from less effort and investment.

Business Impact

- Increased operational efficiency by aligning and integrating previously siloed governance programs such as data quality, MDM, ADM, metadata management, data catalogs and so on.
- Increased return on all D&A investments through more effective and targeted efforts on implementing governance of D&A information assets such as data, analytics, models, etc.
- Reduced complexity and cost across overall information infrastructure and data fabric or mesh.

Drivers

- Demands for seamless data flow across teams, processes and systems in the enterprise, which have increased dramatically in complexity and mission-criticality.
- New demands for consistent and reliable sharing of critical data between the organizations and things that comprise the extended enterprise — for example, in support of Internet of Things (IoT) solutions and new digital products.
- Better collaboration across business-oriented (governance) and IT-centric (integration) roles concerned with delivering data to points of need across the enterprise.
- Longtime and continued frustration of business stakeholders over the lack of consistency and trust of data driving strategic business outcomes — a data hub strategy enables more-focused application of governance controls, as compared with trying to align governance approaches inside many endpoint systems.
- Emerging data fabric design patterns that both need and leverage trusted sources of data and can inform what data should be governed more importantly.
- Growing need for a flexible and governable architecture that complements centralized data stores such as data lakes and data warehouses.
- Desire of many organizations to leverage the concepts and successes of MDM programs toward governance and sharing of other types of critical data. Includes coupling MDM and ADM across the enterprise.

Obstacles

- Inability to modernize D&A governance programs and shift away from legacy domain and data-centric or IT focused efforts to an outcome-based program.
- Resistance from teams or business units that prefer to retain control over their choices regarding how data is shared and governed.
- Inability to enable collaboration and agreement of critical stakeholders on data sharing and governance requirements across boundaries in the enterprise.
- Overreliance on technology and viewing governance and sharing of data as purely an implementation issue.

User Recommendations

- Identify the data that is most frequently used or is most important with most business value, and that requires effective governance and sharing. This might be a lean MDM or ADM/ERP program.
- Design a data hub strategy to understand data and analytics governance and sharing requirements, and to drive integration efforts across multiple use cases.
- Include any master data, application data, reference data, analytics data hubs or other intermediaries (e.g., customer data platforms) in your overall data hub strategy.
- Iterate changes to your data hub strategy as requirements for governance, sharing and integration change.

Sample Vendors

IBM; Informatica; MarkLogic; Profisee

Gartner Recommended Reading

Data and Analytics Essentials: Data Hubs

Data Sharing Is a Business Necessity to Accelerate Digital Business

Use a Data Hub Strategy to Meet Your Data and Analytics Governance and Sharing Requirements

Data Hubs: Understanding the Types, Characteristics and Use Cases

Data Hubs, Data Lakes and Data Warehouses: How They Are Different and Why They Are Better Together

Data & Analytics for Good

Analysis By: Jorgen Heizenberg, Carlie Idoine, Kevin Gabbard

Benefit Rating: Low

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

"Data & Analytics for Good" is a movement in which D&A is pursued because of its ability to deliver a social impact. These initiatives may deliver incidental enterprise value, but this is not their motivation, and they intentionally exclude leveraging social causes to market products or services.

Why This Is Important

"Data and analytics for good" (DA4G) is a movement that transcends organizational boundaries to use data for social impact. Corporate philanthropy is one way commercial organizations can advance societal needs and generate competitive advantage. They can use this human-centric approach as a differentiator in skills enhancement, company engagement and recruiting.

Business Impact

The "Data & Analytics for Good" movement gives resources to a good cause, the public sector or NGOs through free or reduced-cost technology, data and expertise. In the commercial sector, participation in DA4G initiatives can be through philanthropic benefits that attract and retain workers and provide resources. DA4G initiatives can signal social responsibility to investors.

Drivers

Focus on "Data & Analytics for Good" initiatives is growing with increased visibility and understanding of the value provided from these efforts.

- DA4G efforts initially focused on educational enhancement and on providing clean water, reliable food, ecological management, and arts and science community support.
- The number of organizations — from universities and communities to vendors — having a focus on DA4G has increased.
- D&A has the potential to identify, describe, diagnose and address the root causes of human suffering.
- Significant market momentum for DA4G comes from vendors (both software and services) and has been especially evident in relation to sustainability efforts.
- DA4G initiatives within organizations can provide impetus and incentive for attracting and retaining talent as well as incentive to develop new products to alleviate human suffering.

Obstacles

- Justification for “Data & Analytics for Good” is difficult to initiate and maintain because the goals and objectives are considered altruistic, and lose influence relative to business delivery-driven efforts.
- DA4G programs are often dismissed because the funding stream is considered temporary, or at least inconsistent.
- Some DA4G programs are rejected when they seek to qualify inclusion or delivery based on personally identifiable information or data that is considered ethically dangerous.
- Lack of transparency can lead to negative unintended consequences. Data ethics and trust must be incorporated to mitigate the risk of potential data misuse.
- Low levels of data literacy get in the way of effectively using the contributed data and analytics to achieve the desired social impact.
- Legal impediments, technical data format standards and the practical issue of data cataloging and aggregation have also hindered efforts.

User Recommendations

- Leverage free resources (expertise/services, software, technology, data) from organizations that support "Data & Analytics for Good" projects.
- Participate in community events to collaborate on DA4G. Contribute to open data in support.
- Allow employees time to work on philanthropic initiatives. Use this HR benefit as a differentiator in recruiting and skills' enhancement.
- Evaluate internal, external and open data to assess its usefulness for social purpose while adhering to privacy and security policies. Instill data ethics considerations in data use and sharing.
- Drive data literacy to help identify, understand and recommend controls for DA4G use cases in an effort to provide transparency without endangering individuals' privacy or sensitivities.
- Grow awareness about DA4G. Share internal and external case studies as well as resources that demonstrate what DA4Gis and what its impact can be.

Gartner Recommended Reading

Quick Answer: How Can We Start a 'Data and Analytics for Good' Initiative?

Data Fabric

Analysis By: Mark Beyer, Ehtisham Zaidi, Roxane Edjlali, Sharat Menon, Robert Thanaraj

Benefit Rating: Transformational

Market Penetration: 1% to 5% of target audience

Maturity: Adolescent

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, ML 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

Data fabric leverages traditional approaches while enabling the enterprise to adopt technology advances and avoids "rip and replace." It capitalizes on sunk costs and simultaneously provides prioritization and cost control guidance for new spending for data management. It leverages concepts and existing platforms/tools or implementation approaches. It offers flexibility, scalability and extensibility in infrastructure for humans or machines to assure data is consumable across multiple use and reuse cases on-premises, multicloud or hybrid deployments.

Business Impact

Data fabric:

- Increases identification, deployment and availability of data for reuse at scale.
- Provides insights to data engineers by standardizing repeatable integration tasks, improving quality, and more.
- Adds semantic knowledge for context and meaning, and enriched data models.
- Evolves into a self-learning model that recognizes similar data content regardless of form and structure, enabling connectivity to new assets.
- Enables observability across the data ecosystem.
- Reduces maintenance, support and optimization costs associated with managing data.

