DATA-POWERED INNOVATION REVIEW

WAVE VII - 2023

Capgemini

FOREWORD

Get ready to ignite your curiosity with the latest edition of our Data-powered Innovation Review, sparkling with the excitement of a daring adventure.

As for quite a few of us winter's chill sets in, we're thrilled to share with you a selection of data-powered innovation stories that are as invigorating as a walk through Iceland's otherworldly landscapes of ice and fire. It's in these extremes, much like in business, where true resilience and beauty are found. And just like in the country's stunning vistas, the ideas in this edition are powerful and full of contrast.

Our cover captures this perfectly—visions of fire and ice not only shape the land but also mirror our efforts to forge a sustainable future while igniting the sparks of innovation with data and AI. In these pages, you'll find a tapestry of articles that reflect our commitment to a world that is not only technologically advanced but also environmentally sound and ethically grounded.

The most exciting changes tend to happen at the edge, and this is certainly the case for data and AI. As 2024 promises to be another year of space exploration, I was particularly struck by our analysis of all the learnings we can derive from data and edge computing innovations in space technology.

Then there's "Green Data: The Sustainable Foundation of Enterprise Innovation," a piece that champions the idea of eco-friendly progress. It's our way of showing that you can reach for the stars while keeping your operations firmly grounded in sustainable practices.

In our journey through innovation, we also take a pit stop at "Revolutionizing Mobility: The Data-powered Innovation Behind Peugeot's Endurance 9x8 hyper car," showcasing how high-tech can take us to new places, fast and furiously, without leaving a heavy carbon footprint.

This winter edition is not just a read; it's an invitation to explore. Like Iceland's own fusion of hot and cold, we aim to combine the heat of our initiatives with the cool, considered approach of sustainability. We hope the stories within spark your imagination and inspire actions that shape a world where progress and responsibility go hand in hand.

So, grab a warm drink (if you feel like it), settle in, and let's venture together into the landscapes of change.

NIRAJ PARIHAR

CEO, Insights & Data, Capgemini

EDITOR'S NOTE

Welcome to the 7th edition of the Data-powered Innovation Review. Packed with inspiring ideas and best practices, this edition showcases the vast potential of data and AI in technology and business. Think of this magazine as a collective campfire where some bright, innovative minds share strategies and stories that illuminate the path forward. It's a platform for sparking dialogues and igniting digital dialogues, events, and workshops, with a focus on keeping the conversation lively and ideas flowing.

In the first section, **"Data for Edge,"** we explore how data and AI are increasingly present at the periphery of traditional IT, evident in intelligent products, and even extending into space exploration. Indeed, as we approach another year of breakthroughs in data-powered innovations, we also examine how data strategy and governance are becoming integral to business – even when they extend to the very edges of central IT and central Data.

Generative AI continues to be a standout topic, maintaining its position as a bright beacon in the rapidly evolving landscape of data-powered innovation. From igniting creativity to bridging language divides, the **"Data for Augmentation"** section dives into the excitement and challenges of generative AI, touching on ethics, the need for a general ledger for large language models, and introducing our new Generative AI lab. This section also features an article on autonomous intelligence, co-created with our partner Aible, showcasing the phenomenal breadth of AI's capabilities - way beyond large language models.

Our tradition of concluding with a **"Data for Good"** section remains. Here, we present inspiring stories on sustainability and biodiversity, including insights from the winners of our Global Data Science Challenge on monitoring insect diversity, a crucial planetary health indicator. We also highlight our collaboration with the Green Software Foundation, emphasizing the importance of managing the environmental impact of data and AI.

Touching back one more time on generative AI, we assure you that no AI was harmed in creating the visuals of this edition. We did however leverage GPT-4's natural language capabilities for text improvement (but not content creation). This edition is a testament to the combined efforts of Capgemini experts, and partners like Aible, the Green Software Foundation, and Fivetran all aiming to enhance your innovation journey.

Join us at our data-powered innovation campfire for more stories and insights, as we continue to explore the dynamic, rapidly expanding world of data and AI.

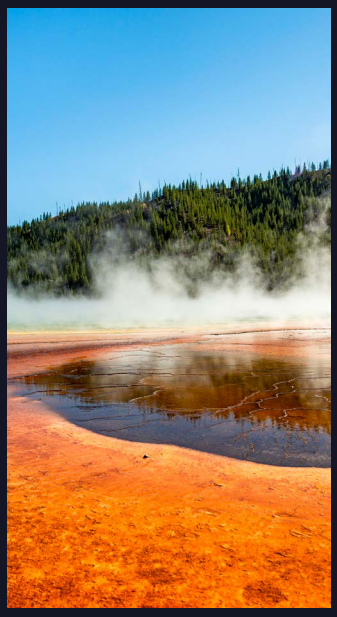
RON TOLIDO

CTO, Insights & Data, Capgemini

DATA FOR EDGE

DATA FOR AUGMENTATION

DATA FOR GOOD



DATA-POWERED INNOVATION REVIEW I WAVE VII

DATA FOR EDGE

CRAFTING AI-POWERED FUTURES THE DATA STRATEGY CANVAS UNVEILED

LIZ HENDERSON

Executive Advisor, Insights & Data, Capgemini

In a dynamic business landscape, generative AI is no longer a mere tech trend – it's the heart of boardroom discussions and pivotal business strategies. There is an urgent need for a fresh strategic approach, one that stands out from past paradigms. Enter the Data Strategy Canvas, an innovative tool tailored for this new AI-powered era. Designed for simplicity yet profound in its impact, the Data Strategy Canvas equips leaders to navigate and harness the immense potential of generative AI. Discover how to pivot, plan, and prosper with this game-changing framework.

A data strategy serves as the overarching blueprint for your data-related endeavors, providing a clear direction that facilitates effective communication and garnering support from your colleagues. In turn, it empowers your organization to align its data initiatives with broader goals, fostering growth and success. However, crafting a comprehensive data strategy is a complex undertaking. It involves navigating a landscape of diverse perspectives on challenges while concurrently envisioning a future where the organization possesses trustworthy, ethical, and shared data. Just imagine the potential if generative AI could assist in creating your data strategy.

Picture a scenario where you engage in chatbot interactions with stakeholders to explore challenges, aspirations, and opportunities. This AI is already well-versed in your organization's business strategy. The outcome? A data strategy generated with the aid of DALL·E, accompanied by a set of prioritized recommendations, each meticulously justified. The possibilities are nothing short of remarkable.

In the interim, we suggest pursuing development of your data strategy using a canvas. This canvas serves as a visual framework, guiding and structuring the critical elements and components that demand consideration during the formulation of a data strategy. To maximize its effectiveness, we recommend a collaborative effort with the business, addressing all building blocks within the Data Strategy Canvas. Here's what it encompasses.

Understand the problem: Identifying the core problems that necessitate resolution is the primary focus in any endeavor. Beginning with a clear articulation of the problem is essential to align stakeholders, delineate project scope, and gauge potential impact accurately. Failing to adopt this approach and instead leaping directly to proposing technological solutions may impede your ability to define and measure success effectively. Consequently, demonstrating progress could become arduous, hindering the approval process.

Business strategy imperatives: A crucial step is mapping ambitions to the data that will serve as a catalyst for achievement. If a business strategy is not in place, one can refer to the business vision or create a vision for the organization's desired state in the next three to five years, not forgetting the opportunities AI can bring to your operations and customer interactions. Subsequently, the elements from this vision can be transformed into specific objectives that data will empower.

For instance, these objectives could encompass endeavors such as enhancing decision-making, elevating customer experience, and fostering a culture of innovation. By aligning data-driven initiatives with the organization's ambitions, the business can effectively leverage data as a strategic enabler to drive growth, competitive advantage, and overall success.

Stakeholders: In the process of formulating and implementing your data strategy, it is vital to identify and engage with key stakeholders who play significant roles in its development and delivery. Understanding the diverse perspectives of these stakeholders, including supporters, resistant parties, and those with negative viewpoints, is crucial in navigating potential challenges and fostering a more cohesive approach.

It is essential to prioritize engaging with stakeholders whose support is critical for the success of the data strategy. For those who might be hesitant or opposed, active

efforts should be made to comprehend their concerns and explore opportunities for them to become more supportive.

Values: The value of the data strategy will be determined by a comprehensive evaluation of the problems it seeks to address. Each identified problem presents an opportunity to generate value through various means, including cost savings derived from improved efficiencies and the creation of additional revenue streams, fostering overall growth for the organization.

"IF EVERY BUSINESS, REGARDLESS OF SIZE, IS NOW A DATA BUSINESS, EVERY BUSINESS THEREFORE NEEDS A ROBUST DATA STRATEGY."

– BERNARD MARR

Data: Having identified the key measures and established objectives to pursue, the next step is to consider what data will be needed. Pinpoint: what is the source? Is the data trusted? How will it be accessed?

Also consider what elements of data management and data governance are required and, most importantly, the benefit each will provide.

KPIs: Establish KPIs and metrics to gauge the effectiveness of data-strategy execution. Measure usage, business outcomes, adoption, and trustworthiness. Align measurements with goals to demonstrate progress and benefits achieved.

Risks: Thoroughly assess potential risks that could impede the successful execution of your data strategy. These risks may involve factors such as skill availability, resource limitations, or resistance to change within the organization. Evaluate how each risk can impact the execution, identify potential roadblocks, and devise suitable mitigating measures to address them effectively.

Actions: Having obtained the necessary insights, the next crucial step is creating your data strategy. We recommend compiling a concise slide pack for your data strategy, encompassing the following key components of: problem statement, challenges and opportunities, high-level phasing of key activities, a detailed plan for the next three months, and a high-level plan from month four onwards.

By encapsulating these key elements, you can effectively communicate your data strategy to stakeholders, ensuring clarity, alignment, and commitment to the successful implementation of the plan.

Change: Change is fundamentally driven by people, making data a "people sport." It is essential to recognize that successful

data initiatives hinge on understanding and managing the human element. This significance is further emphasized by reports from our own research institute and analyst bureaus, indicating that a substantial portion of transformation projects – over 70 percent – fall short of achieving their intended outcomes, primarily due to cultural factors. To ensure success in your data endeavors, careful consideration of how your people will be impacted by the changes is paramount. Implementing measures to support their adaptation to these changes will prove instrumental in overcoming barriers.

Communications: In the context of change activity and facilitating increased adoption, effective communication plays a pivotal role. Clear and concise communication is essential, conveying what changes are occurring, the underlying reasons for these changes, and most importantly, highlighting the benefits for the individuals involved.

While the elements mentioned above provide a comprehensive data strategy, the specific needs of each organization will vary. The canvas should be customized to fit your organization. It serves as a visual guide and reference point for aligning various stakeholders, capturing the most relevant aspects to effectively guide and align your activities.

Until an AI data strategy chatbot is available, your next step is forming your answers into a strategy for approval and communication.

DATA STRATEGY CANVAS

PROBLEM

- WHAT IS THE PROBLEM?
- WHY IS IT A PROBLEM?
- WHOSE PROBLEM IS IT?
- WHO WILL AND HOW WILL THEY BENEFIT FROM SOLVING THIS PROBLEM?
- WHY, WHY, WHY?
- HOW TO PRIORITIZE PROBLEMS?

QUESTIONS TO ASK

- WHAT WOULD YOU LIKE TO DO TOMORROW THAT YOU CAN'T DO TODAY?
- WHAT ARE THE BIGGEST PAIN POINTS/ CHALLENGES?
- WHAT ARE THE OPPORTUNITIES IF YOU HAD GOOD DATA?

STAKEHOLDERS

- WHO IS THE BOARD SPONSOR?
- WHO ARE THE KEY STAKEHOLDERS?
- HOW WILL ONGOING ENGAGEMENT BE GENERATED?

VALUES

- WHAT IS THE SIZE OF THE PROBLEM?
- WHAT IS THE BASELINE?
- HOW WILL IT BE MEASURED?
- WHAT IS THE UPLIFT/SAVINGS/ DIFFERENCE?

DATA

- SOURCE
- TRUST/QUALITY
- ACCESS/AVAILABILITY
- PROCESS/AUTOMATION
- DATA GOVERNANCE
- DATA MANAGEMENT
- CULTURE
- ANALYTICS
- ARCHITECTURE
- GENERATIVE AI

KPI'S

- WHAT DOES SUCCESS LOOK LIKE?
- HOW TO EVALUATE PROGRESS?
- WHAT METRICS SHOULD BE USED?
- HOW WILL THE KPI'S BE DISPLAYED?

ACTIONS

- WHAT ARE THE NEXT STEPS?
- WHO IS ACCOUNTABLE?
- WHO IS RESPONSIBLE?
- HOW WILL PROGRESS BE COMMUNICATED?
- WHAT IS THE ESCALATION PROCESS?

CHANGE

- WHATCHANGEISNEEDED?
- WHO IS IMPACTED?
- WHATTRAININGISNEEDED?
- WHO IS RESPONSIBLE FOR DEVELOPING THE TRAINING?
- HOW WILL YOU MEASURE THE IMPACT OF THE SUPPORT AND TRAINING OFFERED?

RISKS

- WHAT ARE THE RISKS?
- WHAT IS THE IMPACT?
- WHAT ARE THE MITIGATION PLANS?
- WHERE WILL RISKS BE RECORDED?
- WHO OWNS THE REGISTER?

BUSINESS STRATEGY IMPERATIVES

- WHAT ARE THE AIMS/OBJECTIVES OF THE BUSINESS STRATEGY?
- WHO WILL DATA ENABLE THESE OBJECTIVES TO BE ACHIEVED?
- DO YOU HAVE A VISION, DOES THIS NEED DATA TO ENABLE SUCCESS?

COMMUNICATIONS

- HOW WILL THE BOARD PROGRESS BE COMMUNICATED
- HOW WILL THE STAKEHOLDERS BE UPDATED?
- HOW WILL THE WIDER BUSINESS BE UPDATED?
- WHAT DO PEOPLE DO IF THEY WANT TO KNOW MORE/GET INVOLVED?
- WHAT COMMUNICATION CHANNELS WILL BE USED?
- WHO WILL LEAD THE COMMUNICATIONS?
- WHAT WILL BE THE CADENCE FOR COMMUNICATIONS?

INNOVATION TAKEAWAYS

#DATA
#CANVAS
#DATASTRATEGY
#GENAI
#FRAMEWORK

AI-DRIVEN EVOLUTION

Learn how artificial intelligence is revolutionizing data strategy development, enabling businesses to harness the full potential of their data for strategic decision-making and growth.

CANVAS FOR EXCELLENCE

Leverage the Data Strategy Canvas to uncover the key components and challenges of creating your data strategy, setting you on the path to data-driven excellence and competitive advantage in today's rapidly evolving business landscape.

FUTUREPROOFING YOUR BUSINESS

Understand how a data strategy not only drives current success but also prepares your organization for the data challenges and opportunities of tomorrow, ensuring long-term sustainability.

HAMILTON'S SHOT
AGE-OLD WISDOM FOR MODERN DATA GOVERNANCE

DINAND TINHOLT

Vice President, Insights & Data NA, Capgemini

Navigating our data landscape, we're reminded of the Federalist Papers, lessons of these political essays aren't solely American. Just as these papers advocated balance in power, today's data world needs clear rules and unity. Ever notice how data silos resemble old political divides? And just as history urged checks on power, we need safeguards for our data. Let's dive in and see how these age-old principles can guide our techy, data-powered future.

Amidst the swirling mists of America's turbulent birth, Alexander Hamilton emerged: a Caribbean-born orphan turned Revolutionary War hero. Hamilton was a visionary, a contentious person who melded ambition with intellect, shaping America's institutions with fervor, but often clashing with contemporaries, revealing the complexities of revolutionary ideals. These lessons aren't solely American and as an eloquent contributor to the enigmatic Federalist Papers, his life is forever etched in history.

The Federalist Papers were crafted to promote the ratification of the newly proposed U.S. Constitution. They argued for a stronger central government while ensuring checks and balances. These essays offer lessons for modern data governance. Modern data governance refers to the comprehensive management of data availability, usability, integrity, and security within an organization. In a digital age characterized by vast amounts of data, this framework ensures that data is handled efficiently, securely, and in compliance with applicable regulations.

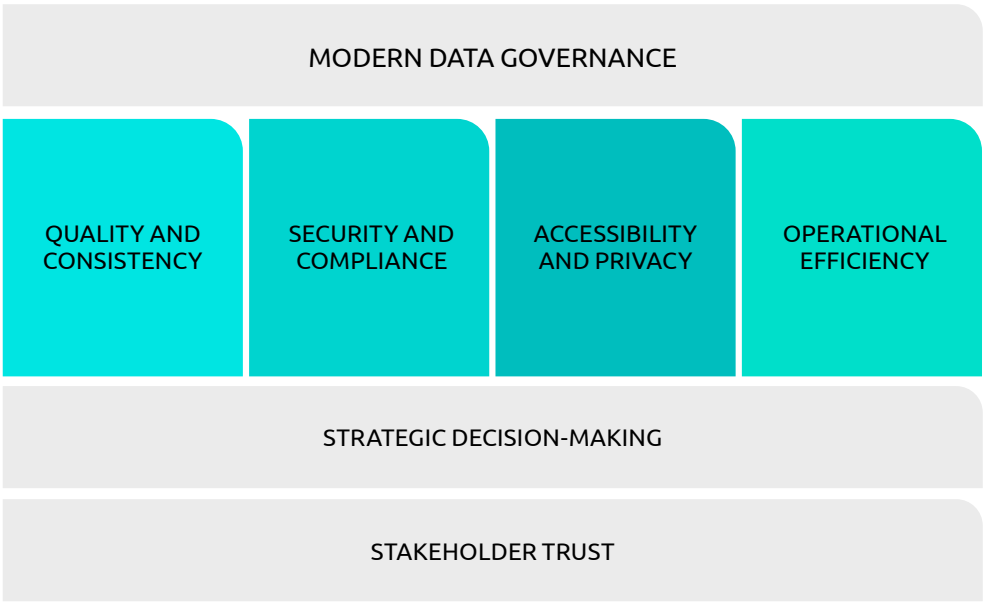
For organizations, modern data governance implies the following:

- **Quality and consistency:** Data governance ensures that data is consistent and reliable across the organization. This is crucial for accurate analytics, forecasting, and decision-making.
- **Security and compliance:** With the rise of data breaches and stringent data protection regulations like GDPR and CCPA, businesses must prioritize data security. Modern data governance helps identify risks and establishes protocols to safeguard sensitive information.
- **Accessibility and privacy:** While data should be accessible to those who need it, there must also be controls in place. Governance policies determine who can access specific data sets, ensuring that employees can perform their roles without compromising data privacy.
- **Operational efficiency:** Proper data management minimizes redundancies and errors. When everyone in an organization refers to a single, unified source of truth, it streamlines processes and boosts productivity.
- **Strategic decision-making:** High-quality, consistent data is a goldmine for businesses. It paves the way for insightful analytics, enabling leaders to make informed strategic decisions.
- **Stakeholder trust:** Customers and partners are more likely to trust organizations that handle data transparently and responsibly. Effective data governance not only ensures compliance but also builds trust, a pivotal factor for customer loyalty and business growth.

