



WAVE II | 2021

# DATA-POWERED INNOVATION REVIEW

# TABLE OF CONTENTS

## THE RISE OF AUTONOMOUS

- |    |  |    |
|----|--|----|
| 01 | <a href="#">Time to Shift the Gear with Software Defined Vehicles</a><br>Rodrigo Maia..... | 06 |
| 02 | <a href="#">The Superpower of Autonomous AI Systems</a><br>Andy Wylie, Kence Anderson..... | 09 |

## THE RISE OF DATA MESH

- |    |  |    |
|----|--|----|
| 05 | <a href="#">Five Pillars for a Data-Powered Future</a><br>Robert Engels, Veronika Heimsbakk..... | 20 |
| 06 | <a href="#">An Engineering Approach to Data Mesh</a><br>Ron Tolido, Stephen Brobst.....          | 24 |

## THE RISE OF DATA ECOSYSTEMS

- |    |  |    |
|----|--|----|
| 08 | <a href="#">Collaborative Data Ecosystems</a><br>Monish Suri, Ron Tolido.....                  | 31 |
| 09 | <a href="#">Unleashing the Potential of Open Data</a><br>Esther Huyer, Gianfranco Cecconi..... | 35 |

## THE RISE OF CREATIVE AI

- |    |   |    |
|----|---|----|
| 12 | <a href="#">Creative AI: Breakthroughs with Generative Models</a><br>Adam Grzywaczewski .....   | 45 |
| 13 | <a href="#">Generative Language Models and the Future of AI</a><br>Rajeswaran Viswanathan ..... | 49 |

## THE RISE OF SUSTAINABLE AI

- |    |   |    |
|----|---|----|
| 16 | <a href="#">The Wave of Data in Oceans</a><br>Espen Johnsen, Robert Engels.....                           | 60 |
| 17 | <a href="#">Make your Brand Stand Out with AI-Driven Eco-Responsibility</a><br>Vincent de Montalivet..... | 63 |
| 18 | <a href="#">Using Data and AI to Crack the Earthquake Code</a><br>Niels Heffinck.....                     | 67 |

- |    |   |    |
|----|---|----|
| 03 | <a href="#">Gesture Recognition for a Safer, More Inclusive Society</a><br>Rajashree Das, Rekha Chandrasekar..... | 12 |
| 04 | <a href="#">Amazon Lookout for Metrics: Chasing the Anomaly</a><br>Rebecca Y. Gonzales, Sriram Kuravi.....        | 16 |

- |    |   |    |
|----|---|----|
| 07 | <a href="#">Data Mesh – Radically Rethinking Data Architecture</a><br>Prajwal Kumar, Vinodh Subramony Iyer..... | 27 |
|----|---|----|

- |    |  |    |
|----|--|----|
| 10 | <a href="#">Data as a Product – New Value from Data</a><br>Chandrasekhar Balasubramanyam.....                    | 38 |
| 11 | <a href="#">Federated Learning – Training AI, with Privacy</a><br>Anupam Srivastava, Padmashree Shagrithaya..... | 41 |

- |    |   |    |
|----|---|----|
| 14 | <a href="#">Creating Intelligent, Engaging Customer Experiences with AI</a><br>Scott Prevost , Pranav Kumar, Punit Santani..... | 53 |
| 15 | <a href="#">Creative AI: From Drug Discovery to Generating Ad campaigns</a><br>Menno Van Doorn.....                             | 56 |

- |    |  |    |
|----|--|----|
| 19 | <a href="#">Battling Deforestation Using Crowd-Driven AI</a><br>Jonathan Walker .....    | 70 |
| 20 | <a href="#">The Importance of Diversity of Minds in AI</a><br>Anne Laure Thieullent..... | 73 |
| 21 | <a href="#">Code of Ethics for AI – Humans at the Center</a><br>Yashowardhan Sowale..... | 75 |

# INTRODUCTION



I am pleased and proud to introduce the next wave of Capgemini's Data-powered Innovation Review to you. Data, analytics, and AI provide an endless amount of innovation opportunities, and this edition shows you many, fascinating sides of it.



**ZHIWEI JIANG**  
CEO, Insights & Data Global  
Business Line, Capgemini

In today's unprecedented times, innovation is no longer optional. And every business needs to embark on it themselves in a unique way, unless they want to end up as the (very) late adaptors of whatever the big technology companies come up with next.

To fully embrace the innovative power of data, it needs to be driven at the core of business - possibly changing entire business models, thriving on a brand-new data economy. And it needs to be done at scale. As we have learned from too many AI proof-of-concepts, data-powered innovation really delivers when it is applied end-to-end, embedded in the foundation of business.

Having said that, we do need to realize that the truly exciting things tend to happen at the outskirts of traditional, central business and IT. 'Edge' AI hence is a topic of great interest, because it brings the power of data right in the middle of where the action really takes place.

Finally, with all the awesome power of data, analytics, and AI in our hands, we need to ensure that the next generation of innovations is sustainable, both in terms of its own energy consumption, and the way it is applied for a healthier environment. Also, innovative solutions need to benefit the entire society, nobody excluded. For that, we actively seek to ensure our innovation teams properly reflect our company's inclusiveness and diversity goals.

In any case, we hope this edition yet again whets your appetite for data-powered innovation – because the waves keep on coming. Happy reading!

# EDITOR'S NOTE



## RON TOLIDO

**Chief Technology  
and Innovation Officer,  
Insights & Data Global  
Business Line, Capgemini**

Innovation is not only a matter of raw creativity, free-format brainstorming, and sailing the disruptive waters of the Blue Ocean – although all of that brings a lot fun. It is also a matter of consistency and persistence, keeping the cadence of innovation processes in terms of reaching milestones, delivering products and results, assessing what could improve, and kicking off the next iteration. Eat, sleep, innovate, repeat, as the scrum master of dance music – Fatboy Slim – once may have said, more or less.

We're therefore happy to introduce the second edition (or 'wave') of our data-powered innovation review to you. Designed to whet your appetite for the phenomenal innovation power of data, analytics, and AI, it contains nothing less than 21 different musings about how data-powered innovation is changing people, businesses, and society. That's a whole lot of innovation. However, not to worry – each article can be digested in just a few Internet minutes. Further, we advise you to roam freely through the

content, picking up whatever topic tickles your curiosity, touches a relevant subject matter area or can prove resourceful for your next data-powered innovation workshop.

In favor of some more structure? The articles have been put in five topic 'containers,' reflecting our current innovation agenda within Capgemini's Insights & Data CTO Office team. 'The Rise of Autonomous' shows the first glimpses of a tantalizing future in which AI-driven systems do all the heavy lifting in business for us and on our behalf. 'The Rise of Data Mesh' explores a quickly evolving architectural concept in which business domains own high-quality data, doing justice to the 'water-like agility, adaptability, and responsiveness that organizations need.' 'The Rise of Data Ecosystems' demonstrates the sheer energy of collaborating on data – whether it is inside the organization or out there, in wider business networks. The 'Rise of Creative AI' describes a mind-boggling revolution in which AI systems augment humans in their most creative activities. Arguably our favorite container, 'The Rise of Sustainable AI,' depicts how data, analytics, and AI create better futures in a sustainable, inclusive, and diverse society.

Yet again, the articles have been crafted by Capgemini experts on several occasions in collaboration with key technology partners. Don't hesitate to contact the contributors. They'll be most happy to discuss further and help you sharpen your innovation skills.

**Enjoy the water, enjoy  
the waves. Surf's Up  
for innovation!**





Time to Shift the Gear with Software Defined Vehicles.....	06
The Superpower of Autonomous AI Systems.....	09
Gesture Recognition for a Safer, More Inclusive Society.....	12
Amazon Lookout for Metrics: Chasing the Anomaly.....	16

# THE RISE OF AUTONOMOUS





# TIME TO SHIFT THE GEAR WITH SOFTWARE DEFINED VEHICLES

Rodrigo Maia

EXECUTIVE BOARD MEMBER  
T&I AT CAPGEMINI ENGINEERING

## Software and Data Drive Change in the Automotive Industry

**Software and data are transforming the automotive industry as we traditionally know it – breaking down barriers in and out of the vehicle through the deployment of cutting-edge technologies. Rodrigo Maia, Executive Board Member at Capgemini Engineering Portugal, shifts you into the not-so-distant future paradigm of vehicles – the Software-Defined Vehicles (SDV) – with an approach focused on the challenges of an emerging automotive ecosystem.**

Software and data define everything nowadays. We see software and data overcome hardware as an innovation driver for vehicles, taking the center stage of new automobile design. Together, they define new functionalities and interact with physical systems and other remote systems. However, the most significant change concerns the hardware no longer driving the function but becoming a part of the software and data platform. Let's take a closer look at emerging key trends to get a behind-the-scenes vantage point.

### Electrifying the fleet

First, there is the drive for electrification. Climate change is an apparent reality. Worldwide, many countries have committed to becoming carbon-neutral will impose bans on internal combustion engines (ICE) as early as 2030 and on hybrids by 2035. As a result, many automakers are committed to selling only zero-emission vehicles in the next decade, with [General Motors](#) being one of the latest examples.

Another major trend is a shared mobility that enables traffic reduction and meets the mobility demands of ever-growing cities. Automation will bring safety to the roads and counteract the number of traffic-related fatalities. It will also boost productivity and make driving time more usable. Finally, connected vehicles will increase the convenience and safety of road trips and optimize traffic flow.

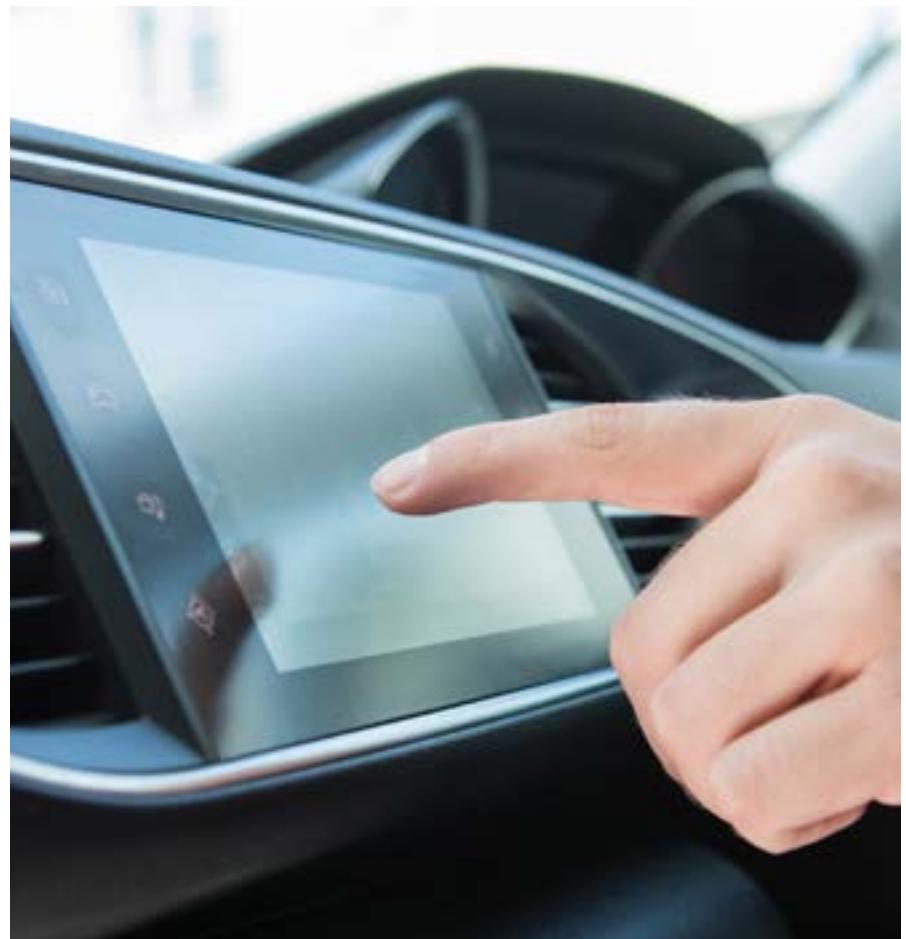
What do these trends have in common? They are powered by software and data. This raises new challenges related to the safety and security of the drivers due to an increased number of vehicle functions and ways of interacting both with the vehicle and others on the road system. However, it won't slow down the adoption of software and data in the vehicle.

Vehicles thus become much more digital through the deployment of cutting-edge technologies such as Edge Computing, Cloud, and AI. This enables building and delivering customer-centric solutions that make for a safer and more personalized ride. Digitalization is progressively changing our vision of vehicles and the way they are used. Undoubtedly, vehicles will become our "third place" and transform time wasted into time well spent.

### Challenges for the new automotive paradigm

For the new automotive paradigm to become a reality, some challenges need to be addressed:

**1. Stay one step ahead of the virtual assistants** – in the future, the automotive industry may have to compete with Siri, Alexa, or Google Assistant in the battle to make vehicles livable one. Adequate integration will be one of the more prominent challenges that SDV will have to tackle.



**2. Overcome obstacles by upgrading and updating** – With the adoption of shorter development cycles and the continuous update or upgrade of vehicle functions, software will expand the adaptability and longevity of cars by constantly updating the current systems in operation. Adapting to a DevOps approach is a necessity for the industry.

**3. Transition from lone shark to shark tank** – established industry players that used to own the value chain face the challenge of no longer being the only "shark" in the tank. New entrants (such as Tesla); tech giants (GAFA and its Chinese counterparts) are jumping into the mobility market. Further, Telco operators intend to use 5G and V2X solutions to create new revenue

streams. Moreover, silicon vendors provide the technology for the electronics deployed in the vehicles of the future. Therefore, the incumbents must decide whether they want to compete, collaborate, or master the software and data platforms that comprise the new vehicle backbone. "Software sharks" are positioning themselves to lead the development of this backbone (e.g., Android Automotive). In other words, whomever (still) wants to rule the tank needs to quickly realize the importance of developing their software and data ecosystem.

**4. Create a balance between digital demand and capacity** - Technology evolution is driving massive demand for the next generation of software and data solutions that the industry

cannot currently deliver. The productivity and efficiency of the Automotive sector are well below the level of demand in this area. How can organizations remain relevant? Here's how. The industry needs to increase its availability of skilled resources, building curricula themselves or through course co-creation with academies and universities. For instance, it needs to find better ways to manage processes, efficiency, and productivity, in a much more agile and more DevOps-oriented way. Also, it must improve access to knowledge and innovation.

In all cases, partnerships are essential. They can be set up through joint ventures between manufacturers, software companies, and research centers to accelerate technology adoption. A good example is Vortex-CoLab, a collaborative laboratory created in cooperation with three universities dedicated to applied research within this exciting, rapidly evolving area.

### Running on software

We see a significant shift in the automotive industry in the way vehicles are designed and delivered. Vehicles went from being mainly hardware-based to being a software and data platform on wheels. In the new automotive paradigm, software and data will account for 90 percent of future innovations for vehicles, which is led by industry leader [Volkswagen](#).

The key to success lies in adapting and being resilient, continuously updating and upgrading the vehicle with software, data, and AI features. This is how you can retain a competitive advantage and gain a privileged and unique position in the continually changing waters of the Automotive shark tank.

#automotive #telco #software #future #cars

## INNOVATION TAKEAWAYS

### Innovate with software and data

It brings a new paradigm – the Software-Defined Vehicle - and a set of unique challenges for every player in the Automotive ecosystem.

### The vehicle is our third place

Use cutting-edge technologies, like AI, Edge Computing, and the Cloud to develop customer-centric solutions for vehicles that can feel like home or work.

### Adapt to the shark tank

Adapt rapidly while staying resilient, continuously updating and upgrading vehicles, leveraging the full potential of software and data as it evolves. Indeed, there are ways to rule the shark tank.



# THE SUPERPOWER OF AUTONOMOUS AI SYSTEMS

Kence Anderson

PRINCIPAL PROGRAM MANAGER,  
MACHINE TEACHING INNOVATION,  
MICROSOFT

Andy Wylie

PRINCIPAL PROGRAM MANAGER,  
AUTONOMOUS AI ARCHITECT,  
MICROSOFT

AI is finally making high-value decisions in the real world

Business decision-making is fuzzy, though, nuanced, and full of strategy trade-offs. Previously, only humans could make some of these decisions well. Still, AI has matured to the point where, for specific conclusions in controlled scenarios, it can imbue process owners with decision-making superpowers that can sense and autonomously act on behalf of the company. In addition, creating these AI systems now is within reach of business experts themselves.

Getting beyond predictions

It's one thing to predict. It's another thing to decide and act. Chatbots, predictive maintenance applications. Further, statistics dressed up as AI may provide valuable insights, perception, and correlations, but it is essential to know that AI can sense and then, in real-time, can decide and act based on what it perceives.

Here's why that's important...



for their processes based on their expertise.

### Making AI specific and real

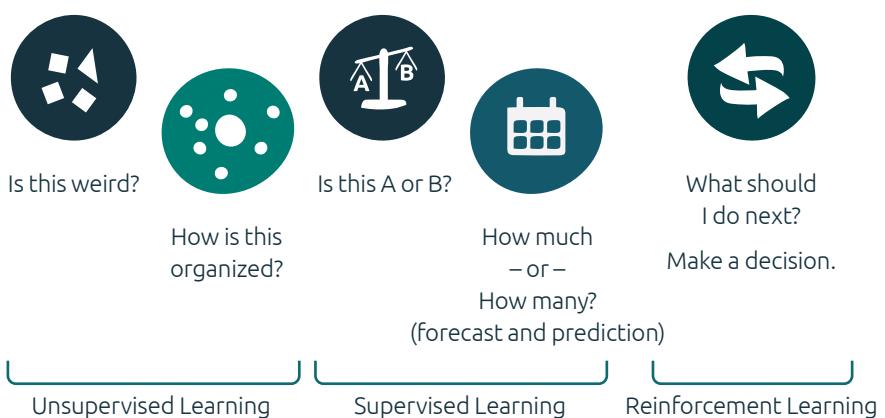
A general, all-knowing AI will be science fiction for a long time to come. It is much better to focus on designing and building AI systems that succeed at specific, high-value tasks. When experts optimize their business processes, they do not leverage the complete human cognitive capability spectrum.

Instead, they direct the mentioned three cognitive abilities into specific tasks in controlled environments. AI can do precisely the same. To generate similar value - whether advising the less experienced or freeing up experts for even higher value tasks – it needs to deliver identical, specific intelligence.

It all boils down to this...

## AUTONOMOUS SYSTEMS = ADVANCED PERCEPTION

### + LEARNED DECISION MAKING



High-quality expertise is a scarce resource. Many companies aim to 'download' their organization's best expertise into AI systems that can advise, upskill and level up their workforce. And when these systems become autonomous, they can even act on their behalf. With low-code AI

development platforms such as Project Bonsai, this is now not only within the realm of highly specialized data scientists. It is actually in the hands of business experts.

#machineteaching #autonomousystems #artificialintelligence #innovation #machinelearning

## INNOVATION TAKEAWAYS

### All by myself

Autonomous systems combine advanced perception, learning-based decision-making, and automated execution. Apply them to processes that require perception, forward-thinking, and strategy.

### Be applied for success

AI is not science fiction, nor is it an all-knowing super brain. Nonetheless, organizations that focus on specific, task-oriented applications will reap early benefits.

### Business tool

Business experts can now be involved in advanced AI design and execution. Autonomous AI can upskill experts as teachers, and others can learn from it. Low-code AI development tools such as Project Bonsai enable scaling across the organization.

# GESTURE RECOGNITION FOR A SAFER, MORE INCLUSIVE SOCIETY

Rajashree Das

SENIOR DIRECTOR, CHIEF ARCHITECT,  
INSIGHTS & DATA, CAPGEMINI

Rekha Chandrashekhar

SENIOR DIRECTOR, ENTERPRISE  
ARCHITECT, INSIGHTS &  
DATA, CAPGEMINI

Gesture control and touchless user interfaces are emerging as hot technologies, paving the way for a low-touch, hands-free society

**AI combined with innovative devices provides the ability to interact with computers and devices without physically touching them. Researchers worldwide innovate in areas as diverse as hand wash monitoring, translation of hand gestures to speech, and wearable biosensors. Gesture recognition can create a safer place, even when the pandemic is over, and enable more people to interact with computers, access information, and communicate with others.**

Gestures are the most natural form of human communication. Gesture

recognition is being used today for Human Computer Interaction (HCI), a means of human-machine interaction using only physical actions without voice. The current focus is primarily on emotion recognition from face to hand gesture recognition, but it can include complete body gestures in the future.

## Gesture Recognition Technology

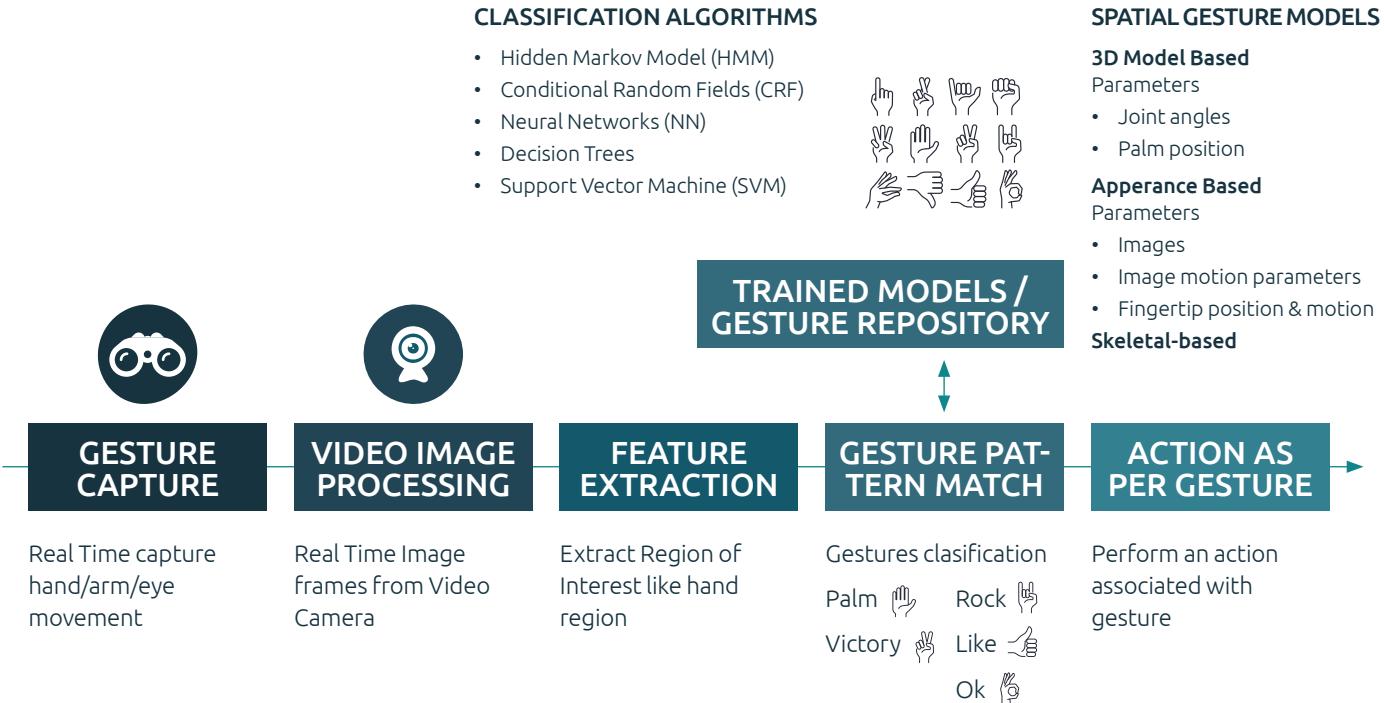
Gesture control, or gesture recognition technology (GRT), is a computer science topic and language technology. It primarily falls under the subdiscipline of computer vision in the

AI world and involves the complex integration of various components.