Drivers

- The dearth of new staffing or personnel seeking data management roles and the attrition of experienced professionals leaving the practice area has increased the demand for more efficient data reuse.
- Demand for rapid comprehension of new data assets has risen sharply and continues to accelerate, regardless of the deployed structure and format.
- Increased demand for data tracking, auditing, monitoring, reporting and evaluating use and utilization, and data analysis for content, values and veracity of data assets in a business unit, department or organization.
- Catalogs alone are insufficient in assisting with data self-service. Data fabrics capitalize on machine learning (ML) to provide recommendations for integration design and delivery, reducing the amount of manual human labor that is required.
- Significant growth in demand and utilization of knowledge graphs of linked data, as well as ML algorithms, can be supported in a data fabric to assist with graph data modeling capabilities and use-case generic semantics.
- Organizations have found that one or two approaches to data acquisition and integration are insufficient. Data fabrics provide capabilities to deliver integrated data through a broad range of combined data delivery styles including bulk/batch (ETL), data virtualization, message queues, use of APIs, microservices and more.

Obstacles

- Organizations will keep applying budget or staff to one-off and point-to-point integration solutions.
- Differing design and semantic standards used by various vendors to document and share metadata create challenges in its integration and effective analysis to support a data fabric design.
- Fabric needs analytic and ML capabilities to infer missing metadata. This will be error-prone at first with staffing and resources assigned to competing demands in advanced analytics, data science and AI near the data consumption layer.
- Active metadata management practices lag behind data fabric adoption but are critical to its implementation.
- Diverse skills and platforms demand a cultural and organizational change from data management based upon analysis, requirements and "design then build" to discovery, response and recommendation based upon "observability and leveraging."
- Improper split from data mesh implies choosing one approach over another and not a complementary relationship.
- Inexperience in reconciling a data fabric with legacy data and analytics governance programs will confound implementers.

User Recommendations

- “Active metadata” and leveraging the inherent practices to it is mandatory in a data fabric (covered separately).
- Invest in an augmented data catalog that permits multiple ontologies over top of business data taxonomies and is alerted to new use cases for data and the related business units utilizing data.
- Deploy data fabrics that populate and utilize knowledge graphs in targeted areas where adequate metadata and metadata management practices already exist.
- Ensure business process experts can support the fabric by enriching knowledge graph capabilities with business semantics.
- Evaluate all existing data management tools to determine the availability of three classes of metadata: design/run, administration/deployment and optimization/algorithmic metadata. When adopting new tools, favor those that share the most metadata.
- Do not permit SaaS solutions to isolate their metadata from access by PaaS solutions that orchestrate across solutions.

Sample Vendors

Cambridge Semantics; Cinchy; CluedIn; Denodo Technologies; IBM; Informatica; Semantic Web Company; Stardog; Talend

Gartner Recommended Reading

Data and Analytics Essentials: How to Define, Build and Operationalize a Data Fabric

Quick Answer: What Is Data Fabric Design?

Emerging Technologies: Critical Insights on Data Fabric

Digital Ethics

Analysis By: Pieter den Hamer, Frank Buytendijk, Svetlana Sicular, Bart Willemsen

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. It applies to areas such as AI, data and analytics, and social media.

Why This Is Important

Digital ethics, especially around topics like privacy, bias, polarization and veracity, is a concern to many. The voice of society is getting louder, with responsible AI coming into sharp focus for individuals, organizations and governments. People, increasingly aware that their data is valuable, are frustrated by lack of transparency, misuse and breaches. Organizations are acting to mitigate ethical risks around data, AI and other digital areas, while more governments are encouraging and regulating responsible use of these in digital society.

Business Impact

Digital ethics strengthens an 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 organization serves all parties and society more broadly, and leads to better business trust and performance.

Drivers

- The media is frequently featuring high-profile stories about the impact of data, AI and other technology on business and society at large. Board members and other executives are increasingly sharing concerns about the unintended consequences of innovative technology use.

- For many technologies, ethics was often an afterthought. However, with the emergence of artificial intelligence, the ethical discussion is now taking place both before and during a technology's widespread implementation. AI ethics aims to establish responsible use of AI and to harness AI's growing powers.

- The current hype around generative AI, including ChatGPT and similar alternatives, is raising awareness about ethical and legal issues surrounding the veracity and (intellectual) ownership of data, including training data. In addition, the potential impact of inaccurate, misleading or insensitive output is fueling ethical concerns.

- Government commissions and industry consortia are actively developing guidelines for ethical use of AI. Examples include the EU's AI Act, the Netherlands' Fundamental Rights and Algorithm Impact Assessment (FRAIA), and the U.S.'s National AI Research Resource (NAIRR) Task Force and National Artificial Intelligence Initiative to advance trustworthy AI in the U.S.

- Over the past few years, a growing number of organizations declared their AI ethics principles, frameworks and guidelines. Many are in the process of going from declaration to execution.

- Universities across the globe have added digital ethics courses and have launched programs to address ethical, policy and legal challenges posed by new technologies.

- Digital ethics is expanding to address concerns about rising energy consumption. In the case of nonrenewable energy, it is focusing on the carbon footprint of digital technology (particularly, machine learning and blockchain).

Obstacles

- Because of the ambiguous, pluralist and contextual nature of digital ethics, organizations often struggle to operationalize it and expend significant effort to implement best practices.
- Organizations see digital ethics as a moving target because of 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, reduced to a checklist, confined to technical support for privacy protection, and/or viewed only as explainable AI.
- AI ethics is currently the main focus of digital ethics. Supporting technology (e.g., to protect privacy or mitigate bias) needs to mature further and apply to the broader scope of ecosystems rather than singular technologies.
- Across people, regions and cultures, opinions differ on what constitutes "good" and "bad" and what doing the right thing means. Even in organizations that recognize ethics as an important issue, consensus between internal and external stakeholders (such as customers) is sometimes illusive.

User Recommendations

- Identify specific digital ethics issues and opportunities to turn awareness into action.
- Discuss ethical dilemmas from diverse points of moral reasoning. Anticipate and account for ethical consequences. Ensure that you are comfortable defending the use of a technology, including any 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 the adoption of new, transformative technology such as AI. Address digital ethics upfront "by design" to create methods that identify and resolve ethical dilemmas as early as possible.
- Organize training in ethics, and run workshops to create ethical awareness within all AI initiatives. These should emphasize 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 Guidelines as Your Starting Point

How to Manage Digital Ethical Dilemmas

How to Operationalize Digital Ethics in Your Organization

Hyperautomation

Analysis By: Frances Karamouzis, Keith Guttridge, Laurie Shotton, Saikat Ray

Benefit Rating: Transformational

Market Penetration: More than 50% of target audience

Maturity: Mature mainstream

Definition:

Business-driven hyperautomation is a disciplined approach that organizations use to rapidly identify, vet, and automate as many business and IT processes as possible. Hyperautomation involves the orchestrated use of multiple technologies, tools or platforms to achieve business results. These include, but are not limited to, AI, machine learning, event-driven software architecture, robotic process automation (RPA), iPaaS, packaged software and process/task automation tools.

Why This Is Important

The primary reason that hyperautomation is critical is the unrelenting demand for accelerated growth through business model innovation or disruption, coupled with the underlying foundation of operational excellence across processes and functions. This is important as organizations continue to focus on business outcomes such as higher quality, more resilient processes, and higher usage due to employee- and customer-centric experiences, among others.

