

WAVE IV | 2022

DATA-POWERED INNOVATION REVIEW



FOREWORD



ZHIWEI JIANG

CEO Insights & Data,
Capgemini



Whatever scenario pans out, data is here to power it.”

When they keep falling on your head, raindrops may not always be a distinct pleasure. But the visual theme of this fourth edition of our Data-powered Innovation Review strikes me in much more positive ways. I see the monsoon bringing much needed relief, after a long period of drought and heat. I see a summer field at its greenest, after the morning rain. In the same way, I believe the many “drops of innovation” in this publication will bring you refreshing, revitalizing perspectives.

And new perspectives we need. More than ever, we see data at the very heart of organizational change. It has been for quite some time to achieve cost-effectiveness, operational excellence, customer satisfaction and growth. But now other objectives have reached the top of the strategic agenda as well. Think about the race for Net Zero and building a more sustainable, more inclusive future. Think about dealing with increasingly scarce natural resources and supply chains that break and re-emerge within the blink of an eye. And talking about scarcity: just look at how the availability of skills, expertise and experience is now determining economic success or failure.

Also – and I really wish this was just a cliché – we are living in a world that is more unpredictable than ever before. We may need to adjust our objectives and strategies overnight, on the fly, over and over again.

Whatever scenario pans out, data is here to power it. You'll find many inspiring examples of it in this edition – ranging from digital twins in the industrial metaverse, “humble” AI, serendipity in user experiences, all the way up to permacomputing and the battle against data waste. And there's so much more. Written both by Capgemini's leading data experts and its innovative partners, each and every item is guaranteed to wow you.

I hope you agree with me when you browse the magazine, whether being on holidays, at work or at home. And in any case, even if rain happens to be predicted, remember that data-powered innovation is just around the corner, promising a sunnier future.

EDITOR'S NOTE



RON TOLIDO

CTO Insights & Data,
Capgemini



Enjoy our shower of refreshing data-powered innovation drops!"

As glossy as this magazine may seem, it is far from the meaning of life to us. To revive that worn-out platitude one more time: it is about the journey, not the destination. The magazine really is our catalyst to brainstorm innovation ideas, reach out to our best experts and partners, create top-quality perspectives and content, and then – after publication – have a long tail of follow-up posts, events, and innovation engagements.

And interesting data-powered innovation ideas we have found yet again indeed. Data, analytics, and AI are at the very center of making innovation real. It is key to addressing some of the most pressing business challenges we face in this era of Uncertainty²: full as it is with challenges, opportunities, and radical change. We are confident this fourth edition of the Data-powered Innovation Review brings you a refreshing shower of different perspectives.

We have 18 inspiring ideas for you, divided into three topic areas.

The Rise of Data Ignition touches on arguably the most important aspect of becoming a Data Master: the ability to activate data for on-the-ground, real-life business purposes. We introduce examples of igniting data in sectors, such as insurance, intelligent industry an automotive, but also explore the innovative potential of data storytelling, self-service BI, and – unavoidably – the Metaverse, with all of its anticipated merits and benefits.

The Rise of AI Augmentation zooms in on the many ways AI augments us in our daily lives. The topic is more relevant than ever, as scarcity of skills, expertise and experience has become a key competitive driver. All articles in this magazine are dear to

me, but the panel discussion about how AI can help to make a more inclusive world is decisively my favorite. Maybe it's because of the excellent panel members Sabrina, Jill, and Kelly ("Good morning, Angels"). Or maybe it's just the AI?

The Rise of Sustainability Data brings back yet again one of our key, recurring topic: the use of data and AI to create a better, more sustainable world. By now, many organizations see the tremendous potential of data and algorithms in battling the impact of climate change. But there is a flipside as well, as storing data and applying machine learning on that data themselves may come with a serious environmental cost. Hence, more about "sustainable data" in this issue as well.

As said, the articles have been crafted by leading Capgemini experts and also by – or in collaboration with – key partners such as Alation, Cognite, Toucan Toco, DataRobot, and The Open Group. Don't hesitate to contact the contributors. They'll be most happy to help you on your innovation journey. Also, look out for a series of follow-up activities, such as additional in-depth articles and live events featuring contributors and guest panelists.



THE RISE OF DATA IGNITION

Data storytelling – making data actionable for business users	06	Turning the underbanked into profitable customers – the sustainable way	17
Baptiste Jourdan, Toucan Toco		Shankar Ramanathan, Capgemini	
The industrial metaverse: more than a glorified digital twin?	09	The role of self-service BI for business agility	21
Petteri Vainikka, Cognite		Myles Suer, Alation	
Reshaping the real and digital world with metamobility	14	AI-infused innovation for automotive data	25
Rodrigo Maia, Capgemini Engineering		Jean-Marie Lapeyre, Capgemini	

THE RISE OF AI AUGMENTATION

Humble AI and the four pillars of trust	31	Using AI to fight cyber threats	42
Ted Kwalter, DataRobot Yashowardhan Sowale, Capgemini		Jerome Desbonnet, Capgemini	
Serendipity systems: architecting personalization systems at scale	34	AI-augmented field service – the future of maintenance	45
Neerav Vyas, Capgemini		Nicolas Claudon, Capgemini Guillaume Gerards, Capgemini	
Quantum computing: the case for machine learning	39	AI panel: How can AI help create a more inclusive world?	49
Nadine van Son, Capgemini Julian van Velzen, Capgemini		OpenAI, AI21 Labs, Cohere Rajeswaran Viswanathan, Capgemini	

THE RISE OF SUSTAINABILITY DATA

Five ways to battle data waste	53	Why businesses need the open footprint standard	62
Roosa Sääntti, Capgemini Ron Tolido, Capgemini		Johan Krebbers, For The Open Group	
Redefining our needs: the promise of permacomputing	56	Data for sustainability: a familiar journey	66
Robert Engels, Capgemini Ron Tolido, Capgemini		Urban Ekeröth, Capgemini Vincent de Montalivet, Capgemini	
Green data engineering: clean water without boiling	59	Reinventing data for the push to net zero	70
Arne Rossmann, Capgemini		Valérie Perhirin, Capgemini Invent	



THE RISE OF DATA IGNITION

**Data storytelling – making
data actionable for
business users** 06

Baptiste Jourdan, Toucan Toco

**The industrial metaverse:
more than a glorified
digital twin?** 09

Petteri Vainikka, Cognite

**Reshaping the real and
digital world with
metamobility** 14

*Rodrigo Maia,
Capgemini Engineering*

**Turning the underbanked
into profitable customers
– the sustainable way** 17

Shankar Ramanathan, Capgemini

**The role of self-service BI
for business agility** 21

Myles Suer, Alation

**AI-infused innovation for
automotive data** 25

*Jean-Marie Lapeyre,
Capgemini*



DATA STORYTELLING – MAKING DATA ACTIONABLE FOR BUSINESS USERS



BAPTISTE JOURDAN

CRO & Co-founder of Toucan Toco

More companies are investing in AI algorithms to become more effective, but the ROI of these high-powered technologies depends on one often-overlooked factor: adoption. The next big challenge for AI is becoming accessible to business users who need the help of algorithms to make better day-to-day decisions. The right data storytelling tool can help.

AI has reached a point of maturity over the past few years, allowing it to effectively help organizations understand their business environment, predict the outcome of strategies, and make better decisions. This is why companies are increasingly investing in sophisticated algorithms that can decrypt their landscape and advise them on the best course of action. But the more powerful the technology, the more important it becomes to bridge the gap between technology and decision making. The results of these advanced analytics

are often not shared or understood beyond the board level, and rarely have a clear impact on the day-to-day operations of a business. The ROI of AI algorithms depends heavily on how widely their results are understood across the company. Companies that want to become truly AI-driven must find the right tools to democratize the results of AI algorithms and make them understandable by everyone, no matter their level of data literacy.





Transforming predictions into day-to-day decisions

Let's say a chain of ice-cream stores has perfected a very simple algorithm for predicting sales according to the weather forecast. It has an effective way to reduce waste and unnecessary expenses by producing the right amount of ice cream. This works well if HQ is solely responsible for supply and can adjust the quantity that gets shipped according to the results of the algorithm. But what if each store is

responsible for its supply? HQ must find a way to share the results of the algorithm in a simple manner that can be understood by store managers, even if they don't have formal data training. The data must be presented in a way that encourages store managers to look up these numbers daily and act upon them.

We encountered a similar case with one of our clients in the catering industry. Its powerful algorithms were able to predict which ingredients needed to be defrosted

on a given day and in what quantity. The trouble was getting all the stores in their franchise to follow these recommendations, and most store managers found it hard to take time in their busy days to look up the data. We plugged our Toucan 2.0 solution on top of these algorithms to combine the power of the AI algorithm and the optimized UX of guided analytics. Store managers received daily recommendations directly on their phones in a simple, contextualized manner that is easy to understand.



A data storytelling tool for all your devices



Toucan 2.0: A solution that gets you from data to action

We built Toucan 2.0 to make it easier for business users to access and understand data. Our aim is to democratize analytics for everyone, even users who don't have data training. Our weapon of choice to do that is adoption: if a tool is user-friendly enough for anyone to understand, a growing number of people will use it and benefit from it. We believe that the battle for ROI in analytics is centered around usage and adoption.

What is true of data analytics in general is particularly true of AI. Even when the algorithms are transparent for data scientists, they remain a black box for most non-technical users. It is important for business users to understand that AI is a tool which allows them to become more efficient in their day-to-day operations.

Our solution centers around readability and user experience. Data storytelling is our specialty: we make sure every number is properly explained and contextualized. We've also packed an "action center" feature in our solution, allowing users to go from data to action. The conclusions that are reached by the

algorithm are repackaged into easy-to-understand (and implement) tips. This can be applied to various industries:

- In manufacturing, AI can help flag productivity problems and push best practices from a network of factories around the world. If the data indicates a problem of hygiene at the end of a factory line, a tutorial about best hygiene practices can be sent to the line manager via an interactive platform.
- In banking, AI can flag suspicious activities or requests emanating from clients. Algorithms are very effective at computing the risk of certain transactions, so bank employees get immediate advice on the best course of action.

Our belief here at Toucan is that democratizing AI is the best way to boost its ROI and make it a worthwhile investment for most organizations. We're also convinced that just as data storytelling can make AI more efficient, AI can make data storytelling more efficient. We're investing heavily in R&D to make our Toucan action center more reactive in the future, allowing dynamic comments, automatic workflow attribution, and root cause analysis.

According to a recent [Forbes article](#), data storytelling is "the action to put data analytics in the hands of less data-literate employees through narrative techniques." Gartner predicts [data storytelling will dominate BI by 2025](#). Analytics will soon become a function anyone can employ, as data narratives and storytelling remake the common understanding of the power of data.



The next big challenge for AI is becoming accessible to business users who need the help of algorithms to make better day-to-day decisions."

INNOVATION TAKE AWAYS

#DATASTORYTELLING
#GUIDEDANALYTICS
#ROIFORAI #AIADOPTION

ROI depends on adoption

Investing in powerful algorithms alone will not make a difference to the bottom line unless the results of these algorithms are shared across an organization and implemented by everyone at every level.

Data storytelling for AI

Data storytelling is a compelling and user-friendly way to present numbers and data. Presenting the results of AI algorithms through data storytelling methods can help boost adoption. The four major trends here are auto-dashboard generation, plain language search, visual datasets, and future prediction.

Guided analytics for business users

Guided analytics tools can make it easier for non-expert users to understand their numbers and take action. These tools can be plugged on top of powerful AI solutions to allow everyone to benefit from AI.

It works both ways

Just like data storytelling tools can make AI more effective and accessible, so will AI enable the next generation of data storytelling tools.



THE INDUSTRIAL METaverse: MORE THAN A GLORIFIED DIGITAL TWIN?



PETTERI VAINIKKA

VP Product Marketing, Cognite

With the metaverse estimated to be worth \$2.5 trillion globally by 2030, according to Bloomberg Intelligence, are industrial companies ditching their digital twins and starting again, or is the industrial metaverse a modern-day Emperor’s new clothes?

Those of us who have been following Industry 4.0 technologies, trends, and hype are all familiar with digital twins: virtual representations of physical equipment, assets, and processes that contextualize relevant industrial data in real-time to enable data-driven decision making, smart maintenance, and overall asset optimization.