Unique interfaces that can capture gesture movements like cameras, use computer vision technology and deep learning algorithms to understand the underlying pattern, interpret sign language, and deploy it to the cloud for enhanced scalability and computing power.

The concept of recognizing gestures using hands and/or other body parts is based on three layers: Detection, tracking, and recognition.

## Gesture Recognition System



Apart from the hand gesture example given above, the identification and recognition of posture, gait, proxemics, and human behavior is the advanced subject of gesture recognition techniques.

There are challenges in designing GRT:

- Standardization of Gesture Language** - The need to bring in consistency concerning gestures globally to mean the same thing across geographies.
- Domain Data** - Creating gesture recognition systems for each domain needs much data to train ML models.
- Accuracy of GRT Systems** - Factors like insufficient background light, high background noise, or something else may lead to inaccurate reading/recognition of motions.

- Diverse and Complex Implementations** - Capturing, processing, training models and recognizing gestures requires advanced machine learning algorithms and intensive computing. Further, various device hardware and sensing mechanisms support different kinds of recognizable gestures.

- Society Acceptance** - User willingness to perform gestures/ use various sign languages keeping in mind social context, acceptability, and the potential implications. We are moving from the world of interaction with people to the world of interactions with systems.

### How is this technology benefiting Industry?

According to Marketsandmarkets, the market for gesture recognition is expected to grow from USD 9.6 Billion in 2020 to USD 32.3 Billion in just five

years. The integration of multiple technologies with gesture control is creating innovation opportunities in many sectors, including:

- Consumer Electronics** - GRT is a natural evolution from touchscreens for the consumer electronics market. Home automation is a big field in which gesture recognition is being employed. For instance, smart TVs can sense finger movements and hand gestures, offer touchless control over lighting and audio systems.
- Automotive** - Car infotainment systems leverage gestures to control music like changing radio channels and incoming calls, among many other things. Less interaction with the touchscreen makes the driving experience safer and more convenient.

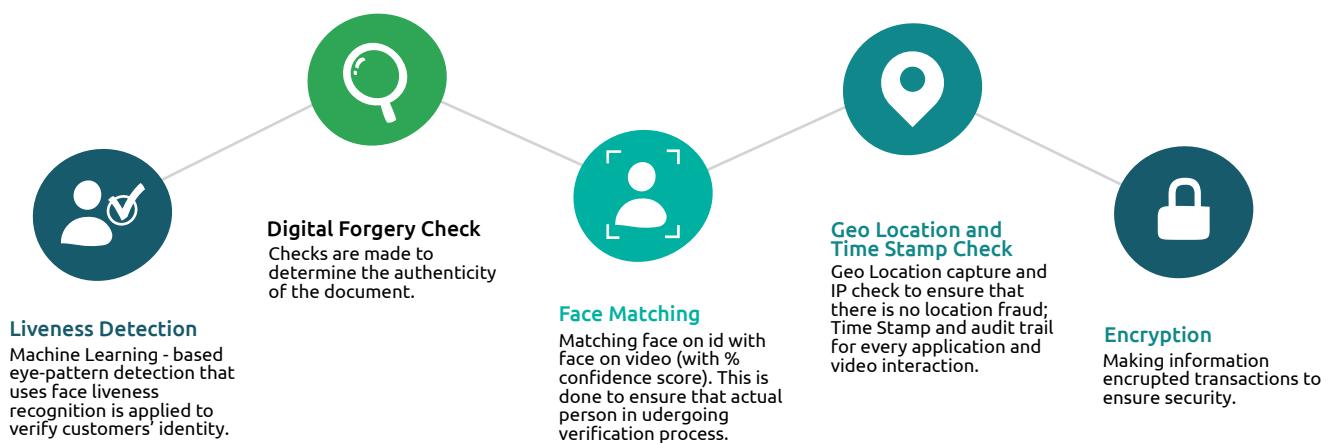
- **Healthcare** - Hand-gesture recognition system enables doctors to manipulate digital images during medical procedures using hand gestures instead of touch screens or computer keyboards.
- **Entertainment/Gaming** - Augmented reality (AR) and virtual

reality (VR) technologies in the gaming and entertainment industries have accelerated the adoption of gesture recognition products and solutions.

- **Banking** - Banks can use gesture recognition for enhanced security to allow only trusted employees/

customers to access secured areas and avoid robberies. KYC has become a challenge for banks and financial institutions during the pandemic. An e-KYC solution can ensure a frictionless customer experience with a combination of facial recognition and live emotion/gesture capture.

## A Contactless KYC



### STEP 1:

A new picture is taken of the 'live' person and matched against the one taken from the government-issued identification document. This is done to ensure that an actual person is undergoing the verification process. A total of 106 facial markers (or landmarks) are used to ensure that the captured image of the person matches the picture in the official document.

### STEP 2:

There's also an option for live motion, which will require the application to blink, smile, nod their head and more to determine that there is a real person present at the time of verification. Machine Learning algorithms are used to enhance the accuracy of this process.



## Facial Recognition and Liveness Test For KYC

There are many more scenarios where you can use GRT, such as in kiosks in the airline industry, contactless store experiences, virtual in-store displays, vending machines, and other similar environments.

#### Key Technology/Market Players accelerating adoption

Today's top researchers and producers of gesture interface products are:

- **Intel** - RealSense Touchless Control Software (TCS) seamlessly converts a touch interaction to a touchless one with an Intel RealSense Depth Camera D435.
- **Microsoft** - has a project to explore camera-based gesture recognition within surgical settings. This would allow a surgeon to view and manipulate a patient's x-rays or lab reports without having to "break out scrubs."

- **Google** - was awarded a patent for gesture recognition systems covering how online shoppers might use augmented reality to examine merchandise remotely.

Apple is increasingly active in the area as well, but so are specialized players such as Tobii Rex (infrared light-based eye-tracking device for disabled people, using it to point and interact with a computer). Another example is Hitachi's finger vein technology, which authenticates a person when they display a hand biometrically..

#### Handwashing

The need of the hour may be found in "Handwash movement recognition Technology" to promote handwashing etiquette. Due to the pandemic, the general significance of handwashing as a measure to protect health has become more evident. There is an urgent need to implement measures to ensure hygiene in food safety standards and ensure it is non-invasive. You can extend this technology to medical facilities, hospitals, schools, hotels, the foodservice industry, and event venues. Automated on-site handwashing recognition eliminates the need for intrusive, time-consuming visual confirmation and manual recording at worksites.

#gestures #gesturetechnology #computervision #gesturerecognition #ai4good

## INNOVATION TAKEAWAYS

### Power to the People

A new AI-driven ability to interact with devices and computers without physically touching them, achieved with simple gestures.

### Fast Adoption

Industries adopt gesture technology to provide for new contact-less business models to assure safety to their consumers, especially during the pandemic

### Societal Impact

Gesture technology comes as a boon to society, providing contact-less, safe, and inclusive experiences. Still, the social and emotional impact of interacting through technology does need to be further explored.

# AMAZON LOOKOUT FOR METRICS: CHASING THE ANOMALY

Rebecca Y. Gonzales

GLOBAL BUSINESS LEADER DATA,  
AI/ML, AWS

Sriram Kuravi

GLOBAL TECHNICAL LEADER DATA,  
AI/ML, AWS

Enterprises either are disrupting, or they are disrupted

In many cases, disruption strikes – by its very definition – unexpectedly and under many mysterious circumstances. Amazon Lookout for Metrics uses the latest in machine learning to detect business anomalies autonomously. It aims to proactively monitor the business's health, diagnose issues as they occur, and find opportunities for a better future for the enterprise.

Today's business landscape is constantly changing. A proven method for retaining a competitive advantage is to thrive on data – not just by gathering data but by understanding it and using it to future-proof the company. Nonetheless, organizations are dealing with more data than ever before at a more rapid pace. Uncovering and reacting to anomalies in data is a critical yet challenging problem to solve. The value of understanding the root causes of these anomalies, staying ahead of issues, and consistently delivering excellent business performance is paramount.

**Autonomous and proactive**

Anomalies are unexpected changes in data that may indicate an underlying issue. In addition to the challenge of detection, it is becoming harder to stay on top of actionable insights based on standard reporting tools. That's where autonomous analytics services using machine learning can play a significant role. They can identify opportunities proactively. Why is this impactful? Well, it means enabling more innovation and having more success at riding disruptive waves.

Digidata is an innovative player in the telco market. They are a

technology and solutions provider, helping telecommunication vendors and operators in their journey to digitally transform their business and networks. Further, they intelligently transform pricing and subscriber engagement for mobile operators, empowering operators to make better and more informed decisions.

"At Digitata, what really matters is getting everyone connected at an affordable price. This requires a deep understanding of economics, specifically supply and demand and customer behavior according to changes in either," said Nico Kruger, Chief Technology Officer, Digitata. Using autonomous analytics, Digitata was able to discover an issue that was negatively impacting pricing for a mobile network operator customer within minutes.

Anomaly detection can lead to deep insights into the customer experience across many industries. Monitoring the customer journey for outliers through enrollment, login, and other key engagement metrics allows businesses to react quickly with new offerings, products, or services. This can drive revenue streams that were previously hidden in the lack of transparency into key anomalies.

According to Amazon's GM of Artificial Intelligence Ankur Mehrotra, "The bar for consumer experiences is continuously being raised. Time is money, and being able to quickly identify a problem or an opportunity and take action can make a big difference in both customer satisfaction, retention, business efficiency, and ultimately business growth."

Most businesses have a business intelligence system in place, but this is not enough to respond at a rapid clip. Traditional anomaly detection methods are typically rule-based,

often missing out on critical anomalies (or seeing too many irrelevant ones), causing negative customer impact, and missed opportunities. And even when anomalies are detected, the root cause is frequently not easy to track down.

### Amazon learnings

At Amazon, we monitor vast amounts of data. We also began with traditional statistical methods of manual BI and dashboards, static thresholds, and rule-based alerts. But over 10 years ago, we started using machine learning and building custom models and algorithms. We moved on to deep learning and causality with anomaly detection integration in analytics platforms like Amazon QuickSight.

Two decades of investment and research have led to some key learnings. ML-based techniques work well, but they need to solve multiple, complex problems. First, (customer) data is unique. An anomaly detection algorithm may not fit all data sets. Secondly, generating actionable results by learning relationships across metrics is critical for faster remediation. Third, it is vital to dynamically adapt to changing business cycles and seasonal trends – helping to reduce false alerts and thus achieving improved accuracy. And finally, building accurate and effective anomaly detection requires specific ML expertise, which is expensive, scarce and difficult to set up.

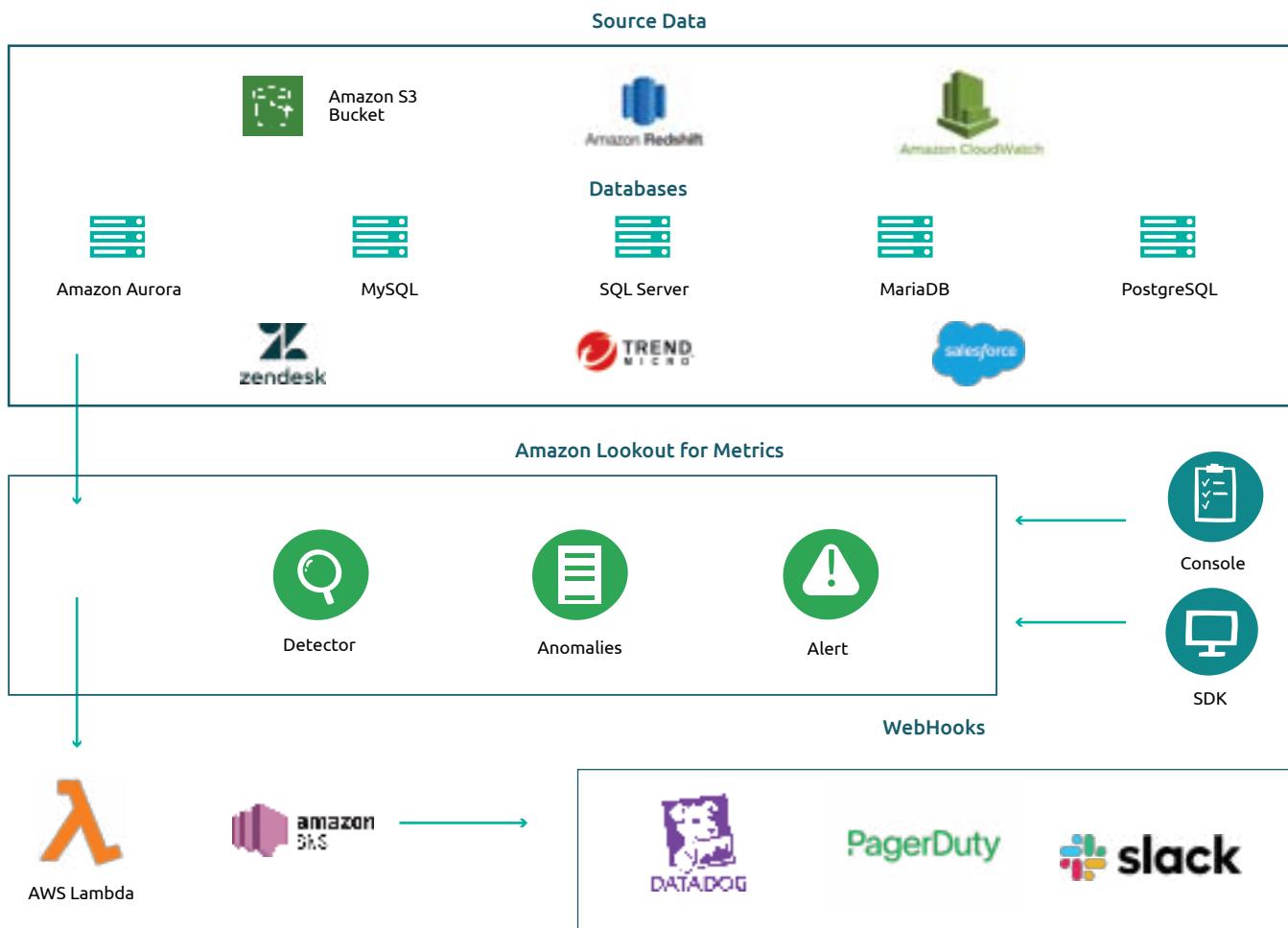
Going back to Digitata, it has many dimensions under which their products are being sold: Thousands of locations, hundreds of products and bundles, hundreds of customer segments, and hourly detection for each day of the week. The amount of data collected is staggering.

Digitata used traditional anomaly detection methods on a measure, such as revenue or purchase count but did not have the necessary insights on a dimension-based level to answer questions such as product comparisons, revenue comparisons per location, or product sales by customer segment.

### Amazon Lookout For Metrics

Digitata implemented autonomous anomaly detection using Amazon Lookout for Metrics to define measures that comprise the key performance indicators on which they wanted to detect anomalies and dimensions, which is the metadata that provides categorical information about the measures. They captured data from call records, data usage records, airtime recharges, network tower utilization, and congestion information. They also fed the data into a centralized data store and detected anomalies in real time.. Digitata was able to identify the first operational issue within two days of using the service and immediately see anomalies that have always been hidden due to a large number of dimensions.

Anomalies are prevalent in all businesses and can happen at any time. Their impact on organizations can range from damaging customer trust, falling revenue and inefficient business operations. AI-driven, autonomous anomaly detection is a simple yet impactful way to create loyal customers and identify new customer behaviors that can add to existing revenue streams. All are leading to increased resiliency over the long term.



Source: Introducing Amazon Lookout for Metrics: An anomaly detection service to proactively monitor the health of your business.

#awscloud #machinelearning #awspartner

## INNOVATION TAKEAWAYS

### Always Anomalies

Outliers hide in all business data. Finding these anomalies and timely responding to them help businesses to address disruption and stay ahead.

### Beyond Tradition

Established ways of detecting anomalies do not often work due to the complexity and volume of data involved.

### Autonomous ML

Trained with the correct parameters, Amazon Lookout for Metrics autonomously detects anomalies using AI and various machine learning options.

Five Pillars for a Data-Powered Future.....	20
An Engineering Approach to Data Mesh.....	24
Data Mesh – Radically Rethinking Data Architecture....	27



# THE RISE OF DATA MESH

A large, powerful blue ocean wave crashing, with a bright blue outline highlighting the text "THE RISE OF DATA MESH" superimposed on it.

# FIVE PILLARS FOR A DATA-POWERED FUTURE

Robert Engels

CHIEF TECHNOLOGY OFFICER,  
INSIGHTS & DATA, NCE, CAPGEMINI

Veronika Heimsbakk

SENIOR APPLICATION CONSULTANT,  
CAPGEMINI

It would help if you had straightforward tactics for your data and AI capabilities to provide you with the tools and organization to utilize your full potential in tomorrow's complex and unpredictable world

Automation, information, data, innovation, autonomous AI, self-service, all these areas require that we build our data future on robust pillars. We describe five that can guide us to the next level and invite you to join us on the journey to AI's next frontier.

Whatever business you run, you may have noticed that your success strongly depends on your ability to harness and utilize your data. So, to be successful: Dig out the data you have, make it machine-readable,

make it fair, make it the "center of your universe," publish, and use it! Whether you want to publish it within or outside your organization is up to you and your needs. However, data no one knows about or is locked up does not exist.

In an organizational landscape where you want to run a plethora of business-critical and enabling digital services utilizing all data you have access to, you need to consider building your house around five pillars:

## 1. Data-centric architecture

Data-centric architectures build on a simple, centrally available, and flexible data model for the whole organization to relate to and use. Put your data at the center of operations and make it available at large, not in an archive in a dusty corner of your system architecture. Doing so will increase decision quality, improve the perception of ownership (especially in a loosely

**coupled, data 'mesh' world), and positively influence data quality. Again, this leads to greater trust in available data, boosting the way data is used for automation and decision-making.**

*"A data-centric enterprise is one where all application functionality is based on a single, simple, extensible data model."*  
– Dave McComb

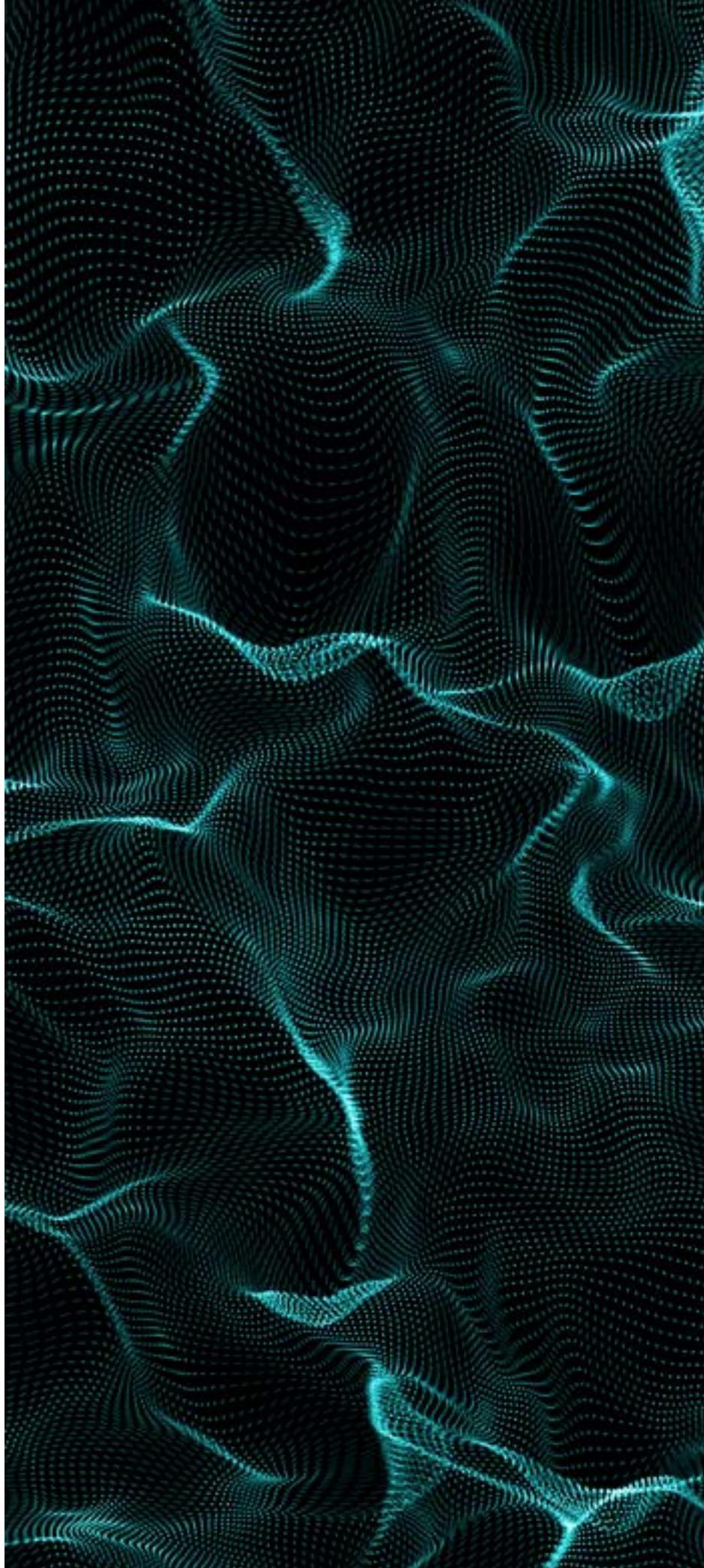
## 2. Data Fabric

Surely a data-centric architecture needs some underlying principles and technologies. Fortunately, many significant developments led the way: Data fabrics, data mesh, and semantics. How can we utilize those for exemplary architectures?

Through data fabric, we need to support architectures that provide uniform access to data in multiple data sources. Sources can be located on-prem or in the cloud. They can be editable and are machine-readable. A data fabric can consist of multi-vendor, multi-platform solutions. Data can even come from open linked data sources like the Linked Open Data cloud, connecting internal data with globally published, machine-readable information.

A data fabric is built around data schemas, where source data models are translated or virtualized and aligned to your organization's core schema (or ontology).

The concept of a core (business) ontology is central in every data-centric architecture and serves as the glue between disparate data and allows one to treat such data as a (distributed) whole.



### 3. Semantics

A powerful method needs practical tools. For this, the W3C developed Semantic Web standards like RDF, OWL, SPARQL, and SHACL, probably the most potent metadata standards available for distributed and machine-interpretable information management. These standards will provide you with the means to add semantics to data and align various data models and allow for reasoning, automated schema alignment, and merging of data and schemas.