Modern data governance is not merely a technical or IT function. It's a strategic imperative that affects every facet of a business. In a data-driven world, effective governance is the bedrock upon which successful, trustworthy, and compliant businesses are built.

Back to the ***Federalist Papers***: the challenge was to balance state and federal rights – just as today's challenge is balancing personal privacy with free-flowing

information. The rapid evolution of technology has thrust data into the center stage of modern commerce and daily life. As we grapple with the challenges of managing, protecting, and utilizing this vast digital wealth, it's intriguing to find parallels in the Federalist Papers that can guide our way.



THE NEED FOR STRONG CENTRAL GOVERNANCE

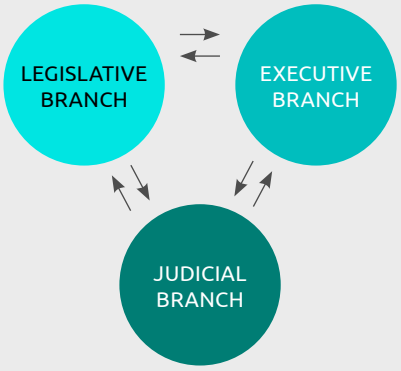
In the ***Federalist Papers***, Hamilton emphasized the importance of a strong central government, warning against the potential chaos of loosely federated states. This mirrors the contemporary need for centralized appropriate governance within organizations. With fragmented data silos, businesses face inefficiencies and inaccuracies. A computational federated governance, guided by unified policies and connected and "harmonized" through API integration while having the freedom for local extensions, can ensure consistent, timely, and quality data access, much in the same way Hamilton's centralized governance aimed to bring order and efficiency to a fledgling nation.

BALANCING POWER AND AUTONOMY

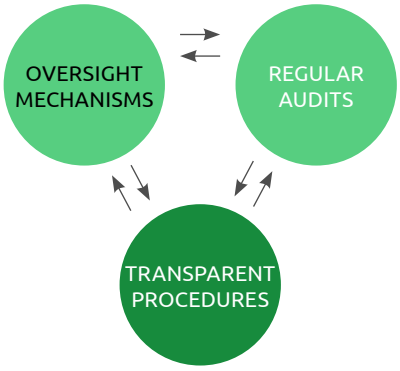
Hamilton often grappled with the balancing act between centralized power and individual rights. Similarly, modern data governance requires a delicate equilibrium between organizational control over data and individual privacy rights. While businesses need data for analytics and decision-making, they must also respect user privacy, making sure not to overstep boundaries and infringe upon personal rights.

A SYSTEM OF CHECKS AND BALANCES

A key feature of Hamilton's vision was the system of checks and balances among the branches of government. This ensures that no single entity wields unchecked power. Translated to data governance, this implies the necessity for robust oversight mechanisms, regular audits, and transparent procedures that ensure data quality, security, and ethical usage. Internal and external checks must exist in the realm of data governance to prevent misuse.



CHECKS AND BALANCES IN THE FEDERALIST PAPERS



CHECKS AND BALANCES IN MODERN DATA GOVERNANCE

"WHEN YOU GOT SKIN IN THE GAME, YOU STAY IN THE GAME. BUT YOU DON'T GET A WIN UNLESS YOU PLAY IN THE GAME."

– ALEXANDER HAMILTON *FROM THE MUSICAL HAMILTON*

EMBRACING ADAPTABILITY

Throughout the ***Federalist Papers,*** the framers acknowledged that the Constitution must adapt to the nation's evolving needs. As technology and digital landscapes shift, data governance policies must be flexible enough to accommodate these changes. Rigid frameworks are likely to become obsolete or even counterproductive. The dynamism that Hamilton advocated in governance is equally crucial in contemporary data strategies.

BUILDING TRUST AND CREDITWORTHINESS

Hamilton believed that for the nation to prosper, it must earn trust and creditworthiness. In the digital age, data is akin to currency, and trust is paramount. Organizations must employ transparent, ethical, and secure data practices to build and maintain this trust. A breach in data integrity can be as catastrophic for a company's reputation and growth as a breach in financial trust would have been in Hamilton's era.

PUBLIC PARTICIPATION AND CONSENSUS

One underlying theme in the ***Federalist Papers*** is the value of public participation and consensus in governance. Several of the essays highlight the importance of understanding and accounting for various stakeholder needs. In the context of data governance, this suggests involving diverse departments – from IT to HR to sales – in framing data policies. A collaborative approach ensures diverse needs are met, potential pitfalls are identified, and broad-based consensus is achieved.

While the digital challenges of the 21st century might seem worlds apart from the political and constitutional challenges faced by Hamilton and his contemporaries, the core principles of governance, trust, and adaptability remain unchanged. The Federalist Papers, with their profound wisdom on governance, can serve as a beacon for modern businesses and policymakers navigating the complex terrains of data governance. It is however crucial to "play in the game" as the quote at the beginning of this article highlights, or in our context: use the data you are governing.

INNOVATION TAKEAWAYS

#DATAGOVERNANCE
#DATATRUST
#DATAQUALITY

MODERN AND STABLE

Modern data governance requires that data is handled efficiently, securely, and in compliance with applicable regulations.

BALANCING ACT

Strong central governance is needed with a balance between power and autonomy and a system of checks and balances.

COLLABORATIVE EFFORT

Organizations should embrace adaptability in their approach to data governance, build trust and creditworthiness, and drive participation and consensus.

INTELLIGENT PRODUCTS BOOST CUSTOMER EXPERIENCES BUT COMPANIES SHOULD BE CAREFUL HOW DATA IS USED

NICOLAS ROUSSEAU

Group Offer Leader - Intelligent Products and Services, Chief Digital Engineering and Manufacturing Officer, Capgemini Engineering

FABIO FUSCO

Data & AI for Connected Products Centre of Excellence Director Hybrid Intelligence, Capgemini Engineering

Intelligent products and services are perpetually evolving, and there is increased demand to use data for real-time innovation to deliver the hyper-personalized experience customers expect. Companies that strategically use data-driven technology such as predictive analytics will gain a competitive advantage in customer experience, but they must also prioritize the challenges inherent in collecting, using, and safeguarding client data.

Track how quickly your heart rate returns to resting after a brisk trail run. Pre-program your car to have the seats already heated for the drive back. Personalize your home security settings with facial recognition to unlock your door and turn on the lights as you approach your house. The data generated and collected from such intelligent products can optimize our lives in dynamic ways, with advancements on the horizon.

"Intelligent products that can adapt their performance based on customer needs will make the competitive difference," said Nicolas Veauville, Product Innovation Leader, Versuni kitchen appliances, formerly Senior Director at Philips Domestic Appliances, according to the 2022 report Intelligent products and services, produced by the Capgemini Research Institute. "I believe that bringing an intelligent product is not the end goal; the end goal is to create a better experience for our customers, so that it drives more revenue over the lifetimes of our brands."

Using artificial intelligence (AI) and machine learning (ML) allows for continuous improvements from data usage, ensuring a product's longevity, not obsolescence. And as products improve their behavior over time, they become more adaptive and responsive to specific user needs. All this power provides opportunities for companies to gain a competitive edge when delivering a hyper-personalized experience for customers, but it also presents three distinct challenges.

ENSURING PERSONALIZATION WHILE SAFEGUARDING CLIENT PRIVACY

Designing intelligent products that enhance users' health and fitness journeys by providing personalized prompts and insights can build customer loyalty, such as a cutting-edge wearable device that cues users to pick up the pace for the last kilometer of a run, guiding them to a personal record. The Capgemini report, which compiled results from a survey of 1,000 global companies, found that 87 percent of organizations say intelligent products and services are crucial to their business strategy. Yet previous research by Capgemini highlighted that 62 percent of the organizations it surveyed cited cybersecurity and data-privacy threats as reasons they struggle to scale up IoT, or Internet of Things, applications.

Identifying a solution to these challenges means finding a balance to help companies achieve hyper-personalization by combining several techniques to preserve user privacy:

- **Federated learning:** AI models are trained across decentralized user data while keeping the information on users' devices. This ensures sensitive information never leaves their control. Biometrics are used to link identity with an action, but the data stays on the local device.
- **Differential privacy:** Differential privacy mechanisms are integrated into data collection and analysis to anonymize and safeguard user data. That means personally identifiable information is never leaked.
- **User consent management:** Users are offered clear choices about data usage and personalized features, respecting their preferences to opt in – or out – and decide what types of data to share.

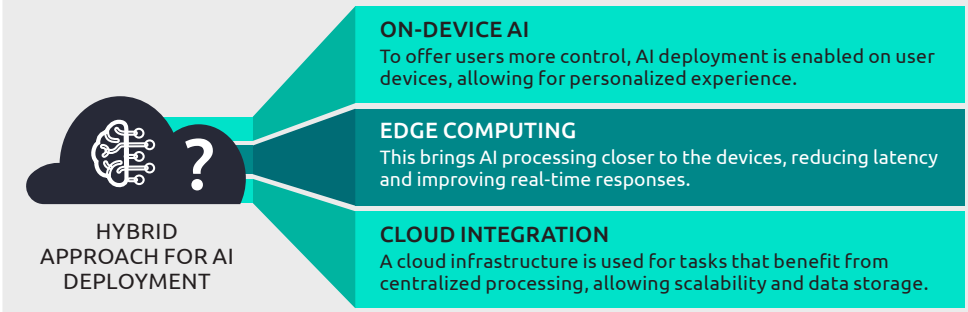
SELECTING OPTIMAL AI DEPLOYMENT LOCATIONS

As touched on above, consider the use case of a company specializing in intelligent home automation devices. It needs to decide where to deploy artificial intelligence within its products, whether to centralize AI processing or distribute it across devices. Other considerations

include performance, security, cost, and sustainability.

The ideal outcome of choosing an optimal AI deployment location includes enhanced product performance, robust security, cost efficiency, and a reduced environmental footprint. However, the wrong choice could lead to poor user experiences and financial setbacks. So, what's the solution? We recommend a hybrid approach that combines three elements:

- **On-device AI:** To offer users more control, AI deployment is enabled on user devices, allowing for personalized experiences.
- **Edge computing**: This brings AI processing closer to the devices, reducing latency and improving real-time responses.
- **Cloud integration:** A cloud infrastructure is used for tasks that benefit from centralized processing, allowing scalability and data storage.



ADOPTING A DATA-FIRST APPROACH IN PRODUCT AND SERVICE DEVELOPMENT

Consider the case of a car maker that wants to leverage data as an asset for value creation rather than merely accumulating it for future extraction. The issue is when it amasses data without a clear strategy for its use. In fact, the 2022 report found that 53 per cent of companies stated they had "critical talent gaps" in "data governance, management, and data science professionals."

In comparison, a car maker that adopts a data-first approach will benefit in several ways, including real-time vehicle performance enhancements, dynamic safety features, and immediate personalization of the driving experience. However, the challenge lies in utilizing data effectively and promptly. Once again, Capgemini recommends combining several best practices rather than pursuing a single approach.

- **Data as a strategic asset:** Data is viewed as a strategic asset and a valuable resource, driving decision-making with regard to product features, design, and improvements that are made based on data analysis, supported by empirical evidence from data.
- **User-centric design:** Start with the customers' preferences and needs, to ensure data-driven enhancements align with their expectations.
- **Impermanent data:** Move from "store all the data" to "store the right data," rethinking practices for cost-efficiency, privacy, and sustainability by emphasizing the efficiency of collecting only essential data.

Intelligent products, whether used by avid trail runners or automotive manufacturers to manage fleet vehicles at scale, are making our day-to-day lives more seamless and helping to solve complex global issues. Companies that use data to improve the customer experience – prioritizing people and embedding privacy in data-driven product design – will benefit from long-term loyalty that drives growth and profitability.

INNOVATION TAKEAWAYS

"INTELLIGENT PRODUCTS THAT CAN ADAPT THEIR PERFORMANCE BASED ON CUSTOMER NEEDS WILL MAKE THE COMPETITIVE DIFFERENCE. I BELIEVE THAT BRINGING AN INTELLIGENT PRODUCT IS NOT THE END GOAL; THE END GOAL IS TO CREATE A BETTER EXPERIENCE FOR OUR CUSTOMERS, SO THAT IT DRIVES MORE REVENUE OVER THE LIFETIMES OF OUR BRANDS."

– NICOLAS VEAUVILLE, VERSUNI

#DATAMANAGEMENT
#PERSONALIZATION
#DATAGOVERNANCE

BALANCE PERSONALIZATION AND PRIVACY

Enable hyper-personalized experiences but integrate privacy mechanisms into the data collection and allow users to retain control of the types of information they share.

CHOOSE AN OPTIMAL AI DEPLOYMENT LOCATION

Use a combination of edge computing, cloud integration, and on-device AI to deliver a solid user experience, robust security, cost efficiency, and a reduced environmental footprint.

ADOPT A DATA-FIRST APPROACH

Strategic planning and decision-making should start with data to develop and enhance intelligent products that put the customer experience and outcomes first.

HARNESSING THE FUTURE OF DATA MANAGEMENT
THE EMERGENCE OF SAP DATASPHERE

MITALEE INGALE

Innovations Lead at I&D
SAP CoE, Insights & Data, Capgemini

ASHISH LAKDAWALA

SAP Datasphere Competency Lead at
I&D SAP CoE, Insights & Data, Capgemini

In the intricate world of data management, a new chapter is being written with the introduction of SAP Datasphere, a visionary technology that marks a milestone in the evolution of enterprise data systems. This innovative platform is not just another addition to the vast arsenal of data solutions; it is a transformative force that is reshaping the landscape of data management. By synergistically merging the robustness of SAP's Business Technology Platform with a seamless data fabric foundation, SAP Datasphere stands as a testament to the future-proofing of data management.

SAP Datasphere is not a mere technological advancement; it is an integrated solution crafted to ensure that legacy Business Warehouse (BW) investments continue to deliver value well into the future. The platform is designed to instill trust and confidence, anchoring itself as a reliable cornerstone in an organization's data strategy. It achieves this by providing a sturdy and flexible infrastructure that adapts to the ever-changing demands of the digital landscape.

STRATEGIC ALLIANCES ELEVATING DATA POTENTIAL

The true might of SAP Datasphere is further magnified through strategic alliances (Collibra, Databricks, Confluent, DataRobot and Google Cloud) that enhance its capabilities. These partnerships are not just a footnote in the platform's capabilities but a significant leap forward, offering a unified experience that is unparalleled in the industry. The platform delivers end-to-end data integration and self-service capabilities that are redefining the very concept of data warehousing. Through these collaborations, SAP Datasphere transitions from a powerful tool to an indispensable ally in data management.

REIMAGINING DATA WAREHOUSING

The innovations brought forth by SAP Datasphere are not merely incremental improvements but a reimagining of data warehousing as we know it. It presents a suite of tools and functionalities that promise a revolutionized approach to handling data. The architecture, designed for simplicity, strips away the complexities that have long plagued data-management systems. By doing so, it opens up the doors to data democratization, allowing for an unprecedented level of access and utilization of data across all levels of an enterprise.

DATA DEMOCRATIZATION AND THE PROMISE OF ACCESSIBILITY

SAP Datasphere is not just about storing data; it's about breaking down the barriers that restrict its flow and utilization. The platform champions the concept of data democratization with its 'Data Marketplace', empowering individuals within an organization to access and leverage data with an ease that was previously unthinkable. This empowerment is a driving force behind the creation of a data-rich future, where decisions are informed by insights drawn from a wealth of data.

AN ESSENTIAL PILLAR FOR THE DATA-POWERED ENTERPRISE

SAP Datasphere acts as the data fabric foundation and provides end to end capabilities for data ingestion, self-service data warehousing, data governance, data trust and consumption. These compelling capabilities of SAP Datasphere firmly position it as an essential pillar for any enterprise that seeks to harness the full potential of its data assets. It is a platform that speaks to the heart of what it means to be data-driven in today's fast-paced, information-centric world. SAP Datasphere is not just a solution for today; it is a strategic investment that prepares businesses for the demands of tomorrow.