Business Impact

The most important business impacts are aligned to business outcomes such as cost optimization, growth, business agility or innovation. Hyperautomation initiatives are fluid enough to align to one or all of these outcomes. Examples of results may be better (higher quality, more resilient) business or IT processes, speed (time to market, cycle time reduction and quicker adoption) or intelligent (data-driven) decision making at scale.

Drivers

- The biggest driver of hyperautomation is funding from business units (as opposed to the IT budget). These business units continue to hire and fund initiatives driven by fusion teams and business technologists.
- The continued unabated spending on hyperautomation initiatives is forecast to exceed $1 trillion in 2023. This includes spending on products (software, platforms and tools) coupled with services spending on consulting, system integration and managed services.
- Additionally, there have been five successive years of capital investment of $1 billion or more in vendors that can be attributed to the various technology categories that enable hyperautomation initiatives.
- The increased investment has fueled the growth of offerings with expanded breadth and depth within the vast vendor landscape (both organic growth and through acquisitions).

Obstacles

- **Lack of measurement of quantifiable value:** Only a few organizations (estimated at less than 20%) have mastered the measurement of hyperautomation initiatives.
- **Lack of planning for total cost of ownership (TCO) or governance:** The explosion of funded hyperautomation initiatives, coupled with the need for speed, often leaves unaddressed the all-important planning for post-production-managed operations and governance structures.
- **"Siloed" approach:** The ubiquity of hyperautomation has led to an incredible volume and velocity of adoption across functions. Unfortunately, the concurrent nature across business functions has been executed via "siloed" or diffuse purchases of technology tools, solutions and platforms.
- **Technology confusion and overspend:** There is no single vendor or technology that will enable hyperautomation initiatives. Highly fragmented and overlapping technology markets have resulted in complex architectures, overspending and lack of enterprise orchestration.

User Recommendations

- Define shared ownership and metrics. Focus on regular intervals for measurement and updates. The leading organizations in the world ensure this involves finance to facilitate public reporting of success.
- Maximize the likelihood of successful hyperautomation initiatives by architecting and planning multiple concurrent initiatives. Demand holistic mapping of collective initiatives, rather than siloes within specific functions.
- Recognize that the technology is not trivial as there is no single vendor or technology that will enable hyperautomation initiative. Focus on modularity and discoverability in the design. Take an API-first approach.
- Ensure appropriate investment in vendor management and risk competencies due to the volume of services and technologies involved.
- Establish and curate an adaptive governance structure with the goal of managing risk, and driving operational resiliency and agility while optimizing TCO.

Sample Vendors

Automation Anywhere; Boomi; Celonis; Microsoft; OutSystems; SnapLogic

Gartner Recommended Reading

The Gartner 2023 Predictions: Hyperautomation (Inclusive of AI, RPA & Low Code)

The Executive Guide to Maximizing Hyperautomation

Future of Work Trends: Hyperautomation Growth Initiatives Delivered by High-Performance Fusion Teams

Information Architecture

Analysis By: Kevin Gabbard, Andrew White

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

Information architecture (IA) is the discipline that formalizes the capabilities needed to analyze and organize the data needed to deliver business value. IA employs design patterns and increasingly active metadata to design emergent state architecture and provide guidance for sharing data assets.

Why This Is Important

Mastery of information architecture practices can define business models and alter an enterprise's competitive landscape. Consistent delivery of business insight requires information capabilities for observation, pattern identification and data sharing. Information architectures must reconcile intentional, as-designed architecture with emergent metadata-driven architecture to enable or drive business value.

Business Impact

CDAOs should be looking to information architecture for:

- Creation of information-centric business models: Monetization of information services and products.
- Strategic decision making: Visibility across the enterprise to make important decisions, which requires investments in common data models and governance.

- Enabling innovation: Access to and use of data are invaluable for enterprises that want to explore data, capture insights and pursue growth opportunities.

Drivers

- Developing a coherent viewpoint across a large number of information silos addresses enterprisewide coordination challenges.
- Active metadata and data fabric design patterns are dramatically changing the economics and work of IA by discovering and exposing established implied taxonomies, useful information assets and as yet unimagined (re)use cases.
- IA practices support the continuous analysis of requirements and enable significant assessment for the evolution of data and analytics (D&A) capabilities maps.

Obstacles

- Although best practices like employing traditional IA design patterns, implementing master data management and delivering common reference artifacts (like information flows) are known, they require long-term commitments to program building, which slows adoption.
- CDAOs and other D&A leaders have not taken accountability for IA, and it is falling into the cracks between traditional IT/CIOs, best-of-breed enterprise architecture (EA) and D&A.
- Successful IA discipline requires the coordination of diverse business and technical domains that may not initially have a consensus view on the strategic importance of information.
- The penetration has remained low because traditional future-state architecture modeling techniques are incapable of accounting for the fast-paced business dynamics surrounding data use.
- Active metadata is underutilized in developing IA. Without active metadata practices, IA will not achieve practical enterprise-scale adoption.

User Recommendations

- Focus as much on reconciling metadata silos as reconciling data silos.
- Dive deep into active metadata and data fabric design patterns. It is the key to developing agile IA practices.
- Clarify what impact information has on the organization's critical business imperatives and set IA requirements. Successful implementation of information-centric initiatives requires collaboration between D&A, IT and EA, including use of business capability models to target investments, build roadmaps, communicate change and deliver projects.
- Use metadata to interpret data user behaviors and anticipate needed solutions.
- Develop data-asset-focused objectives and key results and track progress using quantifiable business key performance indicators (KPIs).
- Explore information-focused business transformation.
- Pilot data monetization efforts.
- Identify potential new information-driven business models.

Gartner Recommended Reading

Tool: Information Architect Hiring Guide

Quick Answer: What Is Active Metadata?

5 Things a D&A Architecture Discipline Does for a CDAO

Assessing the Relevance of Data Virtualization in Modern Data Architectures

Active Metadata Management

Analysis By: Mark Beyer, Guido De Simoni, Ehtisham Zaidi

Benefit Rating: Transformational

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

Active metadata management (AMM) is the continuous analysis of all types of metadata to determine the alignment and exceptions between “data as designed” and “operational experiences.” Its utilization includes operationalizing analytic outputs, operational alerts and recommendations. It identifies the nature and extent of patterns in data operations, which can result in AI-assisted reconfiguration of data processes and use cases.

Why This Is Important

Active metadata management alters/changes the behavior of data tools, platforms and systems in how they respond to instructions from each other in a continuous learning loop. It enables cross-platform orchestration of data tools, cross-industry validation and verification processes, and the identification of flawed data capture, inappropriate usage, logical fallacies and newly encountered data. It leverages machine learning, data profiling and graph analytics to determine data relevance and validity relative to use cases. At mature levels, it can support the evaluation of analytic and data biases, as well as transparency, auditing and DataOps.

Business Impact

Active metadata management can help businesses:

- Support self-service analytics and application development by automating data content, structures, availability and discovery of data assets.
- Identify commonalities among users, use cases, and reporting and analysis models across an organization, and social networks of users based on data needs and operational requirements.
- Automate orchestration for data access, locations, performance, processing requirements and resource allocation.