Much more of a newcomer – having blasted its way across the chasm

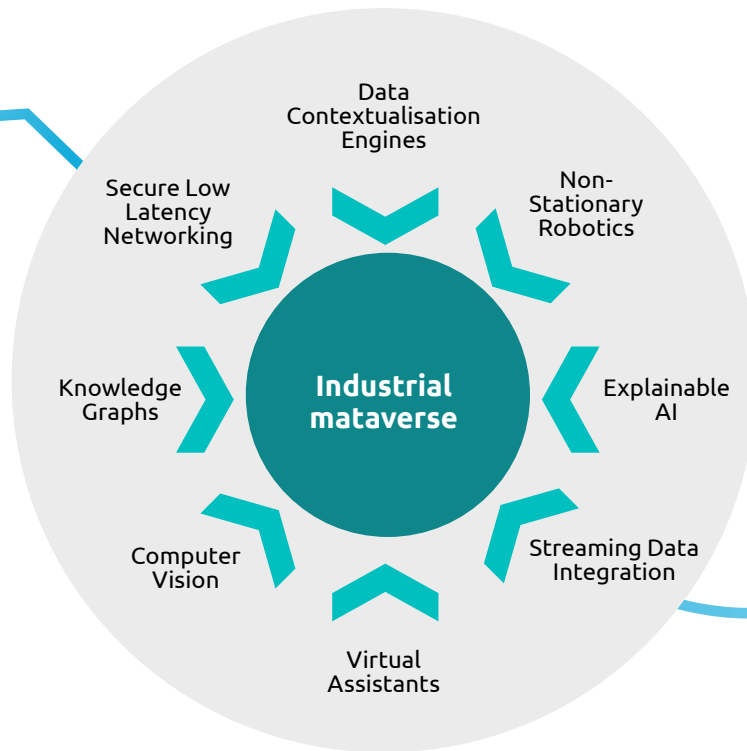
from obscurity into mainstream business consciousness only last year in the tailwinds of social media platform Facebook’s rebranding to Meta, a metaverse company – is the industrial metaverse.

Is the industrial metaverse worthy of the hype, or is it just a glorified digital twin: the Emperor’s new clothes?

Unpacking the industrial metaverse

The metaverse is an evolution of today's internet, enhanced and upgraded to consistently deliver 3D content, contextually organized

information and experiences, and real-time synchronous communication.



Graphic: Industrial metaverse as combinatorial innovation

The metaverse will have both consumer and industrial orientations. In the consumer orientation, metaverse experiences are already playing out quite literally in the form of immersive online gaming universes, with retail and other entertainment experiences set to follow next. Conversely, the industrial orientation is set to impact how – and where – we work.

The industrial metaverse is a prime example of combinatorial innovation: the creation of new things through the mashup of different technologies with varying degrees of maturity. While long-term implementations of industrial metaverse, and industrial metaverse ecosystems, are expected to fuse a dozen or more key digital technologies, the most important technological pillar are:

- Secure low-latency edge and cloud hybrid networking, including public and private 5G networks
- Knowledge graphs to store entity relationships
- Data contextualization engines to automatically populate and refresh knowledge graphs (e.g., by recognizing things and events in the physical world, and linking them to contextually relevant data in the metaverse)
- Event streaming architectures for real-time notifications and ambient awareness based on relevance
- Computer vision (CV) and natural language processing (NLP)
- Explainable AI (while AI techniques will never be fully transparent, advances in physics-guided ML are greatly improving model accuracy as well as trust)
- Non-stationary robotics to automate data collection and routine inspections
- Blockchain and digital currencies for secure and frictionless transactions
- Virtual assistants to help navigate the information overload.

But wait, isn't the above equally applicable to good old digital twins?

Google Maps remains the most impressive example of the metaverse – and digital twins alike

Yes, the industrial metaverse is so close to its better-known twin – the digital twin – that making a formal distinction is splitting hairs.

Central to metaverse experiences is immersion in a virtual environment not only accurately mirroring the real world but providing an analytically enhanced version of it. As such, Google Maps remains the most impressive example of metaverse.

Industrial metaverse outlook based on a decade of digital-twin experiences

As we await the singularity moment – the point in time when we spend more of work life within the virtual than the physical workplace – of the industrial metaverse to arrive, there are some learnings we can confidently extrapolate and layer onto the likely evolution of the industrial metaverse. The industrial metaverse will develop through a series of overlapping stages:



The industrial metaverse is a prime example of combinatorial innovation: the creation of new things through the mashup of different technologies with varying degrees of maturity.”



Metaverse Islands (Miniverses)

2022-

- Use cases mostly isolated
- access limited to certain devices and locations
- No economic dimension beyond own operational cost savings for data & models
- Disjointed edge and cloud

Connected Metaverse Ecosystems

Around 2025

- Fleet level use case scaling commonplace
- Interoperability protocols emerging across miniverses
- Ubiquitous access from any device, anywhere
- Data and model product marketplaces of digital assets being pioneered
- True 2-way connectivity between edge and cloud

Autonomous & Intelligent Unified Metaverse

Around 2030

- Most of physical world has been digitally mapped and contextualized, allowing deep integration of the physical and digital worlds
- Interoperability across miniverses is common, use case scaling across miniverses is out-of-the-box due to interoperability protocols
- Data and model product marketplaces are common, creating direct digital revenue from data beyond own operational improvement
- Ubiquitous access across any connected device, with rich interaction via voice rather than touchscreen part of normative UX
- Non-Stationary robots operate autonomously in the connect physical world

Graphic: Industrial metaverse development outlook 2022-2030



1. Metaverse islands (now)

- Productivity-enhancing use cases are isolated with little interoperability or transferability across assets, even within an enterprise's own industrial metaverse (a "miniverse"); no active connectivity to others' industrial metaverse islands
- Access is limited to certain devices and locations, no real-time remote connection to physical assets beyond upstream sensor data flows
- No economic dimension to industrial metaverse beyond own operational cost savings (no data product, or model, marketplaces)
- Edge and cloud remain disjointed, with industrial metaverse experiences and use cases focused on cloud-based executions

2. Connected metaverse ecosystems (around 2025)

- Use case scaling at the fleet level within miniverses is more and more commonplace
- Protocols start to emerge to facilitate interoperability across miniverses
- Access to industrial metaverses is ubiquitous from any device, anywhere
- Data and model product marketplaces enable monetary assessment and trading/exchange of digital assets becoming common
- Two-way connectivity between edge and cloud, creating a more dense intersection between the physical and digital worlds

3. Autonomous and intelligent unified metaverse (late 2020s through early 2030s)

- Most of the physical world has been digitally mapped and contextualized, allowing deep integration of the physical and digital worlds
- Interoperability across miniverses is common, as is use-case scaling
- Data and model product marketplaces are common, creating direct digital revenue from data beyond own operational improvement
- Access is ubiquitous across any connected device, with rich interaction via voice rather than touchscreen part of normative UX
- Non-stationary robots operate autonomously in the connected physical world



Industrial metaverse in action: a scenario

Even with today's digital twins, engineers and SMEs typically have to travel to the site to perform maintenance or operations on critical assets. This is time consuming, expensive, and often inefficient, as without an accurate understanding of the issue before arrival, the correct tools and parts for repair may be unavailable.

Within the industrial metaverse, domain experts can be called onsite virtually from different locations to collaborate or solve an incident, using headsets to check the condition of assets, monitor machine and equipment data, and take corrective actions in real time, remotely.

Multifunctional teams, from operations, engineering, and procurement, for example, can discuss equipment or production problems, or review the energy consumption and water usage of a facility. The immersive environment of the industrial metaverse will enable visual investigation of an incident as it emerges, taking corrective action before costly downtime or substandard product is

produced. Precise, virtual representations of the plant, asset, or machine can be simultaneously enriched with live feeds from cameras and robots. Access to this real-time data feed can guide the remote control of robots and drones that read both analog equipment (such as gauges) and digital instruments (stream videos or pictures), and perform maintenance operations as required.

The metaverse will evolve into an autonomous and intelligent immersive environment, blending automation and human intuition to quickly solve problems. Industrial companies will gain visibility of and control over operations that contribute significantly to the bottom line, and provide the transparency necessary to measure, monitor, and control their carbon footprint.

Don't repeat the pilot purgatory of digital twins. Design with scalability and money in mind from the start

While the industrial metaverse may not be so different from a digital twin, we should take caution to not repeat the pilot purgatory of digital twins. This means design with scalability and monetization in mind from the start. Most importantly, design for:

1. Ubiquitous access from any connected device for all employees
2. User experience over technical sophistication hardly anyone can use
3. Substantially non-linear scaling of cost of use case
4. Data and model product marketplaces.

INNOVATION TAKE AWAYS

#MAKEDATADOMORE
#DATAOPS
#INDUSTRIALMETAVERSE
#COGNITEDATAFUSION

The metaverse evolves

The metaverse is an evolution of today's internet, enhanced and upgraded to consistently deliver 3D content, contextually organized information and experiences, and real-time synchronous communication.

The industrial metaverse is serious

Its industrial orientation, rather than on a social, consumer, and gaming context, is set to impact how – and where – we work.

The industrial metaverse develops in stages

Development steps will overlap each other: metaverse islands (current status), connected metaverse ecosystems (next two to three years), autonomous and intelligent unified metaverse (late 2020s through early 2030s).



RESHAPING THE REAL AND DIGITAL WORLD WITH METAMOBILITY



RODRIGO MAIA

Vice President, Head of Research,
Technology & Innovation,
Capgemini Engineering

The concept of mobility is increasingly influenced by technology. This dichotomy is evolving as the concept of personal mobility and personal space are merging and replacing the idea of having to be physically present in a place to really influence it. Welcome to the world of metamobility.

Innovations in mobility will define how we move and what we do while we move. Technology is evolving to shape the way we experience mobility. This might come from

physical movement in a real world as well as from experiences enabled by the creation of digital spaces that break down the barriers between physical and virtual worlds.





This article highlights the evolution of the concept of mobility linked to the introduction of a new interconnected world emerging between the digital and the physical that will reshape the idea of travel and personal space. The main drivers of change in mobility include:

- Achieving the goal of a zero-carbon transport system by 2050. We need to cut CO2 emissions from personal mobility in half by 2030 to achieve this ambitious goal. A paradigm shift in the mobility system is needed through innovative and eco-friendly solutions and services.
- The transition of vehicles to automated, connected, electrified, and shared (ACES) mobility for people. This will be the focus in the cities of the future.
- New mobility ecosystem will be quicker, economic, ecofriendly, safer, and efficient, improving people's health, overcoming socio-economic issues and environmental challenges, and creating energy efficiency.

Envisioning the future of mobility

It is easy to imagine a world of shared mobility. Mobility-as-a-Service (MaaS) is an evolving concept focused on moving consumers and businesses away from vehicle ownership to service-based transportation.

Hyundai Motor Company [discussed metamobility at CES 2022](#) and [said its goal](#) is for “robotics to enable all kinds of personal mobility, connected to communicate, move and perform tasks autonomously” to create a network of unlimited Mobility of Things (MoT).

The concept of MoT, in which traditionally inanimate objects gain mobility with connectivity and by harnessing robotics, is planned to be implemented through the Plug & Drive (PnD) and Drive & Lift (DnL) modular platforms as all-in-one solutions for this new Hyundai MoT ecosystem.

Mobility changes will also be personal paradigm shifts. The journey between points could be more than just a trip; it could mean an extension of everyday life, focusing on learning, work, entertainment, or productivity.

The future is now with metamobility

With the growth of the metaverse, new connected concepts are emerging. Hyundai explores metamobility in a recent video called “[Expanding Human Reach](#)” in which robots act as a link between real and virtual worlds. This was strongly influenced by the synergy brought by [Boston Dynamics](#). This will allow users to see the decisions they make in the metaverse reflected in reality and relies on the confluence of robotic technology, autonomous driving, and advances in AI (artificial intelligence).

In this universe of technological mobility evolution, the transport vehicle is seen not only as a mere space to carry people, or community modules (which aggregate a set of personal vehicles), but as an intelligent space that allows a connection to the user's personal life as a platform for interaction with the metaverse.

Ingredients of metamobility

The concept of metamobility is based on the introduction of a digital, intelligent, and connected environment in the vehicle cabin, using virtual reality (VR) and augmented reality (AR) in the construction of multiple layers in the metaverse that enable a truly immersive experience. To extend the richness of this virtual immersive environment, physical ambassadors (PA) complemented by sensors and environment interaction capabilities (e.g., haptic robotic arms) allow scanning of the surrounding real area and make it available in a digital universe. These PAs will also be an extension of the user's own physical senses, able to perceive certain stimulus and express touch and even allow the user to interact with the environment where the PA stands. These units will also transmit the data received to the users, either atmospheric or physical by, for example, recreating the wind present through HVAC equipment.

This interaction between metaverse and robotics, called proxy experience, could have multiple uses. We can drive a car remotely, control a fleet of trucks by a centralized system, take control of machines in smart factories while

doing several tasks and controlling different types of equipment, thus making space for new services to emerge.