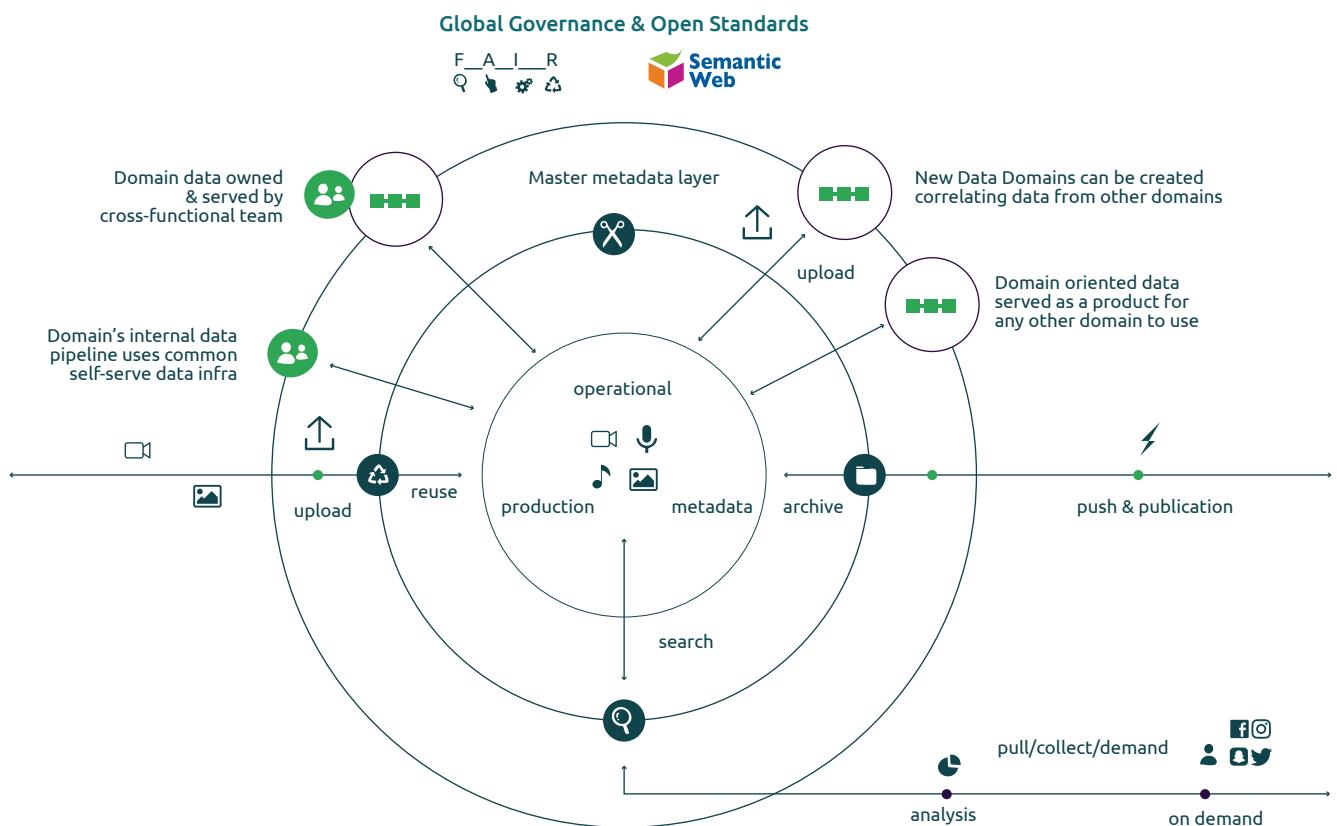
### 4. Data Mesh

Distributed data sources and their pipelines have a property that has often been solved with centralized governance: Bad quality through bad ownership. Yet, data needs love and understanding. Who better can provide this than its creator, a department, organization, or community? Semantic 'Data Mesh' principles guarantee a generation of data pipelines that produce high quality and ensure dedicated

ownership that works through the following:

- Domain-related decentralization of data ownership,
- data served as a product,
- self-serve data infrastructure
- federated governance to enable interoperability between ecosystems
- building on a unifying semantics and metadata model

## Semantic Data Fabric Mesh Approach



Capgemini, 2021. Inspired by [martinfowler.com](http://martinfowler.com) & NRK Origo

## 5. FAIR principles

FAIR principles provide:

- A framework for connecting the dots between a data-centric architecture.
- A data fabric that facilitates distributed access.
- Semantics to allow for interoperability.
- A data mesh for governing decentralization of data responsibility.

Data should always be Findable, Accessible, Interoperable, and Reusable (FAIR). By doing so, your data gains powerful properties. It will be globally unique and resolvable. It will enhance it with rich metadata for automatic discovery, disambiguation,

and alignment. The data is accessible through uniform communication protocols, uses machine-readable formats with formal semantics, uses FAIR vocabularies and ontologies, contains qualified cross-references to other metadata, and comes with clear provenance and user licenses.

Feeling overwhelmed? All points considered, you can conclude that an increasingly complex world does need some clear guidelines in several dimensions to provide us with practical methods for dealing with it. Automation, Information, data, innovation, autonomous AI, self-service: All this requires that we build our future on several robust pillars. We have introduced five principles guiding you to the next level. Will you join us on this journey?



#datapowered #datadentricarchitecture #datafabric #datamesh #fairprinciples

## INNOVATION TAKEAWAYS

### Data-centric architecture

Data is your most crucial core asset for success.

### Data fabric

Ensure consistent technical capabilities across data endpoints.

### Semantics

Ensure consistency in results shared by data endpoints, based on content and meaning.

### Data mesh

Those who know the data best should own its quality and treatment.

### FAIR principles

Ensure that investment in data delivers results for the greater good of your entire organization and its business ecosystem.



# AN ENGINEERING APPROACH TO DATA MESH

Stephen Brobst

CHIEF TECHNOLOGY OFFICER, TERADATA

Ron Tolido

EVP, CTO AND CHIEF INNOVATION OFFICER,  
INSIGHTS & DATA, CAPGEMINI

A data-powered enterprise creates value by making data accessible across the enterprise. Yet, a monolithic approach to building a data estate rarely succeeds. This is where the innovative concept of a Data Mesh comes in. It tackles monolithic data complexity by aligning to the notion of domains. One of the primary advantages is the inherent agility associated with it, with loosely coupled teams focusing semi-independently on their specific areas of business (not technical) expertise. We recommend taking an engineering approach to implementing the Data Mesh concept, balancing different design patterns to actually make the vision come alive.

## Data Products Aligned to Real-World Business

A domain may be aligned to a specific line of business within a bank – such as credit card, direct deposit, or mortgage. A domain may also align to a specific functional responsibility such as customer service or branch operations. Our recommended approach to implementation of the Data Mesh concept is to create separate schemas for each domain. Responsibility for data stewardship, data modeling, and population of the schema content is owned by experts with business knowledge about the specific domain under construction. This approach removes many of the bottlenecks associated with attempting to implement a

centralized consolidation of all enterprise data into a single schema. The domain-oriented schemas provide a collection of data products aligned to areas of business focus within the enterprise.

However, decomposition in alignment to business domains should not imply anarchy across the domains. It is critical to recognize the importance of global standards and interoperability when building a distributed data estate. Global standards include areas such as data typing, naming conventions, and quality metrics. Interoperability across the schemas requires consistency in primary and foreign key relationships across schemas, in addition to within a schema. Global access control and

optimization of cross-domain query execution are also essential.

### Balancing enterprise domains

For realization of enterprise data products, it will often be appropriate to create enterprise domains. Efficiency, consistency, and time-to-market considerations require that these enterprise domains will be constructed with the active support from cross-domain subject matter experts. For example, in banking it would be useful to have an enterprise schema to capture supertype information about accounts across the credit card, mortgage, direct deposit, and other accounts. Subtype information that is specific to a particular business domain (e.g., credit card account attributes not relevant to other domains) would remain in the credit card business domain schema. But supertype information that is common across all account types (such as open date, account status, balance amount) would be “promoted” into the enterprise domain to enable easy analysis across different account types. Similarly, customer information embedded in each business domain increases in value when promoted to an enterprise domain, facilitating a customer 360 view for purposes of enterprise marketing, risk, and other analytics.

In theory, it would be possible to implement a “union” operator across separate business domains to get an enterprise view of the data, but experience shows that the governance and integration of data associated with enterprise domains can have significant value for selected data products. There is additional work of a cross-functional nature to create enterprise domains, but the consistency and quality derived from selective use of enterprise domains is significant. Having robust governance

and an integrated design for these domains creates huge value when performing analysis at an enterprise level.

### Deployment design patterns

Separate schemas aligned to different domains do not necessarily imply a distinct database for each domain. There are different design patterns for deploying schemas within a Data Mesh:

The co-located approach places domains aligned to different schemes under the management of a single database instance. This contributes to better performance when data across multiple domains is combined. Co-location allows for more efficient execution due to improved query optimization and lower overhead than is required for assembling data across multiple database instances (or even multiple clouds). However, there are cases where the co-located approach is undesirable. For example, sovereignty laws may require that data created by a business unit within a specific country must remain in that country. For a multi-national company there need to be multiple schemas deployed across different geographies – even if the database technology used is the same. Other reasons may be a counter-indication as well, such as data gravity created by applications producing data in different clouds or use of fit-for-purpose database technologies which are distinct for unique use analytic cases or data characteristics.

With the isolated approach, the data product is completely self-contained within a single domain. The schemas used with the isolated technique are usually narrow in scope and service operational reporting requirements rather than enterprise analytics. Isolated domains typically have more autonomy in their deployment – both

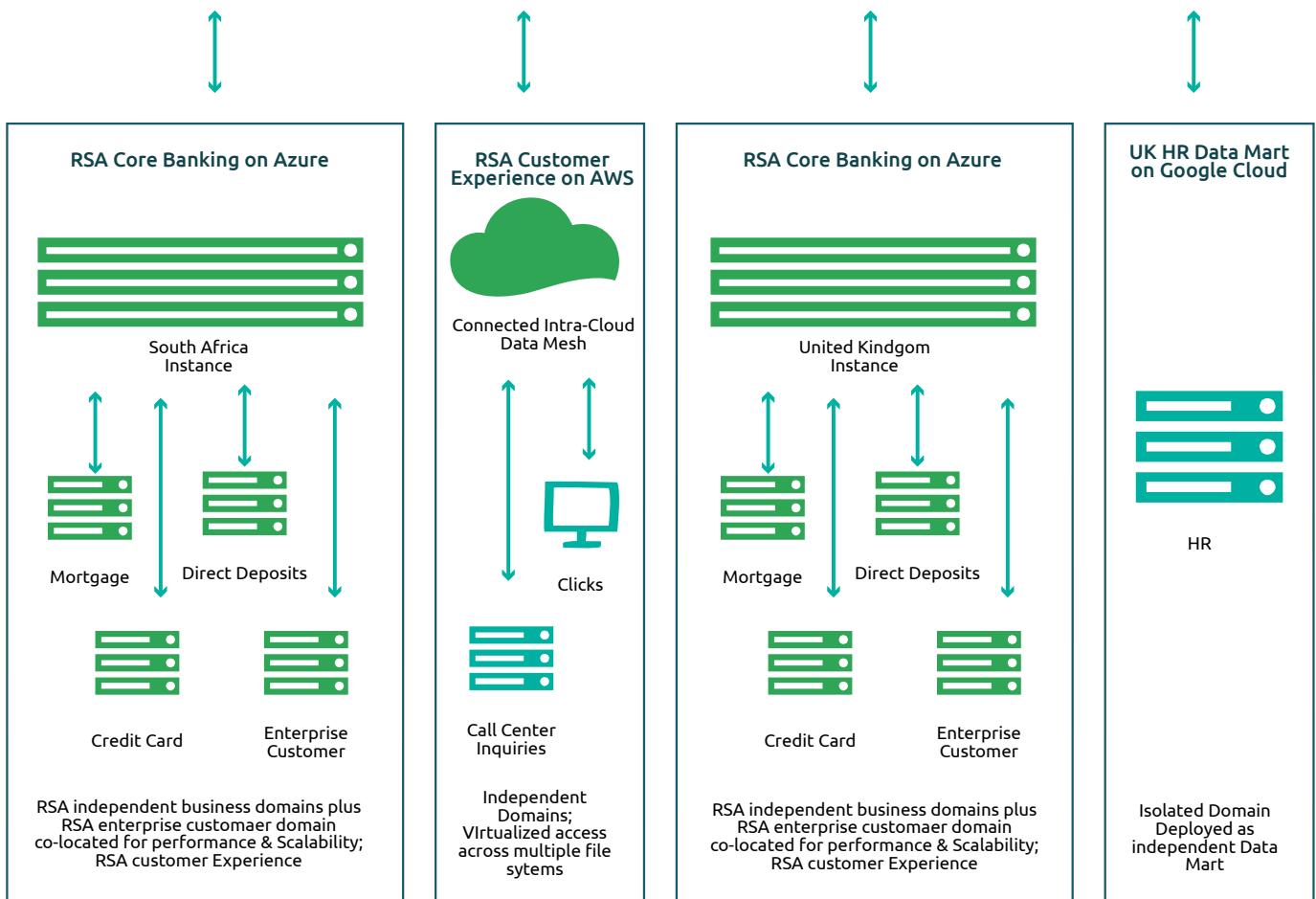
in data modeling and technology selection. Sometimes isolated domains are chosen on the need for strong security. More often though, the “real” reason has to do with politics or the desire for organizational independence.

The connected approach involves query execution across multiple, disparate locations of data. There will typically be collections of co-located schemas that exist across multiple clouds and/or database technologies. It is up to a powerful data fabric software infrastructure (including a global orchestration engine) to decide how the query execution takes place across the multiple collections of co-located schemas, optimizing data movement and efficiency of query processing.

### Best of worlds

Most large enterprises use multiple cloud service providers and operate across multiple geographies, so a connected data warehouse is fundamental to Data Mesh implementation. Within a cloud service provider and within a geography, co-location of multiple schemas aligned to specific business domains within a single, scalable database instance gives the best of two worlds: agility in implementation and high-performance in execution. Applying engineering discipline is critical in high-performance deployment of Data Mesh using a combination of the co-located, isolated, and connected design patterns.

# Connected Multi-Cloud Data Mesh with Query Fabric



## INNOVATION TAKEAWAYS

### Business is a mesh

The concepts of a Data Mesh do justice to the distributed, federated reality of most businesses – and accordingly also their data estates.

### Data is the product

In a Data Mesh, ownership and stewardship of data are assigned to the actual business domains that hold it; data thus becomes a key product, managed by a domain.

### A mixed bag of design patterns

There are multiple ways of deploying schemas within a Data Mesh: co-located, isolated and connected; which pattern is chosen depends on various enterprise considerations

### Engineering approach

A well-balanced combination of different deployment patterns is a matter of engineering a Data Mesh that does justice to both business and technology considerations.

# DATA MESH – RADICALLY RETHINKING DATA ARCHITECTURE

Prajwal Kumar

DIRECTOR, ENTERPRISE  
ARCHITECT - INSIGHTS &  
DATA, CAPGEMINI

Vinodh Subramony Iyer

DIRECTOR, ENTERPRISE  
ARCHITECT - INSIGHTS &  
DATA, CAPGEMINI

The next generation of data platforms that deliver data-powered innovation needs a radically different approach to data architecture

**The current generation of data platforms suffer from excessive centralization, lack of agility, and data trust deficit. The Data Mesh approach solves these challenges by federating ownership of data to domain-aligned teams and leveraging a standardized set of blueprints for self-serve, interoperable, and polyglot data storage and consumption. Therefore, it opens the era of data sharing and data-driven collaboration.**

Data Warehousing drove the evolution of first-generation architectures with database systems to support BI workloads and adopt of multi-dimensional 'OLAP' systems to meet growing data and performance requirements. Further, those systems were monolithic, expensive, and

provided limited support for the ('big') data workloads as they evolved in the business climate.

The last decade saw rapid adoption of second-generation architectures with on-premise Hadoop-based data lakes to support data storage and processing at scale at a low cost per byte. Data lakes also helped organizations deliver new use-cases such as natural language processing, text mining, image processing, and more. Hyperscalers, with their disaggregated storage and compute architectures, on-demand scaling, aggressive pricing, and PaaS/SaaS experiences, have driven the widespread migration of these data platforms to the cloud.

Here's the interesting part...

While the cloud has enabled commodification of hardware, very little has changed in working, organization, and processes around data engineering and data management. Highly specialized technology-focused data engineering teams often disconnect from business domain teams. The horizontal separation of concerns ends up being the bottleneck in data producer/consumer relations.

Like the application development software ecosystem already has done, the data management software ecosystem must adapt and adopt domain-driven design thinking to thrive in a highly volatile, unpredictable business environment.

## Introducing Data Mesh

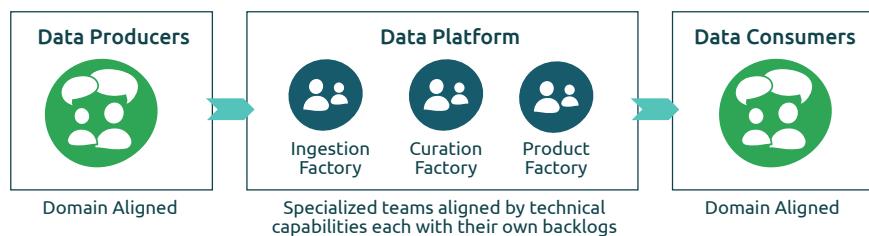
Data Mesh is applying domain-driven-design thinking and agile operations to data platforms, addressing shortcomings in current and previous generations of data lakes and data warehouses. This is done by enabling better connections between producers and consumers. It aligns domain-focused data engineering teams so that data ownership resides with domain experts -- who have the most detailed knowledge of the data requirements and its use. Moreover, data in the data mesh is aligned with [FAIR](#) principles making governance, interoperability, and security first-class concerns.

The concept of a data mesh is not new; in several ways, it is similar to federated 'data hubs' and 'data fabrics' that several organizations have already been deploying as part of their data estate modernization efforts. Unlike Data Warehouses and Data Lakes that were technology-focused, Data Mesh does not advocate any particular technology but instead relies on people and process dimensions to deliver a scalable, future-fit data platform architecture.

Data Mesh embraces domain-specific handling of data ownership and governance and embracing data-as-a-product thinking. The glue connecting domains and their associated data assets is a universal interoperability layer that applies strong governance and data standards.

## Domain-driven Data Ownership

**Ownership** in a Data Mesh resides with the domain experts because they have the most detailed knowledge of the data requirements, are closest to the customer, and are best placed to drive data monetization. Domain team data producers assume full responsibility for ownership, comprehension,



## Classic Data Platform Team Organization

completeness, timeliness, accuracy, findability, and addressability of data. They also enable consumption both within and outside the enterprise. .

By embracing data-as-a-product thinking, domain teams -- via the data product owner -- are responsible for continually improving data quality, making decisions around the vision and roadmap for the data assets, managing concerns about the satisfaction of the consumers, and assuming responsibility for the lifecycle of the domain datasets.

## Data Platform Services

Data Platform Services is a set of scalable and reusable services enabling domain-aligned teams to ingest, curate, transform, store and publish datasets using declarative abstractions underpinned by data contracts and data sharing agreements. These services automatically implement cross-cutting concerns such as findability, accessibility, traceability, security, audit, balance, control and reconciliation, and data versioning capabilities. A group of

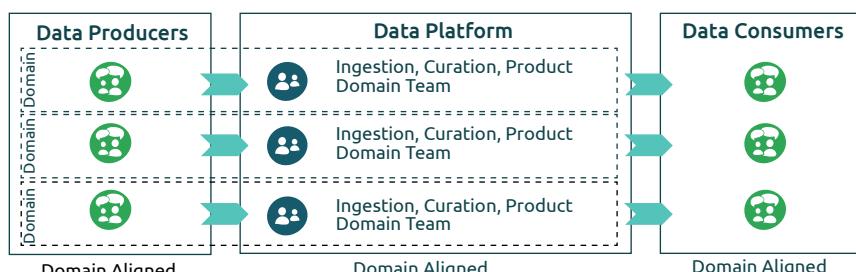
cross-functional engineers manage this self-service layer and reduce the risk of effort duplication when data ownership is distributed.

Delivering a self-serve service catalog-like user experience for domain teams allows leveraging the most suitable tools and templates that best fit the domain use case.

## Organizational Flexibility

Deploying a Data Mesh is a paradigm shift and necessitates the transition in organizational culture to one that embraces openness, encourages experimentation and innovation, and incentivizes cross-domain collaboration. A central team that currently owns and runs the data platform must give up ownership and accountability of data delivery to domain teams and instead focus on delivering cross-functional Data Platform Services.

Domain teams must inject data management skills, including creating new roles such as Data Product owners within their domains to support



## Data Mesh Team Organization

high-quality data delivery. A culture of collaboration where domain teams can create and contribute innovative assets for reuse to the central Data Platform Services becomes a critical value driver.

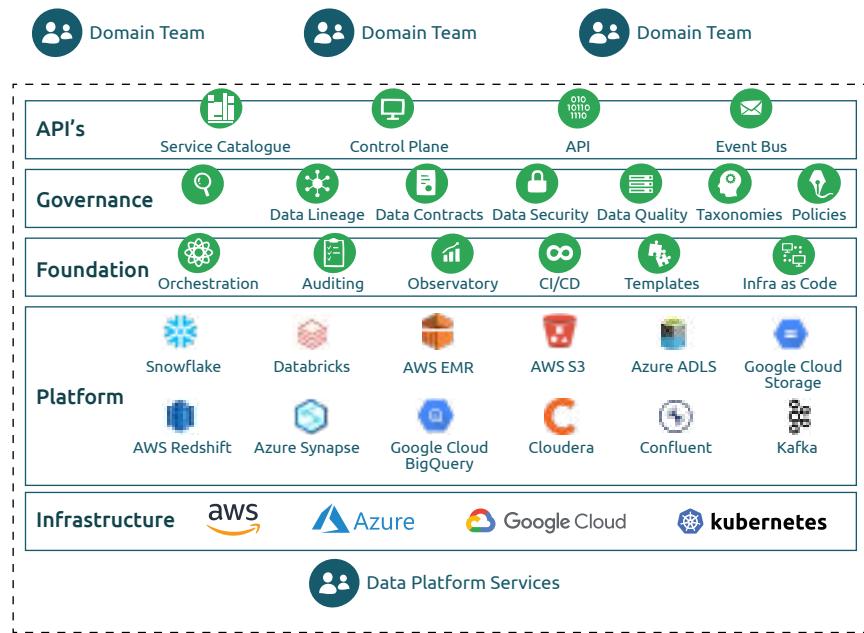
### Data Mesh and the Data Sharing Economy

Data lies at the core of most digital business models and can vastly improve products and services across industries or even create entirely new ones. The rise of cloud-based data platforms makes it easier than ever to transform digital supply chains into complex Data Mesh ecosystems.

Collaboration across organizations and ecosystems with a Data Mesh will create a more open flow of information and ideas across

industries. And, what enables collaboration outside the enterprise surely won't harm data-driven

collaboration inside the enterprise either.



Delivering a self-serve service catalog-like user experience for domain teams allows leveraging the most suitable tools and templates that best fit the domain use case.

#datamesh #datafabric #datahub #datalake #datagovernance

## INNOVATION TAKEAWAYS

### You build it; you own it

Data Mesh provides the autonomy and flexibility for domain teams to deliver data-as-a-product, underpinned by a shared set of data platform services.

### Reimagining metadata

Leveraging knowledge graphs for connected metadata management is ideally suited to flexibly customize and organize the data to enable convenient data sharing and interoperability.