Drivers

- Changing requirements from both business and IT are driving demand for data quality tools, data catalogs, metadata management solutions and data integration tools in one comprehensive solution while also recognizing and supporting data privacy in the midst of data sharing.
- Human-driven data utilization must be augmented to adapt quickly to the demand for the rapid discovery, access and incorporation of new data assets throughout an enterprise or organization.
- Active metadata experienced a significant acceleration in the last 24 months. This occurred in the midst of real progress with regard to quality, master data, integration, governance and even security tools sharing more metadata.
- Data integration is required even from sometimes distantly removed vertical industries but complicated by third-party data and data utilized from enterprise partners.
- Intercloud data demands are increasing rapidly. The large-scale capabilities in cloud-based deployments have enabled the broadest diversity of data structures, processes and use cases to date.
- Demands are emerging for organizations to be able to isolate data anomalies and classify them as errors, outliers or actual data design changes that are undocumented.

Obstacles

- Active metadata is not automated accumulation of metadata for passive utilization. Vendors/suppliers have begun using the terminology to describe any periodic update to passive metadata as "active" (active metadata "washing," sometimes deliberately). This may discredit the approach, forming resistance and barriers that will slow adoption due to skepticism in the market and push active metadata into the Trough.
- Access to all available metadata (such as social, operational, technical and business) is required for active metadata practices to reach full potential. Many existing platforms don't make internal metadata available. Automated cross-platform and tools orchestration will be inhibited as a result.
- Data management solution providers are reluctant to make their metadata assets available to — much less accept — metainstructions.
- Human designers, implementers and users might resist this approach assuming humans are the best interpreters of data value, which also deters metadata management maturity.
- There is a significant lack of metadata management standards in today's market.

User Recommendations

- Introduce an enterprise data catalog strategy, and expand it to ingest metadata beyond your data warehouses/lakes, such as from master data management, data quality, data integration, data preparation tools and analytical tools. Attach it to the catalog entries to begin accumulating metadata for analysis.
- Begin accumulating operational metadata (such as runtime logs and system workloads) from the wider D&A ecosystem. Analyze the logs for patterns of data used together and the frequency of use. Then examine user or connection strings, queries and views executed, and even resource allocation. This creates an operational knowledge graph of which data is used, how often, by whom, for what purpose and on which platform.
- Acquire or deploy at least one prototype combining at least three disciplines from data management to enable metadata notification between tools. Deploy a user interface to reconfigure metadata repositories for analysis by data engineers. This will change the data architecture and design culture to one that is observing and analyzing instead of designing from scratch — over and over again.

Gartner Recommended Reading

Tool: Vendor Identification for Active Metadata Management

Market Guide for Active Metadata Management

Quick Answer: What Is Active Metadata?

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

Deploying Effective Metadata Management Solutions

Enterprise Information Management

Analysis By: Andrew White, Mark Beyer

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Adolescent

Definition:

Enterprise information management (EIM) is an integrative discipline that guides an organization's technology choices and implementation practices to assure key information assets are captured, shared and reused appropriately. EIM can accelerate the deployment and leveraging of a data ecosystem, data fabric, data mesh or microservices approach and assist in multicloud infrastructure designs. EIM helps connect the most important information assets across the enterprise.

Why This Is Important

Data fabric and data mesh are now attracting investment (across all clouds and sources) to help mitigate the challenges of information silos. The most complex business challenges require the business to engage and own the policies and decisions that govern the most important information assets, which cannot be achieved by traditional practices that do not address technology and structural silos. EIM continues to be wanted for its benefits, while challenges persist with implementation. EIM will struggle to get out the trough until organizational maturity in the practice itself improves.

Business Impact

EIM brings the following benefits and impacts:

- As technology can help streamline and automate work, EIM provides the needed business engagement to identify existing organizational and cultural silos to help identify "the least amount of information" across an organization. EIM then introduces the governance and management requirements to meet its most prioritized outcomes regardless of the technology utilized.
- EIM helps connect data, information and analytic silos across an organization with lean business information assets.
- EIM helps coordinate data that drives the biggest business impact while improving productivity across all data and analytics (D&A) initiatives.
- EIM saves time. Connected work, processes and decisions are more effective with shared information assets.

Drivers

- Digital at scale will need critically shared and trusted data at the center of every decision, business process and outcome.
- Most organizations continue with disparate and divergent data and analytics siloed capabilities. Increasingly, business pressures such as scaling digital business or improving decision making lead to a demand to connect these silos.
- Technology-centric approaches that centralize all data (such as those based on a single data warehouse or lake) are time-consuming and resource-intensive, making the prioritization capabilities of EIM compelling.
- New technologies such as the use of ML-enabled data catalogs and design patterns such as data fabric offer promise, but even they are no substitute for identifying and managing the most critical information assets across the organization and developing an appropriate information architecture.
- Whereas master data management (MDM) is in the process of being lean, EIM extends the principles to include all information types, not just master data. This could include content, records, digital, analytics, and so on shared and exchanged across the enterprise.

- Interest continues to grow in connecting governance programs such as MDM or application data management (ADM; for example, for your application data) with data privacy, security, access and analytics governance.
- It is very common for organizations to think about EIM and start with a smaller, focused program such as MDM or ADM to help govern ERP or application data. However, these programs end up becoming silos focused on their own needs. Over time, metadata from each silo might become commonly shared and governed to help reduce impact of the solos. Data fabric would also help identify such metadata for consideration.

Obstacles

- Traditional enterprise and information architectural practices that seek to design, develop and manage all data everywhere as if the variation in use cases is considered a localized issue or not addressed at all. Newer, leaner practices are needed.
- At the same time, Gartner's CDAO surveys, year-in, year-out, suggest that culture and lack of data literacy are the biggest obstacles to effective D&A initiatives like EIM (see CDAO Agenda 2023: Presence, Persistence and Performance).
- The persistent IT preoccupation with taking a system's view of the information remains a significant barrier — instead of taking a business outcome view in which information is malleable and fungible. EIM was a well-known need for some years and firms focus explicitly on EIM. Increasingly more organizations are discovering the need anew, sometimes with new terms. Since it is a major challenge to get right, EIM will likely remain in the trough for a number of years.

User Recommendations

- Leverage outcome-driven initiatives and connect them one at a time using identified, managed and governed enterprise information. Don’t seek to manage all your data equally; use outcomes to prioritize the data and information assets across your business.
- Consider how to sequence all your D&A initiatives over a multiyear horizon. Use prioritized business outcomes to drive that sequence. This will gradually erode the system integration focus of IT practices while simultaneously promoting connecting governance practices (set policy, enforce policy, etc.) across the most widely shared information assets that underpin those prioritized outcomes and support D&A initiatives.
- Use Gartner’s data and analytics strategy and operating model framework to implement EIM (see The Foundation of an Effective Data and Analytics Operating Model — Presentation Materials). The strategy model will prioritize initiatives with business impact. The operating model can be used to connect discrete implementations of MDM, ADM, records management, metadata management, data hub strategy, D&A governance, data fabric, composable analytics, and so on.
- Use augmented capabilities and active metadata to learn patterns, relationships and usability to help prioritize your EIM initiatives.
- Use data fabric to help automate the discovery and seeded data for potential reuse and governance in your EIM program.