How to leverage this new environment market

The concept of mobility will mold itself to the ever-changing needs of consumers and there is enormous growth potential for sustainable change and efficiency, especially with automotive electrification, a subject we have already discussed in [Data-powered Innovation Review | Wave 2](#). Also, software is impacting personal mobility and its disruptive concepts, creating an area of incredible exploration value.

Talking about the future of mobility means talking about autonomous vehicles, clean energy, and sustainable travel at reduced costs. According to a recent study, "\$70 billion (has been) invested in the first half of 2021 alone" in technological innovation for mobility. We foresee that the demand for new innovative forms of mobility will increase, although the timing is uncertain as there is a huge multi-sector potential to be harnessed.

The metaverse is also present in the gaming industry and is starting to

enter retail, automotive, and real estate, where large financial movements are beginning, such as the sale of "The Sandbox" (land in the metaverse) for a record \$4.3 million. This change is likely to affect more sectors, moving from insurance and healthcare to energy and media, in both public and private sectors.

The way value is created in this emerging ecosystem must be reconsidered on a large scale but will be incredibly vast and profitable.



Mobility changes will also be personal paradigm shifts. The journey between points could be more than just a trip; it could mean an extension of everyday life, focusing on learning, work, entertainment, or productivity.

INNOVATION TAKE AWAYS

#MOBILITY #METAMOBILITY
#METAVERSE
#DIGITALTWIN

All about the journey

The journey between points could be more than just a trip; it could mean an extension of everyday life, focusing on learning, working, entertaining, or productivity.

Mobility, the sustainable way

The concept of mobility will mold itself to the ever-changing needs of

consumers and there is enormous growth potential for sustainable change and efficiency.

Don't move so fast

New concepts of mobility might not involve physical displacement to another location, as for instance an alternative physical immersive experience provided by metamobility.

TURNING THE UNDERBANKED INTO PROFITABLE CUSTOMERS – THE SUSTAINABLE WAY



SHANKAR
RAMANATHAN

Senior Director, Head of AI & Advanced Analytics, BCM, Capgemini

Financial service organizations are aspiring to be inclusive in extending credit to the unbanked and underbanked, and an explosion of data sources makes it imperative for banks to build a framework to evaluate data continuously for these decisions. Optimizing the data used and reducing data costs in real time will minimize the carbon footprint of their operations.

Between 15 and 34 percent of US adults were unbanked or underbanked as of 2017, depending on how those groups are defined. Worldwide, an estimated 1.7 billion adults were unbanked. The unbanked are those with no relationship with a bank, while the underbanked are those that may have one account with an institution. This group has traditionally been seen as relatively costly, risky, and unprofitable. Now, these segments are of interest due to an emphasis on financial inclusion and as a business opportunity. Estimates of potential revenue vary but, in the US alone, underserved consumers spent \$189

billion in fees and interest on financial products in 2018.

The key enabling element to the solution is the ability to manage and use data optimally, and effectively automate credit assessment in an environmentally responsible manner. Financial-services organizations will need to use a broader range of data to accurately assess these customers with modified processes and practices, to turn this segment into fully banked, profitable customers.



Finding the right data

The underbanked and unbanked are not represented in traditional data sources. Thus, in order to extend credit and support account onboarding activities (KYC and CIP, etc.), financial service companies will need to pull data from a variety of alternate data sources such as utility payments, rent, buy now-pay later loans, and debit bureaus.

Accessing many sources slows response times, can be expensive, and ultimately increases the cost of originating loans and providing service to the unbanked/underbanked. Financial-services companies need a systematic way to determine which data sources to use, the order in which they should be accessed, and the optimal combination of data sources. That is, they need a process for narrowing

down the huge universe of potential data to focus on the right lean data. Financial-services companies can establish a framework for determining the value of various sets of data and deciding which sources to use. This should be based on two key factors: coverage and filter value. Coverage means the breadth of the data source: what percentage of the targeted segment the source will address. Filter value describes the depth of the data: the degree to which the data source will support effective business decisions.

Filter value is useful to determine the sequence in which multiple data sources are called. For example, if data source A filters out 80 percent of credit applications being considered while data source B filters out only 70 percent, decision makers can call on source A first and

then draw on source B to fill in the remaining 20 percent. (This assumes 100 percent of the segment has similar credit amounts involved; a more complicated framework can be developed if a population segment has significantly different average credit amounts.) Together, coverage and filter value can determine the predictive value of data from given sources or combinations of sources.

This framework can be the basis of an AI-driven process for sorting through potential data sources to develop the required insights using the least number of sources combined in the most effective combinations.

For example, to assess incoming credit applications, a process as defined in the diagram would be needed.

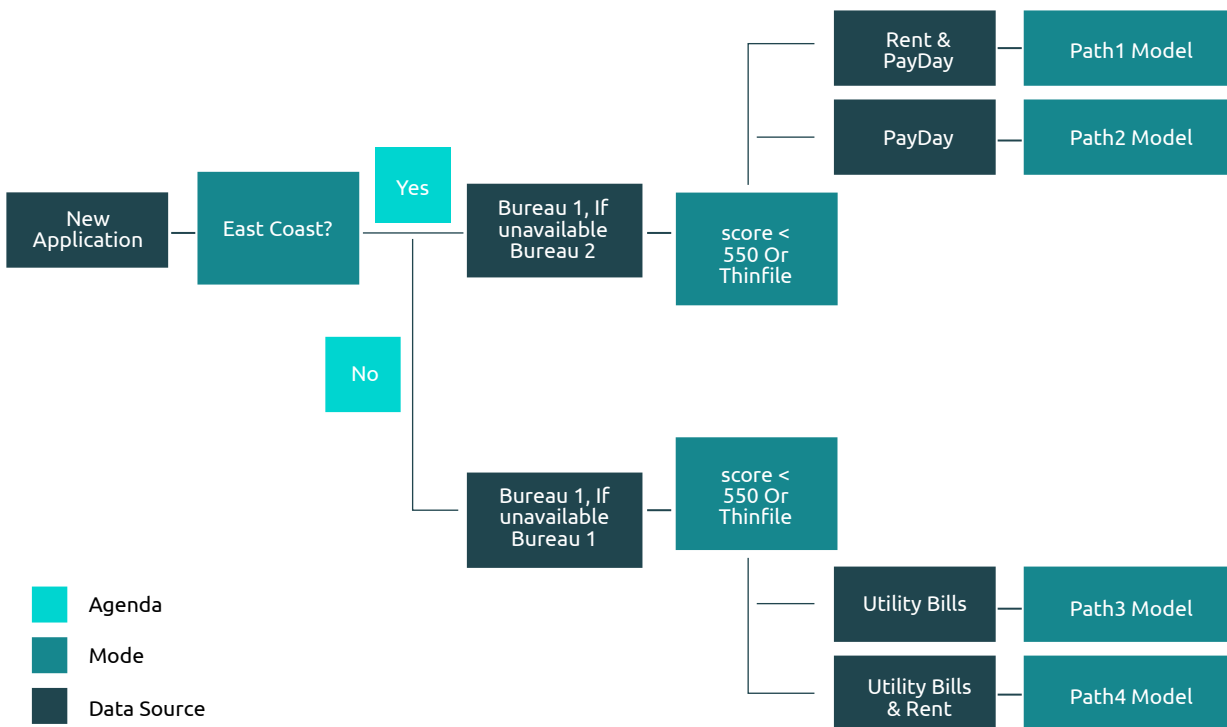


Figure 1: Data-sourcing analysis process, hypothetical example

In addition to reducing data costs, this framework can be used to help negotiate prices with data vendors based on the value to the institution.

The key features of the infrastructure

Financial services companies will need an infrastructure for handling and combining this multisource, nontraditional data, along with traditional data on accounts, demographics, and so forth. This infrastructure should be cloud based because it will need to house data from various and often competing data sources.

Data sources are constantly evolving, so the infrastructure will need to support the continuous monitoring of the value being added by various types of data, along with the evaluation of new data sources and new combinations of data from competing sources, such as various bureaus, and complementary sources that provide additional perspectives, such as utility and rent data.

The infrastructure should enable companies to create models using these combined sets of data, and easily deploy them to assess credit risk, fraud, etc. Machine learning (ML) models are a natural fit for

these tasks, given the large number of attributes involved, and these should be built and executed in a real-time environment.

The data used for these processes is likely to fall under regulations such as the Fair Credit Reporting Act and Regulation B in the United States, which means that data will need to be anonymized to mask personal information. This should be done through a common process across sources and include the creation of a unique ID for each specific consumer to link data drawn from many different sources. (See the figure.)

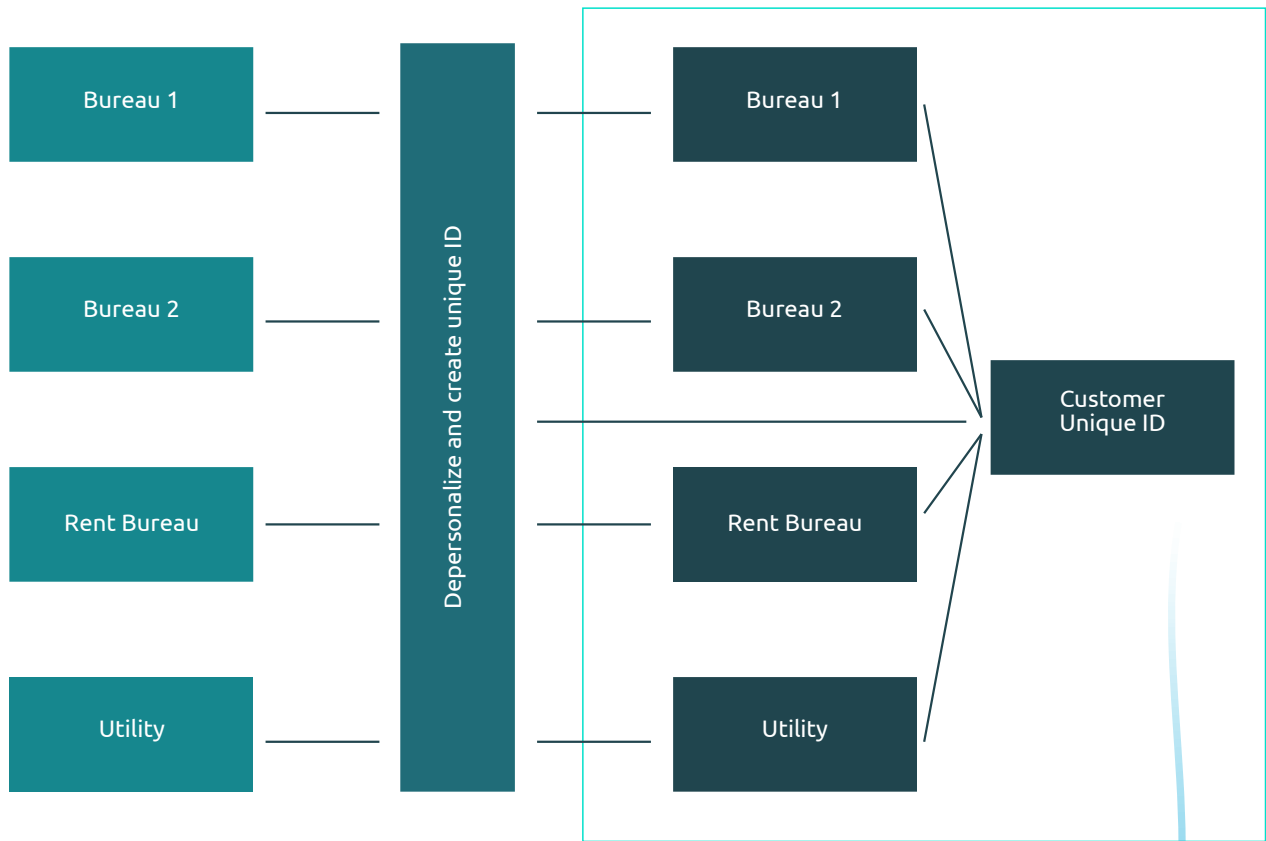


Figure 2: Creating unique customer IDs to link anonymized data



An end-to-end approach to reach the unbanked/underbanked

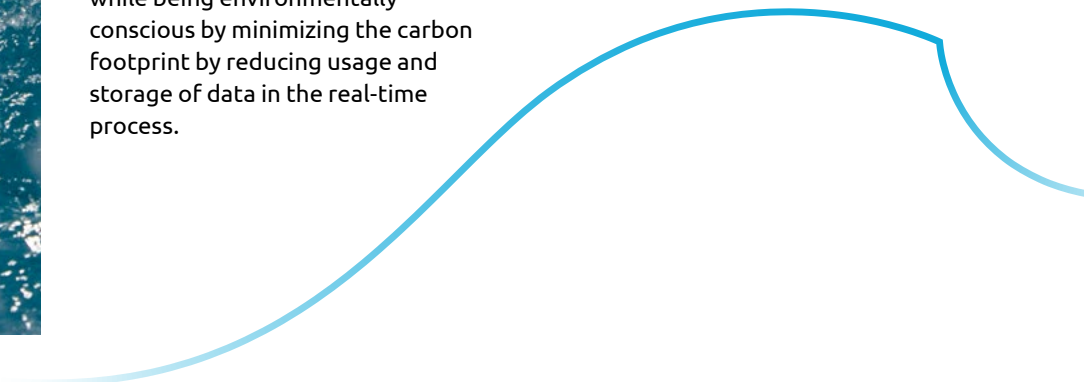
This approach can provide the end-to-end capabilities needed to successfully serve the unbanked/underbanked. With this platform, organizations can:

- Ingest data from many sources, optimize the combination and sequencing of data sources using a value framework and ML
- Monitor performance of the data value continuously and evaluate any new data source
- Continue to adapt to shifts in customer segments, new data sources, and new AI techniques.