### Automation is the secret sauce

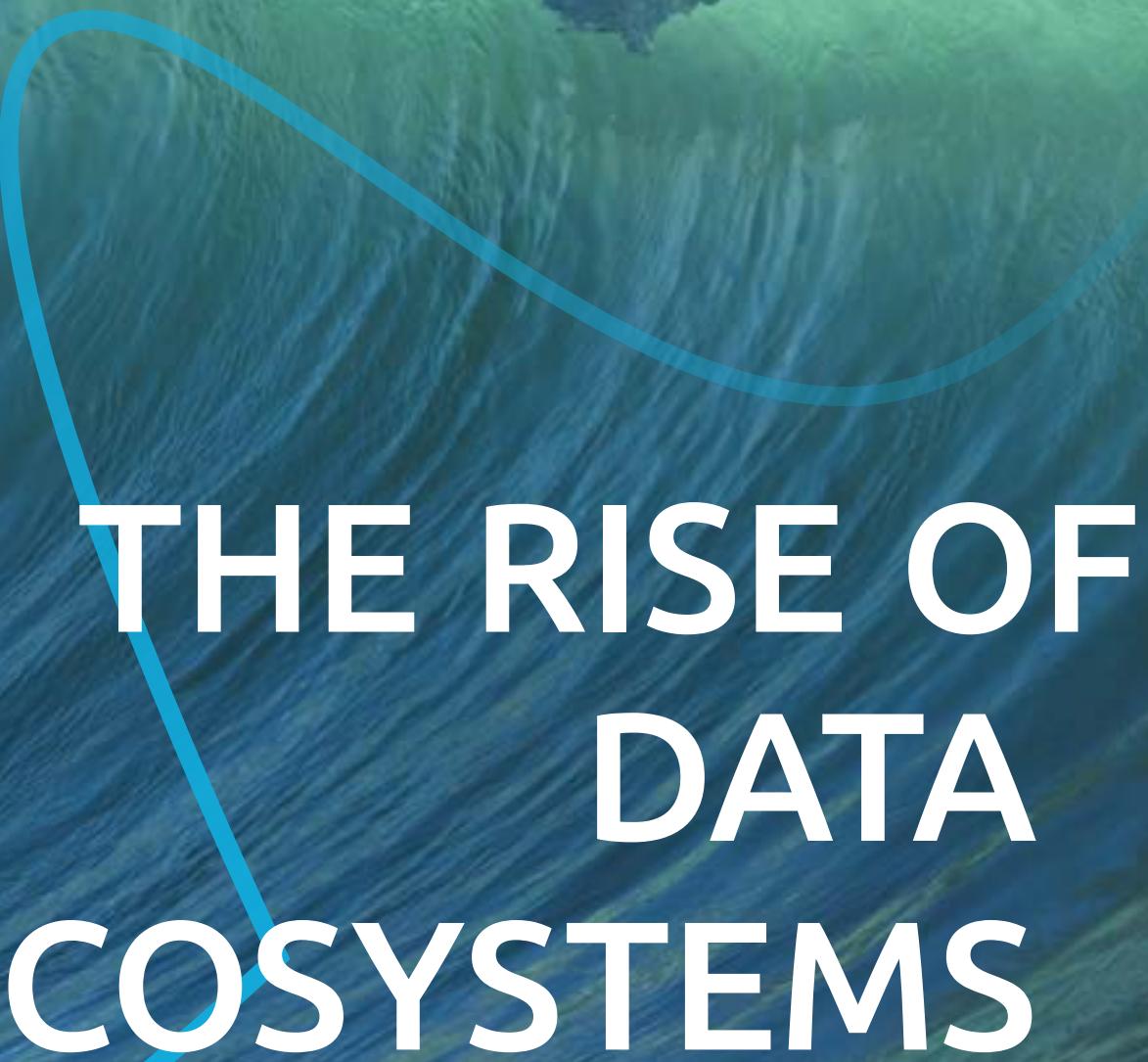
The 'everything-as-code' philosophy to manage operational complexity is a critical enabler for a successful Data Mesh deployment.

### No rip and replace

In large complex organizations, transitioning to a Data Mesh architecture will need consensus building, organizational re-alignment, and iterative migration of technical infrastructure.



# THE RISE OF DATA ECOSYSTEMS



Collaborative Data Ecosystems.....	31
Unleashing the Potential of Open Data.....	35
Data as a Product – New Value from Data.....	38
Federated Learning – Training AI, with Privacy.....	41



# COLLABORATIVE DATA ECOSYSTEMS

Monish Suri

DIRECTOR, GLOBAL INNOVATION LEAD & GOOGLE ALLIANCE LEAD, INSIGHTS & DATA, CAPGEMINI

Ron Tolido

EVP & CHIEF TECHNOLOGY AND INNOVATION OFFICER, INSIGHTS & DATA GLOBAL BUSINESS LINE, CAPGEMINI

## Collaborative data ecosystems

**Data ecosystems are emerging across industries, from financial services to automotive to aviation. Open data ecosystems that share data for societal good are also on the rise. These ecosystems help to create a sustained advantage for all involved. However, the data ecosystem participants need a clear win-win business model and must address key challenges of data ownership, regulatory compliance, ethics, trust, and privacy.**

### Defining data ecosystems

The Capgemini Research Institute (CRI) defines data ecosystem as a partnership between multiple

institutions to share and manage data to create new value that would not have been possible in the previous, siloed system." Data sharing is based on a mutual value exchange, thus making data more beneficial for all participants. Data coming from a variety of sources also enhances the diversity of information and increases inclusivity for all individuals, consumers, and citizens. This sharing happens in compliance with all local regulations and guidelines, and in an anonymous and aggregate manner (especially for personal data).

In a typical data ecosystem, multiple organizations, contributors, aggregators, and user organizations

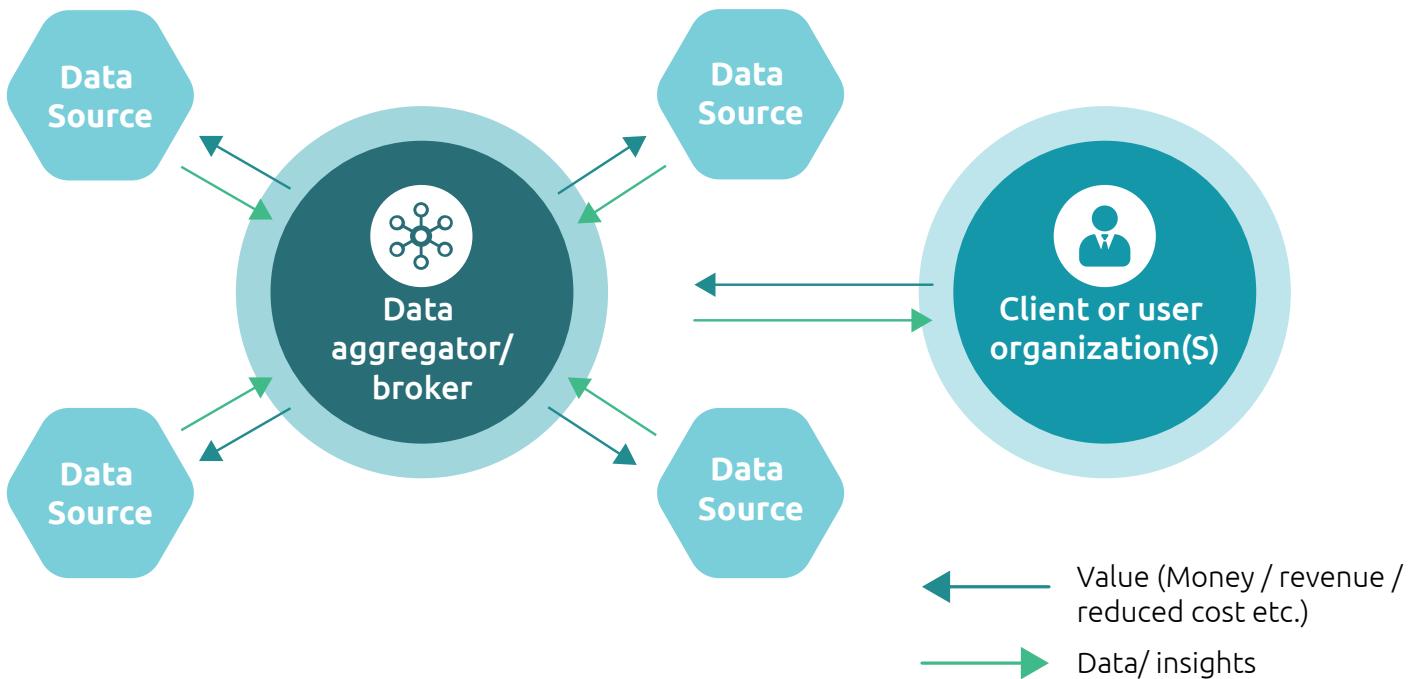
come together to facilitate data sharing and provide value.

### Data ecosystems are on the rise

Data ecosystems help in building economies of scale and a collaboration network. Many organizations with similar interests join ecosystems to pursue common goals.

According to recent work by the Capgemini Research Institute, 77 percent of organizations globally engage in some form of data ecosystem and two in three of these organizations (66 percent) see that there has been a renewed interest in

# Data brokerage and aggregation



Source: [Data Sharing Masters: How smart organizations use data ecosystems to gain an unbeatable competitive edge](#), Capgemini Research Institute, 7/2021

engaging with data ecosystems in the last one to three years. Forty-eight percent of organizations plan to launch new data-ecosystem initiatives in the near future, and 25 percent of organizations will invest upwards of \$50 million in data ecosystems in the next two to three years.

Data ecosystems offer benefits across operating models for positive business impact and societal good. According to the Capgemini Research Institute study, organizations have improved customer satisfaction by 15 percent, improved productivity/efficiency by 14 percent, and reduced costs by 11 percent annually in the last two to three years by engaging in data ecosystems.

## Types of data ecosystems

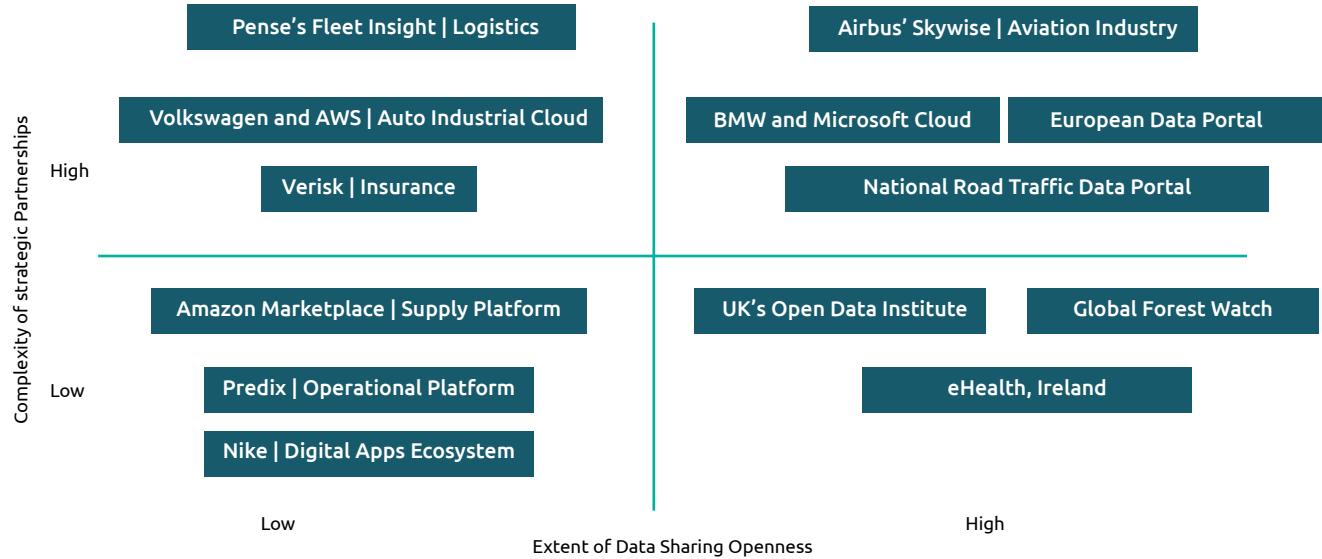
In this article, we focus on two types of data ecosystems:

1. **Industry data ecosystems** – We have seen an emergence of sector-specific data ecosystems across a range of industries, from financial services to auto to aviation.
2. **Open data ecosystems** – Typically, these are organizations that share data in the open for societal benefits.

There are other types of data ecosystems as well. 'The recent study [Data sharing masters: How smart organizations use data ecosystems to gain an unbeatable competitive edge](#) describes all of them in detail:

- Data brokerage and aggregation ecosystems
- Reciprocal data-sharing ecosystems
- Federated analytics ecosystems
- Collaborative data supply-chain ecosystems.

The matrix in the illustration maps examples of specialized industry and open data ecosystems on two key dimensions. The x-axis explores the extent of data-sharing openness. Typically, this translates to the degree to which data ecosystem participants share data openly for societal good. The y-axis explores the complexity of strategic partnerships. Typically, this translates to the number of participants involved in the data ecosystem.



Source: Internal Capgemini analysis and the Capgemini Research Institute

Here are some examples of specialized industry data ecosystems mapped in this way:

- Airbus' Skywise.** This is a leading open data platform for the aviation industry. Its ecosystem of partners includes over 9,000 connected aircrafts from more than 100 airlines from around the world, and over 18,000 unique users. Using data and insights from the Skywise program, a US flagship carrier saved \$13 million in fuel costs per year. Another airline reduced reliability reporting time from three weeks to one day.

- Volkswagen and AWS for Auto Industrial Cloud.** Volkswagen is working with AWS to move its 124 factory sites to a single architecture: the Volkswagen Industrial Cloud. The Industrial Cloud takes transformation beyond automation and consists of three parts: the Digital Production Platform, targeting all Group plants and built on AWS; a Volkswagen Group App Store, with use-case

applications available for all partners; and the company's machines and equipment on the shop floor.

- BMW and Microsoft Cloud.** Microsoft and the BMW Group launched the Open Manufacturing Platform. Companies are partners in a new initiative to drive open industrial IoT development and help grow a community to build future Industry 4.0 solutions.

And here are a few illustrative examples of open data ecosystems:

- European Data Portal.** The EU Data Portal ([data.europa.eu](http://data.europa.eu)) provides access to open data made available across Europe to organizations and the general public. There are over 760,000 datasets from 35 countries, available across all major sectors and 82 catalogues (e.g. portals of open data from various EU countries, government departments, and agencies), ready

for use via various open-data licenses.

- UK's Open Data Institute (ODI).** The mission of ODI is to work with companies and governments to build an open, trustworthy data ecosystem. The ODI's global network includes individuals, businesses, start-ups, franchises, collaborators, and governments who help to achieve the mission.
- Global Forest Watch.** GFW is an open platform that provides data aggregated from various sources to monitor forests. The data is sourced from more than 100 organizations across the world. The platform is used by local law enforcement officers to identify threats to forest covers, analyze tree-cover change driven by external factors, and conserve wildlife.
- National Road Traffic Data Portal (NDW), Netherlands.** The NDW is the national portal for traffic data in the Netherlands. NDW collects current traffic data and situation data (e.g., condition and availability

of a specific road) and offers this as open data and paid services to cities and businesses in Netherlands. One such implementation of this data set comes in collaboration with Utrecht. The city administration uses data from NDW along with other sources to provide up-to-date and accurate information to employees in traffic management.

### Key challenges

Clearly, data ecosystems help to create a sustained advantage for participants. However, some key challenges need to be addressed. First, the right business model needs to be established that contains a unique value proposition and is an obvious win-win for all participants. Second, especially when data is personal or in other ways sensitive, issues around data ownership, regulatory compliance, ethics, trust, and privacy need to be addressed.



#dataecosystems #analytics #ai #ecosystems #opendata

## INNOVATION TAKEAWAYS

### Ecosystem

Industry data ecosystems and open data ecosystems are on the rise, as prime examples of the next generation of value-creation through collaboration on data.

### Win-win

Find the right business model and define a unique value proposition that is a win-win for all data ecosystem participants.

### Trust and ethics at the heart

Address issues of data ownership, regulatory compliance, ethics, trust, and privacy – as a foundational quality of any data ecosystem.



# UNLEASHING THE POTENTIAL OF OPEN DATA

Esther Huyer

MANAGER, CAPGEMINI

Gianfranco Cecconi

DIRECTOR, CAPGEMINI

## Unleashing the Potential of Open Data for Government and Business

You benefit from it every day, whether in business or as a citizen, but you likely don't even know you're using it. Open data is data that anyone can freely re-use for any purpose offered by governments back to their citizens. Its potential economic impact is massive but still largely untapped. Further, the European Union commits to ensuring you can leverage its potential.

The European Union is a unique economic and political union between 27 countries. What began as an economic union in 1958 has evolved into an organization spanning policy areas, from the environment to health,

and from justice to migration. Its main economic engine is the single market that enables goods, services, money, and people to move freely. The EU aims to develop this principle in other areas like energy, knowledge, and data to ensure that Europeans can draw multilateral and optimized benefits.

The commitment to facilitate more and better sharing of data starts from "open data": Data that anyone has the right to re-use, free of confidential and personal information, most often created by government operations and released to their citizens and businesses at no cost.

### Innovating from open data

Open data is not innovative per se. Instead, it enables efficiency and innovation of products and services that re-use it. Capgemini Invent's "Economic Impact of Open Data" study has described the financial and non-financial positive impact of this resource, directly and indirectly extensively. It spans from the creation of new enterprises and jobs to people's time and lives saved, to an entire market depending on it, estimated between 199 (baseline scenario) and 334 (high growth scenario) billion EUR in Europe by 2025.

In 2003, with the first of a series of three “Public Sector Information Directives,” the EU made publishing open data a legal mandate. What came next was a series of “European Data Portal” projects – soon rebranded as “data.europa.eu.” Capgemini Invent in the Netherlands has since led the consortium that implements the EU’s open data strategy. The program was recently extended through 2027, continuing to offer a single, coherent core component of the public sector data infrastructure to the Union’s citizens.

The program offers three key pillars:

- Access to public data resources across Europe via a single point of contact is the portal itself, offering over 1 million datasets across 36 countries, 6 EU institutions, and 79 other EU bodies and agencies.
- Support for EU institutions and Member States by facilitating communities of practice around national open data portals and policy owners, training and consultancy. Why? Well, to improve, sustain, and document publishing practices.
- Evidence of socio-economic benefits concerning re-using public data resources and various stimuli fosters and showcases uptake and value creation.



**USD 224,17tn**

open data EU market size in 2019



**1.09m**

employees in the EU work for businesses that re-used open data in 2019



**1.12m-1.97m**

forecast for 2025

Between

**54-202 thousand**

lives saved in the EU by faster emergency response

**27m hours**

saved in public transport

**5.8m tons**

equivalent of pollutants saved by reducing household energy consumption



**USD 379,25k -  
USD 486.22k**

healthcare budget saved thanks to first aid by bystanders

**USD 16.65tn -  
USD 24.31tn**

labor cost saved by reducing time spent in traffic

How does it work? The three pillars support creating and improving processes, products, and services, as well as re-using public data resources to create economic, societal, political, and environmental impact. Of particular relevance are the project's yearly assessments of open data maturity in Europe and the research around socio-economic implications of open and shared data.

Today, maturity and expectations on data quality, interoperability, and impact creation have evolved. Europe is moving towards a common European "data space" where data re-use - whether in the open or in controlled environments - is enabled and governed for compliance, security, and viability. Therefore, it facilitates even more innovation, equality, and prosperity.

Similarly, the concentration on exploiting open data and data sharing, in general, is becoming the focus, and not just from governments to citizens, but between businesses, in ethical, technologically sound, and commercially viable ways. Capgemini Invent once again supports the EU along this new venture, leading the "Support Centre for Data Sharing": The precursor of more legislation and an endorsement that the Union will provide to European business over the next many years.

## INNOVATION TAKEAWAYS

### Sharing = business

No, data is not "the new oil." Oil is finite and can't be duplicated or shared. But data is replicable and shareable. Open data and data sharing are worth considering in the continuous transformation of your business, whether you're a consumer of it or a provider.

### Discover first

If you don't know what data you have, you can't use it. Before addressing key data issues such as quality or interoperability, the EU's focus has been on making data discoverable. Knowledge about what data is available – and stewarding it – is the critical step towards realizing value.

### Focus on impact

Value and data quality rely on how consumers can use it. It's not about how "good" it is in technical terms such as correctness, completeness, and update frequency. Undoubtedly, whether you work in the public or private sector, move beyond traditional metrics by which data products and services are evaluated and start thinking more along the lines of the impact data can create.



# DATA AS A PRODUCT – NEW VALUE FROM DATA

Chandrasekhar Balasubramanyam

VICE PRESIDENT  
HEAD OF DS&A AND AI COE, CAPGEMINI

More organizations aspire to monetize their data by leveraging their data or analytics as a product towards the inside and outside world. The notion of 'Data-as-a-Product' (DaaP) is also pivotal to the emerging concepts of Data Mesh architecture. Yet, there are multiple challenges in intellectual property, regulatory and compliance requirements, and data stewardship. In this article, the phenomenon of DaaP is further explored, including its importance, future business trends, and platform solutions.

Data can be used to address a business or technical challenge, either directly or through insights and algorithms. Data can be provided as-is to the end-user, or it can be

pre-processed and delivered as insights, algorithms, or intelligent systems. For instance, when delivered as a product, data can instantiate in any of the following forms:

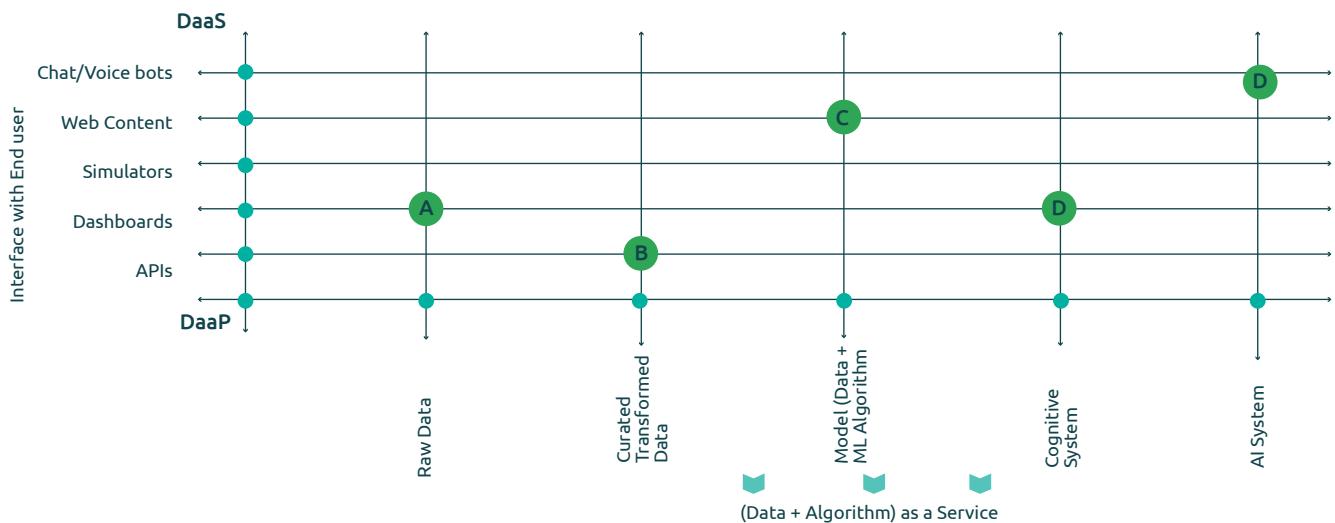
- Raw Data
- Curated or transformed Data
- Model (Data + algorithm)
- Cognitive systems for human intelligence augmentation
- AI Systems for human intelligence replacement

Following the Data-as-a-Product approach, data is provided to a consumer of the product, who will use the product to address an identified challenge, issue or opportunity. Models, cognitive systems, and AI systems are built on (training) data

and hence are considered data products.