Gartner Recommended Reading

5 Things a D&A Architecture Discipline Does for a CDAO

Four Steps to Start an Information Architecture Practice

Strategic Roadmap for Data and Analytics Governance

Tool: How to Connect Data to Business Outcomes

Sequence Your Data and Analytics Investments to Maximize Business Value

Graph Analytics

Analysis By: Afraz Jaffri, Rita Sallam, Jim Hare, Mark Beyer

Benefit Rating: Moderate

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Definition:

Graph analytics techniques allow for the exploration and discovery of relationships between entities and concepts such as organizations, people or transactions. Graph analytics consists of models that determine the connectedness across data points. Graph analytics is typically portrayed via data visualization where surfacing relationships can lead to better-informed insights and decisions.

Why This Is Important

- Graph analytics has proven value in specific use cases (disease tracking, supply tracing, crime prevention).
- Graph analytics can often be the only effective way of analyzing data where connections and links between data items need to be identified.
- Graph analytics is an enabler of knowledge graphs, which are also accelerating in terms of market adoption.
- Graph analytics enable the exploration of connected data without the limitation of legacy data models.

Business Impact

Graph analytics helps in the following ways:

- Analyzes data for insights into relationships in complex, connected data.
- Highly effective at assessing risks to analyze fraud, route optimization, clustering, outlier detection, Markov chains and more.
- Application to digital twin scenarios where network effects and impacts of proposed changes need to be simulated.
- Identifies outlier and unusual patterns that cannot be detected by other methods.
- Augments data discovery capabilities in augmented analytics and business intelligence platforms.

Drivers

- Rapid uptake in use cases that require analysis across complex models or datasets is developed and used within machine learning (ML) with the output stored in graph databases.
- The availability of low- or no-code tools for domain experts and business users to take advantage of graph analytics techniques for complex investigations.
- The increasing maturity of graph databases for storing, manipulating and analyzing the widely varied perspectives in the graph model due to their graph-specific processing languages, capabilities and computational power.
- Established AI techniques (such as Bayesian networks) are increasing the power of knowledge graphs and the usefulness of graph analytics by adding further representational power.

Graph analysis on data can be further augmented by leveraging metadata from unexpected sources adds to the graph analysis capabilities in the following ways:

- Certain evaluations can build data “push” models by analyzing data access logs and users’ analytical model development, graph analytics can track and recommend data based on data’s relationships and users’ acceptance.
- Augmented data profiling combined with graphs can evaluate unfamiliar assets for similarities as compared to currently used datasets — identifying characteristics that are aligned to production AI techniques or ML features.

Obstacles

- Transforming data into graph data models suitable for analysis remains a substantial challenge for large-scale usage. The tooling available is largely concentrated on facilitating end-user ease of use but there is still a need for low code tools that can manage complete graph analytics workflows and life cycles or "GraphOps."
- Graph analytics and closely related graph databases are driving demand for new skills related to graph-specific knowledge, which may limit growth in adoption. Some vendors have created graph analytic solutions that make it possible to execute graph analytics using SQL.
- New skills required include knowledge and experience with graph algorithms and applying the right algorithm to solve a problem.

User Recommendations

- Prototype graph analytics techniques to address use cases that exhibit development, coding and data models that are overly complex using traditional SQL-based queries and visualizations.
- Examine graph analytics to enhance pattern analysis — especially in verticals and core use cases.
- Transition data catalog search and discovery into a graph analysis model to identify user communities' usage patterns and drive personalization applied to shared datasets.
- Implement multiexperience user interfaces with graph elements to find insights and analytic results, and store the outputs/results for repeated use in a graph database.
- Train existing personnel on how to align data assets, statistical processes and algorithms to create training datasets and build identification processes to detect data changes that will drive changes in the analytical models.
- Evaluate existing tools to determine their graph capabilities.

Sample Vendors

DataWalk; Linkurious; Neo4j; Siren; TigerGraph; Virtualitics

Gartner Recommended Reading

Graph Technology Applications and Use Cases

3 Ways to Enhance AI With Graph Analytics and Machine Learning

Use Multistructured Analytics for Complex Business Decisions

Data Engineering

Analysis By: Robert Thanaraj, Ehtisham Zaidi, Roxane Edjlali

Benefit Rating: High

Market Penetration: 20% to 50% of target audience

Maturity: Early mainstream

Definition:

Data engineering is the discipline of translating raw data into usable forms by building and operationalizing data pipelines across various data and analytics systems meeting business requirements, data governance principles and SLAs. Data engineering is a team competency that brings together three different practices — data management, software engineering and I&O — to achieve frictionless, trustworthy data delivery with agreed-upon business and technical SLAs.

Why This Is Important

Data engineering caters to the full data supply chain provisioning trusted, quality data to be used at the right business moments with agreed service levels. It enables the creation, operationalization and maintenance of data pipelines across heterogeneous environments, aimed at delivering integrated data to consumer needs. It also manages data delivery debt around reusability, governance, compliance and operational readiness through data cataloging, data quality and data observability efforts.

Business Impact

Data engineering has the following business impacts:

- Good-quality data to improve data trust among the final consumers.
- Faster time to onboard new data (own or third-party) to existing analytics and data science models with robust data pipeline change management.

- Easier fulfillment of regulatory requirements to meet data transparency expectations through cataloging efforts in tracking sensitive data and enabling data governance enforcement.

Drivers

- **Productivity through automation:** Data management teams spend most of their time on data preparation, data integration and operationalization, and as a result, these are the primary candidates for automation.

- **Agility:** Organizations are forced to change the way they traditionally work with waterfall ways of data pipeline delivery because they are unable to keep up with the demand increase and skills shortage. As a result, they must adopt DataOps approaches to be more agile.

- **Self-service and customer experience:** Organizations seek successful consumer experiences, the last mile in the "data insights decisions" continuum. Data engineering enables self-service data management among citizen users and domain experts; and also provides guardrails and established best practices to follow.

- **Cloud:** As organizations struggle to manage multiple data gravities across their on-premises and multicloud setup, the data engineering practice plays a major role in balancing the collect and connect approaches in distributed architectures.

- **Data science success:** Organizations are bound to fail if they launch data science initiatives without onboarding the necessary data engineering skills because of high-technical debt associated with data management (such as governance, compliance and operations readiness); which must be managed.

Obstacles

- **Inefficient legacy practices:** Baggage around poor integration and operations practices hurts and/or delays data engineering practice adoption.

- **Unrealistic expectations:** Many think a data engineer can "do it all," catering to the full spectrum from data management to software engineering to infrastructure and operations. Data engineering is a team competency. This involves data engineering, data architecture, data modeling, data stewardship, software engineering skills (like DevOps engineer, Python developer, test engineer), domain knowledge and operationalization expertise. Sometimes, organizations might even need adjacent roles for test engineering and infrastructure automation.

- **Inability to scale:** Data engineer roles are a critical part of data engineering teams. However, adding more data engineers in response to increasing data demands is not sustainable. Organizations need DataOps practices to streamline and scale data engineering delivery.

User Recommendations

- Establish a data engineering discipline with roles that support end-to-end delivery of data pipelines aligned to business use cases, SLAs and governance requirements.

- Catalog an inventory of data assets and make them searchable. Use metadata to drive automation of data pipelines and related artifacts. Study data usage patterns among consumers and systems, and employ this metadata to improve efficiency and optimize delivery.

- Introduce targeted, use case-specific tools — such as data warehouse automation tools and data preparation tools — to accelerate data pipeline builds and operations.

- Evaluate progress measures quantitatively — such as time to market, productivity, CI/CD automation of data pipelines, code quality and cost-efficiency of build and operations — regularly and share them with your stakeholders. As a data engineering leader, add this to your communication plan.