With a value framework that helps them sort through a growing universe of data, financial-services institutions can serve this segment while being environmentally conscious by minimizing the carbon footprint by reducing usage and storage of data in the real-time process.



The key enabling element to the solutions that target the unbanked/underbanked is the ability to manage and use data optimally, and effectively to automate decision making about credit worthiness in an environmentally responsible manner.”



INNOVATION TAKEAWAYS

#AI4GOOD #LEANDATA
#EVALUATEDATA
#REDUCEDATAFOOTPRINT
#UNDERBANKED
#FINANCIALINCLUSION

Expand your data universe

Build the ability to handle competing and complementary data sources – as an addition to established data-source management.

Apply a value framework

Continuously evaluate the value of your various data sources.

Leverage AI and ML

Minimize data cost and storage (carbon footprint), and to adapt to shifts in customer segments.



THE ROLE OF SELF-SERVICE BI FOR BUSINESS AGILITY



MYLES SUER

Director of Solutions Marketing at Alation

The move to self-service BI is driven by an organization's need for agility in support of a hybrid workforce. But this requires data accessibility for every worker. Let's look at how to best deliver the potential of self-service BI, demonstrating how an innovative business-centric catalog puts data at the fingertips of decision makers.

The Capgemini Research Institute surveyed 500 organizations and 5,000 employees around the world and spoke with academics and executives, and found that remote working is definitely the new normal: 75 percent of the organizations expect at least 30 percent of their employees to work remotely, while more than one third expect 70 percent of their workforce to

become remote. With such a large portion of the workforce working outside the office, the walking-down-the-hall method for gathering data expertise no longer works. Without those in-person interactions, however, 65 percent of workers now feel less connected to their coworkers. Businesses must recreate that connection virtually, especially where data is concerned.

Everything moves faster

Organizations need to be more agile and reduce decision-making cycle times. To deliver, businesses are enabling workers to make more decisions whenever they are needed. But this requires quick and frictionless access to the right data at the right time. And they also need

to trust that the data is current, relevant, and available.

Easy, right? No, it's not.

Moving a business forward requires fast access to good data for more applications. But that turns IT and analytics teams into the bottleneck.

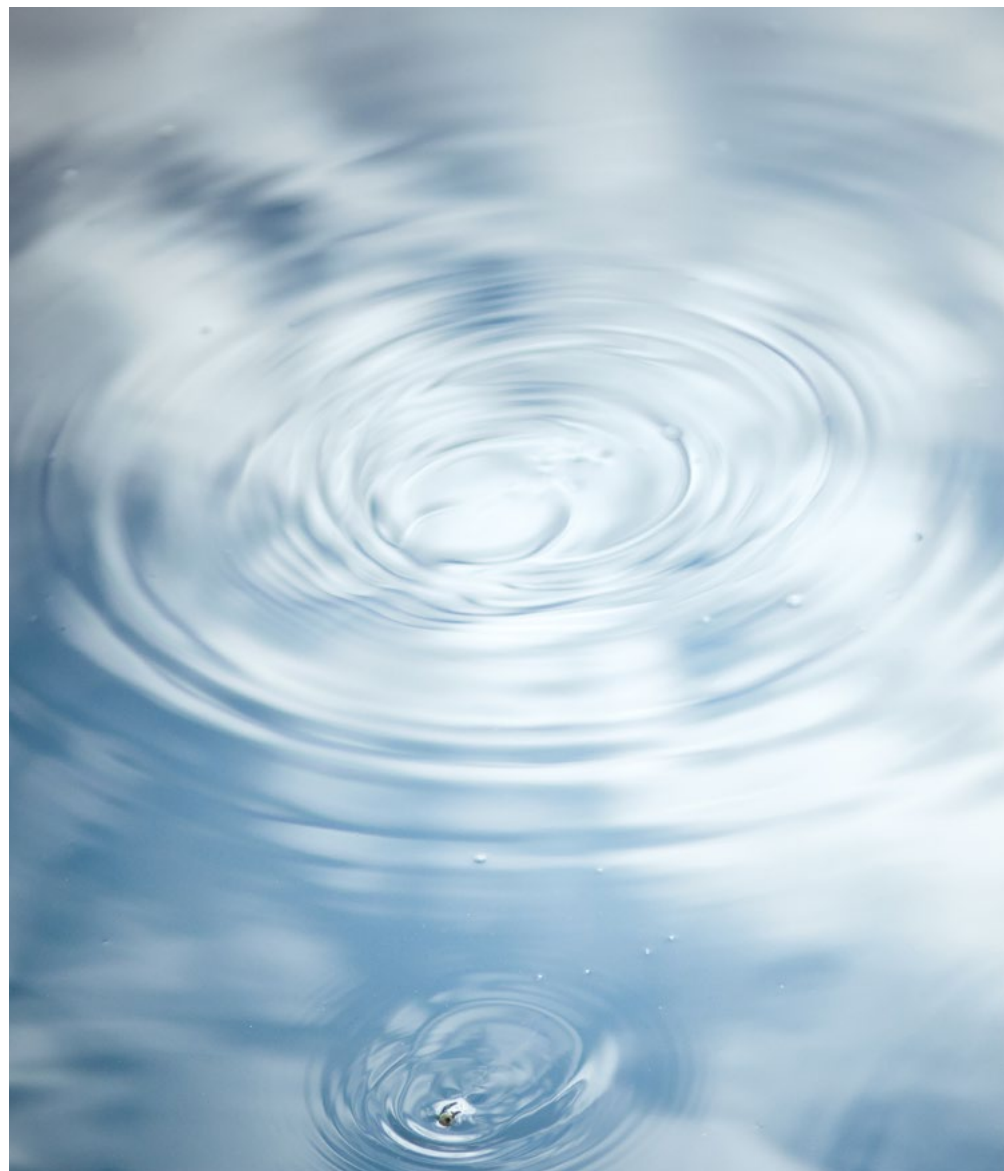
Let's face it, the inquiry model – ask for insights, wait weeks for answers – does not support today's pace of business. That was fine when decisions were made by a few executives at the top of the organization. Today, it's unacceptable.



The inquiry model is outdated

Leading companies, however, empower the middle of the organization to speed time to market, push digital acceleration, and maintain a competitive advantage. The traditional inquiry model simply does not provide answers fast enough. If you're too slow, you lose time, money, customers, market, and maybe your job. The goal should be to allow

Enter self-service BI. It drives faster decisions, more innovation, lower costs, transformation at scale, and improved quality, safety, and efficiency. But how to get there, and become a self-service data master?



Enabling self-service BI

A more participative self-service BI environment needs to be encouraged. Unfortunately, for many organizations, this change is slowed by internal issues, fiefdoms, and siloed data and systems. But for smart organizations that have sorted out data access and sharing requirements, self-service BI drives data literacy. The goal should be to build a culture where people seek to understand data and its context.

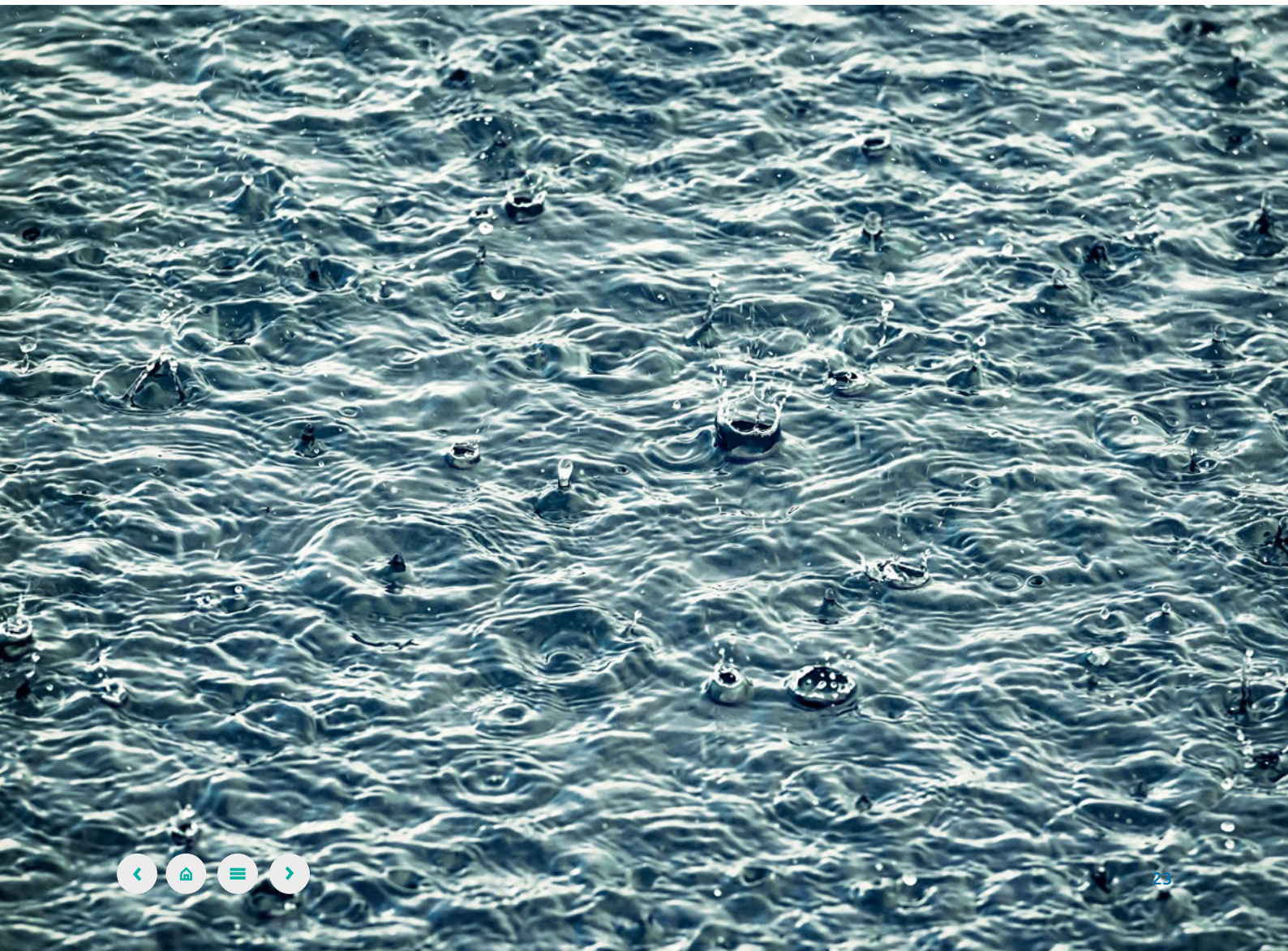
Putting data at every business users' fingertips is the essence of self-service BI. Like the concept of a data mesh, self-service holds that people closest to the data – the data producers – should make the data available. Those who need the data – data consumers – can then access it whenever.

For self-service BI to succeed, however, the entire data value chain may need to be fixed. Data has to be easy to find, understand, access, and use for everyone in the chain: data engineers, analysts, data scientists, and business users. Productivity along the chain can be enhanced with a data catalog, which is a repository of metadata on information sources from across the enterprise, including data sets, business intelligence reports, visualizations, and conversations. It makes the data more accessible and understandable to everyone, especially less-skilled data consumers. It also prevents requests for insights that already exist or questions that have already been asked.