EVOLVING PLATFORM

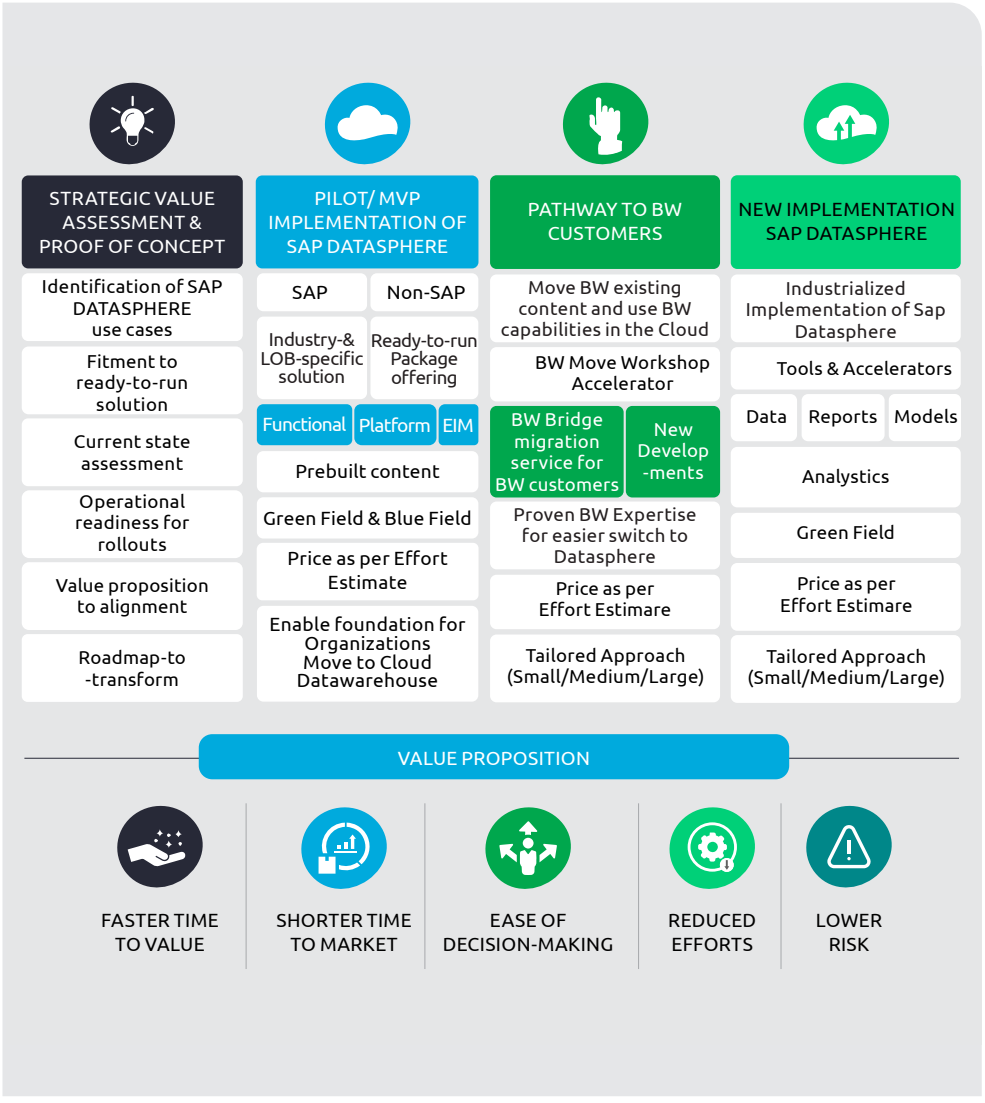
Like any new product in the market, SAP Datasphere has gone through its own maturity curve. Being on SAP's strategic roadmap, it has quarterly releases, and SAP is planning more critical capabilities like real-time connection for SAP Analytics Cloud Planning, integration with SAP's Sustainability Product Portfolio, and enhanced features for embedded data intelligence. We are continuously working with our clients to prove this new technology for better adoption.

"THIS PLATFORM IS NOT JUST A STEP FORWARD; IT IS A LEAP INTO A FUTURE WHERE DATA IS NOT ONLY ABUNDANT BUT ALSO READILY ACCESSIBLE AND TRUSTED."

STEPPING INTO A DATA-RICH FUTURE

The advent of SAP Datasphere is a pivotal moment for data management. It offers a vision of what is possible when innovation meets the need for reliability and simplicity. This platform is not just a step forward; it is a leap into a future where data is not only abundant but also readily accessible and trusted. For enterprises that are poised to embrace this future, SAP Datasphere is more than a tool; it is a transformative partner that will journey with them into the horizon of a data-rich destiny. The approach should be that of a business-driven transformation, to maximize the benefits around data democratization, fueled data trust and self-service data warehousing.

Based on the first implementations by Capgemini at clients and with our trusted partnership with SeaPark, we have developed an accelerated, use-case and value-driven approach to rapidly reap benefits from the innovative power of SAP Datasphere, while harnessing existing investment with 'BW Bridge' and assets. This brings automated tools, accelerators, and a proven framework to migrate safely and effectively towards the new, powerful world of SAP Datasphere. Also, the current SAP BW/HANA workforce at both client organizations and providers can easily be upskilled to SAP Datasphere, ensuring the right skills are available for a large-scale ramp up of this exciting platform.



INNOVATION TAKEAWAYS

#SAPDATASPHERE
#DATAPOWERED
#DATAMASTERS
#DATAMESH

STRATEGIC EVOLUTION OF DATA INFRASTRUCTURE

SAP Datasphere stands as a significant innovation in the data-management sphere by not only enhancing the robustness of SAP's Business Technology Platform but also by establishing a seamless data fabric that future proofs data-management strategies. This evolution signifies a shift towards a more integrated and flexible data infrastructure capable of adapting to the dynamic needs of the digital landscape.

ALLIANCE-DRIVEN DATA MANAGEMENT

The platform's strategic partnerships have broadened its horizon, transforming SAP Datasphere into an essential tool that offers unified and end-to-end data management capabilities. These alliances redefine data warehousing, providing self-service capabilities that allow for a seamless and integrated data-management experience, setting a new standard for industry collaborations in enhancing data solutions.

ADVANCING DATA ACCESSIBILITY AND EMPOWERMENT

By embracing the principles of data democratization, SAP Datasphere has innovated the accessibility of data, allowing for an unprecedented level of enterprise-wide engagement with data assets. This shift ensures that data is not only more accessible but also empowers individuals within an organization to leverage insights, fostering an environment where informed decision-making is the norm and facilitating a business-driven transformation centered on the value of data.

SPACE-DATA-AS-A-SERVICE
THE PROMISE AND POWER OF EXTRA-TERRESTRIAL DATA

SUDARSHAN SAHU

Process Lead, Emerging Technology Team,
Data Futures Domain, Capgemini

Imagine Captain Picard navigating new galaxies, guided by the vast wealth of space data – from Earth-observing satellites, and cosmic telescopes to the International Space Station. The term "space data" encompasses this immense trove of extra-terrestrial information. With an explosion in data volume and the entry of private players in the space arena, satellite launches are skyrocketing while costs plummet. This new era of affordability is set to unlock unprecedented opportunities and spark innovations across various sectors.

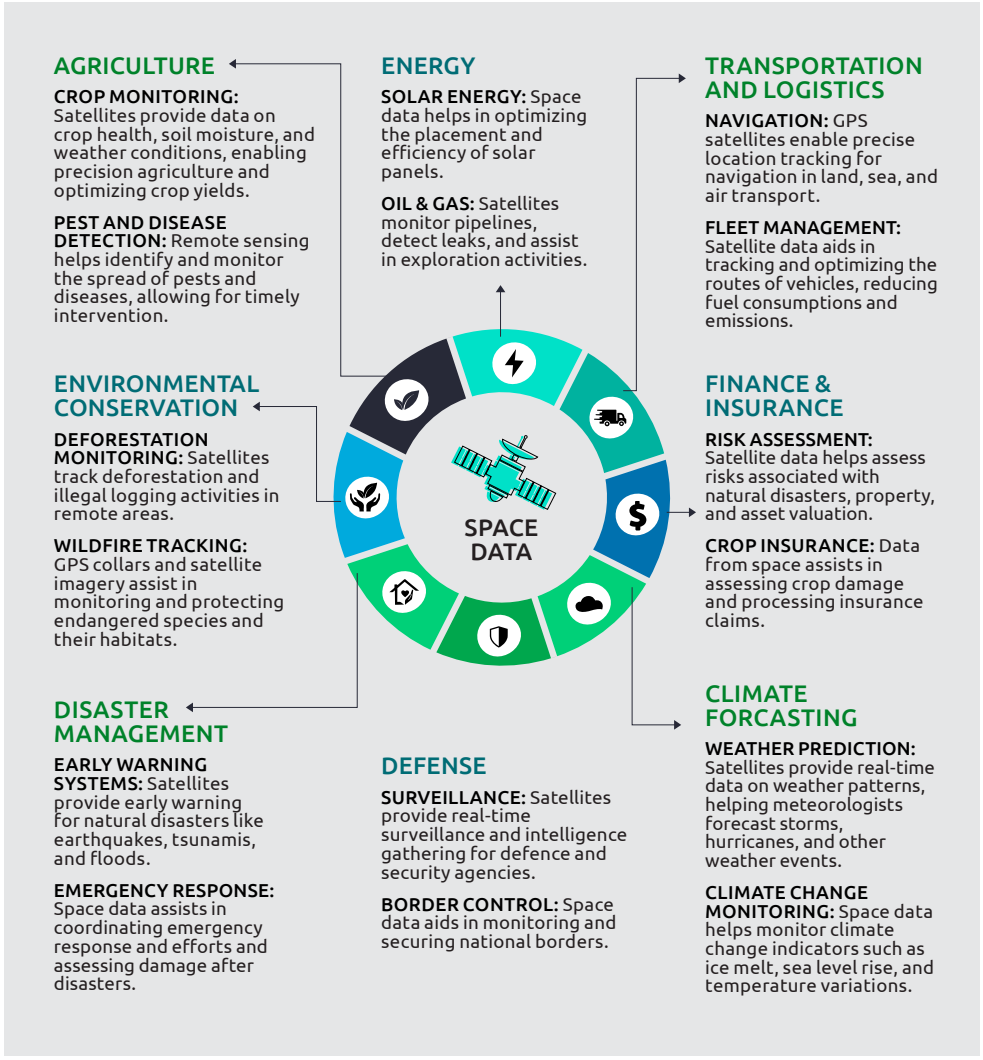
The demand for space data is on the rise, and it's not difficult to understand why. Space data offers immense value for various industries and sectors. From scientific research to commercial applications and governmental policymaking, the demand for space data is driven by its versatility and ability to provide critical insights. Further, with the expansion of the space economy, large and small businesses are embracing new business models, including Space-Data-as-a-Service (SDaaS).

Space data is essential for addressing a wide array of societal, economic, and scientific challenges. It provides a unique vantage point from space that enables us to monitor, understand, and respond to changes and events on Earth and beyond. This can include information about weather patterns, natural disasters, agriculture, urban development, climate change, and much more.

KEY COMPONENTS OF SPACE-DATA-AS-A-SERVICE:

- **Satellites and space missions:** At the heart of SDaaS are the satellites and space missions that capture a vast array of data about our planet and the cosmos. These include Earth observation satellites, weather satellites, space telescopes, and planetary missions.
- **Data acquisition and processing:** SDaaS providers use ground stations and advanced data-processing techniques to receive, process, and analyze the data collected by satellites. This involves converting raw data into usable formats and extracting relevant information.
- **Data storage and management:** Managing the enormous volumes of space data requires robust storage and management systems. Cloud computing and big data technologies are often employed to store, organize, and retrieve data efficiently.
- **Data visualization and analysis tools:** To make sense of the data, SDaaS platforms offer data visualization tools and analytics capabilities. Users can create maps, charts, and models to gain insights from the information.

KEY APPLICATIONS OF SPACE DATA ACROSS VARIOUS INDUSTRIES:



There has been number of applications of SDaaS. Satellite data along with remote sensing technology can support scientists and conservationists to identify, track, monitor, and access wildlife in the most remote areas around the world, including marine habitats. Farmers and agricultural organizations can access satellite imagery and climate data to optimize crop planning, monitor soil conditions, and enhance yield predictions. Space data also aids in disaster preparedness and response by providing real-time information on natural disasters such as hurricanes, wildfires, and earthquakes. Urban planners use space data to monitor urban growth, traffic patterns, and land use, helping cities make informed decisions about infrastructure development. Environmental conservationists rely on space data to monitor wildlife habitats, deforestation, and illegal activities such as poaching and logging.

SDaaS provides access to real-time and high-resolution data, enabling faster decision-making and more precise planning. Organizations can save costs by outsourcing data acquisition and processing, eliminating the need for expensive satellite infrastructure. Space data is not limited by borders, providing valuable insights on a global scale, which is especially critical for international organizations and governments. Space data can also help mining companies map emissions, monitor shipments along the supply chain, and improve exploration efforts by identifying mineral-rich areas. The commercial sector can harness space data for various purposes, including telecommunications, agriculture, and natural resource exploration.

For decades, satellite data was primarily available for government and defense purposes. Further, considerable private-sector investment in space-related

technological innovation has led to unprecedented levels of access to space for organizations of all shapes and sizes. For example, Coca-Cola has been using various data sets such as satellite imagery and weather data to help manage some of its more remote, inaccessible locations, for internet access and supply chain management. Similarly, Scottish companies are planning to achieve net zero by 2045 using space data to create crop maps for agricultural supply chains, using satellite and sensor data to monitor peatlands, and assessing global shipping emissions from space.

The rise in space economy is driven by the introduction of low-cost satellites by increasing accessibility to space for a wide range of organizations and purposes. Further, advancements in miniaturization of satellite components (CubeSats) and the development of reusable launch vehicles (like SpaceX Falcon 9 and Blue Origin) have reduced launch costs from $65,000 per kilogram to $1,500 per kilogram, a greater than 95 percent decrease. According to the United Nations Office for Outer Space Affairs, the number of satellites launched into space has grown by 95 percent in 2022 compared to 2020. In addition, McKinsey estimates there will be more than 27,000 active satellites in orbit by the end of 2030, almost a four-fold increase from today.

Despite its immense potential, space-data-as-a-service faces challenges, including data security, privacy concerns, data latency, and the growing problem of space debris. However, with ongoing advancements in satellite technology and data analytics, these challenges can be addressed.

"SPACE DATA WILL ENABLE US TO SOLVE SOME OF THE WORLD'S BIGGEST CHALLENGES, SUCH AS CLIMATE CHANGE, DISASTER RESPONSE, AND FOOD SECURITY."

CONSIDERING ALL OF THE FACTS

SDaaS has revolutionized the way we use information collected from space, democratizing access to space data and making it accessible to a broader range of users. It has enabled innovative applications across various sectors and has the potential to drive advancements in science, technology, and decision-making processes.

Overall, the future of space data as a service promises to be dynamic and transformative, with increasing opportunities for innovation, economic growth, and addressing global challenges through the insights derived from space-based data. However, it will also require careful consideration of ethical, environmental, and regulatory issues to ensure responsible and sustainable use of space resources.

INNOVATION TAKEAWAYS

#DATAPOWEREDSPACE
#DATA4GOOD
#SPACEDATAASASERVICE
#SPACEDATAECOSYSTEM

GLOBAL SPACE-DATA COLLABORATION

The commercialization of space data has opened up a range of new business opportunities. This model of global cooperation can serve as a valuable example for addressing other global challenges, such as climate change, public health, and disaster response.

REAL-TIME DATA SHARING

Near-real-time data transmission enables quicker responses to emergencies and rapidly changing conditions critical for disaster management, maritime navigation, and aviation safety.

DATA-POWERED DECISION-MAKING

Space data reinforces the importance of data-driven decision-making. The ability to collect vast amounts of data from space and convert it into actionable information has implications for businesses, governments, and organizations across various sectors.

DATA FOR AUGMENTATION

CAPGEMINI'S GENERATIVE AI LAB STEER THE FUTURE OF ARTIFICIAL INTELLIGENCE WITH CONFIDENCE

ROBERT ENGELS

Head of Global
Generative AI Lab,
Insights & Data, Capgemini

MARK ROBERTS

Deputy Head,
Global Generative AI Lab,
Capgemini Engineering

In today's fast-paced business landscape, artificial intelligence has shifted from a buzzword to a crucial part of organizational strategy. Recognizing the imperative to stay ahead of this curve, Capgemini launched its Generative AI Lab, an initiative designed to navigate the evolving topography of AI technologies. This lab is not just another cog in the machine; it's a compass for the enterprise, setting the coordinates for AI implementation and management.

The aim of the Generative AI Lab is straightforward yet ambitious: to understand continual developments in AI. It runs under the umbrella of Capgemini's "AI Futures" domain, tasked with finding new advances in AI early on and scrutinizing their potential implications, benefits, and risks. The lab operates as a dual-pronged mechanism: partly analytical, providing a clear-eyed assessment of the AI landscape, and partly operational, testing and integrating new technologies.

A FORWARD-LOOKING INITIATIVE

The Generative AI Lab also homes in on multifaceted research areas that promise to elevate the field of AI. On the technical front, it ventures into realms like multi-agent systems, where an amalgamation of specialized large language models (LLMs) each excel in particular tasks, such as text generation or sentiment analysis. These are supplemented by model-driven agents rooted in mathematical, physical, or logical reasoning to create a robust, versatile ecosystem.

Also, the lab is innovating ways to make LLMs more sensitive to real-world context, helping to curate AI behavior that's not just intelligent but also intuitively aware. Then, the lab devotes considerable attention to socio-individual and psychological dimensions, exploring how generative AI interacts with and affects human behavior, social norms, and mental well-being.

Finally, it takes a truly global perspective by investigating the geo-political implications of AI, like how advancements could shift power balances or impact international relations. By casting such a wide net of inquiry, the lab ensures a holistic approach to AI research, one that promises to yield technology that's technically advanced, socially responsible, and geopolitically aware.

THE LAB'S FRAMEWORK FOR CONFIDENCE IN AI

A noteworthy achievement of the lab thus far has been the development of a comprehensive framework focused on instilling confidence in AI. According to the lab's research, contemporary technology has only managed to implement three factors necessary for total confidence in AI applications, particularly in decision-making scenarios. The lab has identified an additional eight factors that need to be considered, offering a 11-point blueprint that could serve as the industry standard for the ethical and effective implementation of AI.

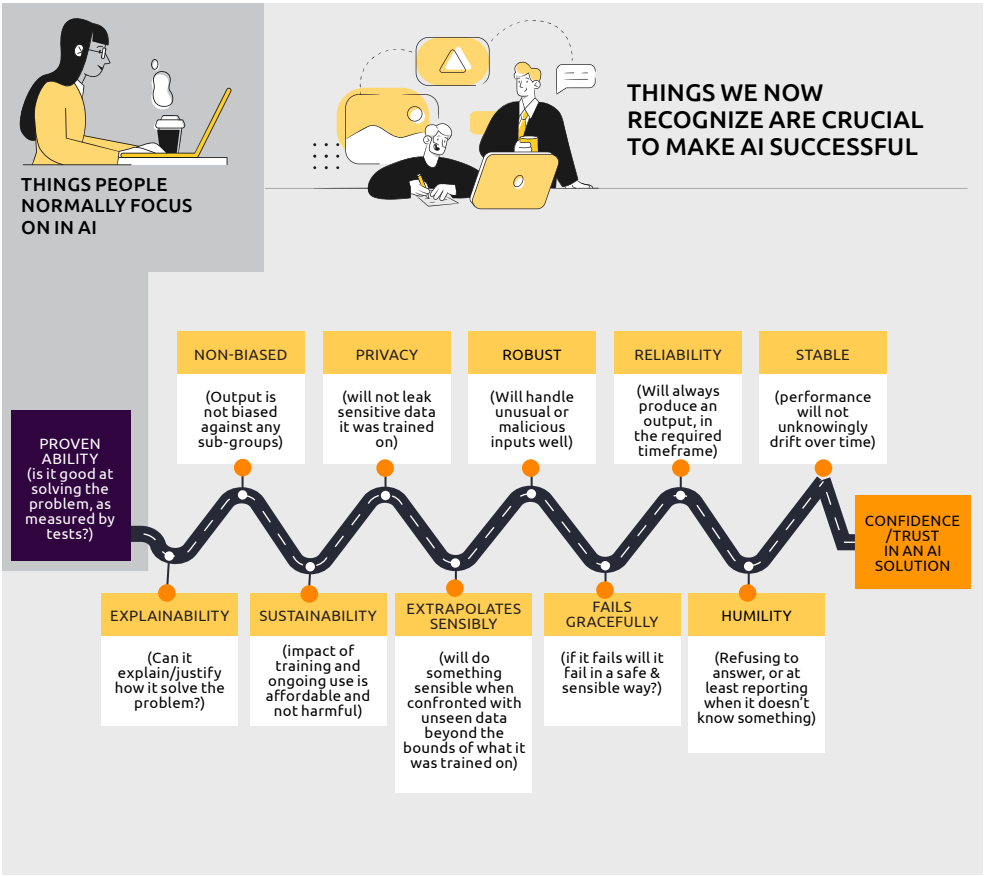
Traditional considerations for AI usually highlight factors like "robustness," ensuring the system performs reliably under a variety of conditions; "reliability," so the performance is consistently up to par; and "stability," meaning the system behaves predictably over time. These factors aim to ensure that an AI system works well, efficiently carrying out tasks and solving problems. However, the lab contends that this isn't enough.