Sample Vendors

Ascend.io; CloverDX; dbt Labs; Nexla; Prophecy; Upsolver

Gartner Recommended Reading

5 Ways to Enhance Your Data Engineering Practices

Critical Capabilities for Data Integration Tools

Data and Analytics Essentials: DataOps

D&A Stewardship

Analysis By: Guido De Simoni, Andrew White

Benefit Rating: High

Market Penetration: 5% to 20% of target audience

Maturity: Adolescent

Definition:

Data and analytics (D&A) stewardship is the analysis, management and control of the operational processes and data needed to enforce D&A governance policies and standards. D&A stewardship may apply to data, analytics, content, algorithms, documents, images and metadata — effectively, any and all data assets as needed.

Why This Is Important

D&A stewardship promotes better behaviors across the enterprise. These behaviors emphasize continuous improvement of D&A hygiene to achieve business benefits. Organizations with established D&A stewardship practices can operationalize D&A governance more effectively.

Business Impact

By promoting and enforcing accepted data policies and standards across the organization, D&A stewardship improves the level of trusted data for business operations. Stewardship of analytics artifacts, such as machine learning (ML) models, ensures proper user behaviors and builds trust in how the business applies analytics and AI.

Drivers

- Effective D&A governance and advocacy are critical for digital transformation programs — including master data management (MDM), application data management (ADM) and business intelligence and analytics. This recognition has resulted in wider acceptance of D&A stewardship.
- When adopted correctly for D&A governance initiatives, D&A stewardship supports day-to-day business operations.
- D&A stewardship helps organizations monitor D&A against policy, identify variances and then resolve them.

Obstacles

- Despite wider acceptance of the need for D&A stewardship, many organizations rely on the reactive and heroic efforts of “citizen stewards” to solve data problems, thus holding back outcomes and decisions.
- D&A stewardship still runs as a secondary function amid the day-to-day responsibilities of business users. It is often not tied to an employee’s KPIs.
- Lack of maturity in D&A governance overall has hindered D&A stewardship, as evidenced by its slow movement through the Hype Cycle.

User Recommendations

- Align data stewardship to operational roles within business areas, as the knowledge needed for business data work might not exist in IT.
- Commit to information stewardship that spans multiple business areas, and potentially identify a lead information steward for areas where strategic programs, such as compliance, are underway.
- Clarify the chief data officer’s relationship with the D&A stewardship process in the business areas, and establish clear reporting lines for D&A stewards for consistency with desired business outcomes. IT can execute the instructions and results of stewardship (for example, data maintenance or policy execution).
- Do not outsource the work of policy enforcement. Outsourcing partners often lack context and have limited business domain knowledge.
- Establish mechanisms to recognize and reward stewardship activities and actions.

Gartner Recommended Reading

A Day in the Life of a Data and Analytics Steward

Tool: Scorecard to Identify Data Stewards in Business Teams

Climbing the Slope

Machine Learning

Analysis By: Shubhangi Vashisth, Peter Krensky

Benefit Rating: Transformational

Market Penetration: 20% to 50% of target audience

Maturity: Mature mainstream

Definition:

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

Why This Is Important

Over the last few years, ML has gained a lot of traction and is entering mainstream adoption because it helps organizations to make better decisions at scale with the data they have. ML aims to eliminate traditional trial-and-error approaches based on static analysis of data, which are often inaccurate and unreliable, by generalizing knowledge from data.

Business Impact

ML drives improvements and new solutions to business problems across a vast array of business, consumer and social scenarios, such as:

- Credit approval automation
- Price optimization
- Customer engagement
- Supply chain optimization
- Predictive maintenance
- Fraud detection

ML impacts can be explicit or implicit. Explicit impacts result from ML initiatives. Implicit impacts result from products and solutions that you use without realizing they incorporate ML.

Drivers

- Augmentation and automation (of parts) of the ML development process has improved productivity of data scientists and enabled citizen data scientists to make ML pervasive across the enterprise.
- Availability of quality, labeled data is driving ML adoption at enterprises.
- Pretrained ML models are increasingly available through cloud service APIs, often focused on specific domains or industries.
- ML education is becoming a standard at many academic institutions, fueling the supply of talent in this space.
- Active research in the area of ML in different industries and domains is driving applicability far and wide.
- Newer learning techniques — such as zero- or few-shot learning — are emerging, reducing the need to have high volumes of quality training data for ML initiatives, thus lowering the barrier to entry.
- New frontiers are being explored, including federated/collaborative, generative adversarial, transfer, adaptive and self-supervised learning — all aiming to broaden ML adoption.

Obstacles

- Conventional engineering approaches are unable to handle the growing volumes of data, advancements in compute infrastructure and associated complexities.
- ML is not the only popular AI initiative to emerge in the last few years. Organizations also rely on other AI techniques, such as rule-based engines, optimization techniques and physical models, to achieve decision augmentation or automation.
- Organizations still struggle to take their ML models into production. MLOps continues to be a hot trend and organizations look to specialized vendors and service providers for support in their journeys of better operationalizing ML models.

- Application of ML is often oversimplified as just model development. Several dependencies that are overlooked — such as data quality, security, legal compliance, ethical and fair use of data, and serving infrastructure — have to be considered in ML initiatives.

User Recommendations

- Assemble a (virtual) team that prioritizes ML use cases, and establish a governance process to progress the most valuable use cases through to production.
- Utilize packaged applications that fit your use-case requirements to derive superb cost-time-risk trade-offs and significantly lower the skills barrier.
- Explicitly manage MLOps and ModelOps for deploying, integrating, monitoring and scaling analytical, ML and AI models.
- Adjust your data management and information governance strategies to enable your ML team. Data is your unique competitive differentiator, and adequate data quality — such as the representativeness of historical data for current market conditions — is critical for the success of ML.

Sample Vendors

Amazon; ClearML; Databricks; Dataiku; Domino Data Lab; Google; H2O.ai; KNIME; Microsoft; MindsDB

Gartner Recommended Reading

Market Guide for Multipersona Data Science and Machine Learning Platforms

Market Guide for DSML Engineering Platforms

How to Improve the Performance of AI Projects

Infographic: Common Layers of Data Science and Machine Learning Activity

Use Gartner’s MLOps Framework to Operationalize Machine Learning Projects

Self-Service Analytics

Analysis By: David Pidsley, Alys Woodward, Peter Krensky, Sharat Menon, Anirudh Ganeshan, Edgar Macari

Benefit Rating: Moderate

Market Penetration: 20% to 50% of target audience

Maturity: Mature 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 fosters agility by enfranchising business analysts. It gives analysts direct access to data, enabling them to blend data, derive insights and collaborate on data visualizations. This approach reduces IT bottlenecks, accelerates decision making and enhances efficiency. While SSA is useful for rapid prototyping, complex scenarios may still necessitate IT support and analytics developer intervention for data integration, cataloging, pixel-perfect reporting or advanced analytics.