## Algorithms versus Data as a Product

When algorithms are used to process the data, patterns emerge from machine learning. From this, it is possible to understand data, root causes, or possible solutions. So, algorithms act on data to bring out value. Like Google's search engine or Uber's matching engine, some algorithms are considered 'Algorithms as a Product.' Data-as-a-Product processes and uses the data to solve a problem with the help of patterns identified from the data. If the product processes data for search and match, it is not considered Data-as-a-Product.



## Types of Data products

In Fig. 1, types of data products are shown along with various interfaces available for a consumer of data products. The X-Axis represents five types of data products: raw data, curated and transformed data, model, a cognitive system, and AI system.. The Y-Axis shows five types of interfaces: APIs, dashboards, simulators, web content and bots.

It's possible to make any data product available through any of these interfaces. Hence there can be twenty-five combinations of data products and interfaces. In the figure, five example combinations are indicated with markers A to E.

To clarify: Marker A indicates raw data is provided to the end user in the form of a dashboard. Insights must be drawn and understood by the end user himself/herself. Marker B indicates the curated data is accessed through an API. Marker C represents a web portal depicting the outcome of a model, with the consumer picking the content relevant to the requirement at hand. Marker D shows the cognitive system is accessed through a simulator.

When a consumer wants to understand a system response, the simulator runs with the cognitive system as an engine in the background with input from different combinations of parameters. This type of DaaP delivery is, for example, helpful for medical practitioners that need to recommend a prescription for specific symptoms. Correspondingly, marker E shows a combination of an AI system and a chatbot. The chatbot can take inputs from the user to settle an insurance claim, analyze the data, and approve or reject the claim..

## Trends

The notion of Data as a Product is becoming more popular, with more organizations aspiring to monetize their data. This can be achieved through selling the data directly, after curating it, or after transforming it into a model or system. When data is sold in the market in the form of a model, the actual data is not going to be with the consuming party. Hence there are no challenges of sharing proprietary information or meeting regulatory and compliance aspects. Therefore, more organizations may prefer to share content or patterns in the data in the form of a model or

system rather than sharing data directly.

In addition, data can also be shared in synthetic, transformed, or masked data to protect intellectual property and avoid issues around data privacy and ethics.. These trends will continue in the future and will grow to a larger scale.

New business models will be based on these specific approaches to sharing (synthetic) data and algorithms. Cloud computing, 'DataOps,' and 'MLOps' are all helping to industrialize and automate the process of continuously developing and deploying data and models into production, making them quickly and reliably available to the consumers of the business' Data-as-a-Product - through any of the identified interfaces.

## 890

[Capgemini's '890'](#) is an example of a platform that can provide DaaP in any form to external or internal consumers (it is good to note that anything that works in sharing DaaP between a provider and an external consumer will work just as well within the same organization). You can

leverage 890 by Capgemini to provide access to data through APIs or as models and systems, all supported by industrialized DataOps, MLOps, and API management.

The platform is easy to configure, designed to kickstart a new DaaP initiative. The 890 Platform can provide DaaP for any combination of identified interface and data products. Many other engaging, innovative data platforms and marketplaces are available. With the growing popularity of Data as a Product – also driven by

advances in 'Data Mesh' architectures, they will thrive even more in the future.



#datapowered #ai4good #datamonitization

## INNOVATION TAKEAWAYS

- The notion of Data as a Product is pivotal to the concept of Data Mesh architecture and is a requirement for successful data monetization.
- There are five different ways to provide data as a provider and five varying ways to consume them from the demand side.
- To avoid issues with data privacy, regulations, and even ethics, delivering data as synthetic or transformed data – or as an algorithm – is preferred.
- Data platforms and marketplaces, such as Capgemini's 890, are the necessary foundation for being successful with Data-as-a-Product.
- What works in providing Data-as-a-Product to external consumers may work just as well for supply and demand within an organization.



# FEDERATED LEARNING – TRAINING AI, WITH PRIVACY

Padmashree Shagrithaya

VICE PRESIDENT, INSIGHTS & DATA,  
CAPGEMINI

Anupam Srivastava

DIRECTOR, INSIGHTS & DATA,  
CAPGEMINI

## AI models are hungry for data

**It is what they are trained with – and the more, the better. With rising awareness around data privacy & compliance, getting access to that data is not exactly a walk in the park. Combine it with the need for ‘data democratization,’ making data readily available across the business. The challenges in this field become apparent – even more with training data that needs to be shared between companies. The AI Death Valley is where AI models cannot be further trained and optimized without compromising the rights of data producers and data owners. Federated learning is a new, critical approach to deal with this conundrum.**

Data about individuals is often needed for training AI systems that aim to deliver highly personalized services. But ‘Personally Identifiable Information (PII) is under the scrutiny of data privacy and security regulations and needs to be carefully managed and safeguarded. This does not only pertain to the consumer market, but it is just as relevant to, for example, at law firms or hospitals.

Further, much of the data needed is collected on the spot, at the ‘mining site,’ and it may be expensive to transfer it to a central data storage. Thus, the future of AI algorithms lies in training data shared between different companies, with apparent restrictions due to confidentiality and

competitiveness. Solutions are needed for AI models that can be trained on relevant data - siloed across boundaries - without outward data movement while solving data privacy challenges.

### Federated Learning

Federated Learning (FL) promises to deliver exactly that. Although still in its nascent stage, it is a promising approach to overcome mentioned obstacles. Federated Learning was coined in 2016 by Hugh Brendan McMahan and is defined as a “Machine Learning (ML) technique where multiple entities (clients) collaborate in solving an ML problem under the coordination of a central server or

service provider. Each client's raw data is stored locally and not exchanged or transferred; instead, focused immediate aggregation is used to achieve learning objectives."

In other words, a baseline model is shared by a server with decentralized heterogeneous devices (like mobile phones or other local devices) to train locally and aggregate only weights. The process is repeated until the desired outcome is achieved. The data continues to reside in the respective devices. But the algorithm is trained and enriched through collaboration between devices and the server.

Google has been one of the pioneers in using federated learning for predicting the next word typed by an Android user in "Gboard for Android." The prediction model is trained during the idle time of the device. Similarly, Apple uses FL for locally training Siri on the iPhone to distinguish the specific user's voice while improving the global model with more general characteristics. Today, there are many more leading players investing in this field, like NVIDIA, IBM, and Intel.

## Applications of FL

In the Life Sciences and drug research space, consortia of organizations invest in joint projects to enhance predictive outcomes. The 'Machine learning ledger orchestration for drug discovery' (or MELLODDY) project is one example in which 10 pharmaceutical companies have come together to enhance predictive machine learning models on decentralized data without exposing their proprietary information to others.

In Financial Services, the 'Open Banking' space - which is an initiative to secure access to financial information to consumers and 3rd party providers - has seen an uptick in Federated learning applications. This is, for example, to predict loan risk or detect anomalous transactions without taking control of the data.

In Manufacturing (or 'Industry 4.0'), there is widespread adoption of machine learning to improve the efficiency and effectiveness of industrial processes while guaranteeing a high level of safety.

Considering the privacy of sensitive data for industries and manufacturers is of paramount importance, FL can be applied to these problems - as they do not disclose any sensitive (sensor) data and still leverage industry-standard robust models. This enhances business process efficiency and failure prediction across the entire industry, with benefits for all involved.

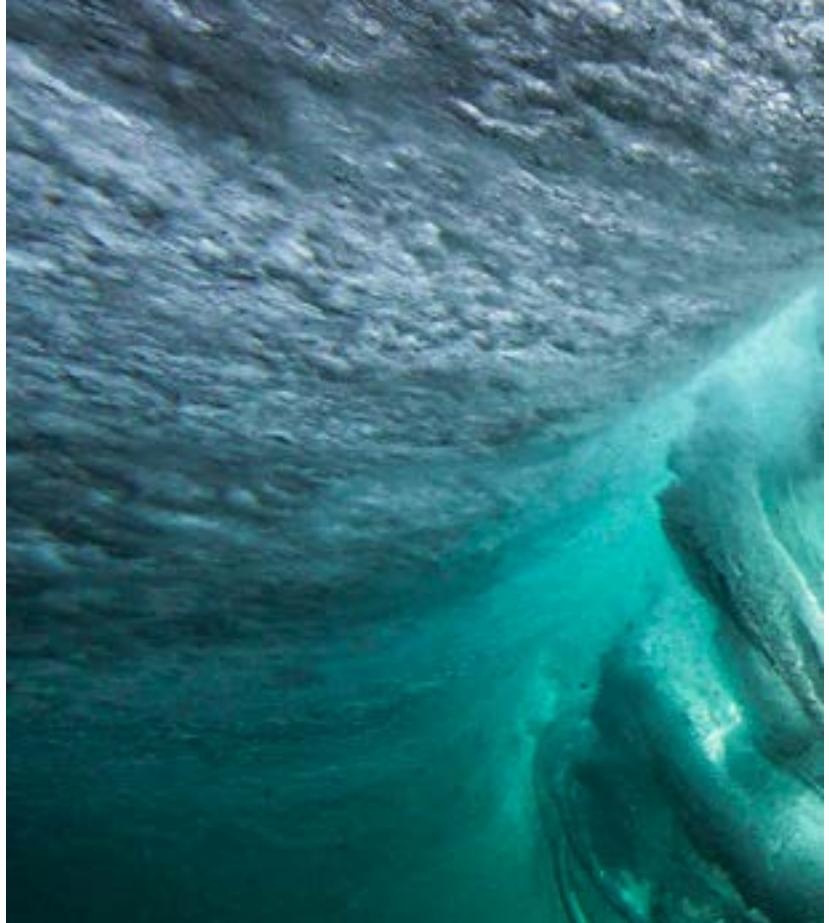
During the pandemic, companies such as Tessella (now part of Capgemini Engineering) and Intel collaborated to create a model to predict lung infection. This was done through analyzing X-ray images, involving various medical institutes and hospitals in Spain and the UK, training models with FL across the parties involved, thoroughly ensuring data privacy.

FL frameworks are evolving in the Open Source community, such as Flower, Pysyft, and HETEROFL. As always, Open Source accelerates innovation in this space.

## Challenges

Indeed, Federated Learning is still in its early days. It has its challenges which – we are optimistic – will be addressed and advanced soon, more in a matter of months than in years. FL cannot yet handle heterogeneous data in its current avatar: Getting the global model to learn from local data is stifled when the data is not homogenous.

Also, the model aggregation is relatively simple, using methods based on averages – this has room for improvement. Then, there are open questions around security. For example: how is FL applied to access private datasets without revealing specific data traits? Concepts such as Secure-Multi Party Computation (SMPC) and Differential Privacy (DP) are explored and already used to fully mitigate the risks of individual data statistics being revealed.



#federatedlearning #ai4ood #datapoweredinnovation #dataalliance #collaborativelearning

## INNOVATION TAKEAWAYS

### Breaking data barriers

With Federated Learning (FL), data needed to train AI from multiple providers can be shared and accessed while ensuring data privacy.

### Industry benefits

By working together – even when being competitors – industry players can jointly build more robust and industry-standard AI models to the benefit of all involved

### Allying on data

Companies can leverage FL by creating data alliances and consortia in sectors as diverse as pharma, manufacturing, consumer goods, and retail.

### Collaborative learning

FL opens up a new era in Artificial Intelligence, further improving results through trusted data collaboration.



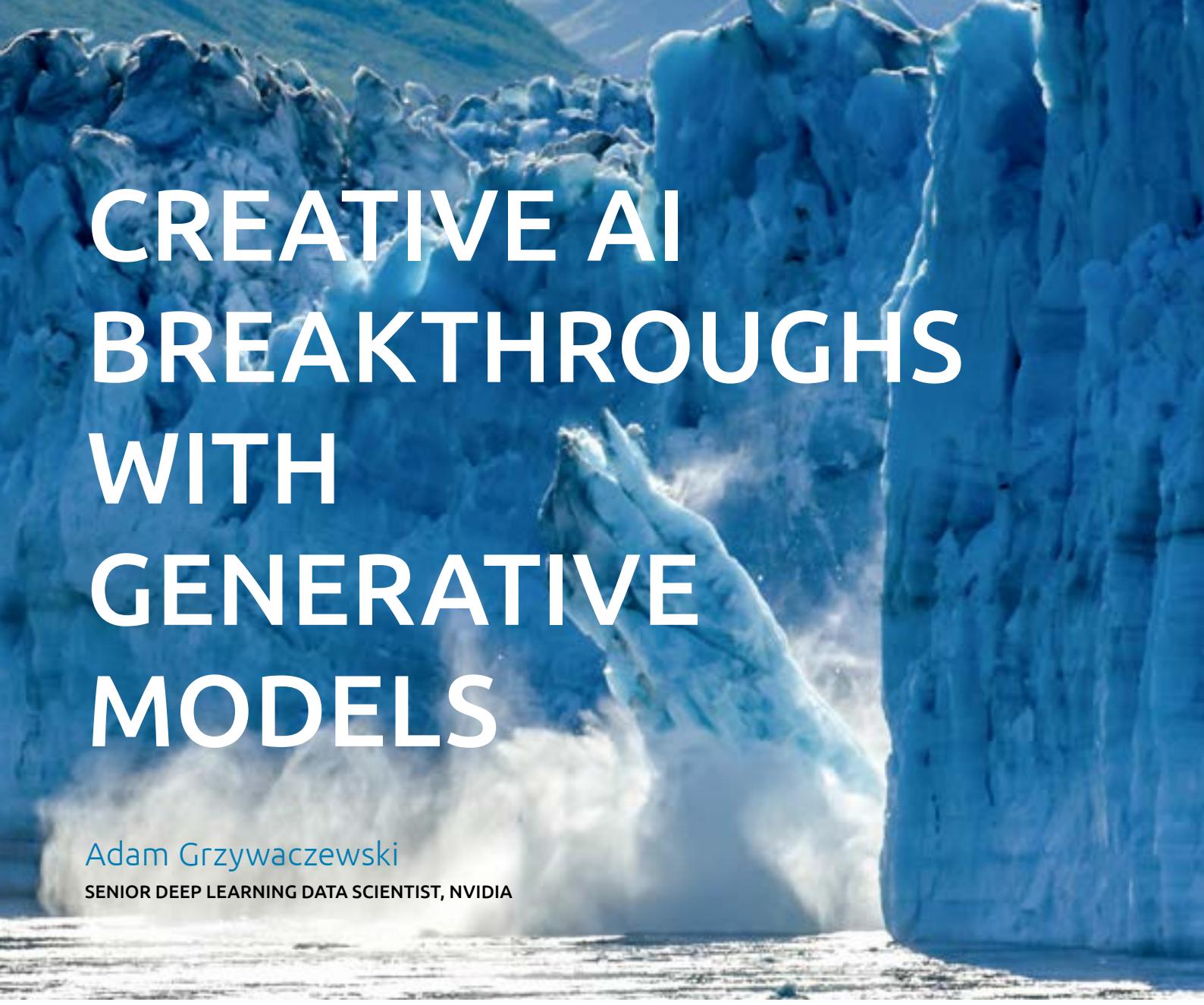
# THE RISE OF CREATIVE AI

Creative AI:  
Breakthroughs with  
Generative Models.....45

Generative Language Models  
and the future of AI..... 49

Creating intelligent, engaging  
customer experiences  
with AI.....53

Creative AI:  
From Drug Discovery to  
Generating Ad campaigns.....56



# CREATIVE AI BREAKTHROUGHS WITH GENERATIVE MODELS

Adam Grzywaczewski

SENIOR DEEP LEARNING DATA SCIENTIST, NVIDIA

Generative AI models, including Generative Adversarial Networks, are changing how we think about creativity and how we create and interact with content (2D/3D images, text, music, speech)

In this article, we will discuss how generative models are made and how they can be applied to solve artistic and design problems but also help to address critical challenges in science and engineering.

Can a neural network be genuinely creative? The first thought that comes to the mind of an AI practitioner is "no." Indeed, the only thing that a neural network can do is to approximate complex functions. You do the math: Given a large enough labeled dataset (let's call it 'x') together with the labels ('y'), the only thing that a neural network can do is to approximate a function mapping data to the labels ( $f(x)=y$ ).

Here's the bottom line...

Regardless of how complex the function approximated by the neural network is (whether it's approximating translation or object detection problems), one would never assign human attributes of intelligence and creativity to it. The neural network does not understand the problem it is solving. It can take the input data and transform it into output using the function approximated during training.

## Generative models

Can we design neural networks to create an illusion of creativity? Can we build a model that given a vector of random numbers could generate new, previously unseen data points that are similar to the points in our training dataset? The answer is yes. In fact, there exists an entire area of research focusing on development of such 'Generative Models' which, given the training data, generate entirely new samples from the same data distribution.

Does that constitute creativity? To some extent it does, since generative models can generate truly new data points that were previously not seen in the training dataset (e.g. when using images of celebrities AI systems generated convincingly looking bearded women – without obviously, any examples in the training dataset).

## Rapid progress

Whether one calls that creativity or not, it is irrefutable that tremendous progress has been made in building generative models. Even though specific models such as 'Autoencoder' were already described in the '80s, the ability to generate high fidelity content is relatively recent.

For example, in the area of AI vision before 2017, it seemed impossible to create high-quality, high-resolution images robustly. Today, the content generated - whether visual, textual, or audio - is almost indistinguishable from the actual samples. It triggered an entirely new AI discipline that focuses on detecting such 'deep fake' content (e.g. DARPA's Semafor program).

Also, the AI field has learned how to control the behavior of a generative

Source A: gender, age, hair length, glasses, pose



Source B: everything else



Result of combining A and B



The image attributes are extracted from images at the top (source A) and combined with the image on the left (source B) to generate new faces combining both features.



model better. For instance, when generating faces, it is possible to control physical properties such as gender, age, and pose. The Figure 1 / Image on the previous page demonstrates the level of control the user has over the generative models. Furthermore it is now much easier to train generative models such as 'StyleGan2', which can be reapplied to countless other creative challenges.

#### Many application areas

They are being adapted to generate full-body images (e.g., fashion model images wearing custom outfits or 'AI video compression') or customized products. They are even applied to the development of self-driving cars, generating 'creative,' unexpected driving scenarios. Case in point: In medical imaging, they generate histopathologic images or CT/MRI scans to handle even the rarest cases. In physics, given an experiment setup as input, they create images resulting from high-energy particle collisions.

#### Here's how it works...

Generative models can plug the holes in a damaged picture (Image Inpainting), generate frames when increasing the video or game frame rate (DLSS), or increase the resolution of TV content via 'AI upscaling.' They help end-users understand microscopy/astronomy images (EM PSSR model) and are the backbone of many scientific applications (Fluid Flow Super-Resolution).

Generative models can also create 3D models and textures from 2D images (GANVerse3D), supporting the development of 3D animation.

For example, Omniverse Audio2Face can manipulate a 3D model of a head, given the recording of speech. Autodesk Generative Design can automatically create 3D engineering

assets meeting physical and budget constraints.

Moreover, generative models can create abstracts from complex articles. They are being adopted for the generation of software code (e.g., designing a user interface or entire games like GameGAN Packman), internet network traffic (to test the quality of network protection mechanisms: SDN-GAN), chemical compounds (AstraZeneca's MolBART). They can also create full news articles (OpenAI's GPT-3). In addition, they are capable of speech synthesis (HiFiGan), audio background noise removal (RTX Voice), or can compose and play music (GTC Intro music). They can even mix text and images to create a new generation of user interfaces, search, or design tools. The image on the next page illustrates how a generative model can create visual content by being provided only the image description.

#### More to come

Undoubtedly, generative models will continue to transform the way we do science, engineering, art, and much more. More extensive models trained on larger datasets will continue to improve in terms of their ability to generate credible, high-fidelity creative content. The image on the next page illustrates how a generative model can create visual content by being provided only the image description"

#### And the best part is...

The emergence of (energy-optimized) tools supporting the development of even larger models - such as Microsoft's DeepSpeed and Facebook's FairScale - combined with robust AI hardware solutions - such as DGX SuperPOD to pave the way towards a bright, creative future.

**TEXT PROMPT** an armchair in the shape of an avocado, an armchair imitating an avocado

**AI-GENERATED IMAGES**



Mixing modalities: generating images matching the text query ([OpenAI Dall-E](#))

#nvidia #nvidiaresearch #deeplearning #artificialintelligence #gans

## INNOVATION TAKEAWAYS

### Significant progress

Within just a few years, AI tools now can generate high-fidelity, creative content – increasingly indistinguishable from original sample content.

### Not only creative arts

The applicability of generative models is not limited to the creative arts at all, as it already solves critical challenges in medicine, engineering, and science, as well.

### Platform power

The next generation of AI hardware platforms (such as those pioneered by NVIDIA) makes generative power available to more businesses at a lower cost and requires less energy consumption.

### The creative future is bright

There is a clear path to larger, more spectacular, innovative capabilities – driving the need for many more people to master and use generative tools and platforms.

# GENERATIVE LANGUAGE MODELS AND THE FUTURE OF AI

Rajeswaran Viswanathan

HEAD OF AI COE, CAPGEMINI

Very large language models like GPT-3 have changed the game for solving Natural Language Processing (NLP) tasks such as summarization, text generation, and even program code creation

**There is no longer model training required for these tasks – but rather ‘prompt engineering’ – articulating the right intention through text to get the best output from the model. This introduces a new concept and approach, a new AI skill that will emerge as language models continue to grow bigger and more powerful.**

From building custom architectures using neural networks to using ‘transformers’, NLP has come a long way in just a few years. To clarify: A

language model is a probability distribution over sequences of words. Language models using neural networks were first proposed in 2001. Since then, meta-learning, word2vec (word embedding), sequence-to-sequence models, and pre-trained language models have made significant progress in the field – enabling tasks such as text summarization, translation, and generation. However, large language models (such as T5 - 11B parameters, Turing-NLG – 17B parameters, and GPT-3 – 175B parameters) are complex

and require expensive, high-end compute resources to train.

## Prompting

With the advent of deep neural networks (deep learning) and hardware improvements (GPU, TPU, and others), you can train more prominent networks. The figure (source: OpenAI) shows the spectacular size evolution.