The lab brings into the limelight other equally important but often overlooked aspects, such as "humility," the ability for the AI to recognize its limitations and not overstep its capabilities. Then there is the idea of "graceful degrading," which refers to how the system handles errors or unexpected inputs; does it crash, start to hallucinate, or manage to keep some functionality? And, of course, "explainability," which revolves around the system's ability to articulate its actions and decisions in a way that's understandable to humans and truthful to its contents.

For instance, a recommendation algorithm might be robust, reliable, and stable, meeting all the traditional criteria. But what if it starts recommending inappropriate or harmful content? Here, the lab's factors like humility and explainability would come into play, supplying checks

to ensure the algorithm understands its limits and can explain its rationale for the recommendations it makes.

The goal is to develop an AI system that's not just effective, but also aligned with human needs, ethical considerations, and real-world unpredictability. The lab investigates and identifies solutions for integrating these more nuanced factors, aiming to produce AI technology that is both high-performing and socially responsible.



OPERATIONAL RESULTS

Since its start, the lab has made significant strides. It has found crucial issues surrounding generative AI, spearheaded the integration of "judgment" in AI systems across the organization, and discovered promising technologies and startups. It has also pioneered methods for using AI for heightened efficiency, superior results, and the facilitation of new organizational tasks. The lab has not only served as a think-tank but also as an operational wing that can apply its findings in a practical setting.

The lab's workforce is its cornerstone. Composed of a multinational, global, cross-group, and cross-sector core team, it has the expertise to cover a wide array of disciplines. This team can dynamically expand based on the requirements of specific projects, ensuring that all needs – however specialized – are met with finesse.

REPORTING AND EARLY WARNING

Transparency and prompt communication are embedded in the lab's operational DNA. It has an obligation for early warnings: alerting the organization about the potential hazards or advantages of new AI technologies. Regular reports support a steady stream of information flow, enabling informed decision-making at both the managerial and executive levels.

Another vital role the lab plays is that of an educator. By breaking down complex AI systems into digestible insights, it provides the leadership teams of both Capgemini and its clients with the tools they need to understand and direct the company's AI strategy. It's not just about staying updated; it's about building a wide perspective that aligns with the broader strategic goals.

FUTURE OF AI

The Generative AI Lab at Capgemini has managed to go beyond the conventional boundaries of what an in-house tech lab usually achieves. It has successfully married theory with practice, innovation with implementation, and foresight with action. As the realm of AI continues to evolve unpredictably, the lab's multi-dimensional approach – grounded in rigorous analysis, practical testing, and educational outreach – stands as a beacon for navigating the complex yet promising future of AI.

With this in-depth framework, the Generative AI Lab isn't just looking at what makes AI work; it's exploring what makes AI work well in a human-centered, ethical context. This initiative is a key step forward, one that promises to help shape the AI industry in ways that prioritize both technical excellence and ethical integrity.

By aligning a high-performing AI system with one that understands and respects its human users, the lab is on the frontier of one of the most groundbreaking technological shifts of our time. And as AI continues to weave itself into the fabric of our daily lives, initiatives like this are not just useful but essential, guiding us towards a future where technology serves humanity, and not the other way around. With its robust framework, multidisciplinary team, and actionable insights, the lab is indeed setting the stage for a more reliable, efficient, and ethically responsible AI-driven future.

"BY ALIGNING A HIGH-PERFORMING AI SYSTEM WITH ONE THAT UNDERSTANDS AND RESPECTS ITS HUMAN USERS, THE LAB IS ON THE FRONTIER OF ONE OF THE MOST GROUNDBREAKING TECHNOLOGICAL SHIFTS OF OUR TIME."

INNOVATION TAKEAWAYS

#GENERATIVEAILAB
#FUTUREOFAI
#ETHICALAI
#AIFORHUMANITY
#CAPGEMINIINNOVATION

NAVIGATIONAL COMPASS FOR AI

Capgemini's Generative AI Lab serves as an enterprise compass, charting the course for AI application and management. This initiative is at the forefront of exploring multi-agent systems and enhancing LLMs with real-world context sensitivity. It's about pioneering an AI landscape that's not only intelligent but intuitively tuned to human nuance, reshaping how AI integrates into the fabric of society.

ETHICS AND UNDERSTANDING IN AI

The Lab's 12-point Confidence in AI framework aims to set new industry standards for ethical AI implementation. By considering factors beyond the traditional – like AI humility, graceful degrading, and explainability – the lab aspires to create AI systems that are not just technically efficient but also socially responsible and aligned with human ethics, ensuring technology that can perform robustly and understand its boundaries.

FROM THINK-TANK TO ACTION

The lab transcends the typical role of an R&D unit by not only dissecting and developing AI advancements but also operationalizing these insights. Its multinational, multidisciplinary team has been crucial in integrating AI judgment across the organization and fostering technologies that enhance operational efficiency, elevating the lab's role to that of both an innovator and an implementer.

CRAFTING CAPTIVATION GENERATIVE AI'S DANCE WITH HUMANITY

PRANAV KUMAR

Senior Director, Customer First & Gen AI for CX Lead India & APAC, Insights & Data, Capgemini

BIKASH DASH

Director, Gen AI India Leader, Insights & Data, Capgemini

Generative AI emerges as a catalyst in the creative domain, not to supplant human ingenuity but to amplify it. This synergy between cutting-edge technology and human creativity is crafting custom experiences that connect with customers on a new level. Forward-thinking businesses are harnessing this potential, marking the dawn of an era in which authentic innovation takes center stage.

Generative AI is a game-changing force that's redefining the way businesses engage with their customers. In a world where technology and human creativity are merging, it's not about machines replacing human ingenuity but rather enhancing it. Generative AI enables businesses to create personalized and authentic experiences that resonate deeply with individuals, setting new standards in customer engagement.

THE NEED FOR PERSONALIZED CUSTOMER EXPERIENCES

Customer expectations have evolved, and one-size-fits-all approaches to engagement no longer suffice. Today, consumers crave personalized experiences that cater to their unique preferences and needs. Generative AI addresses this demand by combining the power of artificial intelligence with the creativity of humans.

Generative AI is a technology that employs complex algorithms and machine learning to create content, designs, and even entire virtual worlds based on specific parameters or inputs. It's capable of producing content that is customized to an individual's preferences, achieving a level of personalization that was previously unattainable.

REAL-LIFE APPLICATIONS

Generative AI for CX is already reshaping various industries. For instance, in the fashion industry, leading brands are utilizing it to design custom clothing based on a customer's style preferences and body measurements. Online retailers are using this technology to provide tailored product recommendations based on each user's browsing history and purchase behavior. These applications illustrate how generative AI enhances customer experience by delivering highly personalized interactions.

It's essential to grasp that generative AI isn't about replacing human creativity with machines; instead, it augments human decision-making processes with intelligent algorithms. This harmony between technology and human talent allows businesses to harness the potential of generative AI while preserving the authenticity of human ingenuity.

REAL-LIFE SUCCESS STORIES

The power of generative AI in enhancing human creativity is evident in various case studies. For example, a leading fashion brand partnered with an AI technology company to analyze customer preferences and create personalized clothing designs. By combining data on color choices, fabric preferences, and style trends with human designers' expertise, the brand was able to offer unique garments tailored to each individual's taste, increasing customer satisfaction and sales.

In another case study, a digital marketing agency employed generative AI to optimize its content creation process. AI algorithms were used to generate headlines, social media posts, and ad copy, which were then refined by human writers. The result was significant improvements in click-through rates, engagement metrics, and conversions for their clients' campaigns, showcasing how generative AI can save time while maintaining authenticity in content creation.

EXPANDING POSSIBILITIES

Generative AI's impact isn't limited to fashion and content creation; it extends to various fields. For instance, it's transforming advertising by analyzing vast amounts of data, including customer preferences and market trends, to generate personalized ads tailored to individual consumers' tastes. This hyper-personalization enhances customer engagement and drives conversion rates.

In the gaming industry, generative AI is being used to create dynamic worlds that adapt in real-time based on player actions and choices. This level of customization provides players with immersive experiences like never before, increasing replay value and overall satisfaction.

Musicians are no longer limited by traditional composition techniques thanks to generative AI tools that assist in creating unique soundscapes, melodies, and harmonies automatically. Artists can explore new genres or experiment with different musical elements effortlessly, leading to fresh compositions that resonate with listeners.

THE MARKET FOR GENERATIVE AI IN CUSTOMER EXPERIENCE

Generative AI has become a driving force in reshaping customer experience. The market is evolving rapidly, offering boundless opportunities for businesses to innovate and deepen their connections with customers. One key trend is the integration of generative AI across multiple touchpoints along the customer journey, from chatbots providing real-time assistance to dynamic product recommendations based on individual browsing behavior.

Furthermore, the technology empowers human creators to push boundaries and explore new frontiers in design, music, art, and storytelling. The collaborative potential between humans and machines, driven by generative AI, is shaping the future of customer engagement.

Looking ahead, it's clear that generative AI will continue to define the future of customer experience. The market holds immense potential for growth as organizations strive to offer seamless interactions driven by genuine personalization. As technology continues advancing together with human ingenuity, we can anticipate even more astonishing developments that redefine what's possible in engaging customers through innovative means.

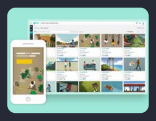
SYNTHETIC DESIGN

Improving scale, consistency, and reach of great creative & campaigns

TODAY: TIME TO MARKET, 2-4 MONTHS

Production is **expensive** and requires rigorous planning to ensure all content is captured

Production **requires creative** expertise from agencies using specialist design tools



Design, all the way through production, can be frustrating – requiring coordination between multiple stakeholders.

Example: Marketing Campaign Builder

Tomorrow Empowered by gen AI: Time to Market, 3-4 Weeks

HYPER-PERSONALIZED, QUICK AND COST EFFECTIVE

CREATIVE: providing creative assistance with image generation

MESSAGING: build contextual messaging

GUARDRAILS: brand guidelines and regulation filters

SWITCHING: generating contextual design

BEST PRACTICES FOR CUSTOMER EXPERIENCE

Implementing generative AI for customer experience is a ground-breaking endeavor that requires careful consideration and strategic planning. To ensure success, businesses should follow best practices that can maximize the potential of this technology.

Define objectives: Clearly define the objectives and desired outcomes of using generative AI in customer experience. Understand your target audience, identify pain points, and determine how generative AI can effectively address these challenges.

Collaboration: Collaboration between data scientists and creative teams is essential. Combining technical expertise with artistic flair enables the creation of truly compelling and personalized experiences for customers. This synergy allows for innovative solutions that strike a harmonious balance between human creativity and technological capabilities.

Regular iteration: Continuously iterate and refine generative AI models. Regularly analyze performance metrics and gather feedback from customers to enhance the accuracy and relevance of generated content over time.

Transparency: Maintain transparency with customers about the use of generative AI in crafting their experiences. Clearly communicate how their data is being utilized to ensure ethical practices while leveraging this powerful technology.

Stay informed: Stay up to date with advancements in generative AI technologies. The field is rapidly evolving, so keeping abreast of new techniques or tools can give you a competitive edge in delivering exceptional customer experiences.

Generative AI is revolutionizing customer experience by enhancing personalization, amplifying creativity, and enabling collaboration between humans and machines. It offers a path to deeply engage customers in a digital landscape where personalization is essential. The market continues to evolve, offering exciting opportunities for businesses to innovate and connect with their customers on a deeper level.

INNOVATION TAKEAWAYS

"THE FUSION OF HUMAN CREATIVITY WITH GENERATIVE AI UNLOCKS ENDLESS POSSIBILITIES FOR CRAFTING MEANINGFUL EXPERIENCES."

#GENERATIVEAI
#CX
#AIFORHUMANITY

HARMONIZING HUMAN AND AI CREATIVITY

Generative AI augments human creativity, enabling businesses to craft personalized and deeply resonant customer experiences.

PERSONALIZED PRECISION

This technology tailors experiences to individual preferences, setting a new bar for customer engagement through hyper-personalization.

INNOVATION AT THE INTERFACE

Generative AI redefines the customer journey, introducing adaptive, real-time personalization at every interaction point.

ETHICAL GENERATIVE AI AT THE CROSSROADS OF INNOVATION AND RESPONSIBILITY

TIJANA NIKOLIC

AI Lead, Netherlands, Insights & Data, Capgemini

YASHOWARDHAN SOWALE

CTIO I&D India, I&D Architecture Head, India Domain Leader for AI, Insights & Data, Capgemini

Generative AI is reshaping business operations and customer engagement with its autonomous capabilities. However, to quote Uncle Ben from Spiderman: "With great power comes great responsibility." While the AI boasts enhanced accuracy, it can also amplify societal biases and misuse. Ensuring it's trained on diverse, unbiased, and ethically sourced data is crucial. Ethical generative AI isn't just a buzzword; it's a commitment to uphold principles and prevent invasions of personal privacy. Organizations stand at a crossroads: embrace this transformative tech, but tread ethically.

Managing generative AI has been challenging as generative AI models are outperforming humans in some areas, such as profiling for national security causes. Sometimes, anti-principles clearly explain why ethics must be enforced, so it is important to understand the following challenges:

- Generative AI can assist in managing information overload by helping extraction and generating meaningful insights from large volumes of data but, at same time, information overload can dilute precise messaging.
- A lack of domain-specific knowledge or context leads to inaccurate information and contextual errors in addition to bias and subjectivity.
- There may be limited human resources to oversee training and regulate output, due to a lack of experienced personnel.
- Stale data may be used in training.
- Elite and/or not always ethically sourced data may be used for training.
- There may be a lack of resilience in execution.
- Scalability and cost tradeoffs may cause organizations to consider a shortcut.

Although complex, these challenges can be alleviated on a technical level. Monitoring is a good example of ensuring robustness and observability of the behavior of these models. Additionally, since generative AI capability is exposing business to new risks, there is a need for well-thought-through governance, guardrails, and the following methods:

- Model benchmarking
- Model hallucination
- Self-debugging
- Guardrails.ai and RAIL specs
- Auditing LLMs with LLMs
- Detecting LLM-generated content
- Differential privacy and homomorphic encryption
- EBM (Explainable Boosting Machine).

It is crucial that generative AI design takes care of following aspects of ethical AI:

- Ensuring ethical and legal compliance - Generative AI models can produce outputs that may be biased, discriminatory, or infringe on privacy rights.
- Mitigating risk - Generative AI models can produce unexpected and unintended outputs that can cause harm or damage to individuals or organizations.
- Improving model accuracy and explainability - Generative AI models can be complex and difficult to interpret, leading to inaccuracies in their outputs. Governance and guardrails can improve the accuracy of the model by ensuring it is trained on appropriate data and its outputs are validated by human experts.
- Ethical generative AI approaches need to be different based on the purpose and impact of the solution, so diagnosing and treating life-threatening diseases should have a much more rigorous governance model than using generative AI to give marketing content suggestions based on products. Even the upcoming EU AI Act prescribes risk-based approaches, classifying AI systems into low-risk, limited or minimal risk, high-risk, and systems with unacceptable risk.
- AIs must be designed to say "no," a principle called "Humble AI."
- Ethical data sourcing is particularly important with generative AI, where the created model can supplant human efforts, if the human has not granted explicit rights.
- Inclusion of AI: most AIs today are English-language only or, at best, use English as a first language.

USING SYNTHETIC DATA FOR REGULATORY COMPLIANCE

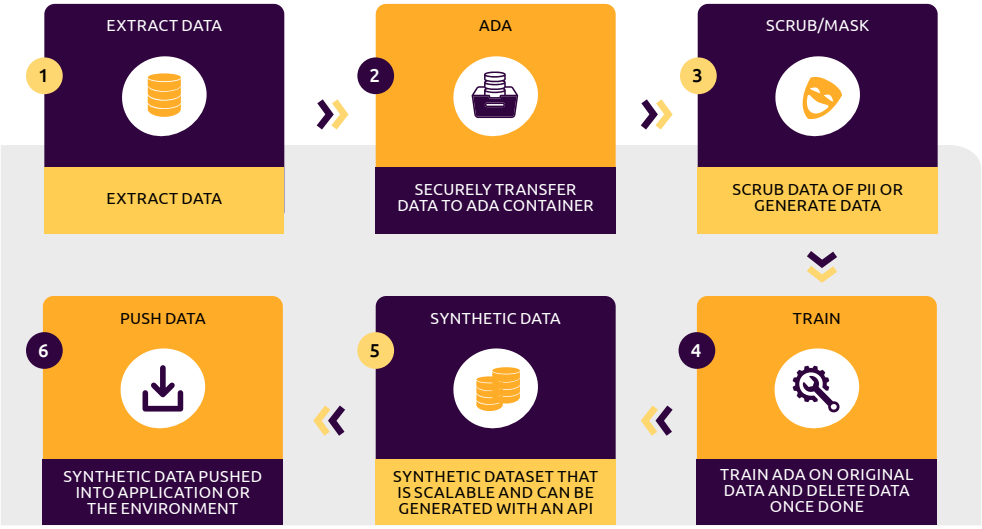
Försäkringskassan, the Swedish authority responsible for social insurance benefits, faced a challenge in handling vast amounts of data containing personally identifiable information (PII), including medical records and symptoms, while adhering to GDPR regulations. It needed a way to test applications and systems with relevant data without compromising client privacy. Collaborating with Försäkringskassan, Sogeti delivered a scalable generative AI microservice, using generative adversarial network (GAN) models to alleviate this risk.