Business Impact

Self-service analytics is critical to 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 business technologists 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

- **Enhanced vendor offerings:** Analytics and business intelligence (ABI) platforms and vendors in adjacent markets continue to improve SSA capabilities, ensuring alignment with the abilities of less technical users, such as business analysts.
- **Evolving business-user needs:** As business users' information requirements advance, they expect SSA to extend into data management. Tasks such as adding data sources, selecting from data catalogs and integrating external data sources are anticipated capabilities for advanced business analysts (power users or citizen developers).
- **Decentralized budgets and spending patterns:** Compared with central IT teams, lines of business allocate a larger proportion of their overall IT budgets to D&A, emphasizing the need for self-service solutions that cater to their specific requirements.
- **Demand for timely insights:** Business users require prompt insights, but centralized teams may struggle to provide the necessary support. This support gap drives users to seek modern BI tools enabling SSA.
- **Decision-making empowerment:** SSA allows business users to access critical information and make data-driven decisions faster, uncovering valuable insights that might have been overlooked by centralized teams.
- **Analytics collaboration:** Organizations are increasingly seeking to provide environments where a diverse range of users can simultaneously co-produce analytics projects. This collaboration enables users to share knowledge, streamline workflows and drive collective decision making, further boosting the adoption of SSA.
- **Metrics stores and governance:** A virtualized layer that allows users to define and manage metrics as code supports governing metrics from data warehouses and servicing all downstream SSA, data science and business applications.
- **Generative AI:** ABI platforms are increasingly integrating large language models like GPT, which can be leveraged in data preparation, code generation, debugging, and creation of data stories and visualizations. Generative AI accelerates SSA, allowing newer users to enter this workflow. However, intelligent prescriptive applications lessen the need for visual SSA.

Obstacles

- **Governance challenges:** Inadequate user enablement and training often lead to overwhelming governance issues, hindering self-service tools' effectiveness.
- **Struggles between agility and control:** Organizations grapple with striking the right balance, risking either stifled innovation or jeopardized data integrity.
- **Intense data engineering collaboration:** The increased need for data engineering involvement creates collaboration requirements, potentially disrupting workflows and causing metric inconsistencies.
- **Cumbersome DataOps practices:** DataOps introduces complex processes that challenge organizations to adapt effectively, making analytics collaboration more difficult for business analysts.
- **Persistent data quality issues:** Organizations continue to battle poor data quality, risking misunderstandings and detrimental misuse of data.
- **Overhyped vendor claims:** Many exaggerated claims have yet to be fully realized in products, necessitating advancements in augmented analytics and data literacy programs.

User Recommendations

- Segment your users by their ability and inclination to become self-servicing, and deliver to the most prepared users first. Build data literacy and certification programs to ensure users are best prepared to add value from self-service without mistakenly delivering bad or siloed information. Success often compounds and drives further successes, and aids in improving D&A maturity over time.
- Evaluate analytics catalogs and SSA capabilities to allow business users to add curated or external sources to their data landscapes.
- Form communities (analytics franchises) consisting of both business analysts doing self-service and augmented consumers. Self-service should not be self-serving. Communities where sharing, collaboration, education, project overviews and success evangelism occur are critical as analytics audiences grow.

Sample Vendors

Domo; Microsoft; Oracle; Pyramid Analytics; Salesforce (Tableau); TIBCO Software

Gartner Recommended Reading

Critical Capabilities for Analytics and Business Intelligence Platforms

Toolkit: Create a RACI Matrix for Self-Service Analytics

Infographic: Self-Service Analytics and BI Adoption Roadmap

How to Balance Control and Agility in Your Self-Service Analytics

Rethink Self-Service by Establishing Analytics Franchises to Drive Adoption and Break Bottlenecks

Master Data Management

Analysis By: Sally Parker

Benefit Rating: High

Market Penetration: More than 50% of target audience

Maturity: Mature mainstream

Definition:

Master data management (MDM) is a technology-enabled business discipline in which business and IT work together to ensure the uniformity, accuracy, stewardship, governance, semantic consistency and accountability of the enterprise's official shared master data assets. Master data is the least number of consistent and uniform sets of identifiers and extended attributes that describe the core entities of an enterprise.

Why This Is Important

MDM is a cross-organizational collaborative effort that focuses on the consistency, quality and ongoing stewardship of master data. Master data is the subset of data that describes the core entities an organization requires to function — customers, citizens, products, suppliers, assets and sites. Master data sits at the heart of the most important business decisions, driving a need for a consistent view across business silos.

Business Impact

MDM initiatives are progressing as a foundational component of digital transformation. Leading organizations draw a causal link between their master data (parties, things and places) and the business outcomes it supports, including customer retention, supply chain optimization, and risk and regulatory compliance.

Interest in MDM extends to a broad range of vested-interest stakeholders, including finance, marketing and supply chain. MDM is now mainstream. Organizations seeking a single view of their master data recognize it as a necessity.

Drivers

- MDM is not a new concept, but adoption varies across geographic regions, with North America the most mature region, followed by Western Europe. The rest of the world is earlier in the maturity cycle and representative of markets primed for growth.
- Business process integrity eludes organizations with complex or heterogeneous application and data landscapes. Such organizations can suffer from inconsistent master data and/or a lack of trust in their master data. Organizations are increasingly recognizing the direct and causal link between this data and business outcomes, which MDM is designed to address.
- Rapidly evolving business needs, particularly in uncertain times, translate into greater demand for the benefits afforded by MDM — notability agility. The COVID-19 pandemic, which initially stalled projects, ultimately served to fast-track a broader realization of the causal link between trusted and connected master data and business resilience.
- Interest levels are increasing across a broader range of stakeholders (beyond technology), in both private and public sectors.
- A prior hesitance to embark upon MDM initiatives, due to complexity and cost, is easing.
- The barrier to entry has dropped significantly over the past two years with the broader availability of cloud-based and subscription-based MDM vendor offerings, which are now the most dominant offerings for net new clients. This lowering of the barrier to entry renders MDM viable for a broader target audience that comprises small and midsize organizations.
- A shift in mindset toward a more granular and business-outcome-led MDM program is reflected in the MDM vendors' "land and expand" strategies, where clients start small and progress toward incremental mastery of use cases and domains.
- Digital transformation requirements are forcing organizations to either start or modernize their MDM programs to leverage more recent cloud-based offerings and new augmented MDM capabilities.

Obstacles

- **Lack of consistent vendor presence:** Coverage is weaker outside North America and Europe.
- **Technology blinkers:** The prevailing pitfall remains the instinct to treat MDM as a technology initiative in isolation. Technology alone won't solve a challenge that traverses people, processes and technology.
- **Human factors:** Organizations that fail to proactively engage business stakeholders in scoping struggle to meet expectations of value and to establish an operational governance structure in service of MDM.
- **Goals:** MDM is still too often seen as an IT project. When MDM is a data or IT project that doesn't align to business outcomes, it fails.
- **Perceived complexity:** The MDM solutions market only recently shifted toward subscription pricing, cloud-based offerings and simpler products, which contribute to more approachable solutions and shorter deployment times.
- **Skills:** Successful MDM implementations require business acumen, technology and governance capabilities. Finding the right balance and availability of these skill sets remains problematic and is driving a need for third-party services as the norm.

User Recommendations

- Use business outcomes to identify the least amount of data with the greatest business impact.
- Approach MDM as a technology-enabled business-led initiative.
- Secure executive sponsorship to facilitate cross-organizational collaboration.
- Ensure that the causal link between the MDM initiative and the business outcomes it supports is clearly understood and articulated.
- Keep your master data attributes lean and focused.
- Leverage third-party services to fast-track time to value. The majority of organizations leverage external support with their MDM strategy and/or implementation. Third parties offering industry expertise and accelerators can greatly impact time to value.