# Generative language models and the future of AI

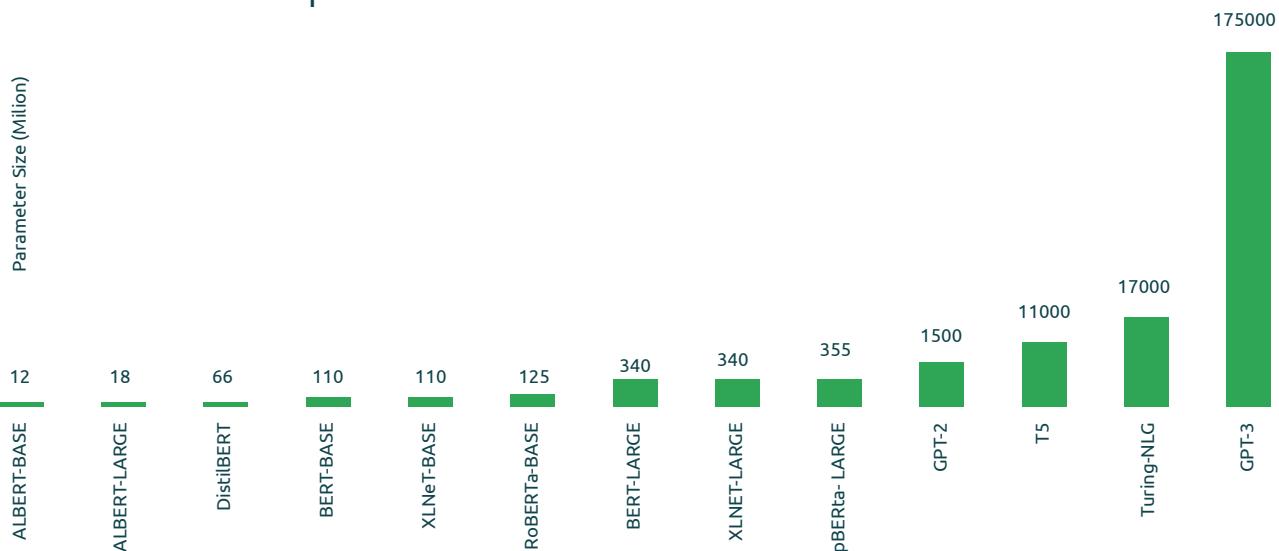
Very large language models like GPT-3 have changed the game for solving Natural Language Processing (NLP) tasks such as summarization, text

generation, and even program code creation. There is no longer model training required for these tasks – but rather ‘prompt engineering’ – articulating the right intention through text to get the best output from the model. This introduces a new concept and approach, a new AI skill that will emerge as language models

continue to grow bigger and more powerful.

From building custom architectures using neural networks to using ‘transformers’, NLP has come a long way in just a few years. To clarify: A language model is a probability distribution over sequences of words. Language models using neural

## Comparision: NLP Pre-Trained Models



networks were first proposed in 2001. Since then, meta-learning, word2vec (word embedding), sequence-to-sequence models, and pre-trained language models have made significant progress in the field – enabling tasks such as text summarization, translation, and generation. However, large language models (such as T5 - 11B parameters, Turing-NLG - 17B parameters, and GPT-3 - 175B parameters) are complex and require expensive, high-end compute resources to train.

### Prompting

With the advent of deep neural networks (deep learning) and hardware improvements (GPU, TPU, and others), you can train more prominent networks. The figure (source: OpenAI) shows the spectacular size evolution.

Concerning the training phase, the only effort is to engineer the prompts in such a way as to maximize the quality of output and tune the API parameters. In traditional machine learning, we spend time identifying the exemplary architecture, selecting

the library (such as Tensorflow or Pytorch), writing code to train the model, evaluating the model performance - iterating this process until we are happy with the model performance.

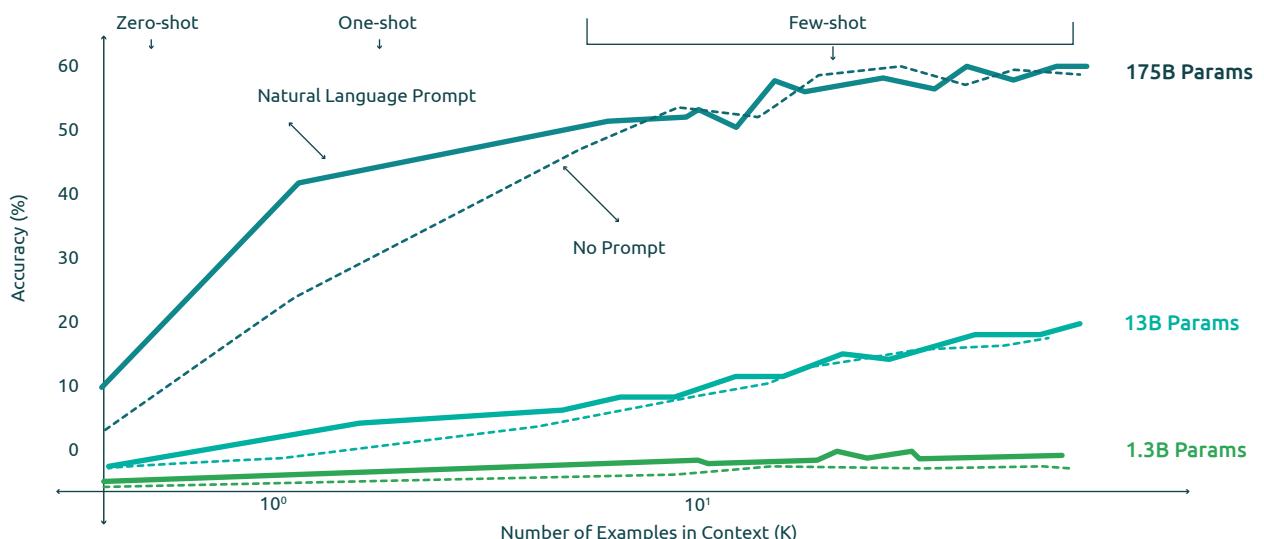
Simple prompt engineering and API parameter tuning replace the entire cycle. Also, in the classical deployment phase, we have to take care of all sorts of infrastructural considerations (e.g., container use). This is all replaced by just content-filtering the output of GPT-3.

## No semantics

GPT-3 API identifies the best pattern that fits our intentions. It applies 'fuzzy matching' to an enormous

model in the abstract sense. It will find the closest match, and specific parameters can control its ways. It does not have any semantic understanding of language, though.

For example, when asked, "I have a heart attack. What medicine should I take?" GPT-3 has been reported to confidently respond, "Aspirin." Hence, content engineers need to carefully



filter GPT-3's responses before taking them to production to mitigate any obvious risks of reputation loss or even legal issues. GPT-3 by itself does not have an "I don't know" answer.

It just ingests the prompt and provides the best pattern match. Further, it requires human oversight and engineering to create the best results. Once deployed, however, there is no concept of monitoring since the model is not trained anymore and its performance is constant.

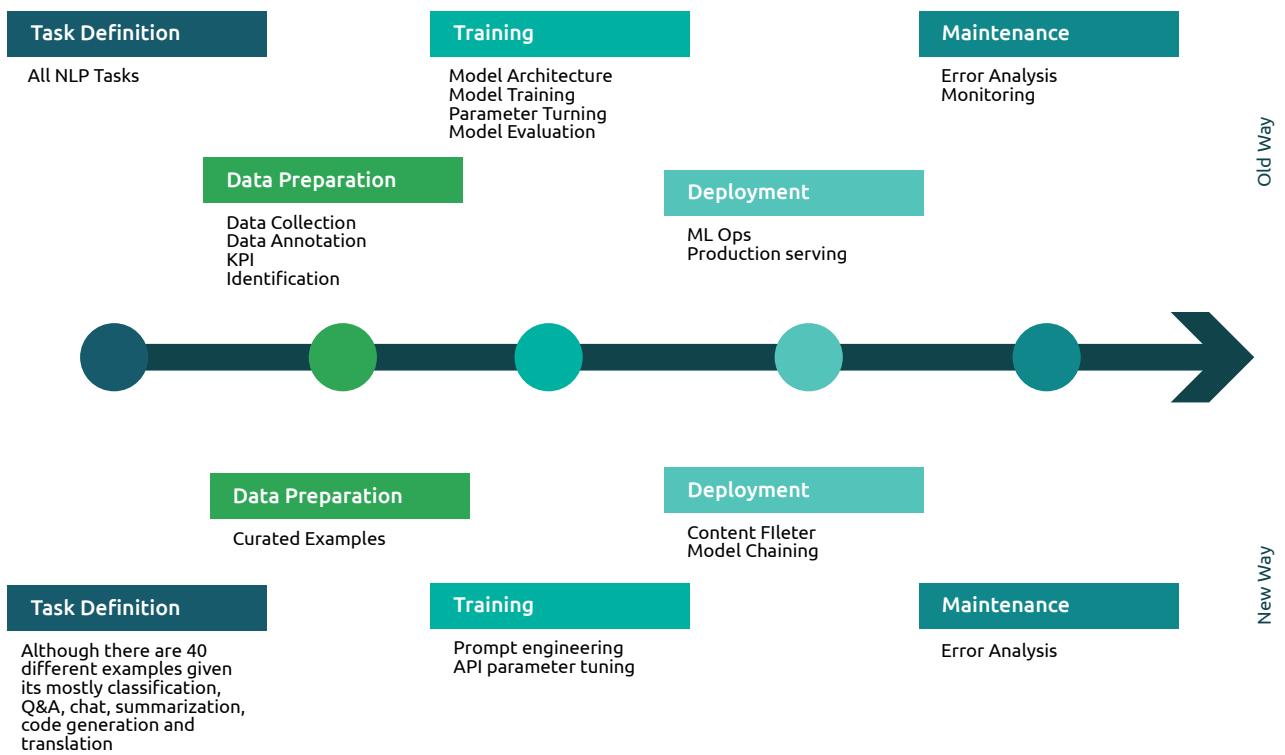
## Best practices

Some best practices for getting results with language models such as GPT-3:

- Look for initial use cases which are not yet customer-facing since developing appropriate content filters will be more complex and take more time.
- For customer-facing applications, keep the text internal and carefully controlled. The GPT-3 API provides probabilities for the words, giving insight into the content analysis

and weightage insights for those developing and overseeing the content.

- Create a multi-disciplinary team of "prompt engineers" that understand the areas of bias (on gender, race, etc.), fairness, legal, and various other ethical considerations – combined, of course, with excellent language skills.
- Use automated tools for content filtering: technologies to ensure responsible and fair AI are increasingly available.



#datapowered #ai4good #datamasters #gpt3 #languagemodels

## INNOVATION TAKEAWAYS

- Large language models such as GPT-3 provide state-of-the-art capabilities for a variety of Natural Language Processing tasks.
- Most companies will need to access these models through pre-defined APIs rather than developing and training these models themselves – a stark difference with more ‘classic’ machine learning.
- Carefully expressing the correct textual input for the API (a ‘prompt’) is key to getting the correct text output – and thus value – from the language model.
- The new AI skill of ‘prompt engineering’ is not so much technical skill – it requires deep language and subject matter understanding, plus a solid mastering of ethical considerations and model understanding.

# CREATING INTELLIGENT, ENGAGING CUSTOMER EXPERIENCES WITH AI

Punit Santani

SENIOR DIRECTOR, AI & ANALYTICS,  
CAPGEMINI

Pranav Kumar

DIRECTOR, AI & ANALYTICS,  
CAPGEMINI

Scott Prevost

VP, ENGINEERING, ADOBE  
SENSEI & SEARCH

In today's digital-first world, consumers expect more personalized experiences

Yet, some businesses have been painfully slow in delivering these experiences, impeding or even breaking the creation of deeper customer engagement. However, with augmented AI mature creative processes, we can overcome these hindrances, paving the way towards intelligent and engaging customer experiences.

With the explosion of content, formats to deliver media are growing increasingly more sophisticated to improve the likelihood of. Today's digital experiences require the ability

to stitch together content and data. There are an ever-increasing list of channels. In addition, consumers expect seamless content delivery across these channels via varying formats and sizes, thereby requiring automation and enhancement of content velocity. This calls for higher efficiency and effectiveness across content creation, approvals, management, and delivery throughout the entire content lifecycle.

By using AI and mature creative processes, we can overcome these obstacles on the path towards

creating engaging customer experiences. Leveraging Artificial Intelligence (AI) and Machine Learning (ML), we can train the algorithms to understand the content and associated text, generate insights from the content, and provide recommendations, boosting and augmenting the productivity of the creative teams.

## Content Velocity Through Creative AI

Through deep learning models, Digital Asset Management (DAM) tools can understand and decipher images uploaded to the content repository and assign relevant metadata (tags) – making it much easier to catalog and search. With metadata assigned, the integrated marketing automation solution and content repository can produce relevant image recommendations based on email texts or mobile campaigns in development.

When delivering personalized experiences at scale, there is a need to understand the personas, segments, behaviors, and expectations of consumers. The next step is to display relevant content based on where they are in their customer journeys. The activity of matching the content pertinent to the correct customer profile becomes a herculean task when dealing with millions of customers. However, with AI and ML, it's feasible to automate tasks such as personalization tactics, A/B testing, and multi-variate testing.

Enterprises have accumulated vast content repositories, often existing in silos. Retrieving and preparing the right information quickly for further processing comes with challenges such as:

- The variety of both structured and unstructured content
- The lack of systems to index content intelligently and understand the meaning of critical keywords for industries or domains.
- The dependency on people to manage and edit documents, FAQs, and more.

New advancements in AI are designed for automating the most mundane, repetitive tasks involved in the content lifecycle. And it's not just

about accomplishing those tasks but doing it at a breakneck pace, managing content volumes without compromising quality. It also frees up room to inject more creative components into the content lifecycle.

Several AI/ML capabilities that enhance content velocity, include:

- **Smart Tags:** Automatically analyzes a set of images in the DAM to identify crucial elements and subject matter to allow specific photos to be found quickly.
- **Smart Crop:** Detects and crops to the focal point in any image, capturing the intended point-of-interest across multiple screen sizes.
- **Automatic Text Summarizations:** Creates shorter versions of a broader text for various marketing channels to speed up processing.

## Examples

### Capgemini KIS

Using AI/ML capabilities and natural language processing, Capgemini's Knowledge Insights Services (KIS) provides an enterprise search engine that helps users get accurate and personalized results to their search

query quickly and correctly. Essential features include:

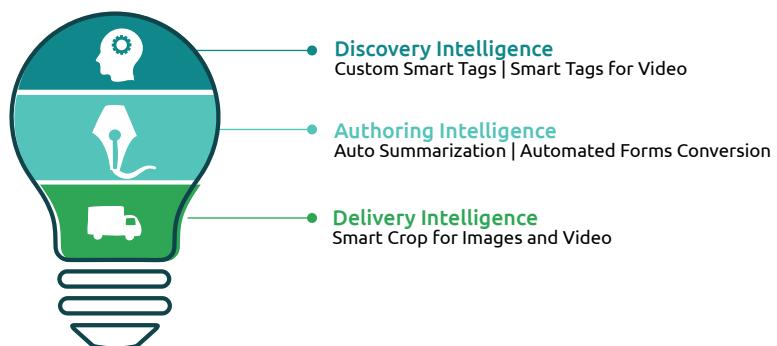
- Search on various file types (.docx, .ppt, .txt, .pdf, jpg, .mp3\*, .mp4\*) as well as tickets.
- Provision to ingest Q&A data and search.
- Intuitive dashboards offer various insights.
- Custom NERs targeting specific industries.
- Custom training on client data.

### Adobe Sensei

Adobe Sensei is Adobe's AI/ML-driven technology platform that powers the creation and delivery of digital experiences by blending the art of human creativity with the science of data.

In addition, it helps creators deliver more intelligent, more efficient experiences at scale while handling mundane and time-consuming tasks to free them to focus on what matters most. Adobe Sensei powers hundreds of capabilities across the company's product portfolio.

Also, Adobe Experience Cloud leverages Adobe Sensei to expedite the time between marketing ideation



Augment human creativity with AI to speed content supply chain

and execution. Adobe Sensei seamlessly connects to all of Adobe's cloud services, helping to deliver better digital experiences to customers.

#### KIS + Adobe Sensei

Integrating KIS with Adobe Sensei creates a solution that enhances and reduces the turnaround time for the creative life cycle of content management. The diagram shows the integration and solution flow.

#### Unlocking the magic of AI

AI is transforming creativity and business — and Adobe Sensei enables the interpretation and categorization of visual assets in real-time. This extensive categorization extends human capabilities, picking up where they left off and catching things they may have missed.

Moreover, incorporating AI into animation has the potential to accelerate workflows that are currently tedious and slow. AI can determine the video content that generates the best response and deliver more on top of it. UX and UI designers are increasingly aware of the

potential of AI to improve their overall productivity and capabilities.

Adobe Sensei integrated with Capgemini's text and image algorithms reduce time to deploy campaigns, deliver greater personalization at a large scale, garner insights from consumers, and enable expedited fail-fast testing.

AI will augment human creativity – not replace it – by driving more efficiency and effectiveness for creatives and marketers. By intelligently automating traditionally time-consuming and mundane tasks, they free up time to focus much more on understanding their customers and create experiences that matter the most.



#ai #aiethics #ethicalai #sustainabledevelopment #ai4good #datapowered

## INNOVATION TAKEAWAYS

### Massive You Experience

Today's consumers need highly relevant, highly personalized experiences; the scale of this brings challenges to the current content lifecycle and platforms, which have not been designed for this.

### AI and ML come to the rescue

AI and ML can significantly increase the productivity and effectiveness of the content lifecycle by intelligently automating repetitive and cognitive tasks, benefiting from breakthroughs in audio/video/image recognition, context-sensitive search, and natural language understanding.

### You Ain't Seen Nothing Yet

Already at this stage, AI augment humans, freeing up space for more creative, high-value activities; and with AI systems getting more potent by the day, there seems no end to what man and machine jointly can create.

# CREATIVE AI: FROM DRUG DISCOVERY TO GENERATING AD CAMPAIGNS

Menno van Doorn

RESEARCH DIRECTOR, SOGETILABS, CAPGEMINI

Breakthroughs in AI technology introduce a novel breed of possibilities

We can start to dream of creativity coming from a machine. From making music and paintings to designing homes and inventing new medicine, generative AI systems can do it all.

It's getting increasingly difficult to invent a new drug. But what if AI comes up with good, creative suggestions? What if AI could create new drugs? We're talking about the announcement of the biotechnology company Insilico Medicine in the *Nature Biotechnology* publication on 2 September 2019.

The answer?

Thanks to generative AI, this company has succeeded in designing a molecule that is medicine for the prevention of fibrosis and several related conditions. They did not present the announcement without a sense of drama: "It's the AlphaGo moment of

the pharmaceutical industry." What's underneath all of this? Generative AI.

## What is generative AI?

The critical acronym is GAN: Generative Adversarial Network. The magic word is "generative." To put it simply, AI can generate something that can be a solution to the problem. Generate a molecule, a medicine, a piece of music, a painting, an ad campaign, the design of a house, or a human face that looks like an actual human. Keep reading for a detailed explanation.

## 1. Generative:

The AI model generates new – synthetic – data based on a given set of training data. For example, if the training set consists of molecules, the AI model generates synthetic molecules. The goal is that the synthetic data has the same properties and correlations as the training data and therefore forms an excellent artificial representation of reality.

## 2. Adversarial:

This refers to the aspect of competition between the two algorithmic models: The creator and the critic. The creator creates, and the critic decides whether it should pass the test as a possible solution.

## 3. Networks:

Depending on GAN complexity, it is a relatively simple neural network – a feed-forward – or a convolutional network that is mainly used to process images.

## 4. Synthetic data:

Computer-created artificial data, and the output of GANs. Synthetic data can generate ideas, accelerate learning models, bypass privacy issues, and overcome language barriers between different media forms.

## Computer creativity

You can define creativity with only two words: Originality and effectiveness. Computers can be creative, but that does not give free rein to all sorts of anthropomorphism projections about the world of experience of these machines.

Here's how it works...

Consistently, human beings set the objectives and the preconditions and

determine the question that needs to be answered. The GAN then analyzes, iterates, varies, generates, and produces it all. Therefore, it would be nice to see computer creativity as just another complementary form of creativity, instead of an enemy or a creative god. We define computer creativity as follows:

Computer creativity is the soulless creative power (of synthetic data) that requires originality and effectiveness.

## Some other examples

The number of GAN applications and GAN networks is growing day by day, and we are seeing continuous improvements in the basic model and specializations for niche applications. We now count more than five hundred, brought together in the so-called 'GAN Zoo.' These applications show the variety of how GAN-tech can be made applicable:

- The way living spaces are designed has long since shifted from paper to computer screens and pencil to CAD software. NVIDIA now has a step-by-step plan in place to let GANs do the work. It's called Michigan: a Generative Stack for Apartment Building Design.

- Another example: high definition images on Instagram that show all sorts of surroundings, buildings, and furniture that look real but are, in fact, unreal. Coined as "Renderporn," it domesticates the aspiration and surrealism of the digital age.

- Researchers from Microsoft and Tencent, among others, developed a way to let GAN tell a story in pictures by just typing in some words. They appropriately called it "StoryGAN." Scriptwriters for animated films will soon generate a (concept) film directly from their written text. There's already a



Flintstone StoryGan, where you can enter a text like “Wilma is entering the room with a nice meal,” and the GAN generates the image (video) as if you were watching The Flintstones.

- Trypencil’s ([trypencil.com](http://trypencil.com)) pay-off is “Made by machines, Loved by humans.” Their service is creating advertisements within five minutes by using their creative AI platform. They call it MGC. “MGC is a new type of content. It’s generated automatically by an algorithm to achieve specific goals set out by a human brief.” Their pricing scheme starts at \$250 per month with an enterprise account at \$5,000 per month.

### What the generative age means for business

The entire concept of generating stuff (drugs, designs, music, advertisements, and so on) is setting its mark on what to expect. An era of “generation” lies in front of us. From solving detrimental issues like discovering drugs or solutions for climate change to something trivial as creating Flintstone animations: AI creativity has no limits.

What are the business implications? There’s one that stands out. If your business depends on good ideas, new products, new designs, new anything - and what business isn’t - you better prepare for fierce competition. Old fashioned human creativity loses when computer creativity and human creativity join forces. And that’s precisely what is happening now. Join forces with AI to take the future into your own hands.

#ai #gan #creativity #machinecreativity #inventions



## INNOVATION TAKEAWAYS

### Synthetic Data is the new creative oil

Computer-generated data lies at the basis of AI creativity. All kinds of data (images, video, text, music) can be generated; as it is computer-generated, it surpasses privacy issues.

### The era of computer co-creativity

AI, like humans, can generate compelling and original “ideas.” Therefore, we can call them creative. It’s the co-creativity - humans, and machines that work together - and deliver value simultaneously.

### Crucial and trivial solutions from GAN technology

We need all the support we can get for pervasive problems. GAN technology can be of help in finding solutions for those problems. Still, creativity is applied to all dimensions of human life - also the trivial ones - and so will GAN technology.

This text is based on a research paper called ‘Infinite Machine Creativity’ and can be downloaded here:  
<https://labs.sogeti.com/research-topics/infinite-machine-creativity/#infinite-machine-creativity>

The Wave of Data in Oceans.....	60
Make your Brand Stand Out with AI-driven Eco-Responsibility.....	63
Using Data and AI to Crack the Earthquake Code.....	67
Battling Deforestation using Crowd-Driven AI.....	70
The Importance of Diversity of Minds in AI.....	73
Code of Ethics for AI – Humans at the Center.....	75

# THE RISE OF SUSTAINABLE AI



# THE WAVE OF DATA IN OCEANS

Espen Johnsen

PROJECT LEADER LOVE - ECOSYSTEM ACOUSTICS

Robert Engels

CHIEF TECHNOLOGY OFFICER,  
INSIGHTS & DATA, NCE, CAPGEMINI

## Sensing the life with AI

Oceans are changing fast, and we do not always understand the how or why. Yet, oceans play a crucial role in our global ecosystem, and these changes impact us all. The problem is, we have less insight into our ocean's inner operating mechanisms in many respects than we have on our galaxy.