This solution involved feeding real data samples into the GAN model, which learned the data's characteristics. The output was synthetic data closely mirroring the original dataset in statistical similarity and distribution, while not containing any PII. This allowed the data to be used for training AI models, text classification, chatbot Q&A, and document generation.

The implementation of this synthetic data solution marked a significant achievement. It provided Försäkringskassan with realistic and useful data for software testing and AI model improvement, ensuring compliance with legal requirements. Moreover, this innovation allowed for efficient scaling of data, benefiting model development and testing.

Försäkringskassan's commitment to protecting personal data and embracing innovative technologies not only ensured regulatory compliance but also propelled it to the forefront of digital solutions in Sweden. Through this initiative, Försäkringskassan contributed significantly to the realization of the Social Insurance Agency's vision of a society where individuals can feel secure even when life takes unexpected turns.

HOW DOES ADA WORK?



MARKET TRENDS

The market for trustworthy generative AI is flourishing, driven by these key trends.

1. **Regulatory compliance:** Increasing government regulations demand rigorous testing and transparency.
2. **User awareness:** Growing awareness among users regarding the importance of trustworthy and ethical AI systems.
3. **Operationalization of ethical principles:** Specialized consulting to guide AI developers in creating ethical risk mitigations on a technical level.

RESPONSIBLE USE OF GENERATIVE AI

Ethical considerations are at the heart of these groundbreaking achievements. The responsible use of generative AI ensures that while we delve into the boundless possibilities of artificial intelligence, we do so with respect for privacy and security. Ethical generative AI, exemplified by Försäkringskassan's initiative, paves the way for a future where innovation and integrity coexist in harmony.

INNOVATION TAKEAWAYS

"ETHICAL GENERATIVE AI IS THE ART OF NURTURING MACHINES TO MIRROR NOT ONLY OUR INTELLECT BUT THE VERY ESSENCE OF OUR NOBLEST INTENTIONS AND TIMELESS VALUES."

#GENERATIVEAI
#ETHICALAI
#LLM

TRANSPARENCY AND ACCOUNTABILITY

Generative AI systems should be designed with transparency in mind. Developers and organizations should be open about the technology's capabilities, limitations, and potential biases. Clear documentation and disclosure of the data sources, training methods, and algorithms used are essential.

BIAS MITIGATION

Generative AI models often inherit biases present in their training data. It's crucial to actively work on identifying and mitigating these biases to ensure that AI-generated content does not perpetuate or amplify harmful stereotypes or discrimination.

USER CONSENT AND CONTROL

Users should have the ability to control and consent to the use of generative AI in their interactions. This includes clear opt-in/opt-out mechanisms. Respect for user preferences and privacy and data protection principles should also be upheld.

FROM INFORMATION TO IMPACT
THE RISE OF AUTONOMOUS ANALYTICS

RAJESH IYER

VP & Global Head of ML & Generative AI, Insights & Data FS, Capgemini

ARIJIT SENGUPTA

Founder and CEO, Aible

Traditional BI and AutoML platforms enable self-service access to mountains of high-fidelity data, but they fail to deliver actionable insights to drive better business outcomes. The platforms struggle to take all data into account and narrow down anomalous trends and KPI drivers, and neglect the link between operational performance attributes and KPIs. Enter autonomous analytics, such as Aible, which can surface anomalous KPIs and trends and the key drivers as actionable insights. They complement popular BI tools to guide analysts to swift, precise insights.

For decades, firms have struggled to make BI work to drive business outcomes. Today, end users have access to more data than ever before but not the actionable insights necessary to help steer the business to the best possible outcomes. As Herbert Simon, Nobel Prize and Turing Award winner, noted back in 1971, "Wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it."

In 2013, Clayton Christensen (of Innovator's Dilemma fame) et al. wrote in the Harvard Business Review article "Consulting on the Cusp of Disruption" that, "The big data company BeyondCore can automatically evaluate vast amounts of data, identify statistically relevant insights, and present them through an animated briefing, rendering the junior analyst role obsolete." BeyondCore eventually became Salesforce Einstein Discovery and inspired the modern augmented analytics wave. The team behind BeyondCore has now started Aible, which takes autonomous analytics to the next level by marrying it with generative AI.

GENERATIVE AI-POWERED INSIGHTS DRAMATICALLY FASTER, WITH MINIMAL COST

Aible's patented technology automatically explores millions of cuts of data in minutes, at a fraction of the cost of legacy analytics. Customers from a wide variety of industries have published case studies showing the scalability, cost-efficiency, and speed of the Aible platform. A Google blog post entitled Aible's serverless journey to challenge the cost vs. performance paradigm explains how Aible delivered such analytics efficiencies on BigQuery, but such efficiencies can be expected on any platform.

AIBLE CUSTOMER	VARIABLE COMBINATIONS (# questions ChatAible) could answer	TOTAL COMPUTE COST FOR AUTOMATED ANALYSIS	AUTOMATED ANALYSIS TIME/DATASET
Ciena 700+ DATASETS, 60M ROWS	150M+	< $50	< 5 Mins
Cisco 120 COLUMNS, 164M ROWS	4.6M	< $15	< 5 Mins
OpenX 21 DATASETS, 2B+ ROWS	20M+	$27	< 10 Mins
NYC Health + Hospitals 81 DATASETS, 119M+ ROWS	48M+	$105	< 10 Mins
UnitedHealthcare 75 DATASETS, 100M+ ROWS	150M	< $80	< 10 Mins

"AUTONOMOUS ANALYTICS PLATFORMS CAN WORK IN STANDALONE MODE BUT WORK BEST AS COMPLEMENTS TO POPULAR BI PLATFORMS."

As an illustration of this capability, consider data for an outbound call-center that makes calls to offer co-branded credit cards to prospects. Aible enables autonomous analytics to use transaction data, enriched with raw operational data such as agent attributes like education, experience, call quality, and scores, to understand how and to precisely what extent they drive KPIs like conversion rates and offer a focused view into what can be done to address opportunities for improvement.

For example, it is helpful to combine agent attributes in the credit card call center illustration above into agent segments that can be used as engineered features in the analysis to better understand the precise extent to which cohorts drive outcomes. Aible automatically generates and evaluates such cohorts to determine the "net effect" of each combination on the KPI of interest. The most significant drivers of the tracked KPIs are reflected in a circular Sankey visualization, which shows the net effects of all variables at a glance. Alternatively, for business users, Aible can auto-generate traditional dashboards with the key charts organized in order of their impact on the KPI.

AUGMENT BI PLATFORMS

Autonomous analytics platforms can work in standalone mode but work best as complements to popular BI platforms. Aible automatically evaluates raw and engineered data, determines key insights, and auto-generates the KPI driver view. It can even export native BI tool dashboards, to be embedded into yet other BI tools such as Tableau and Power BI. In this design, we retain all the capabilities of the BI platform, with Aible's analytics engine helping us find the key statistically-sound insights behind the scenes.

This also suggests a new way for working for analysts and business leaders. The BI platform should be configured to monitor KPIs and alert analysts and/or business leaders about key insights related to the KPIs with additional information on which of the patterns surfaced are credible. The analysts can use these reports as starting points to pull further data. When used in this manner, the Aible engine can be thought of as driving BI for enterprise performance analytics. The BI platform should also be leveraged when analysts are looking to get reports in broader contexts than understanding KPI drivers.

DATA GRANULARITY IS ESSENTIAL

To get the most out of autonomous analytics platforms, the transaction data must be at the most granular level possible and then tagged with all raw and engineered attributes, such as from interaction and segmentation analysis, that make sense for that level. The autonomous analytics system can also auto-generate engineered features. Aible uses this data to generate key insights from the data and enable users to ask business questions such as "How can I improve sales to Gen Z customers?" instead of just analytical questions that must be translatable to SQL.

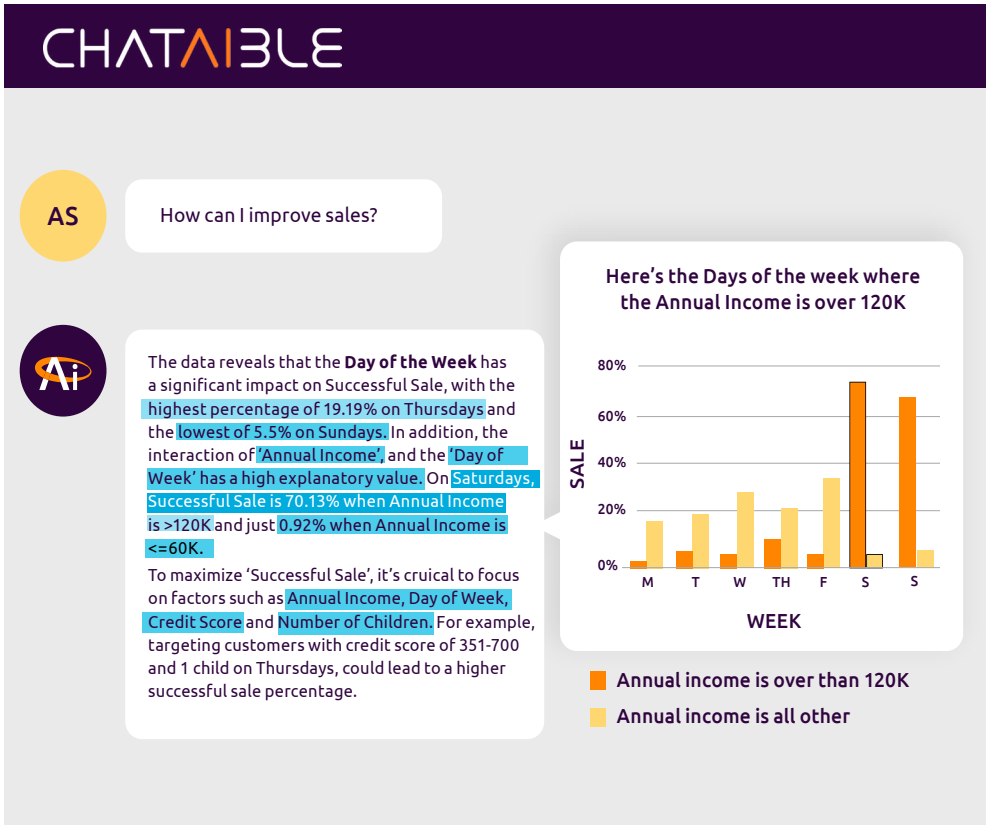
The engine will monitor KPIs, identify significant trends and shifts in the KPIs, and highlight statistically-credible alerts. It also generates visuals to explain the single or multi-variate patterns in a matter suited to the user persona – from circular Sankey charts and mind maps for expert analysts, to dynamic dashboards and generative AI storytelling for business users.

The Aible AI engine provides a list of drivers in a circular Sankey chart, with any overlap clearly indicated. The same view also provides ordered lists of drivers and corresponding charts showing the exact impact of each on tracked KPIs. In addition to this, the AI engine also provides a view into the behavior shift and population shift for each driver for period-to-period results, where the rate effect reflects change attributable to the average change in the value of a cohort, whereas the mix effect reflects the change in the proportion of that cohort.

Aible includes a generative AI platform that uses foundation models such as PaLM 2 and GPT-4 to allow users to ask the Aible engine questions about the drivers and their precise extent of impact behind KPIs in plain English; these generate a well-articulated response, also in plain English. Such a system gives everyone at firms the power to interrogate the engine about business questions related to the KPI and the drivers behind KPIs or KPI changes.

A properly implemented platform can provide insights into which customer segments and/or employee segments are

driving observed KPIs as applicable. Aible provides the necessary guardrails for enterprises to securely scale insights from generative AI responses with its ability to automatically double-check the output to reduce hallucinations (where generative AI creates inaccurate facts).



Aible can deliver insights in near real-time, allowing firms to respond to market threats and opportunities much faster, and in a very surgical fashion to optimize outcomes.

INNOVATION TAKEAWAYS

#AUTOBI
#TALK2DATA
#TALK2INSIGHTS
#GENBI
#GENINSIGHTS

AI FIRST

An AI-first approach automatically analyzes raw data across millions of variable combinates – group-by and drill-down charts – in a matter of minutes and costing cents.

THE ART OF STORYTELLING

Generative storytelling automatically highlights key insights in the data while double-checking the generative AI for hallucinations.

HAVE IT YOUR WAY

Insights can be consumed in multiple ways, from conversational interfaces to dashboards and mind maps.

GEN GARAGE
AN INNER-SOURCE ENGINE FOR DATA-POWERED INNOVATION

SUBARNA BHATTACHARYA

Director, Insights & Data FS,
Capgemini

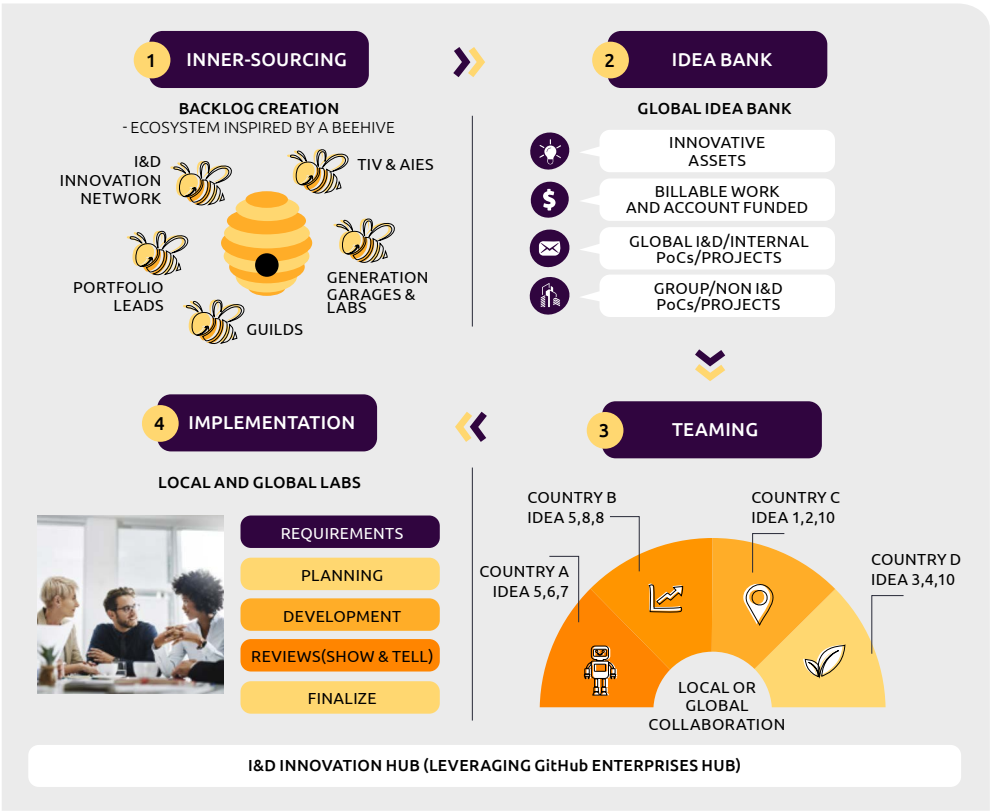
Ever wonder where the next big tech breakthrough will come from? Welcome to Gen Garage, Capgemini's data-powered innovation hotbed, mixing fresh talent with seasoned experts. Here, global collabs yield pioneering solutions, from AI-powered document analysis to virtual humans for video chats. It's not just a tech playground; it's an engine for global innovation, leveraging "inner-source style" collaboration. From now on, in each edition of this magazine, we'll show you some of the best ideas and innovations that recently popped up. Enjoy our first take.

Steve Jobs once said, "I sat in a garage and invented the future." That is a profound statement that well encapsulates the fact that innovative thoughts often originate in unconventional places.

With this enduring notion in mind, Capgemini Insights & Data launched its own Gen Garage, a dynamic platform that brings together a diverse array of talents and resources to cultivate innovation. It is an innovative platform that combines new hires, interns, mentors, industry leaders, and infrastructure support from across the globe to create a melting pot of ideas, innovation, and assets.

Gen Garage aims to pioneer data-powered innovation through a highly collaborative innersourcing approach (modeled on the Open Source way of working) of ideation and implementation for greater visibility, enhanced collaboration, and reduced silos. Leveraging platforms such as GitHub Enterprise, teams can generate multiple ideas, discuss and debate the solutions, reuse existing solutions to minimize redundancy, and leverage talent across the globe.

GEN GARAGE - INNER-SOURCE ENGINE FOR DATA-POWERED INNOVATION



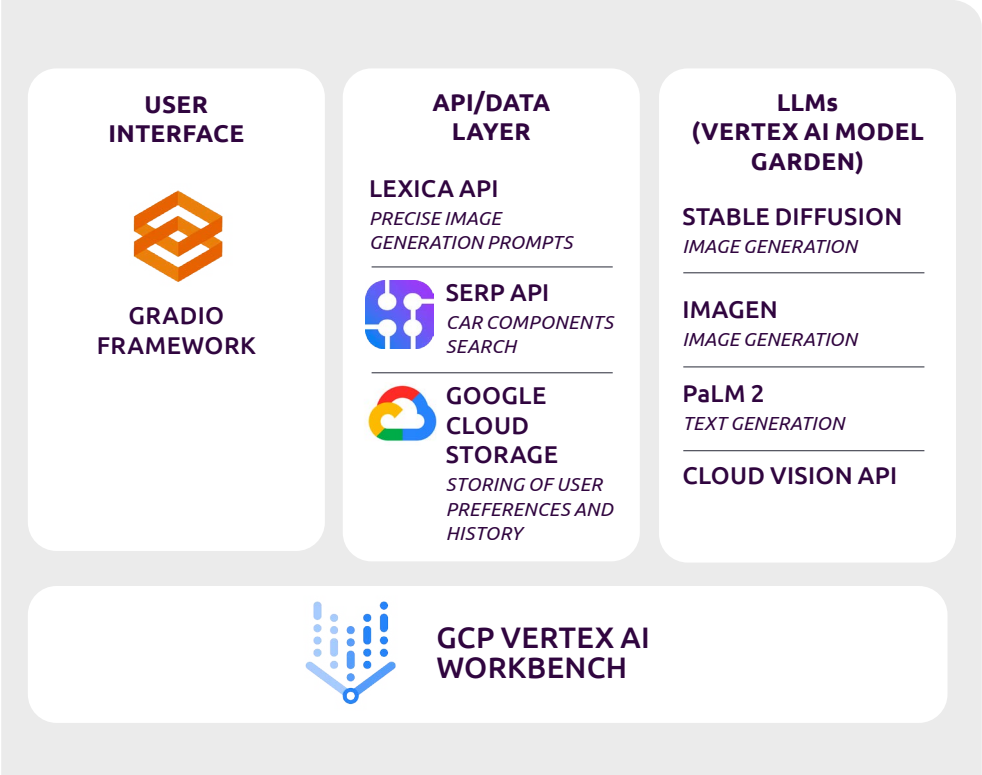
In recent times, our teams have explored a wide array of technology trends, including artificial intelligence, machine learning, generative AI, natural language processing, sustainability, and cloud computing. Let's take a glimpse at some of the ideas that have emerged from the Gen Garage.