One insurance customer determined that just having a data lake didn't by itself generate important business

insights. It needed to teach people to fish in the data lake. They needed to move from a culture of analysis to a culture of reporting. A key element of doing so was implementing a data catalog. The company is in the process of using catalog to data nearly 12,000 employees so they can answer questions that drive to the best next action. Who are the customers that I should call today? And what tasks should I complete today? With claims employees, they saw increased efficiency with self-service information. And on the business analyst side, they saw 25 percent time-savings due to decreased data inquiries.



A data catalog for trust

Many technologies are needed to deliver self-service BI. First, though, users need to find the data. That data discovery – and understanding the context of the data you do discover – is critical for the typical data consumer. It's also important to clearly understand when data is not available, and when data is old, incomplete, inaccurate, or otherwise questionable. For this reason, it is not surprising that recent research by Capgemini has found a massive trust gap between the IT-facing arm of organizations and business units.

CIO David Seidl says, “as a user, a highly usable data portal or access tool including data discovery and contextualization is critical for more casual, non-power users. I think that’s the real destination of self-service BI in the long term.”

A data catalog does this all by delivering data discovery, contextualization, and user-friendly tools for casual, non-power users. Specifically, a data catalog enables any user, regardless of skill set, to find and understand data via natural language instead of SQL queries.

At the same time, a data catalog provides a business glossary to convert technical jargon, obscure field names, or complex database nomenclature into easy-to-understand business terms.

What a data catalog really delivers to data users is trust. It does this by:

- Surfacing learned and collaborative data recommendations
- Flagging potentially sensitive data
- Integrating data quality scoring
- Collecting popularity rankings, user recommendations, and usage recommendations
- Flagging data health and policies to avoid misuse and compliance issues.

A data catalog enables self-service BI, seamless data collaboration, integrated communication, and the sharing of internal expertise, all built on a foundation of trust. This empowers workers to explore data and discover the answers on their own. And, when they can't, it points them to the resources and people that can help.

The next wave of catalog and self service

The next wave should aim to unlock all the enterprise data using AI. This will empower knowledge workers to explore data before a report or analysis exists, and then drill into data and discover the answers on their own. And when they can't, it points them to the resources and people that can help. This step will sustain knowledge workers in an increasingly hybrid work modality. The combination lets AI analyze data and surface insights while Natural Language Processing (NLP) allows users to build on top of those surfaced insights and ask the next questions. This will give these workers immense power.



“This not only empowers the business but also makes them more accountable for driving data-powered decision making.”

INNOVATION TAKE AWAYS

#SELFSERVICE
#DATACATALOG
#DATACOLLABORATION
#TRUST

Self-service leads the way

An increasingly distributed and digitally enabled workforce needs self-service BI, rather than solely relying on central services.

Understanding data is key

Those serving themselves need not only access to data, but also to the data about the data (metadata).

The data catalog is foundational

A data catalog provides a trusted, empowering foundation for self-service BI.

AI drives the next generation

Artificial intelligence augments current data-catalog functionalities, making self-service BI even more accessible to more people.



AI-INFUSED INNOVATION FOR AUTOMOTIVE DATA



JEAN-MARIE LAPEYRE

Chief Technology & Innovation Officer, Global Automotive Industry, Capgemini

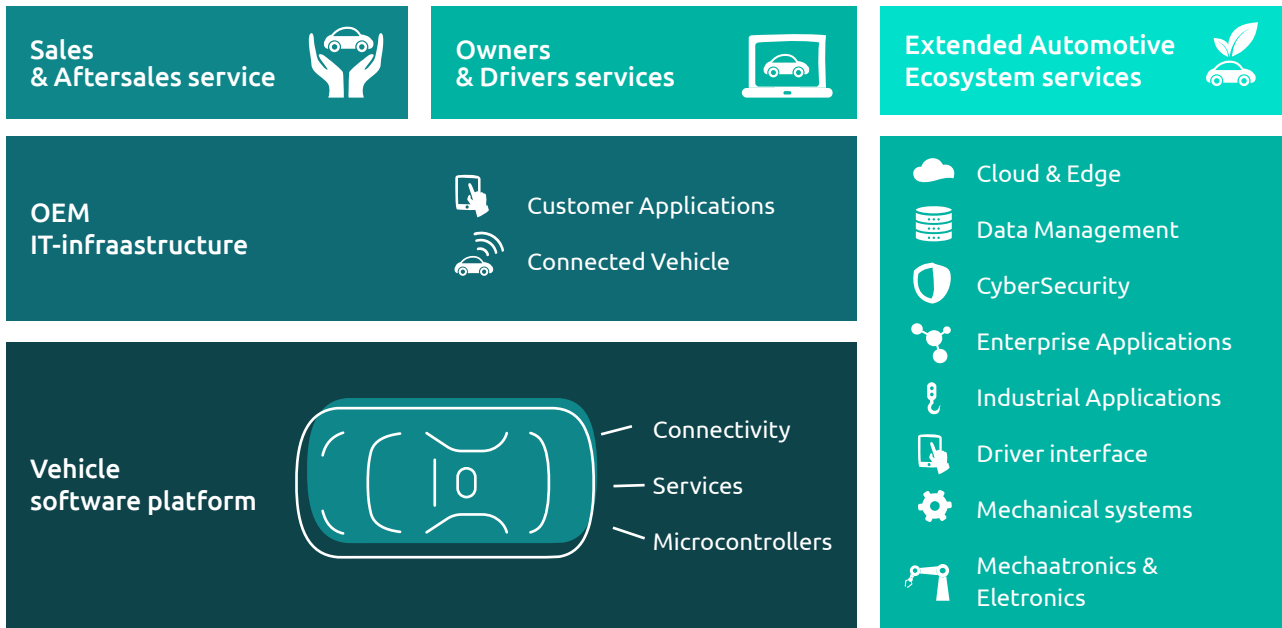
The automotive industry is encountering unprecedented disruptions, and it is rapidly transforming in response. A frugal approach to innovation is essential to this – and applying AI to existing data is a prime example.

The automotive industry is innovating rapidly in response to multiple disruptions. To focus first on the customer perspective, vehicle buyers increasingly want the whole purchasing experience to happen online, from initial research right through to buying. The ownership model, too, looks set to change, with

increasing use of carpooling, ride-hailing services, short-term rentals, and community fleets. And relationships between manufacturers and customers are being extended: drivers will continue to receive services and over-the-air updates throughout their vehicle's life.



Customer First



Intelligent Industry

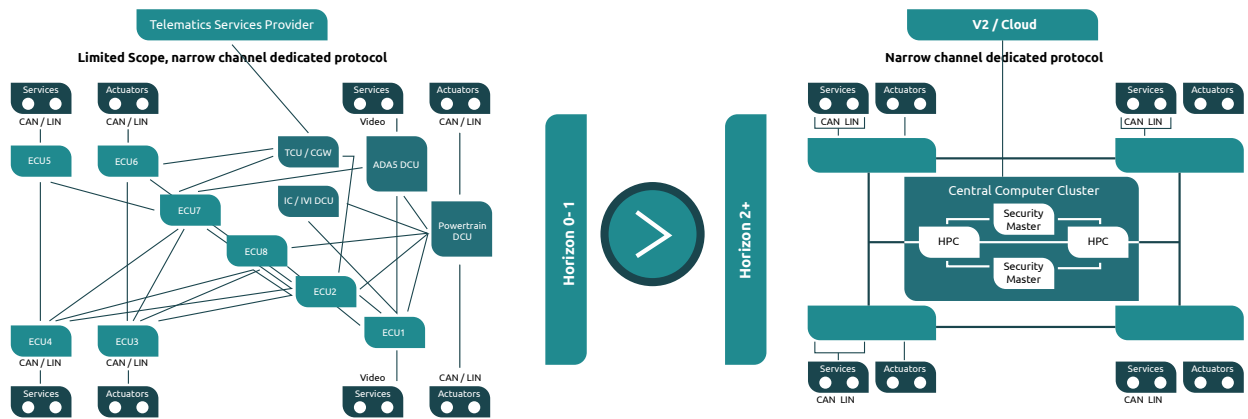
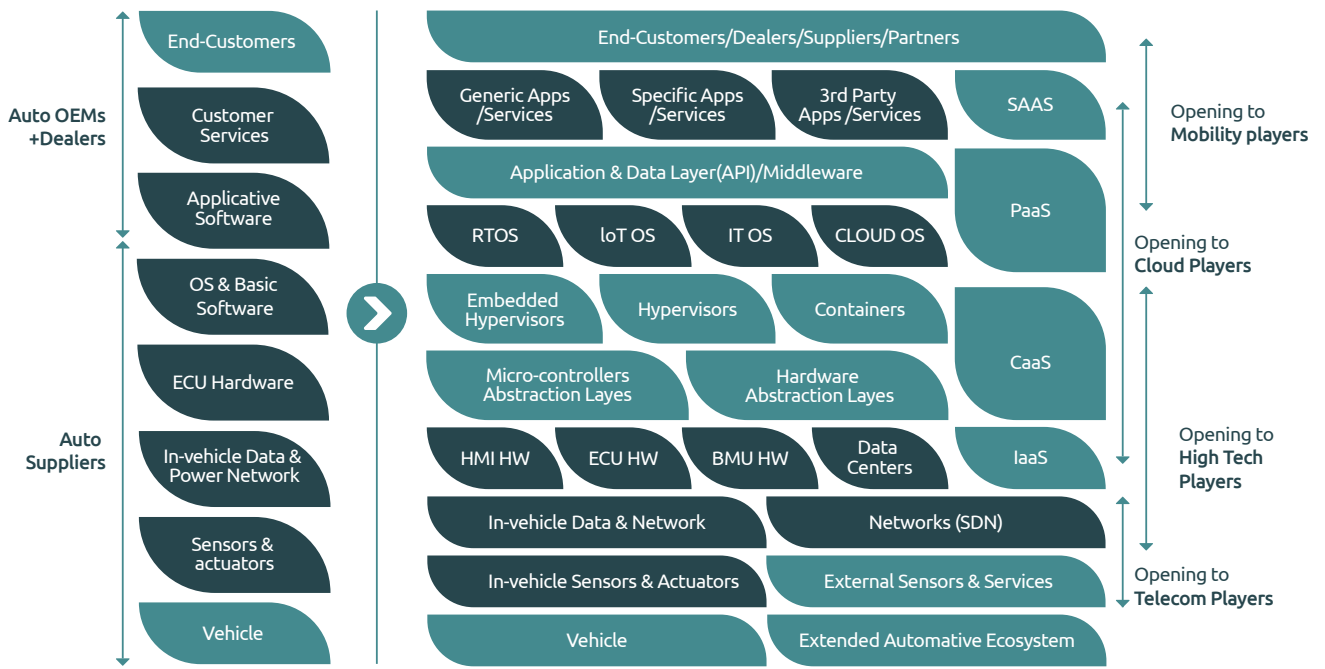
The three playing fields in vehicle software development

Turning to the product, electric vehicle (EV) sales are booming, and additional ways are emerging to make the product more sustainable, such as rejuvenating older models by fitting electric engines. Individualization is increasing, together with a trend to integrate the car into the customer's digital life. More tasks are being automated and user interfaces are becoming more intuitive and streamlined.

In terms of technology, products, industrial environments, and enterprise processes are increasingly defined, designed, and operated around software. Soon, vehicle capabilities will be determined by software running on generic, integrated hardware. Vehicles will be permanently connected and will interact with infrastructure elements such as traffic lights. In the factory, industrial processes are harnessing computing power attached to each device, as well as

high-speed connections to centralized capabilities.

Meanwhile, the industry's ecosystem is being disrupted by new entrants, including EV-only startups and established technology companies from other industries. Supply-chain disruptions and resource shortages arising from the pandemic are being worsened by the tragic events in eastern Europe, and companies are having to rethink their supplier dependencies.



Vehicle data flow

Looking finally at the culture perspective, manufacturers are having to adopt Agile processes end-to-end to keep pace with today's dynamic environment. They also need to blend skills from traditional engineering and IT, competing with other industries for the relevant – scarce – talent.