So what's the solution? Well, Capgemini and the Norwegian Institute of Marine Research have taken on the challenge to use machine learning and AI to read, analyze and interpret vast amounts of data collected hundreds of meters below sea level, thus gaining a better understanding of events and inner

workings of the ocean's mechanisms. Further, what works for the oceans indeed might work for businesses too.

Despite the significance of our oceans, vast parts of the marine ecosystems remain unexplored. There are tremendous unknowns in understanding biophysical processes, recruitment processes, and the vulnerability of these systems. Blanket darkness in deep waters is challenging as it makes it impossible to use optical techniques for long-range observations. Nonetheless, the 3D-underwater world is highly dynamic with tidal currents, mixing water masses and nutrients, and large

vertical and horizontal movements of plankton and fish.

Here's the big secret. In contrast to light, sound may propagate long distances in water, and several fishes and, in particular, whales are using sound for communication, navigation, breeding, and identification of prey. Humans can mimic, however still poorly, the sound expertise of whales by developing tools to observe the underwater world.

### Sensing the LoVe Ocean Observatory

LoVe (Lofoten-Vesterålen) Ocean Observatory situated in the ecological

hotspot off the Northern Coast of Norway has placed sensor platforms across the continental shelf, the dominant currents and whale and fish migration routes, e.g., the world largest cod spawning migration to and from spawning grounds are passing the transect to explore the unknown.

Equally important, the observatory uses subsea sensor equipment to perform accurate time recording of many advanced underwater sensor data. Broadband echo-sounders are used to identify organisms in the water column, while hydrophones are used for listening to natural and manufactured sounds. In addition, a hydroacoustic current meter monitors water currents. Other sensors at LoVe Ocean Observatory measure CO<sub>2</sub>, chlorophyll, oxygen, temperatures, salinity, pressure, and more. Moreover, up-close images are taken of deep-water corals and other seabed habitats.

These measurements provide fundamental new insights into marine processes, including the behavioral dynamics of organisms and variation concerning physical and chemical properties. It gets better. They enhance understanding of relationships between physical drivers and biological response, a prerequisite for models that reflect actual dynamics in the marine ecosystem.

Real-time streaming of data supports better modeling, a fundamental tool for handling hazards like oil spills and evaluating impacts of human activity and exploitation of marine ecosystems. Invariably, detailed information helps to understand the consequences of climate variation and associated trends.

#### **Help from machine learning and AI**

The growing ability to measure an increasing number of dynamic

characteristics within marine ecosystems in real-time drives the need for ever more intelligent data analysis methods and techniques. Issues ranging from vast amounts of multi-dimensional, real-time data streams to sensor-drift, incomplete data, replaced sensor output shifts or addressing re-calibrated equipment elegantly.

Thus, there is a crucial need to harness the power of machine learning and AI techniques to analyze sensor data quality, detect sensors, identify an occurrence of ecosystem events and provide alerts for rare and harmful events.

Understanding our oceans is vital. Despite the complexity and differences between marine ecosystems worldwide, global solutions are needed to identify abrupt events or, even more importantly, to recognize gradual changes in marine ecosystems



affected by worldwide warnings and the cumulative effect of man-made stressors on the ecosystem.

### Tackling data science challenges

Capgemini's yearly Global Data Science Challenge has gathered over thousands of consultants and data scientists to contribute to the Norwegian Institute of Marine Research (NIMR), helping them to create a step forward towards the goal of an improved understanding of our Oceans and contributing to the United Nations decade for Ocean science for sustainable development. Together we build solutions for accessing, interpreting, analyzing, and reporting on information contained in terabytes of data collected hundreds of meters below sea level.

Case in point, insights gained range from ways of processing data collection under extreme circumstances, analysis of vast amounts of time-series data on many diverse groups of granularity, unraveling patterns of bio-marine life

migration related to many different features of underwater life, and more.

The result? Insights, data, and methods are shared with the community for the greater good.

Learning about the oceans by inducing events and theories in a data-driven manner is challenging.

Not only is monitoring a complex ecosystem in change a problematic task, inducing system behavior from vast amounts of sensor data is not trivial either. Observable and provable through multi-dimensional data are events such as algae-blooming due to seasonal effects or cod migration patterns in the northern Atlantic.

You can correlate these local events with expected seasonal influences and locations for events to support conclusions on the absence or presence of climate change, the status of the gulf stream, or other super-events. Those events are published and eventually correlated with other analyzed events across the globe.

Also, mechanisms to (semi-)automate such event publishing and cross-correlation are yet to be implemented. Currently, we experience a global uptake of monitoring oceans. UNESCO, the Global Ocean Observing System, and LoVe observatory and private-owned foundations like C4IR Ocean show the way.

A beneficial choice for global and open publishing of analyzed events using semantic web standards would make event reports machine-readable and allow for automated alignment and merging of found trends and facts. Opening and sharing all our data on our oceans and co-operating on their analysis and interpretation globally is crucial for sustainable co-inhabiting species on our planet. Indeed, it's an effective form of guidance for many business endeavors as well.

## INNOVATION TAKEAWAYS

### The big unknown

Oceans remain largely unexplored and still hide many secrets, but increasing amounts of relevant data are becoming available.

### Understanding is critical

Without deeper insights into the mechanisms of balancing and keeping oceans as healthy and self-maintaining ecosystems, we will not be sustainable over the long term.

### Information

The Lofoten-Vesterålen Ocean Observatory generates tons of data. Capgemini and the Maritime Research Institute co-operate using state-of-the-art ML and AI techniques to unravel raw data for information modeling.



# MAKE YOUR BRAND STAND OUT WITH AI-DRIVEN ECO-RESPONSIBILITY

Vincent de Montalivet

EXPERT IN AI AND SUSTAINABLE DEVELOPMENT,  
CAPGEMINI

You're likely already aware of AI's strategic opportunities, but you may not realize AI's infinite potential for transforming your environmental impact. AI offers massive computing power to help produce more sustainable products and services efficiently.

## Take a closer look to carbon emissions

Large organizations tend to have a massive carbon footprint, so it's essential to use innovative technologies to drive a greener planet. One of the primary reasons for higher carbon footprints is high energy consumption in industrial applications.

Unfortunately, many companies cite challenges around absorbing the costs

associated with replacing their current infrastructures with a low-carbon emitting infrastructure. AI addresses this key issue by using data analytics to optimize operational efficiencies while delivering deep insights into carbon emissions, reducing expenses, and facilitating sustainable transformation. With this in mind, AI is a key driver for eco-friendly brand evolution.

Many consumers are also paying attention to the impact of their

purchasing and consumption habits. In fact, 57% of consumers prefer brands that make a positive impact on societal issues. When consumers demand brands prove their value and commitment to social, economic, and environmental issues, this figure also conveys that consumers would be indifferent if 77% of brands disappeared altogether.

## Build richer environmental impact models with AI

Awareness is increasingly essential (seven out of ten people in France consume organic foods at least once a month), but it's still tricky to benchmark market demand concerning sustainability. Whether the topic is organic, carbon-neutral, recycling, recycled, zero-plastic, circular economy, green, vegan, natural products, and something else – it's not always easy to deploy an eco-responsible strategy. However, there are several proven strategies to ensure an eco-responsible approach to organizational operations and product design.

## Use AI to do good and to ensure the common good.

More importantly, you can pursue sustainable development while improving profitability. Invariably, nearly 80% of brands say that doing so increases customer loyalty, and 63% report that it directly contributes to revenue increases.

Not to mention, artificial intelligence has already reduced greenhouse gases by about 13% among manufacturers and retailers. It can also help brands reach 45% of their carbon reduction target by 2030.

AI makes data analysis more manageable and more effective; it democratizes access. Visual platforms, with drag-and-drop functionality, helps to close the giant carbon data loophole. You can certainly use an AI-driven solution to monitor, predict, and reduce emissions. Therefore, the question remains: How can AI solidify and accelerate your eco-responsible initiatives? Consider adopting the tips

below and tailor them to fit your unique business needs:

1. Invest in optimizing resource use and add global logistics operational excellence.

At the most basic level, manufacturing firms aim to produce only the exact number of products they sell without any excess. How can AI help? With the correct data, you can decrease the number of unsold items and reduce your carbon footprint.

From supplying materials to shipping products, it's important to incorporate AI across multiple stages of the value chain. Once you produce each unit, you can send the right products to the right stores while mitigating errors and optimizing supply chain logistics.

For example, AI solutions have helped Amazon reduce packaging requirements by 33%, saving more than 915,000 tons of packaging materials, equivalent to eliminating 1.6 billion shipping cartons. In addition, Capgemini offers a tool that generates a complete cost-benefit analysis on the shipping side, which details estimated fuel consumption, fuel costs, and CO<sub>2</sub> emissions while depicting a variety of potential scenarios to optimize delivery strategies.

Forecasting is a vital component of the process. Yet, most sales forecasts still rely on legacy or historical projections. In contrast, new algorithms can achieve unprecedented accuracy translating into solid gains in economic and environmental performance.

Let's take a look at an example from H&M. In early 2018, the Swedish multinational clothing company announced their stock of unsold clothing exceeding \$4 billion in costs.

The conundrum highlights the increasingly complex sales forecasting with tight-flow logistics.

Continuous investment in AI makes it possible to predict better in-store demands based not only on a post-sales perspective but also by using external data such as weather forecasts, scheduled sports, or cultural events. In essence, your organization can make consistently reliable data-driven decisions down to the item and store level.

For instance, Carrefour partnered with Capgemini to integrate an AI-driven SAS solution for supply chain management. As a result, Carrefour was able to optimize inventory management and reduce waste. By collecting and processing data from stores, warehouses, and e-commerce sites, Carrefour can now consolidate the right data to anticipate demand and refine incoming supplier orders. As a result, Carrefour reduced the number of breaks and overstock in their stores and warehouses.

Let's chew over the agri-food industry: stakes in this sector are vast. Food waste is a prominent issue. Moreover, reducing food waste using AI algorithms can save companies around \$127 billion. Regarding agri-specific AI implementations, consider the following benefits:

- Use image recognition to determine when fruit is ripe.
- Effectively match food supply and demand.
- Augment the value of food by-products.

The good news? Large organizations now have the right tools and data to improve their economic scale and environmental impact. By implementing eco-responsible tactics, brands can progressively improve their carbon footprint

starting with optimizing packaging. To illustrate, decreasing transport saves fuel while data-driven inventory management maximizes heating and cooling efficiencies. Optimizing stock also eliminates unnecessary production while enhancing resource utilization.

## 2. Create fresh eco-responsible products to fortify a circular economy.

The development of sustainable products is a challenge for companies that pursue traditional business objectives focused entirely on cost optimization. The product design phase is a significant driver to prevent reuse, repair, refurbishment, and recycling materials. Considering tools to assess the environmental impact of eco-responsible products are complex and not user-friendly, designers and marketers often have to outsource critical analysis.

On the other hand, data mining makes it possible to recover and automate vital lifecycle data in the decision-making process. With AI, you can predict product and carbon costs right from the design phase to ensure optimized scenarios. For example, you can source local products and reduce the carbon footprint associated with transport or product substitution during the manufacturing phase.

## Achieve your sustainability goals and eco-responsibility vision

AI can also enhance and accelerate new product development, components, and suitable materials aligned with a circular economy. How? Well, by using machine-learning-assisted iterative design processes that enable rapid prototyping and continuous testing. Additionally, AI can facilitate the implementation of new economic models relevant to the circular economy.

For example, AI makes it possible to simplify the resale process on the second-hand market. Capgemini and its partners have developed a "circular" customer journey: Customers bring in used clothing and photograph the items with a store camera; the solution scans and analyzes the clothing and automatically generates a product page, listing defects, holes, stains, or scratches. Next, it estimates the value of the item based on brand, authenticity, description, size, condition, and materials.

Decision-makers are increasingly relying on the circular economy and sustainable development. Recently, Ikea announced that it was creating a store 100% dedicated to second-hand sales. In northern Europe, they built a shopping mall to focus solely on second-hand resale.

## 3. Sustainability as financial scorecard.

It is increasingly vital to understand your company's approach to eco-responsibility. The rapid clip of the innovative changes mentioned above



has mushroomed to regulators who strive to increase transparency around consumer consumption's ecological impact.

For example, Goldman Sachs has made "sustainable finance" a foundation of its business. Consequently, we can no longer afford to evaluate companies based primarily on extra-financial criteria. Instead, it is essential to integrate the carbon standard into the entire value chain so that consumers can choose brands respectful to the planet.

Similarly using a French example, nutritional qualities are evaluated with a Nutri-score, or a "carbon score." The score informs consumers about the environmental effects of their products and helps drive environment-conscious purchasing behavior. Undoubtedly, resource-saving behavior is no longer a trend. It is the standard.



#### Final thought

As a society, we face tremendous challenges around ensuring ecological balance. AI will not save the planet alone. It is up to us to leverage its capabilities. Brands that want to

improve customer retention over the next decade should take note. With the examples shared above, let's continue to discuss the variety of ways AI implementation can positively impact your eco-responsibility initiatives.

#sustainability #brandexperience #ai #ecofriendly #carbonemissions

## INNOVATION TAKEAWAYS

Large organizations have a massive carbon footprint

Use innovative technologies like AI to drive a greener planet.

Consumers pay attention to brands' sustainability actions

Many consumers make purchase decisions based on brands' impact on societal issues.

Improve profitability by pursuing sustainable development

Customers prefer eco-friendly brands. Forward-thinking companies must follow suit.

AI makes data analysis more manageable, effective, and accessible

AI-powered platforms are intuitive and can democratize access to key insights organization wide.

# USING DATA AND AI TO CRACK THE EARTHQUAKE CODE

Niels Heffinck

PRODUCT MANAGER, AIE SHENZHEN,  
CAPGEMINI

## Using Data and AI to crack the earthquake code

The AETA (Acoustic Electromagnetic to AI) project is one of the world's leading projects in earthquake forecasting. Using sensory systems to collect data, then feed this data through an AI algorithm, the AETA project aims to accurately forecast earthquakes days before they occur. Over the past 12 years, a team from the Peking University has been working on a better understanding of earthquake forecasting, including hosting global competitions between data science teams.

Building on an accumulated data pool of over 42TB+, the AETA team hosted a competition in 2020, inviting teams from all over the world to participate in developing an accurate predictive algorithm. Out of the more than 150 participating teams, 10 prevalent teams achieved an over 70% accuracy rate with their machine learning-based algorithms. Approaches included LSTM neural networks, Nested-LSTM, lightGBM, random forest, CatBoost, logistic regression and Spatio-temporal graph convolutional neural networks (STGCN).

### New Competition

The AETA team has begun a new competition for 2021, inviting international institutions to participate in refining the algorithm, with the hopes of achieving an over 90% accurate result. If the competition can achieve this goal, the AETA project may be the only research institute to finally 'crack the code' of earthquakes, bringing about a new perspective for mankind.

The algorithm focuses on 3 key predictive factors: 1. Whether an earthquake actually will take place (whereby one of the 2020 teams achieved a 100% accuracy rate). 2. The

epicentre of the earthquake, and 3. The magnitude of the earthquake.

Although the further development of the algorithm still has hurdles to overcome, the AETA team and their partners are optimistic that the project will achieve its goal, creating a safer world.

The AETA project was founded by Professor Wang Xin'an and Dr. Yong Shanshan, shortly after the 2008 devastating earthquake that hit Sichuan, affecting the lives of over 400,000 people. Both are experts in the field of seismology and believe there should be better predictive measures in place, preventing the loss of life.

### Sensors

The technology involved includes SVV-made 3-part sensory systems between 20 – 100 km apart. Out of this grid of sensors, AETA collects 91 variants of seismic data, using electromagnetic sensors and acoustic sensors. This data is then sent to the transmission box where it is uploaded to a central database, ready for data processing and analysis. What makes these sensors unique is that they collect a huge amount of valuable data and cover a broad frequency range. Furthermore, it is quite simple to deploy the sensors, which are very durable, cost-effective and energy efficient. Through these sensors, AETA is already able to offer a unique capability: reliable earthquake forecasting up to a week before the actual event happens.

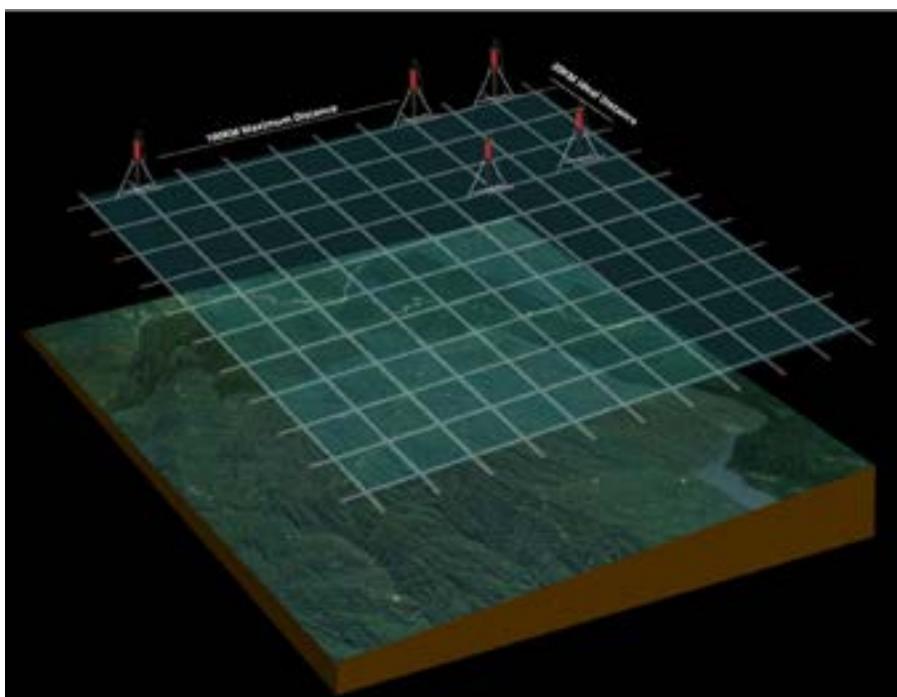
Already back in 2018, the AETA team successfully detected 2 earthquakes, one being a magnitude 5.0, informing the local government timely of its potential danger. This triggered the seismic community to take the AETA forecasting system more seriously than ever. Since 2018, no other earthquake

(of 5.0+ magnitude) has been detected or occurred within the region.

### Better predictions

More research groups are focusing on earthquake predictions, but none have yet proven results that can compare to those of AETA. Many institutions use alert systems, as a temporary solution to warn citizens when an earthquake happens. These systems are only able

to give a less than 1-minute warning compared to AETA's 'days before' prediction forecasting system. With the spread of sensors and the current algorithmic model, earthquakes can be forecasted just like the weather is forecasted. This gives governments the time to prepare, take preventive measures such as closing public transport, securing nuclear plants, ensuring continuity for hospitals and warn its citizens. It reduces material

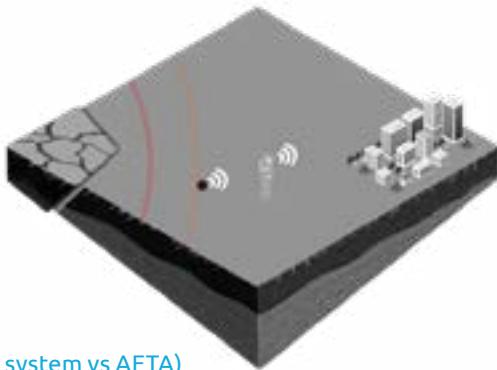


and physical damage, providing much better chances of safe survival when the earthquake finally hits.

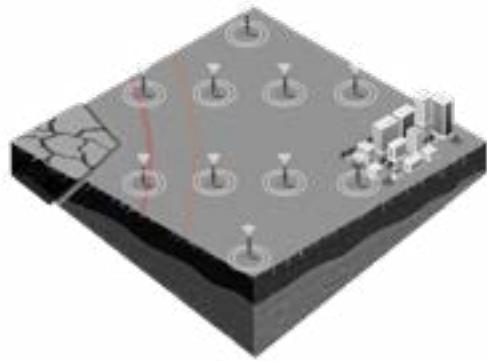
And crowdsourcing works: previously the global accuracy standard for predicting earthquakes was less than 30%. Now, after the 2020 AETA competition, this standard has been increased to 70%, paving the way for even more revolutionary breakthroughs in the field.

Currently, the sensors are starting to be deployed globally. This will greatly increase and enhance the data pool available, and may bring surprising new information about how earthquakes evolve. Seen as one of the most ambitious projects of 2021, creating this global network of sensors will inspire even more data scientists across the world to tackle one of the most daring challenges of our time.

The AETA team and its partners are adamant that they will eventually unveil the full mystery behind forecasting earthquakes. And with the sensor platform and algorithms becoming available across the world, it may save numerous lives for many years to come.



(Alert system vs AETA)



#scienceandenvironment #bigdata #datascience #predict #earthquakes

## INNOVATION TAKEAWAYS

### Making sense

High tech, yet easy-to-deploy sensors serve as a vehicle for collecting and transferring tremendous amounts of valuable data

### Doing the math

Advances in AI enable a much better forecasting of the three key elements of earthquakes (timing, epicentre and magnitude)

### Power of the crowd

Crowdsourcing (through a global competition) unleashes a unified data science power, producing improved results year by year

### Data & AI for Good

Nothing inspires more than leveraging data and algorithms for a better, safer planet



# BATTLING DEFORESTATION USING CROWD- DRIVEN AI

Jonathan Walker

ARTIFICIAL INTELLIGENCE SPECIALIST, SAS

Climate change's devastating impacts continue to threaten our planet and its people

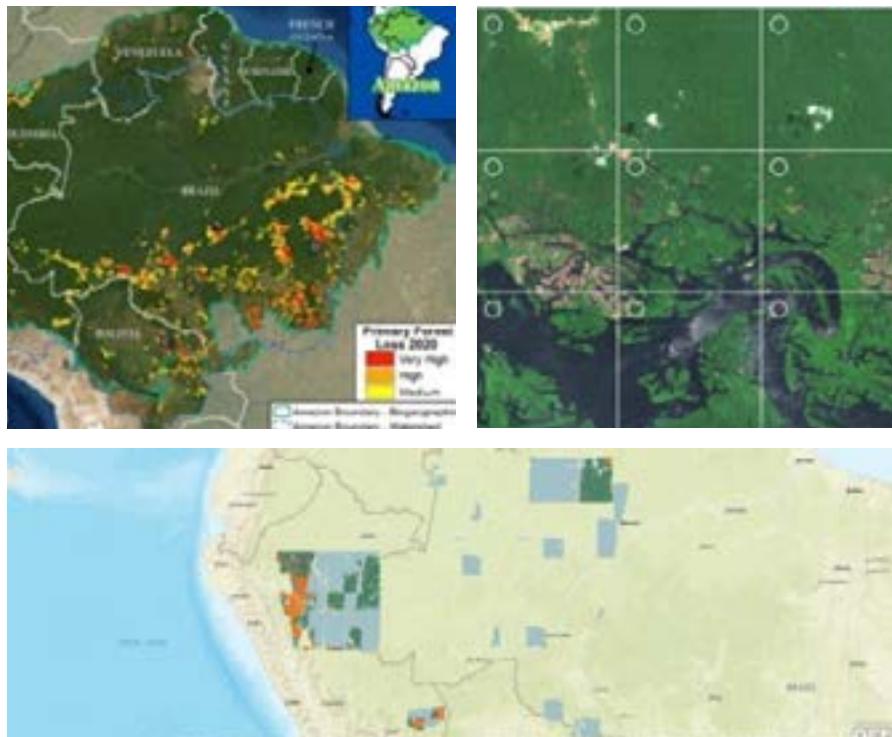
In the Amazon rainforest, deforestation destroys over 800 km<sup>2</sup> of land each month. To address this, SAS and the International Institute for Applied Systems Analysis (IIASA) developed a model to detect the earliest signs of deforestation in the rainforest by combining the power of crowdsourcing and artificial intelligence (AI). This effort will allow decision-makers to detect change better and drive policy responses quicker.