REVOLUTIONIZING AUTOMOBILE DESIGN

Current design processes offer limited freedom to automobile enthusiasts to fully harness cutting-edge AI capabilities for exploring and creating innovative automobile design solutions.

The team conceived a generative AI-based platform for automobile design which empowers consumers to explore multiple customized designs for their cars based on personal preferences.

The platform leverages image and text generation models from Model Garden, including Imagen, Palm, Stable Diffusion, and ControlNet from Google Vertex AI. To ensure accuracy and relevance, it was integrated with Lexica API, to fetch precise prompts, and SerpApi, to scour the internet for information on car components. Gradio was used to develop an intuitive front-end interface. The solution was hosted on GCP Vertex AI Workbench.



The innovative solution has gained recognition, being shortlisted for prestigious events like Google Cloud Next 2023 in the US and TechCon Conference in Bangalore. Future enhancements include material, model, and engine-design options, along with integration with Google Glass-based VR.

VIRTUAL PRESENTER: BRINGING VIRTUAL HUMANS TO LIFE

For large organizations like Capgemini, communication is one of the key pillars and a significant part of this involves someone presenting almost repeatable content to a set of audiences through a virtual mode, e.g., employee onboarding, HR communications, etc. Such sessions demand time and effort while also making the process monotonous.

Virtual Presenter creates realistic virtual humans for professional video communication, leveraging generative AI. The platform will deliver pre-defined content and also be trained on custom datasets, FAQs, and process documents relevant to the context to help answer live queries during virtual interaction. The system will incorporate Tonal Analytics features to make the interaction more human-like.

The solution envisaged here is primarily focused on HR and learning and development functions, though virtual AI-driven avatars or digital humans have recently been implemented for news anchors, online influencers, patient communications, etc.

CLINICAL TRIAL ACCELERATOR: STREAMLINING CLINICAL TRIALS

Due to the uncertainties and risks involved, clinical trials have always been an expensive and time-consuming process that may stretch for decades. COVID vaccines were a rare occasion in which multiple stakeholders across the globe worked together to speed up the journey.

The team has envisioned a generative AI-powered solution which aims to streamline the clinical trial process by better planning candidate identification, risk assessment, and event probability insights. The proposed features of the solution include the following:

- **Patient recruitment and eligibility:** Streamlines candidate identification based on pre-set criteria using advanced technology.
- **Survival analysis:** Utilizes time-to-event analysis to inform treatment decisions with event probability insights.
- **Patient engagement and support:** Employs virtual assistants to provide information, reminders, and participant assistance.

Though generative AI offers promising solutions for enhancing clinical trials, it's important to acknowledge that acquiring suitable datasets for modelling these applications may pose challenges due to data-privacy concerns.

BUSINESS AS USUAL SUPPORT: KEEPING THE LIGHTS ON FOR THE ORGANIZATION

While the Gen Garage explores cutting-edge technologies, it also lends support to various operational initiatives for its own internal organization and beyond. This includes operational dashboards, employee feedback, FinOps solutions, technology-driven CSR Initiatives, and more, demonstrating its commitment to driving innovation across the organization.

Welcome to **Gen Garage**, where innovation meets ignition, and the future is ours to create.

"GEN GARAGE AIMS TO PIONEER DATA-POWERED INNOVATION THROUGH AN INNER-SOURCING APPROACH OF IDEATION AND IMPLEMENTATION FOR GREATER VISIBILITY, ENHANCED COLLABORATION, AND REDUCED SILOS."

INNOVATION TAKEAWAYS

#DATAPOWEREDINNOVATION
#INNERSOURCE
#GENGARAGE
#GENAI

INNER-SOURCE IS THE FUTURE OF ENTERPRISE INNOVATION

An inner-source approach to data-powered innovation helps teams break down silos and promote collaborative innovation globally.

OPEN-SOURCE PLATFORMS FOR INTERNAL COLLABORATION

Platforms such as GitHub Enterprise, well-known within the open-source community, can drive internal collaboration on innovation as well.

COLLABORATION ACROSS GENERATIONS DRIVES FRESH IDEAS IN CONJUNCTION WITH LEARNING FROM EXPERIENCE

Gen Garage is the perfect platform to bridge the wisdom of experienced people with the fresh perspectives of young talent.

GENERATIVE AI IS ONLY AS GOOD AS THE DATA YOU FEED IT

YOUR DATA IS YOUR COMPETITIVE ADVANTAGE

TAYLOR BROWN

COO & Cofounder, Fivetran

Generative AI is the pinnacle of data science. It will boost profits, reduce costs, and help you expand into new markets. To take full advantage of generative AI's capabilities, train your models on all your data.

The world is being transformed by AI-assisted medicine, education, scientific research, law, and more. Today, researchers at the University of Toronto use generative AI to model proteins that don't exist in nature; pharmaceutical giant Bayer now uses generative AI to accelerate the process of drug discovery; and education provider Khan Academy has developed an AI chatbot/tutor, Khanmigo, to personalize learning. And with each passing day, the list of AI use cases across all industries only continues to grow.

According to the Capgemini Research Institute, nearly all (96 percent) of executives cite generative AI as a hot topic of discussion in their respective boardrooms. Generative AI is not just used as an aid to surface information the way a search engine does; with generative AI, organizations can combine their proprietary data with foundation models that have been pre-trained on a broad base of public data to create a sustainable competitive advantage.

Generative AI then becomes the most knowledgeable entity within your organization.

However, as with all analytics, generative AI is only as good as its data. To fully leverage AI, an organization needs a solid data foundation and organizational norms that facilitate responsible and effective use of data.

Data readiness for generative AI depends on two key elements:

1. The ability to move and integrate data from databases, applications, and other sources in an automated, reliable, cost-effective, and secure manner
2. Knowing, protecting, and accessing data through data governance

Automated data pipeline platforms, like Fivetran, allow enterprises to capture all of their data, irrespective of the source platform. These automated tools reduce the friction and overhead required to maintain the flow of data to continuously train generative AI applications.

OPERATIONALIZING GENERATIVE AI

To operationalize generative AI effectively, organizations must establish a solid foundation of automated, reliable, and well-governed data operations. Generative AI requires a modern and scalable data infrastructure that can continuously integrate and centralize data from a variety of sources, including both structured and semi-structured data.

However, as businesses start to operationalize generative AI, they may encounter a number of challenges.

- **Data quality and preparation:** Generative AI models are only as good as the data they are trained on. It is important to ensure that the data is high-quality, clean, and well-organized. This includes identifying any potential biases in the data that may distort the outputs of any model trained on it.
- **Security and governance:** Security and governance in the context of generative AI concern masking sensitive information, controlling data residency, controlling and monitoring access, and being able to track the provenance and lineage of data models.
- **User experience:** It is important to design user interfaces for your model that make it easy for people to interact with your models.
- **Scalability:** It is important to choose a generative AI platform that can scale to meet your needs at a reasonable cost.

Generative AI models are trained on massive datasets of text, code, images, or other media. Foundation models, which are off-the-shelf generative AI models that are pre-trained on large volumes of (usually public) data, may be specialized by industry or use case. Choosing the right foundation model can have a significant impact on performance and capabilities. For example, a foundation model that specializes in code generation will do so in a more comprehensive and informative way than a model that is trained on a general dataset of text. Other specialties of foundation models may include sentiment analysis, geospatial analysis, image generation, audio generation, and so on.

While you can easily make use of pre-trained, publicly available AI models, your data is a unique asset that differentiates your organization from the competition. To make the most of it, you must additionally supply foundation models with your business's unique context.

With access to your organization's accumulated data, a properly tuned generative AI model can become the most knowledgeable member of your organization, assisting with analytics, customer assistance, sales and marketing, software engineering, and even product ideation.

The Fivetran product team leverages generative AI and natural language processing technologies to develop Fivetran Lite Connectors in a fraction of the time of Fivetran's standard connectors, while ensuring the same high quality, data integrity, and security customers expect from Fivetran.

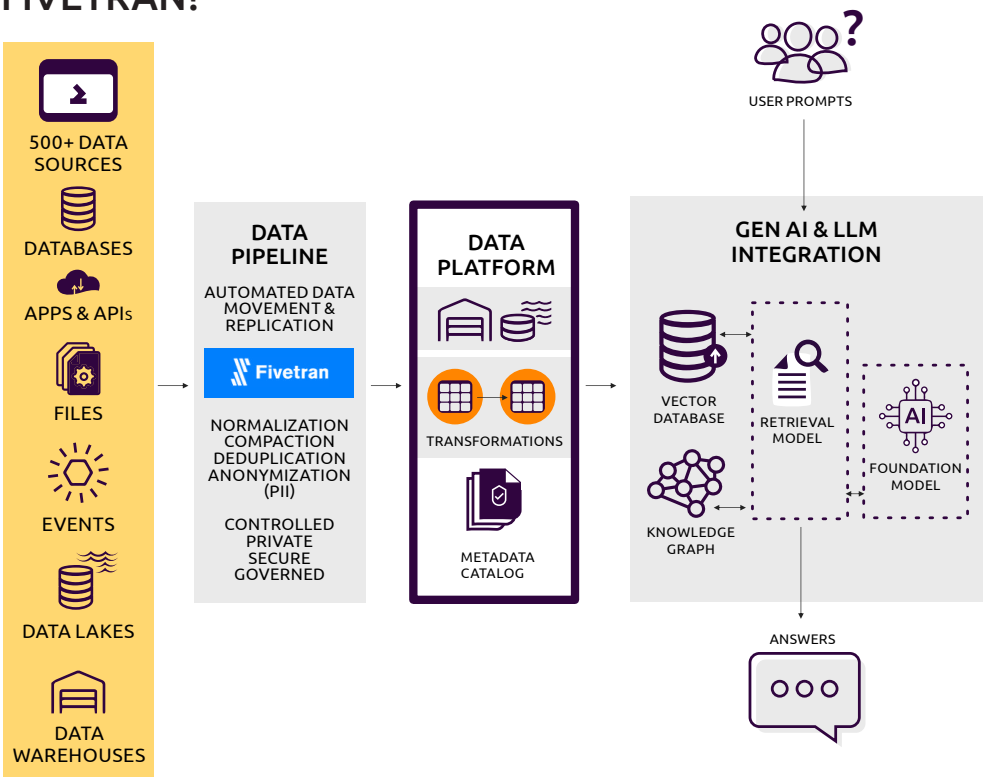
In addition, several notable organizations have already found practical ways to use generative AI. Global commercial real estate and investment management company JLL recently rolled out a proprietary large language model that employees access through a natural language interface, quickly answering questions about topics such as an office building's leasing terms. Similarly, the motor club in the US, AAA, now uses generative AI to help agents quickly answer questions from customers. Of the 100 tech companies profiled in the Forbes Cloud 100, more than half use generative AI.

According to Carrie Tharp, VP Strategic Industries, Google Cloud, "Generative AI opens up a new avenue, allowing people to think differently about how business works. Whereas AI and ML were more about productivity and efficiency – doing things smarter and faster than before – now it's about 'I can do it completely differently than before.'"

Until enterprises get the data right, the nirvana of asking generative AI app-specific and contextual organizational questions in a "Siri-like" way will remain elusive. Get the data right, and it opens up possibilities for all analytics workloads, including generative AI and LLMs.

To make full use of an ever-expanding roster of powerful foundation models, you must first ensure the integrity, accessibility and governance of your own data. Your journey into generative AI and the innovation and change it can bring will be fueled by high-quality, usable, trusted data built on automated, self-healing pipelines.

WHERE'S FIVETRAN?



INNOVATION TAKEAWAYS

"GENERATIVE AI APPLICATIONS ARE ONLY AS GOOD AS THE DATA THAT POWERS THEM."

#GENERATIVEAI
#FIVETRAN
#LLM
#GENAI
#DATAPIPELINES
#ARTIFICIALINTELLIGENCE
#AUTOMATION
#MACHINELEARNING

OPERATIONALIZE GENERATIVE AI

Operationalization begins with centralizing data and modernizing the data stack to include all available data.

AUTOMATED DATA ACCESS

By automating data pipelines, enterprises can focus on improving data models and algorithms to accelerate the efficacy and ROI of investing in a generative AI application.

CREATE AN UNFAIR ADVANTAGE

Generative AI trained on your data will provide insights and guidance driven by your data, creating a unique competitive advantage that cannot be replicated.

GENERATIVE AI UNLEASHED
PIONEERING A GLOBAL REGISTRY FOR RESPONSIBLE EVOLUTION

ARUNA PATTAM

Head Generative AI, Analytics, and Data Science, Insights & Data APAC, Capgemini

Embarking on the journey through the landscape of artificial intelligence, we are companions to the rise of large language models (LLMs), the newest prodigies in the technological arena. It's a thrilling time, reminiscent of the early web and bursting with potential and fraught with the inevitable question: how do we harness this power responsibly? Enter the concept of a global LLM registry, an idea as novel and necessary as the AI it seeks to catalog.

From the outset, the registry presents itself not as a stifling ledger of dos and don'ts but rather the backbone of a thriving AI ecosystem that champions creativity while keeping an eye on the ball: responsibility. Picture a bustling digital gathering space, where the keystones of AI – data sovereignty, privacy, and ethical use – are not just footnotes but the headlines of every conversation.

The dynamism of LLMs is undisputed; they're revolutionizing industries and accelerating innovation with the voracity of a Silicon Valley startup. Yet, amidst this brilliance, the potential for misuse lurks, like a plot twist in a tech thriller. It's a double-edged sword, demanding a wielder skilled in both innovation and ethics. The question that echoes in the halls of this AI renaissance is: can we balance these twin imperatives on a global scale?

GLOBAL REGISTRY

Enter the global LLM registry, the harbinger of this balance, a beacon of structure in the fluid world of AI. Its importance cannot be overstated. As things stand, the LLM arena is a bit like the Wild West: a few big guns rule the roost while the rest of the town scrambles for a place. This registry aims to democratize the frontier, providing an equitable platform for all stakeholders, from tech moguls to policymakers, to sing from the same hymn sheet.

Imagine an online compendium, accessible with the click of a mouse, detailing the inner workings of these digital behemoths. It's an initiative that borrows the best notes from other successful oversight systems: think financial regulations or even nuclear controls. These aren't just checkpoints; they're the lighthouses guiding us through murky waters, ensuring we don't crash against the unseen rocks of malpractice.

The foundation of the registry would be cast on solid principles: comprehensive disclosure of an LLM's genesis and operational ethics. The pillars – registration criteria, processes, and protocols – would ensure a robust framework where developers and users dance in step to the tune of transparency and accountability.

The registry's processes would marry simplicity with security. Developers would have confidential channels to divulge the nuts and bolts of their creations, while users could verify sources and track updates, all within a fortress of data protection. This is the kind of transparency that doesn't shatter trust but reinforces it, solidifying a communal spirit in the AI cosmos.

Of course, no structure, no matter how grand, can stand without a legal scaffold. Here, the registry shines again, enforcing compliance with an iron fist in a velvet glove – rewarding rule followers while ensuring rule breakers don't get off scot-free. Yet, this isn't about hampering innovation with red tape; it's about creating a fair playing field where the mavericks of code can still shoot for the stars.

BUILDING ON WHAT IS ALREADY THERE

A comprehensive LLM registry comes loaded with benefits – and its fair share of challenges. It's about creating a symphony from a cacophony, uniting diverse voices in a chorus of innovation and trust. It won't be a walk in the park, considering the varied interests at play, but who ever said that pioneering work was a piece of cake?

It's also about improving upon what we already have. Current AI acts are well-intentioned but are like gardening in your backyard while the forest remains wild and untamed. They are geographically limited, while AI knows no borders. A global registry, in contrast, has its gaze firmly fixed on the horizon, encompassing the vast expanse of digital landscapes.

The AI acts and the registry, when working in tandem, could form a formidable alliance.

AI IS NOT JUST A TOOL; IT'S A FORCE THAT CAN RESHAPE OUR WORLD. WITH A GLOBAL LLM REGISTRY, WE CAN SHAPE IT RESPONSIBLY AND ENSURE IT SERVES HUMANITY'S BEST INTERESTS."

Where regional acts provide a scaffold, the registry could add the walls and roof, creating a sturdy home for AI governance. This partnership could pave the way for a more holistic approach, aligning regional specifics with global objectives.

The road ahead for establishing a global LLM registry will require a caravan of collaboration and phased strategies. It's about global consensus, where tech tycoons and international stewards converge to script the next chapter of AI. Pilot programs would serve as the scouts, charting the terrain and flagging potential pitfalls.

Let's not forget, AI doesn't recognize borders; it's a citizen of the world. It's high time our regulatory measures matched that ethos. Organizations like the UN or global tech coalitions could spearhead this initiative, setting the standard for international cooperation and oversight.

As we stand at this crossroads, the question is not just whether we can but whether we must. The establishment of a global LLM Registry is the call to action echoing across digital valleys and virtual mountaintops.

Such a registry isn't just a bureaucratic necessity; it's the cornerstone of a future where innovation doesn't outpace accountability. It's a testament to our commitment to steward AI through its adolescence into a mature, responsible adult in the society of technology. Here lies the heart of the matter: It's about ensuring the AI revolution uplifts rather than undermines, that it remains a force for good in a world ever-hungry for progress.