Frugal innovation through applying AI to existing data

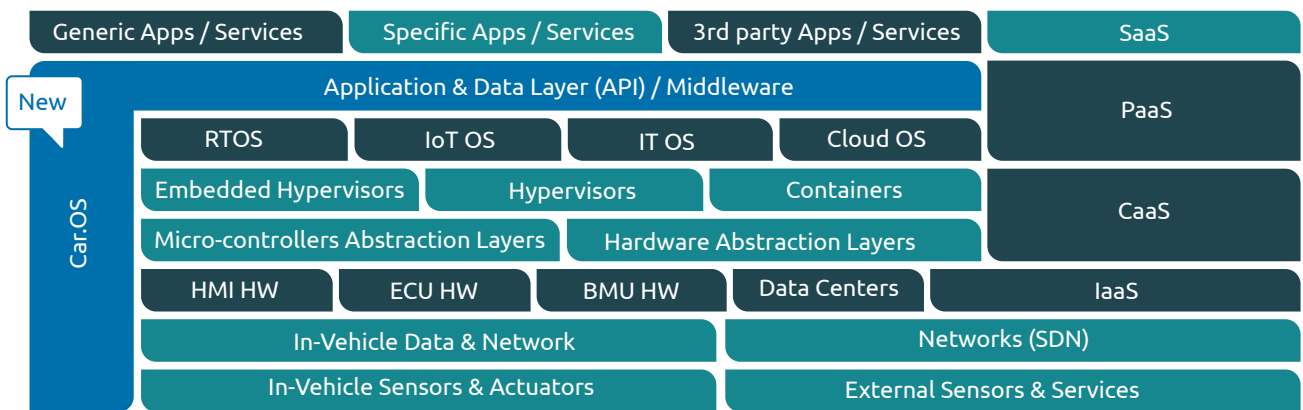
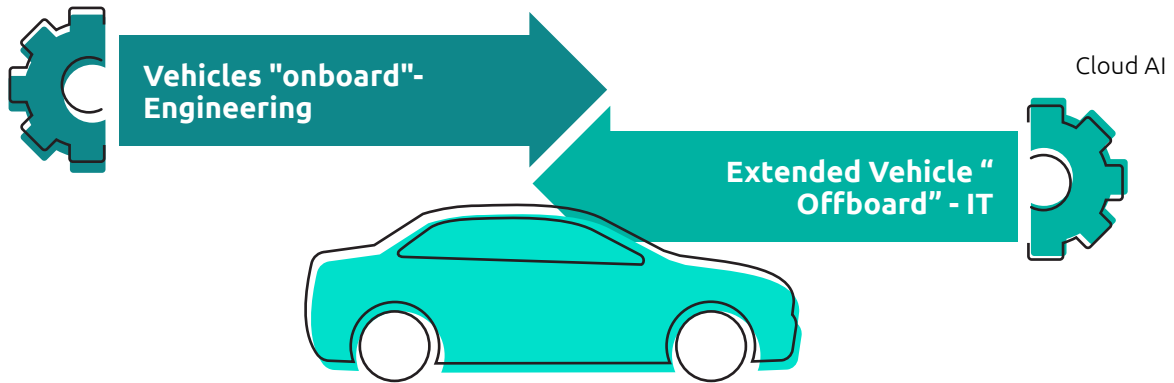
Innovation is critical to dealing with these changes, but innovation does

not necessarily mean invention; often, it's about reusing an idea or a resource in a different context. This frugal approach to innovation, called "Jugaad" in India, is a theme of Capgemini's [TechnoVision for Automotive 2022 playbook](#).

The application of AI is a prime example of how the automotive industry is innovating by making better use of what it has. Of course, AI needs data, and this is where the frugality comes in, because much of the necessary data often exists

already. For example, data generated by designing and building a vehicle was often discarded after the product's completion, but with software increasingly determining which options an individual vehicle offers, its value is clear. AI can help companies make the most of data in a range of contexts.

In vehicle AI



Data service models

Inside the vehicle:

Perhaps the best-known use of AI in automotive is to automate the task of driving. Even if fully autonomous vehicles are still a few years away, Advanced Driver-Assistance Systems (ADAS) features are already appearing. AI can open up a whole world of seamless driver interactions and can support safety, reliability, and robustness.

On the factory floor:

AI is integral to many current production-line automation initiatives designed to increase efficiency and quality. AI-based systems can help to analyze camera outputs, carry out shop-floor quality checks on the assembly line, optimize truck loading to improve space utilization, or power augmented-reality goggles to minimize operator errors.

In the design workshop: With AI, revolutionary propositions can emerge from data, including new elements for use by human designers. AI can also evaluate solutions generated by humans or machines and recommend the most promising.

In the back office:

For strategic planning purposes, AI-enabled processes can assist with rationalizing the choice of vehicle configurations. AI could even help the human resources function because, when talent is scarce, AI can make the most of the people you have.



AI individualizes driver experience

Using AI technology from Anagog, ŠKODA AUTO Digilab has developed Citymove, an app to ensure that each customer gets the right offers. Before a journey, the app helps them decide between using their own car and mobility options such as buses, trams, rental bikes, taxis, and ride-hailing cars – all of which can be chosen, booked, and paid for via the app. The app also proposes personalized offers via notifications pushed to smartphones: anything from local deals on favorite drinks to discounted car washes.

Driving into the future with data and AI

For all this to happen, the industry needs to make sure that the data required to power AI is available in the right form, at the right place, and at the right time. Various data-sharing initiatives are underway to help this happen.

Once we organize the data correctly, and make it available for the right AI applications, the sky's the limit – perhaps literally. AI-enabled progress in drones could lead to the development of cars that are not only autonomous but also hover in the air. Watch this space!



Once we organize the data correctly, and make it available for the right AI applications, the sky's the limit.”



AI helps optimize ADAS

General Motors is assessing the potential of an AI-enabled pattern-recognition technology to accelerate the design of ADAS. The Multi-node Evolutionary Neural Networks for Deep Learning rapidly evaluates convolutional neural networks for use in pattern recognition. This approach could, for instance, reveal ways for cars to quickly and accurately assess their surroundings in order to navigate safely through them.



Sharing safety data

Capgemini has been working with Volkswagen and Audi to demonstrate the value of the German Federal Government's Mobility Data Space, of which Volkswagen Group is a founding member. An early use case is Local Hazard Information, which provides aggregated event data on traffic hazards collected from vehicle sensors in the Audi fleet. This data could be used by a navigation service to warn road users of upcoming danger spots in near real time.

INNOVATION TAKE AWAYS

#AIENABLEDINNOVATION
#AUTONOMOUSDRIVING
#ADVANCEDDRIVER
ASSISTANCESYSTEMS
#FRUGALINNOVATION

Automotive Innovation should be Frugal


Find new ways to use ideas and resources you already have. Applying AI to existing data is a great example.

Powered by Automotive Data, AI knows no limits

Data-powered AI can transform both vehicles and the companies that make them.

The right automotive data at the right time

Much of the data's already there, but it needs to be made available where and when it's needed for AI to use.



THE RISE OF AI AUGMEN TATION

**Humble AI and the four
pillars of trust** 31

Ted Kwalter, DataRobot
Yashowardhan Sowale, Capgemini

**Serendipity systems:
architecting
personalization
systems at scale** 34

Neerav Vyas, Capgemini

**Quantum computing:
the case for
machine learning** 39

Nadine van Son, Capgemini
Julian van Velzen, Capgemini

**Using AI to
fight cyber threats** 42

Jerome Desbonnet, Capgemini

**AI-augmented
field service – the future
of maintenance** 45

Nicolas Claudon, Capgemini
Guillaume Gerards, Capgemini

**AI panel: How can AI help
create a more inclusive
world?** 49

OpenAI, AI21 Labs, Cohere
Rajeswaran Viswanathan,
Capgemini



HUMBLE AI AND THE FOUR PILLARS OF TRUST



TED KWALTER

VP of Trusted AI, DataRobot



YASHOWARDHAN SOWALE

VP, I & D Architecture Head, India Domain Leader for AI, Capgemini

The quest for trusted AI is becoming critical as the technology is becoming more entwined with business operations. But how do you build AI systems that can be trusted? How do you keep AI humble and cognizant of its own limitations and caveats? It turns out there are four key pillars to consider.

AI has undeniable benefits to business and organizations, in disciplines as diverse as supply chain, food security, reduced carbon emissions, and others. But more needs to be done to ensure AI and machine learning aligns to organizational and societal ideals. Multiple media outlets have covered AI's missteps, including unfair hiring practices, over scheduling low-wage workers, or inadvertent race-based digital advertising for housing. The main cause of such glaring incidents is that AI solutions do not admit that

a problem is outside their sphere of action – hence we need to find new ways of 'taming' the AI, making it more "humble".

Whether it's called "#humbleAI", "#responsibleAI", "#trustedAI" or #ethicalAI, the aim is to mitigate risk in AI systems. For AI practitioners, the hurdle is not putting a model into production, it is doing so with an eye to stakeholder impacts. This requires organizations to think in multiple dimensions.

An important aspect of AI-driven insights is a trade-off of machines getting more intelligent versus enabling them to take more control and overstepping their authority. While everyone agrees that AI will be transformational for businesses, organizations are still figuring out how to build and rely on it.

Successful AI implementations need AI to be tied to business processes. However, most AI systems give a level of confidence that is based only on narrow data. Humble AI will therefore leverage historical knowledge and learnings from human intervention and integrate new information periodically. It is based on the following four pillars:

1. People

People are the foundation of a successful, trusted AI implementation. All AI implementations involve people. There are four different personas in a successful implementation of trusted AI.

Business leaders must understand the benefits of AI to their business. These AI Innovators may not know how to code but they understand how the technology fits into a process, or new market. In contrast, the AI Creators understand data, modeling, and computer science explicitly. AI Creators write code, test models, and tune them before production. Next in the implementation workflow are the AI Implementers who generally work within information technology to ensure the model architecture. This is often referred to as “MLOps” and requires a slightly different skill set. Lastly, AI Consumers range from internal risk and compliance personnel to employees receiving predictions and even external stakeholders like regulators, applicants, or customers.

In small organizations, the AI Innovator, Creator, Implementor, and Consumer can be a handful of people, while in others entire teams



can be devoted to a single task. All of these parties have to understand the aim of the project and ultimately agree on risk-mitigation efforts to have a trustworthy system. Without alignment among the four, one or more of the parties will mistrust the model, be frustrated by opacity or, worse, the organization will not realize the potential value of the AI project.

Humble AI works best with human-over-the-loop systems. The aim is to balance automation and AI creator participation or intervention, by enabling the system to perform automatically within a zone of competency. To be successful, AI creator intervention requires a constant feedback loop with stakeholders.

2. Process

Humble and trusted AI projects have defined processes. Processes that provide clarity on expectations and

roles will increase an organization’s modeling throughput. Machine learning projects need proper guardrails to identify risks, rate them in an adaptive policy framework according to severity and likelihood, and then have the personas agree on mitigations.

System safety engineering has been employed in many industries where negative outcomes can be tragic. For example, aviation has employed risk mitigation processes for decades with the use of preflight checklists, to be performed by the pilot, while the less technical passengers must listen to standardized emergency landing procedures. Some organizations use the concept of “AI confinement” where AI will use only constrained channels of learning and carefully screened datasets so that safety mechanisms are created based on process guidance.

3. Technology

As organizations seek to deploy AI according to the former stakeholder expectations, they will need a technology that is flexible yet supportive. Custom model development in various languages like R, Python, Scala, and Julia make model auditing and deployment difficult.

An organization should choose technology that can standardize model evaluation across languages during ML development, then enforce workflow approvals during ML evaluations, and finally actively monitor production models in ML operations. Providing a standardized model evaluation, documentation, and “single pane of glass” for all production models is currently beyond many organizations but it is possible with the right ML technology.

4. Regulation

Lastly, organizations need to acknowledge the emerging regulatory landscape for automated decision systems. There are numerous proposals that aim to limit AI’s scope and ensure it is built with societal ideals. Organizations need

to proactively prepare for regulations that may limit AI’s use or require a yet to be determined compliance document to be submitted to a regulator.

At this point, there are three far-reaching AI proposals gaining traction.

- The US House of Representatives has twice proposed an Algorithmic Accountability Act requiring model documentation.
- The European Union is currently revising the proposed AI Act to protect EU citizens against algorithmic victimization. It is expected that the EU’s AIA will lead the way in requirements for building trustworthy AI.
- DARPA is developing algorithms which make decisions aligned with human experts.

Gone are the days of popping champagne when a model was put into production or merely “thrown over the fence to IT.” AI-mature companies must align the right people and have thoughtful procedures to identify and mitigate risks whether to the brand, people, or the bottom line.



Machine learning projects need proper guardrails to identify risks, rate them in an adaptive policy framework according to severity and likelihood, and then have the personas agree on mitigations.

INNOVATION TAKE AWAYS

#HUMBLEAI #TRUSTEDAI
#DATAREGULATION
#AILEARNING #ETHICALAI

People for humble AI

There are four personas that need to align for trusted, humble AI: AI Innovators, AI Creators, AI Implementers, and AI Consumers.

Process for humble AI

Structured process and well-defined guardrails are foundational to trusted, humble AI.

Technology for humble AI

Not all AI technology is created equal when it comes to trust; choose the right, automated tools.