The Amazon rainforest, spanning 5.5 million km<sup>2</sup>, is home to the world's greatest variety of plant and animal species, and absorbs billions of tons of CO<sub>2</sub> from our atmosphere every year. Unfortunately, deforestation is wreaking havoc on the Amazon. More than 800 km<sup>2</sup> are destroyed each month as forests are cleared for timber extraction, crop production, and infrastructure development. In 2020, deforestation levels spiked while enforcement efforts lagged due to the pandemic's ripple effect.

Detecting deforestation one image at a time

Step one in rainforest protection is pinpointing where deforestation occurs. In an area as large as the Amazon, this is challenging, but SAS and IIASA knew AI could help.

IIASA conducts policy-oriented research on global problems that are too large or complex to be solved by one country or academic discipline. IIASA knows today's world systems are approaching levels that exceed existing human experiences, scientific understanding, and traditional



CREDIT: Amazon Conservation's MAAP Program

modeling capabilities, making it increasingly difficult for humans to intervene rapidly and effectively.

To drive deforestation policy responses, SAS and IIASA jointly built an AI platform to analyze satellite images of the rainforest to show the location and magnitude of damage effectively. Then, they asked an army of data scientists to review the images, which show land untouched or possibly affected by human development, to better train AI models.

From middle schoolers to professional data scientists, anyone can help. These efforts improve the algorithm and expedite the process – one click at a time.

"Since image analysis is an important first step for the accurate assessment and projection of deforestation, we are asking people to see what

technology can't see, by examining and reporting on images of smaller incremental forestry incursions worldwide," said Albert van Jaarsveld, Director General and CEO at IIASA. "Crowdsourcing helps improve the AI algorithms, expediting what used to take years to analyze, thus helping to drive vital policy responses to protect our forests more quickly."

### The process

Tens of thousands of satellite images are continuously collected, and the crowd labels images that show signs of deforestation. SAS applies statistical techniques to determine whether the group has reached a consensus on each image. Further, IIASA experts review images that proved challenging for accurate labeling. Behind the scenes, computer vision models trained to recognize areas of deforestation are learning at every iteration. They test models for

accuracy and deploy high-quality models while lower-quality models keep iterating.

Training the data is crucial for the success of this project. Approximately 80% of the images make up the training data set, while the other 20% are used as a test set to measure model accuracy. Therefore, they can continually refine the models for maximum efficiency.

### Humans amplifying AI

AI is not magic; it relies on the expertise of people as the basis for its training data.

Tapping into collective knowledge to achieve results faster is critical given the unprecedented deforestation rate. But it also shows the human side of AI. People see AI behind the scenes and understand how human-driven it is.

For our industry to keep moving forward, we need people to recognize and trust the technology that will help drive real societal change. Giving people a seat at the analytics table, we help people witness the magic behind AI, and equip the next generation with the data confidence needed to succeed in our digital world.

### Looking ahead

This project was launched on Earth Day 2020, and by March 2021, citizen scientists from across 95 countries classified more than 845,000 km<sup>2</sup> of the Amazon.

By using image data from this ecologically diverse territory, the model is given various examples so it can eventually learn to detect human impact anywhere in the Amazon. Suppose researchers can show their predictive model is successful at identifying areas most at-risk. In that

case, it could be helpful for forest monitoring bodies to track and respond to forest changes carefully.

SAS is expanding its work on this critical issue, recently collaborating with Amazon Conservation to boost efforts for identifying and tracking illegal deforestation and expedite intervention.

SAS and Amazon Conservation will incorporate deforestation alerts from the University of Maryland's Global Analysis and Discovery (GLAD) lab, which uses satellite imagery to collect weekly data on deforestation across the tropics. By prioritizing the highly-threatened, protected areas and indigenous territories, the project will use crowdsourcing and AI to help automate the process of determining whether deforestation is natural or man-made. In addition, building upon current GLAD alerts, officials can use the additional insights to understand

the source of deforestation and determine its legality in real-time.

SAS is committed to using technology to solve our world's most pressing issues, whether working to fight deforestation, analyzing the impact of climate change on girls' education, or uncovering factors contributing to racial disparities in homeownership. By combining the passion of people and the power of technology, we know we can accelerate global progress and secure a more sustainable future.



#dataforgood #aiforgood

## INNOVATION TAKEAWAYS

### AI for good

Time is not on our side as we tackle our world's crises. But we can better protect our planet with the power of a robust, scalable AI platform, giving accurate, fast results.

### Power of the crowd

AI is powerful, but when you combine technology with the passion and empathy of humans, progress happens even faster.

### AI demystified

When we show people the behind-the-scenes of AI, they see just how human-driven it is, and they'll better trust the technology driving real change for our society.

# THE IMPORTANCE OF DIVERSITY OF MINDS IN AI

Anne Laure Thieullent

VICE PRESIDENT, ARTIFICIAL INTELLIGENCE AND ANALYTICS GROUP OFFER LEADER, CAPGEMINI

As this emerging technology reshapes the world, it must represent a range of perspectives and experiences. In order to build truly powerful AI-driven solutions, the tech sector needs “diversity of thought.” The industry must encourage and foster people of all genders, ages, cultures, backgrounds, and abilities to ensure everyone feels they can and should have a voice in how AI is developed and used.

**AI-powered innovation holds significant promise for tackling the world’s problems – from immediate priorities such as fueling the economic recovery from the global pandemic to long-standing issues such as addressing climate change. To best deliver on this promise, companies using AI need to recruit passionate people from a diversity of backgrounds, experiences, and perspectives.**

AI-powered innovation holds significant promise for tackling the world’s problems – from immediate priorities such as fueling the economic recovery from the global COVID-19 pandemic to long-standing issues such as addressing climate change. To best deliver on this promise, AI companies need to recruit passionate people from a diversity of backgrounds, experiences, and perspectives.

That will only happen if, as an industry, we create opportunities for talented people from underrepresented groups. The need is well documented. For example, in researching The Global

Gender Gap Report 2018, the World Economic Forum discovered that women hold roughly 50 percent of all science and engineering degrees yet make up just 22 percent of AI professionals.

To address this, we must encourage young people from all backgrounds to consider STEM (Science, Technology, Engineers, Mathematics) careers.

## Promoting AI careers to young women

The Girls In AI Global Hackathons are a prime example of how to do this. Launched by Teens In AI founder Elena Sinel in 2018 after her daughter was discouraged from taking computer science and math courses in secondary school, these events introduce young women from around the world to the potential of AI to create meaningful change. The hackathons are designed to inspire the next generation of AI researchers, entrepreneurs, and leaders – even as they instill the importance of ethics and diversity in the teenagers who take part.

Earlier this year, as part of our partnership with Teens in AI, I was honored to help judge entries in the most recent Girls in AI global hackathon. This event took place online starting in March. It involved almost 1,000 participants from 140 cities in 23 countries, making it the program’s most diverse hackathon to date. Significantly, 78 percent of participants were young women, 90 percent were taking part in their first hackathon, and 72 percent had little or no prior coding experience.

## Powering the next generation of thought leaders, innovators and technologists in AI

*“As a judge, I was amazed at the challenges these young people tackled – proposing creative, AI-driven solutions for issues related to education, health, and the climate. The team members did a terrific job of defining their projects and created impressive prototypes to showcase their solutions. Their passion was inspiring, and matched by impressive technical skills in people so young.”*

These events benefit the girls – but also the adults who help run them.

This is not just inspiring: It's essential. In an interview with Capgemini after the event, Sinel noted that as countries around the world grapple with the fallout of the COVID-19 pandemic, the novel approaches that young minds bring to the challenges and their fresh perspectives on the power of AI will be critical to our success.

"I don't see the economic recovery happening if we don't invest in education and if we don't invest in young people from an early age," Sinel said. "It's so important because they will be the ones to help bring us out of the mess we're in right now."

#### Diversity of thought delivers better results

The future is AI-driven. Soon, AI will be a core enabler of every industry and every job will include some element of AI. Therefore, we need to ensure the AI-driven solutions we create are ethical and trusted – that they fix problems instead of exacerbating them.

The challenge today is that the AI sector remains not only male dominated, but also doesn't bring full representation of the world we are living in – across all the dimensions. This causes problems for the solutions we create. They may be built on data that is biased, or they will be dominated by a single perspective or cater to a limited demographic.

The industry is addressing this and progress is being made, but more needs to be done.

The answer is to foster "diversity of thought" in AI. We must engage with people – especially young people – from a range of backgrounds, experiences, and abilities. We need to demonstrate to them the power of AI to solve problems and encourage them to join us in that effort.

#### Palestinian team uses AI to tackle bullying

During the 2021 Girls in AI Global Hackathon, the Ataria Team from Palestine created "Express Yourself" – an app that asks a child to answer a few questions and to sketch drawings, then uses AI to determine whether that child is being bullied. The four girls on the team – Aya, Naden, Nour, and Rahaf – won the Social Impact & Innovation Award and an opportunity to pitch their solution to leaders from Capgemini's Innovation Community. In addition, we are providing mentorship to this team through the Capgemini AI Academy program.

Speaking through an interpreter, the girls report that the experience taught them about working in teams and helping others learn new skills, how to code in Python, how to use tools like business-model canvases, and how to build prototypes. They all enjoyed the experience, would like to participate in the next hackathon, and are already thinking about how to use AI to design other health and wellness applications.

#girlsinaid #hackathon #diversity

## INNOVATION TAKEAWAYS

### Diversity of thought makes AI better

It's important that AI companies build teams that are diverse – in gender, age, ability, culture, and backgrounds – to ensure applications reflect all users and their needs.

### It's a two-way street

AI benefit when they get involved in programs to promote AI careers to young people and underrepresented groups – because these people bring passion and novel approaches to the task of applying AI to solve problems.

### Action is needed now

In many countries, the pandemic wiped out years of progress towards creating more equitable futures for girls and other people who are underrepresented in the AI sector. Programs that help everyone explore career opportunities in AI are one way to rebuild that progress, even as they make the AI sector stronger.



# CODE OF ETHICS FOR AI – HUMANS AT THE CENTER

Yashowardhan Sowale

VP, ARCHITECTURE LEADER, INDIA DOMAIN LEADER  
FOR AI, INSIGHTS & DATA, CAPGEMINI

Artificial Intelligence is creating innovative breakthroughs in both the technological world and society

With AI advancing, the role of humans in a rapidly shifting context becomes a focal point. To most organizations, an ethical approach to AI is a requirement for further adoption of AI, and hence also a driver for innovation with AI. Capgemini, consistently ranked as one of the world's most ethical companies, thus places ethics at the center of all AI activities. Its Code of Ethics for AI provides 7 guidelines for delivering tangible benefits from AI within a trusted framework, always putting humans at the center.

AI is a general purpose, pervasive technology area, potentially affecting individuals, businesses, and the society at large. It comprises of a variety of technologies and capabilities, but has seen its recent breakthroughs mainly due to spectacular advances in data storage and access, new machine learning models and highly specialized hardware components. It is spreading from corporate and government domains to various areas of daily life.

## Conundrum

Organizations are racing to become data-driven businesses to capitalize on new opportunities for innovation and growth. They however must balance leveraging the opportunities presented by AI with addressing the risks that it may generate. The trust in AI-based systems is pivotal to success, and hence the ethical, human, economic, societal, and environmental implications of emerging AI are becoming critical topics.

Deep learning, reinforcement learning, and other innovative AI components are exploring the next frontiers. They make systems perceive, decide and act on their own - much like humans, albeit in a decisively 'artificial' way. Such advances, fascinating as they may be, can increase specter of AI systems taking undesired decisions and actions. AI experts and practitioners need to address this ethical conundrum, balancing the needs of businesses and customers. Strong Ethical AI practices need strong leadership, governance, audits, security, and operationalization.

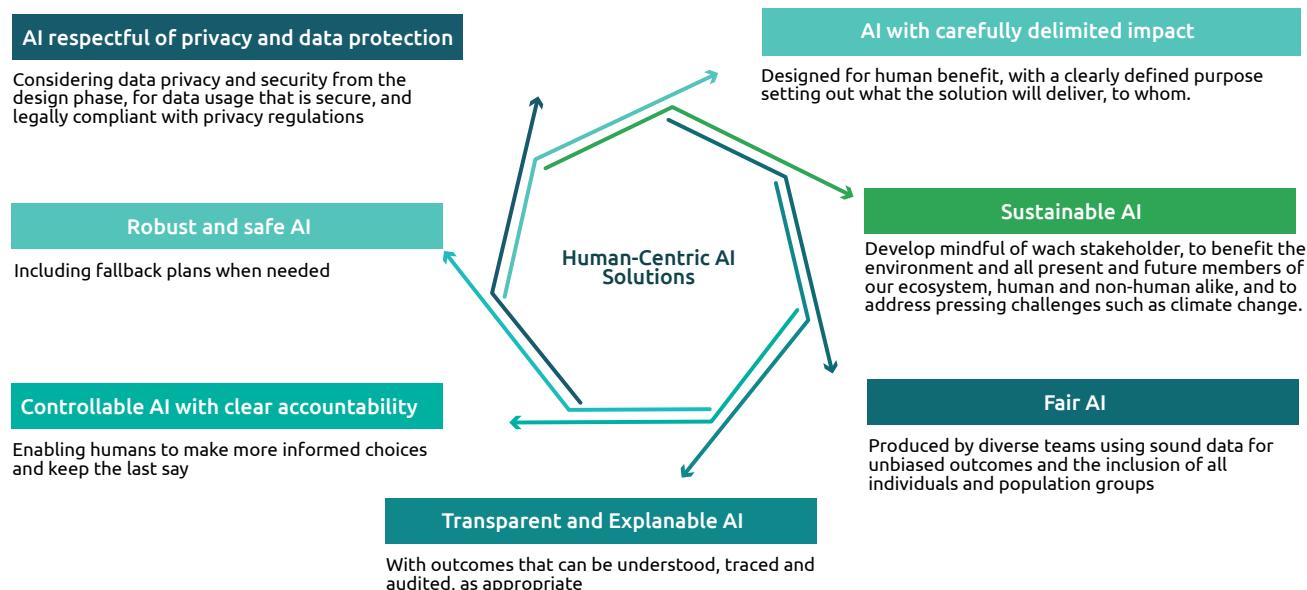
An ethical approach to AI needs to provide guardrails for fair, transparent, explainable, controllable, and accountable AI implementations. Capgemini's openly available Code of Ethics for AI describes 7 guidelines to achieve this.

## Bias

One key aspect of Ethical AI is to emphasize and embed diversity and inclusion principles throughout the lifecycle of AI projects to enforce fairness aspects. It is already well-documented that AI systems can be biased – through the historical training data that its algorithms are based on. This needs to be addressed, especially when the decisions and actions of AI systems affect public trust or society. One famous example of bias in AI systems is the COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) algorithm used in the US court systems to predict the likelihood that a defendant becomes a recidivist. A study showed that the algorithm showed prejudice and recommended too severe sentencing.

Within the context of societal impact, deployment of AI implementations can impact workforce, economic growth planning, and productivity adjustments. On one hand, AI can dramatically improve the efficiencies in the workplace, on other hand it can cause workforce rationalization. Ethical AI provides guidance on how human developers, software providers, SMEs and architects should architect and design to minimize the ethical harms that can rise from bad design and imprudent implementations not adhering to the Ethical AI principles. 'AI with carefully delimited impact' is one of the seven core principles of Capgemini's Code of Ethics for AI.

## Code of Ethics for AI - Our 7 Principles



## Code of Ethics for AI - Our 7 Principles

## Human impact

The principles are relevant for organizations as they impact top and bottom line by mitigating reputation risk. They help to secure beneficial reputation in society and avoid loss of reputation due to data protection breaches, misuse of AI and risk exposure. They help empower humans to make informed decisions through AI oversight and control mechanisms. As AI technologies are evolving continuously, continual analysis of human impact of AI technologies is important to help identify the overall impact of likely impacts against prospective risk arising from undesired side-effects and subsequent ethical risks.

Ethical AI practices also help build resilient and secure AI systems that ensure accurate, reliable, and reproducible algorithms that guarantee repeatable results. Ethical AI guidelines establish adequate data governance considering quality, integrity, and legal equitable access rights of data. Ethical AI systems and business models must be transparent in terms of implementation.

## Explainable

AI system's capabilities and purpose should be transparent. Decisions by AI systems should also be explainable, with the degree of 'explainability' dependent on the context and severity of the consequences if the result is erroneous. The data sets and processes used for the AI solution should also be documented to allow for traceability and, as required, for auditability.

AI systems also need to attribute ownership to specific human actors in case responsibility or overriding action is needed. The design of AI solutions should protect the human's autonomy and decision making. AI solutions should help humans make more informed choices. Humans should be part of the AI governance mechanism in such a way that they always keep control over AI.

Appropriate measures should be implemented from the AI solution's design phase, with the appropriate level of human oversight depending on the AI solution application area and potential risks.

Accountability helps define and clearly identify together roles and responsibilities amongst the different actors involved in the design, code build, integration, deployment, and operationalization, including the data provider, and the company that adopts the AI solution or the final user. This would enable appropriate allocation of liability and effective recourse when needed. Ethical AI systems need to also account for non-discrimination in all aspects of life cycle right from procurement of data to testing.

## Society

AI systems must be sustainable and environmentally friendly to contribute to overall societal benefit (for example, they should not require huge amounts of energy to train deep learning models). Expensive and unsustainable machine learning options have given rise to techniques such as federated learning, transfer learning, low data learning – all helping to achieve full benefits from AI without excess power consumption.

#ai #aiethics #ethicalai #sustainabledevelopment #ai4good

# INNOVATION TAKEAWAYS

## Balance act

AI systems provide great innovation opportunities for individuals, businesses and society; but as its powerful capabilities overlap more and more with those of humans, the benefits and potential risks must be carefully balanced

## Trust Thrust

AI systems that have been built on a foundation of trust will not only mitigate the identified risks that come with AI systems, but will also positively drive adoption by its intended users and target audience

## Code of Ethics for AI

Capgemini's published Code of Ethics for AI describes 7 guidelines for creating AI systems that deliver full benefits, while ensuring a trusted, ethical foundation that puts and keeps humans at the center

# AUTHORS



SENIOR DEEP LEARNING  
DATA SCIENTIST, NVIDIA

**Adam Grzywaczewski**



DIRECTOR,  
CAPGEMINI

**Gianfranco Cecconi**



PRINCIPAL PROGRAM  
MANAGER, AUTONOMOUS AI  
ARCHITECT, MICROSOFT

**Andy Wylie**



ARTIFICIAL INTELLIGENCE  
SPECIALIST, SAS

**Jonathan Walker**



VICE PRESIDENT, ARTIFICIAL  
INTELLIGENCE AND ANALYTICS  
GROUP OFFER LEADER, CAPGEMINI

**Anne Laure Thieullent**



PRINCIPAL PROGRAM  
MANAGER, MACHINE TEACHING  
INNOVATION, MICROSOFT

**Kence Anderson**



DIRECTOR, INSIGHTS & DATA,  
CAPGEMINI INDIA

**Anupam Srivastava**



RESEARCH DIRECTOR,  
SOGETILABS, CAPGEMINI

**Menno Van Doorn**



VICE PRESIDENT/ HEAD OF  
DS&A AND AI COE

**Chandrasekhar  
Balasubramanyam**



DIRECTOR, GLOBAL INNOVATION  
LEAD & GOOGLE ALLIANCE LEAD,  
INSIGHTS & DATA, CAPGEMINI

**Monish Suri**



PROJECT LEADER LOVE -  
ECOSYSTEM ACOUSTICS

**Espen Johnsen**



PRODUCT MANAGER, AIE  
SHENZHEN, CAPGEMINI

**Niels Heffinck**



MANAGER,  
CAPGEMINI

**Esther Huyer**



VICE PRESIDENT, INSIGHTS  
& DATA, CAPGEMINI

**Padmashree  
Shagrithaya**





DIRECTOR, ENTERPRISE ARCHITECT - INSIGHTS & DATA, CAPGEMINI

**Prajwal Kumar**



EVP, CTO AND CHIEF INNOVATION OFFICER, INSIGHTS & DATA, CAPGEMINI

**Ron Tolido**



DIRECTOR, AI & ANALYTICS, CAPGEMINI

**Pranav Kumar**



VICE PRESIDENT, ENGINEERING, ADOBE SENSEI & SEARCH

**Scott Prevost**



SENIOR DIRECTOR, AI & ANALYTICS, CAPGEMINI

**Punit Santani**



GLOBAL TECHNICAL LEADER DATA, AI/ML, AWS

**Sriram Kuravi**



CHIEF ARCHITECT, INSIGHTS & DATA, CAPGEMINI

**Rajashree Das**



CHIEF TECHNOLOGY OFFICER, TERADATA

**Stephen Brobst**



HEAD OF AI COE, CAPGEMINI

**Rajeswaran Viswanathan**



SENIOR APPLICATION CONSULTANT, CAPGEMINI

**Veronika Heimsbakk**



GLOBAL BUSINESS LEADER DATA, AI/ML, AWS

**Rebecca Y. Gonzales**



EXPERT IN AI AND SUSTAINABLE DEVELOPMENT, CAPGEMINI

**Vincent de Montalivet**



SENIOR DIRECTOR, ENTERPRISE ARCHITECT, INSIGHTS & DATA, CAPGEMINI

**Rekha Chandrasekar**



ENTERPRISE ARCHITECT - INSIGHTS & DATA, CAPGEMINI

**Vinodh Subramony Iyer**



CHIEF TECHNOLOGY OFFICER, INSIGHTS & DATA, NCE, CAPGEMINI

**Robert Engels**



VP, ARCHITECTURE LEADER, INDIA DOMAIN LEADER FOR AI, INSIGHTS & DATA, CAPGEMINI

**Yashowardhan Sowale**



EXECUTIVE BOARD MEMBER, T&I AT CAPGEMINI ENGINEERING

**Rodrigo Maia**





## 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 everyday by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of 270,000 team members in nearly 50 countries. With its strong 50 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 2020 global revenues of €16 billion.

Learn more about us at

[www.capgemini.com](http://www.capgemini.com)

## For more details contact:

### Ron Tolido

EVP, CTO and Chief Innovation Officer, Insights & Data, Capgemini

[ron.tolido@capgemini.com](mailto:ron.tolido@capgemini.com)

### Monish Suri

Director, Global Innovation Lead & Google Alliance Lead, Insights & Data, Capgemini

[monish.suri@capgemini.com](mailto:monish.suri@capgemini.com)

### Johanna Sundh

Global Marketing Communications Director, Data, AI & Analytics, Capgemini

[johanna.sundh@capgemini.com](mailto:johanna.sundh@capgemini.com)

### Laura Immonen-Beatty

Global Content Marketing Lead, Insights & Data, Capgemini

[laura.immonen-beatty@capgemini.com](mailto:laura.immonen-beatty@capgemini.com)

Interested in learning more? [Download Wave 1](#)

**GET THE FUTURE  
YOU WANT**

The information contained in this document is proprietary. ©2021 Capgemini.  
All rights reserved.