INNOVATION TAKEAWAYS

#AI
#RESPONSIBLEAI
#LLMS
#GLOBALREGISTRY
#INNOVATION
#ACCOUNTABILITY

AI'S DUAL NATURE

AI possesses a dual character, presenting both transformative benefits and potential risks due to deceptive content generation.

IMPORTANCE OF A GLOBAL LLM REGISTRY

A global registry for LLMs would promote transparency, ethics, and responsible governance in the AI field.

COLLABORATIVE APPROACH

This must be a united, collaborative approach involving nations, tech giants, and global bodies to ensure responsible AI advancement, emphasizing the synergy of innovation and accountability.

INTELLIGENT MESH ELEVATING THE DATA MESH WITH GENERATIVE AI

WEIWEI FENG

AI Nordic Lead, Insights & Data, Capgemini

In our rapidly evolving landscape, where AI and generative AI are fueling innovation at an unprecedented rate, and the adoption of data-mesh principles has shifted from being an optional choice to an absolute imperative. The intricate interplay between data and AI is undeniable, and seamlessly integrating data mesh principles has become a crucial stride in fully unleashing their collective potential.

DECENTRALIZED OWNERSHIP FOR ENHANCED AI GOVERNANCE

The fundamental principle of the data mesh framework – decentralized ownership – can be seamlessly extended to AI management, including the intricate task of overseeing complex services involving multiple AI models and various other tools. In the domain of AI and generative AI, where a diverse array of models and techniques are employed, decentralization becomes a powerful mechanism for empowering specialized teams to take ownership and accountability for their specific AI domains.

This decentralized approach ensures that teams possessing expertise in particular AI areas have the responsibility for managing their models and overseeing the orchestration of complex services, which often entail multiple AI models and the integration of various tools. By decentralizing AI management in this way, organizations can reduce bottlenecks and enhance decision-making within the complex services landscape. The result is a dynamic ecosystem in which AI experts actively steer the development and maintenance of models, leading to higher-quality outputs and optimized performance across these intricate and multifaceted AI solutions.

BRIDGING DATA MESH AND AI WITH ETHICAL AI GOVERNANCE

One of the key strengths of the data-mesh concept lies in the creation of data products that promote data reuse and accessibility. This idea can be translated to AI through the development of AI domain products – well-defined assets like pre-trained models, pipelines, re-useable prompts, and AI services that teams can readily access. These AI domain products serve as building blocks that teams across the organization can leverage, fostering collaboration and reducing the need for redundant development efforts. Just as data products streamline data access, AI domain products streamline model access and deployment, promoting consistency and efficiency.

Data mesh emphasizes data quality and ethical considerations, which aligns well with the challenges posed by AI, especially generative AI. These models generate content, making ethical considerations paramount. By extending data-mesh principles to AI, organizations ensure that AI governance frameworks are in place to address biases, ensure fairness, and uphold ethical standards in AI and generative AI outputs. Additionally, the complexity introduced by generative AI's output diversity and prompt management can be managed effectively through a data mesh-inspired approach. Structured methodologies for combining, versioning, and managing diverse foundational models can ensure reliable and ethical content generation.

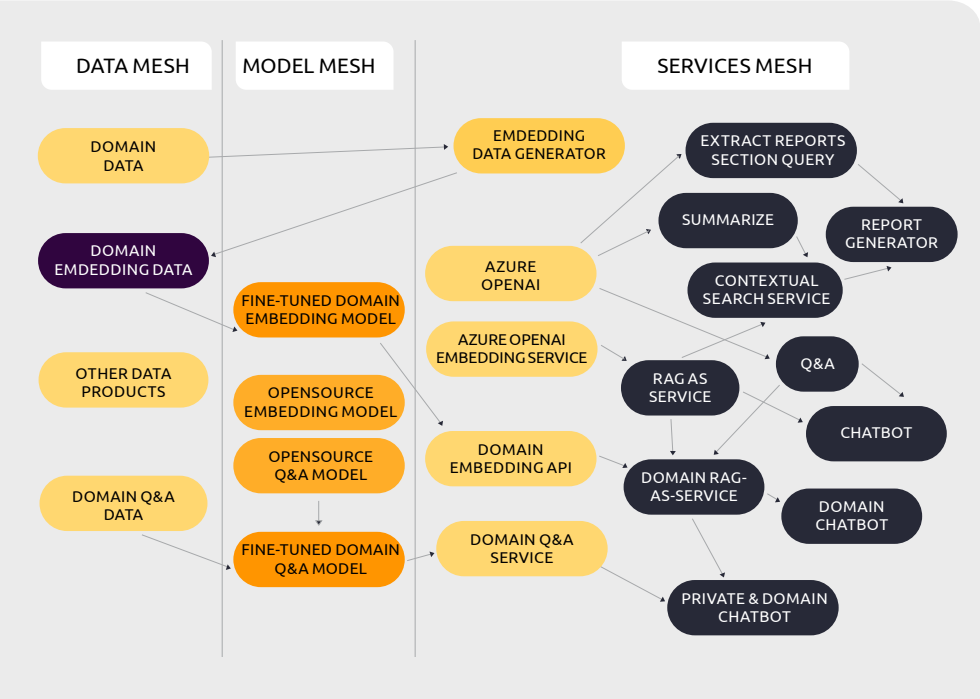
CRAFTING THE INTELLIGENT MESH

An intelligent mesh, in practical terms, comprises three crucial layers: the data mesh, model mesh, and services mesh. A standout advantage of this intelligent mesh setup lies in its ability to enhance data, model, and services reusability and accessibility. The services mesh, forming the top layer of this ecosystem, offers its own set of advantages. It enables the rapid creation of AI services to orchestrate model services and other services.

As an example, Databricks can serve as the foundational infrastructure for constructing the intelligent mesh. Within this ecosystem, MLflow functions as an invaluable tool for logging essential metadata related to data, models, and services. Concurrently, the AI Gateway assumes a central role in facilitating the creation of endpoints, catering to both CPU and GPU serverless environments. These endpoints not only accommodate large language models but also extend their support to a range of other services that depend on them.

Building upon these capabilities, Mosaic and MLflow are employed to harness the flexibility needed for training specialized models, or fine-tuning models for specific purposes or domains. Subsequently, these models can be seamlessly deployed as endpoints within the mesh.

To establish robust governance over our data, models, and services, Unity Catalog plays a pivotal role. In conjunction with Databricks models, it preserves vital lineage information, not only for data and models but also for the interconnected services within the ecosystem.



This journey begins with the deployment of diverse open-source models within the environment, effectively transforming them into private large language models (LLMs) and embedding services. Additionally, there exists the capacity to fine-tune or train specialized models using the governed data within the same system, deploying them seamlessly as services. Subsequently, advancements are made by developing specialized services, such as "RAG-as-a-Service," which interact seamlessly with endpoints provided by the embedding services and LLM services. (RAG is retrieval augmented generation.) These foundational components form the basis for constructing more intricate services that combine various models, embeddings, and pre-existing services to fulfill specific and intricate requirements. These complex services can continuously serve as building blocks for other future services. When updates are introduced to any of these components – models, embeddings, or pre-existing services – our interconnected system ensures that other services built upon them are automatically updated. This keeps the ecosystem agile, adaptable, and aligned with the latest advancements, meeting evolving requirements.

THE PATH TO ADVANCED AI AUTONOMY

Beyond the existing advantages, the implementation of consistent governance within the intelligent mesh framework paves the way for future advancements. The connections between services are recorded in the intelligent mesh and are particularly valuable, as they could serve as learning material for a large language model (LLM) agent and enable the agent to acquire the knowledge needed to orchestrate solutions to meet the requirements. Furthermore, these agents can acquire the aptitude to innovate and create new services when confronted with novel challenges, expanding their problem-solving capabilities. These newly created services can then be leveraged for various other AI agents.

This forward-looking vision also encompasses the potential for multiple AI agents to autonomously collaborate in tackling complex tasks. Through seamless interactions and integration with the services in the ecosystem, these agents can collectively enhance efficiency and usher in more advanced and automated problem-solving techniques. This not only boosts

productivity but also ushers in a realm of more sophisticated solutions.

The dynamic constructive collaboration between data mesh and intelligent mesh principles represents a pivotal moment in the evolution of AI and generative AI. By embracing decentralized ownership, ethical governance, and collaborative innovation, organizations can harness the true power of these technologies. As we move forward, we must remember that the journey towards a fully interconnected intelligent mesh is not just a technological pursuit; it is a commitment to shaping a future where AI serves as a force for good, driving progress across diverse domains. With thoughtful governance and a collaborative spirit, we stand on the brink of a transformative era where AI seamlessly integrates with our lives, creating solutions that are not only intelligent but also ethical, efficient, and endlessly innovative.

"BY EMBRACING DECENTRALIZED OWNERSHIP, ETHICAL GOVERNANCE, AND COLLABORATIVE INNOVATION, ORGANIZATIONS CAN HARNESS THE TRUE POWER OF AI AND GENERATIVE AI."

INNOVATION TAKEAWAYS

#GENAI
#DATAMESH
#GENERATIVEAGENT

DECENTRALIZED AI STEWARDSHIP

Granting specialized teams authority over their AI domains streamlines management, fostering innovation with accountability and reducing operational logjams.

ETHICAL AI INTEGRATION

Adopting data-mesh principles for AI, including generative AI, enhances ethical governance, ensuring responsible and unbiased AI solutions while maintaining quality and efficiency.

ECOSYSTEM FOR AI AUTONOMY

Constructing an intelligent mesh with robust governance mechanisms sets the stage for AI solutions to self-orchestrate, enhancing problem-solving and enabling continuous innovation.

DATA FOR GOOD

GREEN DATA
THE SUSTAINABLE FOUNDATION OF ENTERPRISE

ARNE ROSSMANN

Innovation Lead,
Insights & Data, Capgemini

ASIM HUSSAIN

Executive Director, Green
Software Foundation

Imagine a future where enterprises don't just aim to become data powerhouses but do so sustainably, ensuring both technological advancement and planet preservation. Tapping into the vast data sources for improved products and digital services is crucial. Yet, in our race for innovation, sustainability emerges as a pivotal cornerstone, safeguarding both our planet and a company's future relevance. The secret sauce? A sustainable data value chain. Dive in as we explore the essence of green data, drawing insights from the Green Software Foundation.

In recent years, enterprises followed the goal of becoming data-powered enterprises to leverage the full potential of data for their value chain. Levering the insights from all the sources of relevant data to create improved and new products and additional (digital) services is the top priority for enterprises. But sustainability has become a main goal for businesses too, to preserve the planet and the company's relevance in the next decades.

The Greenhouse Gas Protocol defines three scopes (scope 1, scope 2, and scope 3) to delineate direct and indirect emission sources, improve transparency, and provide utility for different types of organizations and different types of climate policies and business goals. Through this framework, companies can define and manage their emissions.

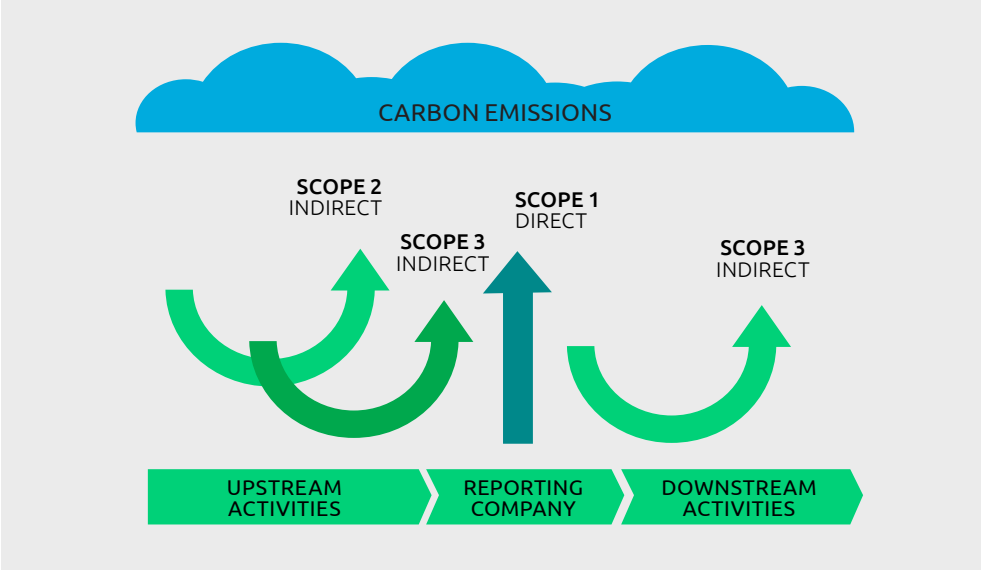


FIGURE 1: GHG FRAMEWORK

But this is only the first step. With clarification of emissions within the three scopes, companies get transparency on what's happening within the value chain and clarity on where to reduce. But the main challenge is to know how to reduce their emissions.

Here, the Green Software Foundation has defined six principles to be applied in software development:

1. **Carbon efficiency:** Emit the least amount of carbon possible.
2. **Energy efficiency:** Use the least amount of energy possible.
3. **Carbon awareness:** Do more when the electricity is cleaner and do less when the electricity is dirtier.
4. **Hardware efficiency:** Use the least amount of embodied carbon possible.

5. **Measurement:** What you can't measure, you can't improve.
6. **Climate commitments:** Understand the exact mechanism of carbon reduction.

For each of the six principles, examples on achieving them are available, especially in the area of carbon awareness. Not only have the big hyperscalers (AWS, Azure, Google) made this topic a top priority, but dedicated smaller solutions can also be found. Two innovative examples are Green Mountain, which provides 100-percent renewable energy sourced data centers for co-location in Norway, and windCORES, which helps companies deploy small co-location data centers in wind turbines, powered by 100-percent renewable energy and maximizing the space used by the wind turbines. With Green Data Engineering, a first view on how to apply these principles to data engineering has been laid out.

But one question remains: how can companies aiming to be data-powered enterprises do this in a sustainable way? The answer sounds simple and complex at the same time: apply the principles of the GHG framework towards the data value chain and make the carbon footprint of data products and use cases transparent.

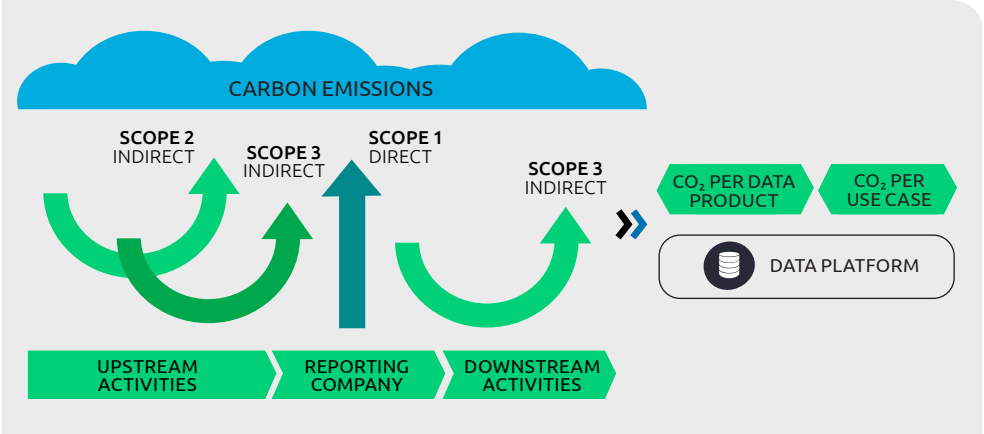


FIGURE 2: APPLYING THE GHG FRAMEWORK TO DATA PRODUCTS AND USE CASES

This is not as complicated as it sounds; most information is already available.

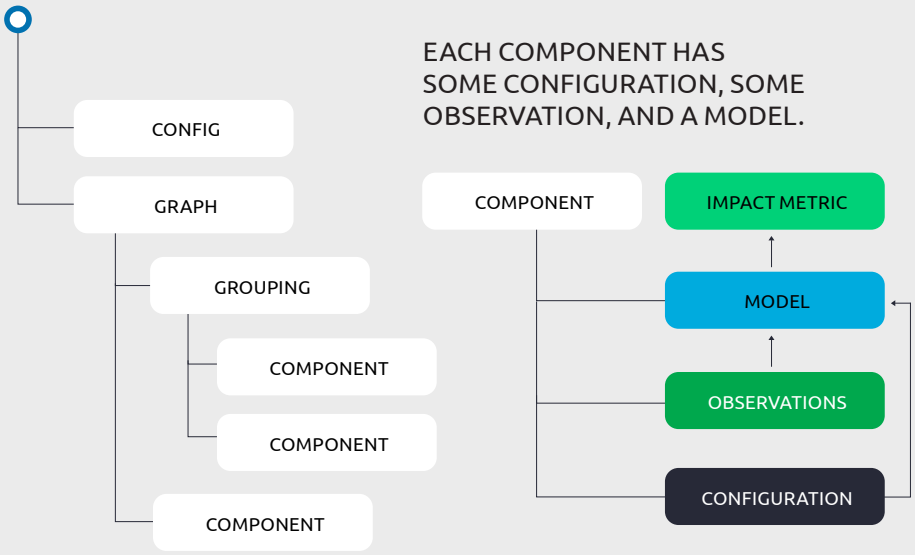
With the Carbon Aware SDK from the Green Software Foundation and the Sustainability APIs, SDKs, and dashboards from hyperscalers, it is possible to calculate the carbon footprint of applications and processes. As an example, the Azure Sustainability Manager provides a comprehensive overview with multiple reports on the customer landscape running on Azure. But this is limited to one cloud. What about the more common example of customers running a multi-cloud environment strategy?

Modern applications are composed of many smaller pieces of software (components) running on many different environments, for example: private cloud, public cloud, bare-metal, virtualized, containerized, mobile, laptops, and desktops.

Every environment requires a different model of measurement, and there is no single solution to calculate the environmental impacts for all components on all environments.

To achieve this, the Green Software Foundation has incubated the Impact Framework (IF). The IF is a framework to model, measure, simulate, and monitor the environmental impacts of software. It allows you to define a calculation manifest file, a YAML file which describes the calculation of emissions. So rather than just saying "Carbon is X" you can say "Carbon is X and here is all the data, all the working out, and all the assumptions and models that we used." You can run the YAMLs to confirm a claim and if you don't agree with some of the data, models, and assumptions, you can change and run it again to see how that alters the value.