Regulation for humble AI

Regulation increasingly sets directions and perimeters for the responsible use of AI systems



SERENDIPITY SYSTEMS: ARCHITECTING PERSONALIZATION SYSTEMS AT SCALE



NEERAV VYAS

Head of Customer First, Insights & Data North America, Capgemini

Personalization systems are all about the right advice at the right time. When it's spot on, that advice leaves us stunned as to why we never thought of it before. If it's more than spot on, it makes us feel eternally grateful. Welcome to serendipity systems, and the way next-generation personalization engines aim to consistently deliver it.

Think of the last innovative enterprise you interacted with. Can you think of one where some form of recommendation or personalization was not part of the experience? Recommendations are no longer product features, service attributes, gimmicks, or nice to haves. They are the central organizing design principle of modern experiences.

Personalization wasn't an add on for Amazon, Uber, Netflix, or Airbnb. It was core to their experiences, because personalization was core to their business model. We're moving to a world where not doing personalization is a recipe for guaranteeing underperformance and obsolescence. So, how do we do this well?

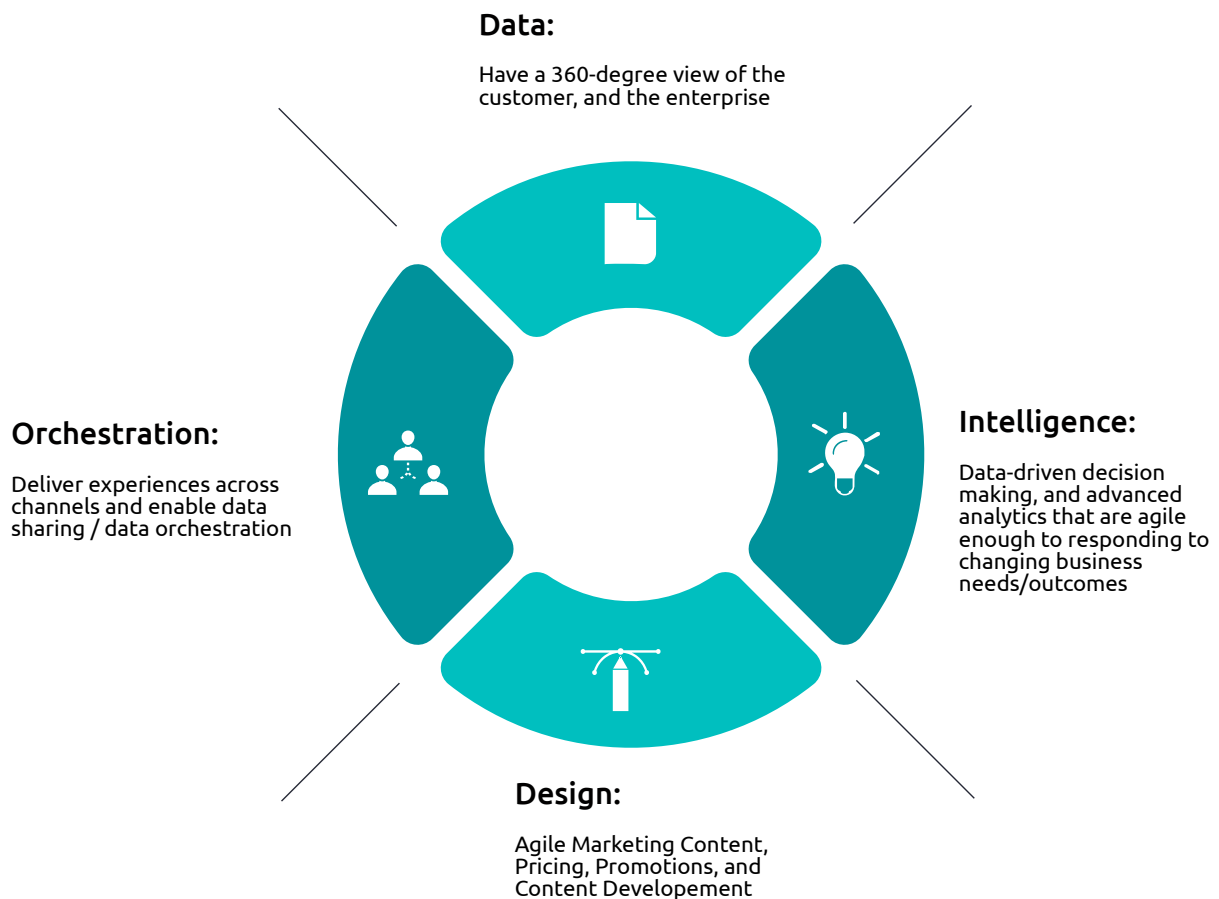
Serendipitous influence

Architecting modern personalization systems isn't just about data and algorithms. Recommendation architectures are those that persuade and influence our choices. The choices they present and how they are presented change not only

how we discover products and experiences but how they promote internal self-reflection. Done well, recommendation and personalization are systems of "serendipitous influence." The best recommendations are those that inspire you in a way you didn't expect. They make you wish you'd

gotten that recommendation earlier. Fostering serendipitous delight consistently becomes the hallmark for enterprises that want to provide delightful experiences.

In our experiences in building such systems, we find there are four key dimensions that help firms elevate from personalization to serendipity:



Graphic: The cycle of activities

Data

1. A consolidated view of "customers" and the enterprise is critical for consistent experiences. Data silos are likely to result in incongruent experiences as customers traverse channels and inconsistent recommendations between content, products, and services.

2. Proliferation of data across applications, devices, and brands makes it critical that data is connected, otherwise the customer is more likely to ask "Why don't you know who I am?"

3. Resolution of unknown to known users is also critical in maximizing the total market of customers that you can personalize for.

producer consumer "data mesh" to democratize its insights on customers. This enables Amazon teams to use data across the entire ecosystem to recognize customers across channels and devices (e.g. from web to mobile, to Prime Video, and to Alexa devices).

Amazon's internal data lake (Project Andes) was the start of creating a

Intelligence

1. The best personalization systems are built from strong data-powered cultures. Literacy in data science, AI, and ML is necessary, but not sufficient. Teams need to be literate in utilizing insights to drive action.
2. Linking analytics to outcomes and a holistic set of KPIs is needed to monitor the health of recommendation systems and understand their impact (intentional or otherwise). YouTube found optimizing on clicks created click-bait recommendations that resulted in poor experiences and less user engagement versus optimizing for time spent watching videos.
3. Rules-based strategies help to “fake it till you make it” but they struggle at scale and often underperform due to biases in rules and an inability to iterate quickly. Personalization systems should allow employees to build campaigns to optimize based on business goals as opposed to rules (e.g., maximize revenue, minimize my carbon footprint through fewer shipments). This allows employees to be more strategic and creative while permitting underlying analytical models to make decisions on data volumes and patterns that would otherwise go unnoticed.

Amazon.com is a different store for every customer. Its Personalization Platform (P13N) allows business teams to set strategies and filters to leverage personalization algorithms optimized against business outcomes (e.g. improve conversion, drive engagement). The system also understands when signals like customer intent are changing. If I searched for shoes in the morning but now I’m searching for toothpaste, my recommendations should understand my intent and needs have shifted, and consumer goods and staples are more relevant than athleisure products. These signals are optimized against business or experience KPIs which allows for dynamic

recommendations to improve business outcomes while boosting the customer experience.

Design

1. Experimentation leads to better personalization. Design is key to better experimentation. The best recommendation cultures view experimentation of systems as an end-to-end exercise from data and analytics to UI/UX development to creative development and qualitative research.
2. Recommendation Experience Design (RXD) as a competency is integral to designing systems that create the proper nudges and in understanding the intended and unintended consequences of personalization. Done well, these teams are a blend of technologists, behavioral scientists, creatives, and experience designers.
3. Volume of experimentation can be a valuable KPI unto itself. “If you double the number of experiments you do per year, you’re going to double your inventiveness.” – Jeff Bezos.
4. Without trust in the system, it’s hard to get adoption. Without adoption the system will not survive. There’s a fine line between helping to make someone aware of a system and navigating their choices and manipulating them. Systems that manipulate are unlikely to survive.

Through user experience research, Stitch Fix found that people couldn’t judge what they would like from clothing images alone. Customers might say from an image they don’t like something, but when they interacted with it or put it on, they would find they loved it. This gave encouragement to be more aggressive in sending products that algorithms suggested a customer would like even if the customer stated they weren’t interested in those types of products.



The best recommendations are those that inspire and delight you in a way you didn't expect.”



Orchestration

1. Orchestrating intelligence and actions for interventions across the customer journey is essential to consistent experiences. Enterprises need a view of the critical journeys (if not all journeys), and this requires organizations to easily orchestrate data and analytics internally and externally to empower employees to improve customer experiences.
2. Employee experiences can be as critical as customer experiences. If employee experience is an afterthought (and poorly done) then adoption of the system will be low and the corresponding pace of experimentation will suffer (if it occurs at all).

Stitch Fix's orchestration of customer and merchandising data enabled the development of Hybrid Designs, its internal AI-driven design group. The apparel designs are a true collaboration built on the orchestration of human intelligence and artificial intelligence that helped Stitch Fix generate 2021 revenue of \$2.1 billion, with over four million active users



Recommendation systems that are thoughtfully architected across these four dimensions drive differentiated and innovative experiences through more experimentation and greater degrees of adoption within the enterprise and by customers. This results in systems that evolve from a goal that's transactional – “Will I buy this?” – to those that make us wonder “How did I live without this?”

What does the future hold? The expectation from consumers is some form of personalization. Moving forward, discovery should be like talking with a friend who knows you so well that they can anticipate your needs. This is a world where we've democratized access to serendipity, and firms should provide such experiences or be left behind. It is either serendipity at scale or obsolescence with haste.

INNOVATION TAKE AWAYS

##PERSONALIZATION
#SERENDIPITOUS
PERSONALIZATION
#MACHINELEARNING
#ARTIFICIAL INTELLIGENCE
#INNOVATION #RESPECTFUL
PERSONALIZATION
#ROOTCAUSEANALYSIS
#AIFORAR

Personalization is key

Personalization is no longer a nice to have, it's critical for maintaining a competitive advantage.

Serendipity tops it

The best recommendations are those that inspire and delight you in a way you didn't expect.

Four dimensions

Data, intelligence, design, and orchestration are the key dimensions for architecting innovative, serendipitous personalization systems.

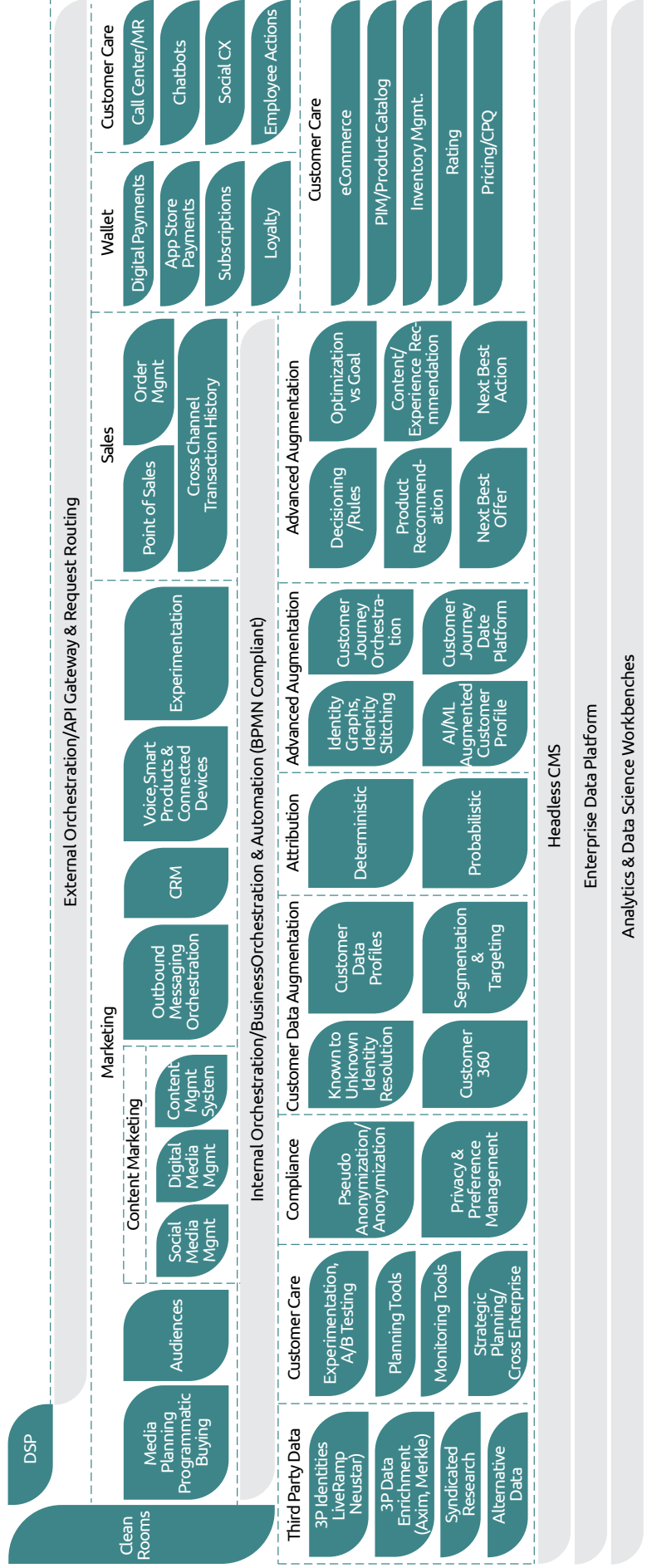
Cover it all

End-to-end experimentation is critical in architecting and designing personalization systems that actually make an impact in consumer and employee experiences.