Gartner Recommended Reading

3 Essentials for Starting and Supporting Master Data Management

Create a Master Data Roadmap With Gartner's MDM Maturity Model

Data and Analytics Essentials: Master Data Management — Presentation Materials

Appendixes

See the previous Hype Cycle: Hype Cycle for Data and Analytics Programs and Practices, 2022

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

2023 CEO Survey — The Pause and Pivot Year

2023 Gartner CEO and Senior Business Executive Survey: This survey was conducted to examine CEO and senior business executive views on current business issues, as well as some areas of technology agenda impact. The survey was conducted from July 2022 through December 2022, with questions about the period from 2022 through 2024. One-quarter of the survey sample was collected in July and August 2022, and three-quarters was collected from October through December 2022. In total, 422 actively employed CEOs and other senior executive business leaders qualified and participated. The research was collected via 382 online surveys and 40 telephone interviews. The sample mix by role was CEOs (n = 277); CFOs (n = 95); COOs or other C-level executives (n = 19); and chairs, presidents or board directors (n = 31). The sample mix by location was North America (n = 169), Europe (n = 105), Asia/Pacific (n = 102), Latin America (n = 29), the Middle East (n = 11) and South Africa (n = 6). The sample mix by size was $10 million to less than $50 million (n = 3), $50 million to less than $250 million (n = 51), $250 million to less than $1 billion (n = 102), $1 billion to less than $10 billion (n = 190) and $10 billion or more (n = 76). Disclaimer: 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.

Gartner Chief Data and Analytics Officer Agenda Survey for 2023: This study was conducted to explore and track the business impact of the CDO role and/or the office of the CDO and the best practices to create a data-driven organization. The research was conducted online from September through November 2022 among 566 respondents from across the world. Respondents were required to be the highest-level data and analytics leader in the organization: chief data officer, chief analytics officer, the most senior leader in IT with data and analytics responsibilities, or a business executive such as chief digital officer or other business executive with data and analytics responsibilities. The survey sample was gleaned from a variety of sources (including LinkedIn), with the greatest number coming from a Gartner-curated list of more than 4,499 CDOs and other high-level data and analytics leaders. Disclaimer: Results of this study do not represent global findings or the market as a whole but reflect sentiment of the respondents and companies surveyed.

2022 Gartner CEO and Senior Business Executive Survey, Wave 2: This survey was conducted to examine CEO and senior business executive views on current business issues, as well as some areas of technology agenda impact. The survey was conducted from July 2021 through December 2021, with questions about the period from 2021 through 2023. One-quarter of the survey sample was collected in July and August 2021, and three-quarters was collected in October through December 2021. In total, 410 actively employed CEOs, and other senior executive business leaders qualified and participated. The research was collected via 382 online surveys and 28 telephone interviews. The sample mix by role was CEOs (n = 253); CFOs (n = 88); COOs or other C-level executives (n = 19); and chairs, presidents or board directors (n = 50). The sample mix by location was North America (n = 176), Europe (n = 97), Asia/Pacific (n = 86), Latin America (n = 40), the Middle East (n = 4) and South Africa (n = 7). The sample mix by size was $50 million to less than $250 million (n = 58), $250 million to less than $1 billion (n = 81), $1 billion to less than $10 billion (n = 212) and $10 billion or more (n = 59). Disclaimer: 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.

2022 Gartner CEO and Senior Business Executive Survey, Wave 1: This survey was conducted to examine CEO and senior business executive views on current business issues, as well as some areas of technology agenda impact. The survey was conducted from July 2021 through December 2021, with questions about the period from 2021 through 2023. One-quarter of the survey sample was collected in July and August 2021, and three-quarters was collected in October through December 2021. In total, 410 actively employed CEOs, and other senior executive business leaders qualified and participated. The research was collected via 382 online surveys and 28 telephone interviews. The sample mix by role was CEOs (n = 253); CFOs (n = 88); COOs or other C-level executives (n = 19); and chairs, presidents or board directors (n = 50). The sample mix by location was North America (n = 176), Europe (n = 97), Asia/Pacific (n = 86), Latin America (n = 40), the Middle East (n = 4) and South Africa (n = 7). The sample mix by size was $50 million to less than $250 million (n = 58), $250 million to less than $1 billion (n = 81), $1 billion to less than $10 billion (n = 212) and $10 billion or more (n = 59). *Disclaimer: 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.*

Document Revision History

Hype Cycle for Data and Analytics Programs and Practices, 2022 - 28 July 2022

Hype Cycle for Enterprise Information Management, 2021 - 2 August 2021

Hype Cycle for Enterprise Information Management, 2020 - 31 July 2020

Hype Cycle for Enterprise Information Management, 2019 - 25 July 2019

Hype Cycle for Enterprise Information Management, 2018 - 31 July 2018

Hype Cycle for Enterprise Information Management, 2017 - 3 August 2017

Hype Cycle for Enterprise Information Management, 2016 - 13 July 2016

Hype Cycle for Enterprise Information Management, 2015 - 17 July 2015

Hype Cycle for Enterprise Information Management, 2014 - 6 August 2014

Hype Cycle for Enterprise Information Management, 2013 - 9 August 2013

Hype Cycle for Enterprise Information Management, 2012 - 26 July 2012

Hype Cycle for Enterprise Information Management, 2011 - 29 July 2011

Hype Cycle for Enterprise Information Management, 2010 - 28 July 2010

Hype Cycle for Enterprise Information Management, 2009 - 27 July 2009

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

Top Trends in Data and Analytics, 2023

Over 100 Data and Analytics Predictions Through 2028

How to Overcome the Top 6 Roadblocks to D&A Leader Success

How CDAOs Must Lead Data Literacy and Data-Driven Culture

Data Sharing Is a Business Necessity to Accelerate Digital Business

---

© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by Gartner's Usage Policy. Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "Guiding Principles on Independence and Objectivity." Gartner research may not be used as input into or for the training or development of generative artificial intelligence, machine learning, algorithms, software, or related technologies.

Table 1: Priority Matrix for Data and Analytics Programs and Practices, 2023

Benefit	*Years to Mainstream Adoption*			
↓	**Less Than 2 Years** ↓	**2 - 5 Years** ↓	**5 - 10 Years** ↓	**More Than 10 Years** ↓
Transformational	Hyperautomation	Data Literacy Decision Intelligence Generative AI Machine Learning	Active Metadata Management Adaptive D&A Governance Augmented FinOps Composable D&A Data Fabric Data Sharing Responsible AI	
High	Digital Ethics	AI for Sustainability D&A Stewardship Data Engineering Data Hub Strategy Data Marketplaces and Exchanges Data Observability DataOps Data Storytelling Information Architecture Master Data Management Multistructured Analytics	Data Ecosystems Enterprise Information Management Enterprise Metadata Management Sovereign Data Strategies	Connected Governance Decision Engineer

Benefit	*Years to Mainstream Adoption*			
↓	***Less Than 2 Years*** ↓	***2 - 5 Years*** ↓	***5 - 10 Years*** ↓	***More Than 10 Years*** ↓
Moderate		Data Monetization Graph Analytics Self-Service Analytics	D&A Product Management	
Low		Data & Analytics for Good		

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)