IF represents the carbon footprint of different components in a graph to aggregate the information and draw dependencies and interconnections.



- Configuration describes shared information regarding this component and, most importantly, the parameters required by this model.
- Observations are a time series of data points used as inputs to the model.
- Model is a plugin which when given some configuration and a series of observations can calculate the impact, e.g. carbon impact, from an observation of CPU utilization.

With this approach, it's possible to aggregate up the carbon footprint of software components of applications easily.

And with proper application of portfolio management, the mapping of application-based carbon footprint along the value chain comes down to mainly pure calculations.

We might wonder: how about data products and use cases? Didn't we want to be data-powered? Sure, just as a small recap: the data product, the architectural quantum is the node on the mesh that encapsulates three structural components required for its function, providing access to the domain's analytical data as a product, as Martin Fowler mentions in his article. They are:

- Code
- Data and metadata
- Infrastructure

Sounds familiar? Right, it's easily comparable to any other application.

"LEVERING THE INSIGHTS FROM ALL THE SOURCES OF RELEVANT DATA TO CREATE IMPROVED AND NEW PRODUCTS AND ADDITIONAL (DIGITAL) SERVICES IS THE TOP PRIORITY FOR ENTERPRISES. BUT SUSTAINABILITY HAS BECOME A MAIN GOAL FOR BUSINESSES TOO, TO PRESERVE THE PLANET AND THE COMPANY'S RELEVANCE IN THE NEXT DECADES."

Therefore, the transfer of the Impact Engine Framework towards a data mesh approach is not as hard as it sounds.

And with that, companies have the right tooling in place to ensure ESG compliance for their data-powered enterprise journey. And as the whole value chain transformation towards more digital services and products continues, the importance of mapping their carbon footprint along the data value chain is essential. Not only to be compliant with ESG reporting, based on the scope 3 disclosures required under the European Union's Corporate Sustainability Reporting Directive, which comes into force in January 2024, but also to maintain the Race to Zero. The race is still on, and it's a data-powered race.

INNOVATION TAKEAWAYS

#SUSTAINABILITY
#IMPACTENGINEFRAMEWORK
#DATAPOWERED
#DATA4NETZERO

COMPLIANCE = TRANSFORMATION

Enterprises need to comply with the EU's Corporate Sustainability Reporting Directive, which has an impact on the transformation towards more digital services and products.

THE SOLUTIONS ARE IN THE CLOUD

Hyperscalers provide solutions within their environments to tackle carbon footprint.

AN OPEN FRAMEWORK

With the IEF by the Green Software Foundation, a framework for overarching carbon impact calculations exist.

REVOLUTIONIZING MOBILITY

THE DATA-POWERED INNOVATION BEHIND PEUGEOT'S ENDURANCE 9X8 HYPERCAR

PIERRE-DENIS AUTRIC

Engagement Manager 9x8 project, Insights & Data, Capgemini

ALEXANDRE DOUMBIA

Data Architect 9x8 project, Insight & Data, Capgemini

The Peugeot Endurance 9x8 hypercar doesn't only look like it's from the future, it's also built like cars in the future: completely data-powered with a closed-loop approach. Capgemini has industrialized the data gathering of this unique vehicle and made the data available to train AI models. Those algorithms are the basis for the fine-tuning work of engineers to improve combustion and electric-engine performance and pave the way for greener mass-market vehicles. Efficiency is the driving force behind the commitment to sustainable mobility.

Peugeot Sport, with support from Capgemini's software engineering team, joined the FIA World Endurance Championship in 2022 in the Hypercar LMH category. Involvement in this competition requires a long process of developing innovative and high-tech solutions, in an environment that is by definition highly competitive.

Team Peugeot TotalEnergies is crafting the Peugeot 9X8, featuring a 2.6-liter V6 rear axle with over 500 kW and a 200 kW front motor-generator (MGU-K). While competitors like Porsche, Cadillac, Acura, Lamborghini, and BMW are in the LMDh championship with common chassis and hybridization, Peugeot, Toyota, and Ferrari have chosen LMH, requiring in-house development of the chassis and GMP. The process demands extensive work to ensure speed and reliability in races lasting 6 to 24 hours.

Lab developments inform dynamic circuit progress. Over 500 recording channels and thousands of sensors gather data from all components. In post-track testing, engineers scrutinize vehicle performance and component functionality.

Engineers can access analytical tools, including the ECU's proprietary software, but handling the vast data volume critical for car performance and reliability is challenging. Capgemini was enlisted by Peugeot Sport to provide comprehensive analysis capabilities across extensive test days and multiple seasons.

Engineers can access their database to locate past scenarios like weather conditions, day-to-night variations, speed profiles, aerodynamics, and driving styles over a history of several years. This process takes minutes, significantly reducing the time and effort needed compared to hours without the platform.

Peugeot Sport has seen significant improvements, as this platform brings gains in speed of data analysis and boosts efficiency during practice sessions and races.

Data analysis facilitates the development of intricate operational strategies for hybrid cars. The electric motor's power is capped at 200 kW, and the GMP unit, combining combustion and electric engines, should not exceed 500 kW. Precise control over power distribution and energy recovery offers a range of racing possibilities, with strategies adapting to circuit layouts and weather conditions.

The platform shares data across Peugeot Sport services, aiding experts and enabling information cross-referencing. The team's departments notes that departments focusing on performance, reliability, and simulation-related operations can access this valuable database, with many performance indicators extracted from it. This approach proves especially advantageous for complex systems with extended development periods. Also, users save time as the platform streamlines information synthesis and sharing.

The data processing time, between the readings from the car and their availability on the platform, is about 10 minutes.

PLATFORM DEVELOPMENT

The project, involving a project manager and around a dozen Capgemini collaborators, aimed to convert data into a unified, efficient, and scalable format. They selected cloud-based components available on the market to fulfill various requests within Peugeot Sport. These components enable specific queries and large spreadsheet pivot tables, automatically generating easily readable graphs, streamlining data analysis.

The management of big data leverages existing cloud components, eliminating the need for new developments and reducing costs. Capgemini's focus is on software, with hardware characteristics determined based on software requirements. Scalability is made easier through cloud components. This eliminates the need for provisioning extensive hard disk capacities for the future, allowing for a budget that adapts to current needs, reflecting a more flexible and cost-effective strategy.

This cloud operation provides the necessary level of security for this application. If even more confidential data is used, it is possible to increase security but probably with less flexibility, as is often the case.

VIRTUAL SENSORS

Data analyses drives equipment development, including spoilers, suspensions, and cooling systems. Additionally, numerical estimators, known as virtual sensors, are developed and integrated into calculators. These virtual sensors are especially valuable because competition vehicles are restricted to about 40 onboard sensors, and certain real sensors, like cylinder pressure sensors, are forbidden in racing.

We're creating a virtual sensor using extensive test data collected and stored on our platform. During races, this supplements data from authorized sensors, aiding the vehicle's systems. Developing the model involves significant upfront effort before integration into the vehicle's ECU, with the flexibility to adjust parameters during a race weekend.

The cloud-based database accelerates neural network learning for virtual sensors. Cloud components are available for this purpose, and with Capgemini's support, they have been gradually implemented in 2023. The approach mirrors the one used for the Peugeot 9X8, where data is collected and processed for potential applications in Stellantis group production vehicles to enhance redundancy of valuable information.

"PEUGEOT SPORT HAS SEEN SIGNIFICANT IMPROVEMENTS, AS THIS PLATFORM BRINGS GAINS IN SPEED OF DATA ANALYSIS AND BOOSTS EFFICIENCY DURING PRACTICE SESSIONS AND RACES."

This project is a very good example of software engineering for mechanical engineering. It reduces the time spent searching for numerical values, comparisons, and consolidations of information. The product can be developed with better data visibility and more reliably. Thanks to the data platform, all the data can be analyzed, compared, and processed in a few minutes. This saves time, opens new possibilities, and allows engineers to ask deeper question with the data, which means a better understanding of the extreme behavior of the car in its cutting-edge driving envelope and results in better settings for performance and reliability.

This method can be applied to the development of all types of mechanical/mechatronic components and systems, and this platform offers great flexibility of use as well as freedom of evolution, without limits in the life cycle of the product or the number of users.

INNOVATION TAKEAWAYS

#AUTOMOTIVE
#DATA
#AI

DATA-POWERED PERFORMANCE ENHANCEMENT

Real-time data analytics optimize Peugeot's hypercar, improving efficiency and speed, showcasing how data is the new pit crew in racing innovation.

CLOUD-POWERED AGILITY

Cloud-based platforms empower engineers with swift data access and analysis, proving that the race for performance is as much in the virtual world as on the track.

VIRTUAL SENSORS AND FUTURE INSIGHTS

Integrating virtual sensors with AI enriches racing strategies and extends learnings to mass-market vehicle innovation, fueling the future with past race data.

THE GLOBAL DATA SCIENCE CHALLENGE TACKLES BIODIVERSITY BUZZ

HELPING THE WORLD OF INSECTS TELL ITS STORY

LUCAS UNTERBERGER

Data Engineer, Insights & Data, Capgemini

LUKAS KEMETINGER

Student Data Scientist, Insights & Data, Capgemini

DOMINIK LEMM

Data Scientist, Insights & Data, Capgemini

The 2023 Capgemini Global Data Science Challenge attracted more than 1,500 people, the biggest hackathon to date. Participants used data-science technologies to develop a solution for monitoring insect diversity – a key indicator of the health of the planet.

Insects may be tiny, but they play a massive role in the world's ecosystems. Insects perform many essential functions – including pollinating crops, recycling nutrients, and acting as a high-protein food source for other species. The loss of even a few species of insect can devastate an ecosystem, so protecting them is critical.

But helping these tiny creatures comes with some huge challenges due to the size and diversity of the insect realm. It's estimated insects make up more than 80 percent of animal life on earth, and they're by far the most varied creatures on the globe. Entomologists have identified more than one million species of insect, but that's only a fraction of the estimated 5.5 million types on the planet.

In any given location, the sheer number of insects and the broad range of species present make it impossible for researchers to accurately monitor insect diversity – or does it? Could technology help? That's the question the 2023 Global Data Science Challenge set out to answer.

THE 2023 GDSC: BIODIVERSITY BUZZ

Capgemini launched its flagship hackathon in 2016 as a means to provide developers with hands-on training in the use of AI and machine learning while devising solutions that address the many environmental and social challenges facing the planet. For each event, organizers identify a challenge and then invite teams to develop a solution using artificial intelligence, machine learning, and other state-of-the-art data-science tools. Having used the 2019 GDSC to help researchers protect whales, organizers went to the opposite end of the animal kingdom for 2023, with a focus on insects.

For the 2023 challenge, Capgemini partnered with the Naturalis Biodiversity Center in Leiden, Netherlands – a research institute focused on biodiversity – as well as long-time challenge partner Amazon Web Services.

The result was Biodiversity Buzz, which attracted more than 1,500 participants organized into more than 400 teams. We formed our team – called ***It Is A Bug*** – along with Lukas Kemetinger, a student consultant with a focus on AI, and another person who has since left Capgemini for another opportunity.

Naturalis provided each team with approximately 24 hours' worth of audio recordings, which featured the sounds of 66 different species of cicadas and grasshoppers. The challenge was to create an AI model that could improve how researchers identify these insects from the sounds they make. This intrigued us as soon as we heard about it, since applying AI to audio files was not something we had done before. Therefore, we started researching the issues even before the challenge's official launch.

FROM AUDIO CLIP TO COMPUTER VISION

We learned that audio files are usually presented as waveforms – which represent the amplitude of a sound over a period of time. But these waveforms also implicitly represent the frequency of the sound. This provides insights into the pitch of the sound.

Waveforms can be translated into spectrograms – visual representations of the amplitude, frequency, and duration of the audio. These can then be converted into a Mel Scale – which represents a sequence of pitches spaced at equal distances from each other. When learning to read music, people are often taught to sing the Solfège (such as "do-re-mi-fa-so-la-ti-do"). Each note is equally spaced, making this an example of a Mel Scale.

By converting audio files into Mel Scale spectrograms, we could then analyze them using deep-learning algorithms for computer vision, such as convolutional neural networks.

THE WINNING SOLUTION

The ***It Is A Bug*** team's approach was based on four pillars:

Fast prototyping. To achieve this, we developed our own machine learning workflows based on PyTorch Lightning – an open-source Python library that acts

as a high-level interface for the PyTorch deep-learning framework.

Preprocessed audio clips. To leverage the audio data most effectively, we split it into five-second audio clips as an input to our model. These were of sufficient length to reliably identify the species of insect on the recording.

On-the-fly data augmentation. To make sure our solution could work in real-world situations, we randomly added noise, changed the volume, or made other modifications to our audio clips. This helped train our model to identify and learn the underlying sounds the insects make. This made our model more robust and able to identify insect species from audio clips of varying quality.

Lightweight convolutional neural networks. We relied upon EfficientNet-V2 – a family of convolutional neural networks known for fast training speeds and excellent performance predictions. When used for audio classification, the EfficientNet-V2 architecture has demonstrated impressive performance in predicting bird songs, so it seemed like an excellent starting point.

Our final model improved the baseline performance on the dataset supplied by Naturalis from 79 percent to 92 percent accuracy – and earned us the top score in this competition.

We were thrilled to win the GDSC Biodiversity Buzz challenge. But we also recognize that in data science there's always opportunity for improvement, so we have made our solution – including the model's code, the training workflow, and our documentation – openly available via GitHub for others to use.

"THE LOSS OF EVEN A FEW SPECIES OF INSECT CAN DEVASTATE AN ECOSYSTEM, SO PROTECTING THEM IS CRITICAL."

THE BENEFITS OF COMPETING

As the winners of this year's GDSC hackathon, our team is looking forward to visiting with the researchers at Naturalis and exploring how our solution can enhance their work. While that's the grand prize, it's only one of the many benefits of taking part in these competitions.

The GDSC is an opportunity to improve data-science skills, experiment in a safe environment, learn about the Amazon Web Services platform, and build connections with like-minded individuals – all while working towards a solution that can help the planet. Those are great incentives to make this an annual activity, and we're already looking forward to the next challenge – at GDSC 2024.

INNOVATION TAKEAWAYS

#GDSC
#BIODIVERSITY
#AIFORGOOD

THINK OUTSIDE THE VOX

Don't let the nature of the dataset limit your approach to the problem. One of the keys to training the winning solution was to translate the audio files into a visual format that deep-learning algorithms could analyze.

EMBRACE THE IMPERFECT

Training models with perfect data results in models that can only analyze perfect data. Adding noise and other changes to the inputs made the winning solution more robust.

HIVE MIND

GDSC's benefits extend beyond creating a solution to help the planet. They are also excellent networking opportunities, a chance to experiment in a safe environment, and a fun way to learn about data-science technologies.

INNOVATION MOVERS & SHAKERS

ANANDA DESHMUKH

AI/ML, Analytics, Data Science, Insights & Data, Capgemini

ARNE ROSSMANN

Innovation Lead, Insights & Data, Capgemini

CHRISTOPHER SCHEEFER

Intelligent Industry Lead, Insights & Data, NA, Capgemini

DANIELA RITTMEIER

Head of Data & AI Center of Excellence, Insights & Data, Capgemini

JAMES WILSON

I&D Advisory, Insights & Data, Capgemini

LIZ HENDERSON

Executive Advisor, Insights & Data, Capgemini

MARIJN MARKUS

AI Lead, Managing Data Scientist, Insights & Data, Capgemini

DR.AKHAURI PRAKASHKUMAR

Chief Strategy Advisor, Insights & Data, North Central Europe, Capgemini

ARUNA PATTAM

Head Generative AI, Analytics, and Data Science, Insights & Data APAC, Capgemini

DANIEL KÜHLWEIN

AI & Analytics Portfolio Lead, CoE for AI, Insights & Data, Capgemini

DINAND TINHOLT

VP, Insights & Data, NA, Capgemini

KUMAR CHINNAKALI

Managing Delivery Architect, Insights & Data, Capgemini

MARK OOST

Generative AI Global Lead, Insights & Data, Capgemini

MARK ROBERTS

Deputy Head, Capgemini Global Generative AI Lab, Capgemini Engineering

INNOVATION MOVERS & SHAKERS

MONISH SURI
Director, Insights & Data GBL, Capgemini

NICOLAS CLAUDON
CTO, Insights & Data, South Central Europe, Capgemini

OLIVIER PERHIRIN
CTO, Insights & Data France, Capgemini

RAJESH IYER
VP & Global Head of ML & Generative AI, Insights & Data FS, Capgemini

ROBERT ENGELS
Head of Capgemini Global Generative AI Lab, Insights & Data, Capgemini

RON TOLIDO
CTO, Insights & Data Capgemini

ROOSA SÄNTTI
Head of Insights & Data Finland, Capgemini

DR.SERGEY PATSKO
Data & AI Deputy Group Offer Leader Capgemini

SHYAMSREE NANDI
VP, Insights & Data FS, Capgemini

VINOD JAIN
COO & Head of Delivery, Insights & Data India, Capgemini

URBAN EKEROTH
Head of Portfolio, Insights & Data Nordics, Capgemini

WEIWEI FENG
AI Nordic Lead, Insights & Data, Capgemini

YASHOWARDHAN SOWALE
CTIO I&D India, I&D Architecture head, India Domain Leader for AI, Insights & Data, Capgemini

CORE TEAM

RON TOLIDO

CTO, Insights & Data
Capgemini

ROBERT ENGELS

Head of Global Generative AI Lab,
Insights & Data, Capgemini

DEBLEENA MISRA

Innovation Marketing Lead,
Insights & Data, Capgemini

ARNE ROSSMANN

Innovation Lead,
Insights & Data, Capgemini

About Capgemini

Capgemini is a global leader in partnering with companies to transform and manage their business by harnessing the power of technology. The Group is guided every day by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of nearly 350,000 team members in more than 50 countries. With its strong 55-year heritage and deep industry expertise, Capgemini is trusted by its clients to address the entire breadth of their business needs, from strategy and design to operations, fueled by the fast evolving and innovative world of cloud, data, AI, connectivity, software, digital engineering, and platforms. The Group reported in 2022 global revenues of €22 billion.

Get the Future You Want | www.capgemini.com

For more details contact us:

RON TOLIDO

CTO, Insights & Data
Capgemini

DEBLEENA MISRA

Innovation Marketing Lead,
Insights & Data, Capgemini