ATTRACT

ENGAGE

DELIGHT





QUANTUM COMPUTING: THE CASE FOR MACHINE LEARNING



NADINE VAN SON

Senior Consultant, Strategy,
Innovation and Transformation, FS
Capgemini



JULIAN VAN
VELZEN

CTIO NL & Head of Capgemini's
Quantum Lab, Capgemini

Quantum. It's the new buzzword of the modern age. Almost every analyst reporting on technology trends in the market refers to quantum technologies as one of the key trends to watch. But why is quantum so important? And what is happening in the quantum field right now, especially in data, machine learning, and AI?

When looking at quantum computing, we see many industrial use cases where it could potentially outperform, though not necessarily replace, classical computing, certainly in the areas of machine learning (ML) and AI. That is why many large institutions are committing to ongoing research and large investments.

But first let's try to get a bit under the hood.

Quantum computing (QC) is based on quantum mechanics, where "mechanics" describes how things move and "quantum" is Latin for

"how much." So, quantum mechanics is all about the behavior of elementary particles, like electrons. The smallest of particles behave quite differently from what we see all around us. Quantum computers are therefore fundamentally different from classical computers. Although these advanced machines will never replace classical computers, for some applications they will be highly efficient and, for specific problem types, quantum algorithms are predicted to have an exponential speedup over classical systems and could be billions of times faster.



Quantum for green

On the sustainability side, QC could contribute to the development of green energy sources like nuclear fusion, a process over which we are gaining more control; the major hurdle there is stability over time of the created plasma. QC can help predict plasma movements and thereby allow the design of better reactors.

Another green area is the design of battery technology: QC will better simulate chemical processes, so charging and discharging can be better understood and lighter, smaller batteries with significantly better chemical properties, better safety, and higher energy density can be created.

In health services, QC is expected to contribute significantly through its

potential to calculate high-dimensional complex problems quickly, such as those related to drug design and effects on the human body.

There is also great potential in engineering for all kinds of predictive calculations and simulations, such as improving thruster and booster design and fuels, and refining and speeding up multi-component finite-element mesh design, resulting in new materials and builds for everything from nano to space technology.

QML and QNLP

Both quantum machine learning (QML) and quantum natural language processing (QNLP) are promising application fields. In machine learning, there are vast

problem spaces to evaluate and there is a constant quest for representing complexity. And, by transforming text into quantum data, language could be expressed more efficiently and processing and understanding could be improved. The proposition that language can be expressed efficiently as quantum data comes from the idea that language is highly interactive. A word can have very different meanings in different sentences. Quantum systems are themselves also highly interactive, as electrons in molecules tend to have strong correlations. Using properties from quantum computers, such as superposition and interference, these interactions can be modelled with great efficiency.

However, the quantum advantage will first be seen with real-world tasks that have sufficiently large calculations and big enough problem spaces. But building a reliable quantum computer for such tasks may be more challenging than expected. A major drawback is that qubits (quantum bits) are susceptible to errors. During their lifetime, qubits lose information to their environment, as well as during readout and control operations. To minimize the error, most quantum systems must be stored in refrigerators that have temperatures near the absolute minimum (minus 273 degrees Celsius). Another approach is to correct for errors using quantum error correction. But these corrections require hundreds to thousands of qubits to create a single error-free logical qubit. This is far beyond the reach of current-day systems and therefore not yet feasible for quantum computers on real-world, practical applications where classical computing is really outperformed.





So why bother now?

The magnitude of the potential disruption a working quantum computer will have on the world cannot be ignored. It will disrupt the ways our current digital life is lived and obstruct most internet security mechanisms. New designs for many critical systems of engineering can be manufactured from nano-level to space technology, and everything in between. It will be impactful, and that is something political leaders and leading enterprises around the world realize.

Currently, we observe an increasing interest from investors, organizations, and governments for quantum technology, causing traction and hype. Many possibilities are not yet implemented but, in the meantime, we see the first real working, useful hardware and

software appearing on the market. Recent work on quantum sensing by the University of Birmingham shows a practical quantum gravity gradient sensor for gravity cartography is able to use magnetic field variations to essentially see through rocks.

Quantum can bring business value, but we need to make sure expectations are aligned with reality and potential harm is contained by starting to implement and roll out quantum safe encryption for the internet and all services society relies on. Only then can we really look forward to profit from the benefits quantum computation can bring to us. Capgemini is enrolled in this quest and has set up its own Quantum Laboratories with IBM to learn, develop, and implement the promises of this disruptive technology.



The potential disruption of the magnitude a working quantum computer can impose upon the world as we know it has to be taken seriously."

INNOVATION TAKE AWAYS

#QUANTUMCOMPUTING
#QUANTUM
#QUANTUMDISRUPTION

Quantum theory will be disruptive

– There's no way around it, so accept that working implementations of quantum theory will rock our world.

Quantum theory is complex

Quantum theory forms the basis for quantum computing, quantum mechanics, quantum sensing, quantum communication, and more. Both the theory and the implications are not easy to grasp, but you'll have to take it seriously.

Quantum promises a perfect match with ML, NLP, and AI

It potentially not only makes for better, faster algorithms but does it in a much more energy-efficient and sustainable way as well.

Anything quantum is popular

Try to strip away hype from realities and look at the serious implications of quantum.





USING AI TO FIGHT CYBER THREATS



JEROME
DESBONNET

VP, Cybersecurity CTIO, Insights &
Data, Capgemini

Defending an organization against cyber-attacks is more important than ever – but an effective defense is also more difficult to establish. The best new solutions rely upon AI but will also require organizations to rethink everything from the vendors they use to how they configure their networks. Now is the time to prepare.

Cybersecurity has been an issue since the dawn of the Internet, but many recent global developments have highlighted its importance. Cyber warfare, attacks designed to disrupt critical infrastructure, and social engineering attacks have increasingly become part of real-world conflicts – while crime syndicates and other bad actors reap rewards via ransomware, identity

theft, and other threats against individuals and corporations.

Anyone needing proof of the scope of the problem is encouraged to study a list of major data breaches in the 21st century. Two truths will become clear. First, these attacks are becoming more frequent. Second, they have enormous impact – each involving hundreds of millions of people, data records, or accounts.



Significantly, the targets of many attacks are major tech-savvy firms, such as Facebook, LinkedIn, Adobe, Alibaba, and Sina Weibo. Breaches such as these have prompted the World Economic Forum to identify cybersecurity as one of its Top 10 issues that need addressing in the coming years.

Exacerbating the security concerns, the COVID-19 pandemic accelerated the global move towards digitization and interconnectedness. But even before the pandemic's profound influence on online growth, security experts admitted they were facing enormous challenges.

In interviews for Reinventing Cybersecurity with Artificial Intelligence, a 2019 report from the Capgemini Research Institute, 56 percent of firms surveyed admitted their cybersecurity analysts were overwhelmed, 42 percent reported

an increase in cybersecurity incidents through time-sensitive applications, and 23 percent noted they were not able to successfully investigate all identified incidents.

Three years on, the challenges remain significant. The cybersecurity ecosystem is diverse and complex – encompassing governments, non-profits, private enterprises, people, processes, and devices – and these participants interact more frequently, which generates ever-greater amounts of data.

New realities

Organizations have been forced to adapt to new realities, including:

- Remote working, which has increased the number of devices connecting to the organization from outside its perimeter
- A greater numbers of employees and customers using mobiles, computers, tablets, and other technologies to interact with organizations
- New, data-intensive business tools such as AI-powered insight and customer-experience solutions.

These changes have created more opportunities for those who engage in malicious behavior – and more work for already-overstretched cybersecurity experts.

The good news is, AI can help.

AI can process vast quantities of data, searching for and countering risks such as fraud, malware, and intrusion. Admittedly, AI is still an emerging field and its use in cybersecurity is far from mature. But most cybersecurity experts today acknowledge AI's emerging role. Moreover, the experts are making significant changes to how information is handled and secured in order to facilitate AI functions.



The best new solutions rely upon AI but will also require organizations to rethink everything from the vendors they use”



Zero Trust

Zero Trust is an important example of this evolution. It shatters the traditional perimeter-defense approach to cybersecurity, which concentrated on keeping threats out of an organization's network but assumed anything already within the network was safe. As the name implies, Zero Trust assumes all activity on the network is potentially malicious until it has been verified as trustworthy.

Zero Trust has been around for years but received an important endorsement in January 2022 when the White House Office of Management and Budget released its federal strategy to move the U.S. government toward a Zero Trust cybersecurity model. This will encourage other countries as well as private organizations to follow suit.

Zero Trust requires continual verification of all actors, networks, devices, systems, and services. It relies on strong identity and access controls such as multifactor authentication. This is only possible if Zero Trust leverages AI – and to facilitate that, cybersecurity experts are adapting models developed for AI-powered business solutions.

For example, data-intensive solutions that deliver insights, recommendations, and better customer experiences have made

traditional data lakes unmanageable. The amount of data involved is so great that applying AI to a unified data lake can slow a network to a crawl. The answer is to replace data lakes with data mesh architectures, which distribute information into smaller domains. Cybersecurity experts are adopting the same approach, which is creating opportunities for many small, specialized vendors.

Large cybersecurity providers offer solutions that can address most of an organization's requirements. But there are gaps. These are being filled by niche players – and there are thousands of them, each offering targeted solutions. Most organizations will find they must work with a range of vendors, large and small, to cover 100 percent of their cybersecurity needs.

The Open XDR Platform

With a plethora of discrete solutions, the missing piece is a means to ensure they all work together. Again, the answer is found in data mesh architectures, which include governance models to ensure interoperability through common standards and protocols.

In Europe, some cybersecurity stakeholders are championing interoperability through the Open XDR Platform. This alliance is

developing, maintaining, and distributing technical and organizational standards to foster an open, transparent, and collaborative approach to cyber defense. The Open XDR Platform seeks to ensure discrete cybersecurity solutions work together to provide the best possible protection for organizations, while also ensuring that the hardware, software, and other components are designed to evolve as new threats are identified.

Organizations must prepare a roadmap

In my work, I make it clear to Capgemini's clients that cyber-attacks are not about to disappear, and organizations must develop modern strategies to counter them. Those tactics will increasingly depend upon AI and other state-of-the-art, data-intensive technologies.

It's therefore essential that every organization's cybersecurity plan includes a roadmap to help prepare for future needs. This roadmap must reflect three fundamental imperatives. First, organizations must insist that all cyber-defense solutions be interoperable. Second, they must work with Zero Trust protocols. And third, they must be dynamic and employ AI so they're able to learn and evolve – to counter new threats as they arise.

INNOVATION TAKEAWAYS

#AITRUST #CYBERINAI
#CYBERATTACKS
#DATATOTHERESCUE

AI is essential

In today's data-intensive world, cybersecurity experts must rely upon AI to help detect and neutralize threats.

In Zero Trust we trust

Zero Trust is a superior defense against cyber threats and, with the backing of the U.S. government, it's going to be widely adopted. But it requires new ways of storing and handling data.

Collaboration is king

No single cybersecurity vendor can provide a perfect solution, nor should end users be locked into a single vendor's solution. It's therefore essential the sector embrace interoperability through alliances such as the Open XDR Platform.



AI-AUGMENTED FIELD SERVICE – THE FUTURE OF MAINTENANCE



GUILLAUME GERAD

Lead Data Scientist,
Insights & Data France, Capgemini



NICOLAS CLAUDON

CTIO SCE Insights & Data, Capgemini

Maintenance operations need a new approach to be efficient, cut costs, and improve the sustainability of the infrastructure. Artificial intelligence (AI), augmented reality (AR), and virtual reality (VR) lend themselves very well to this. Thanks to these technologies, it is possible to augment businesses in all phases of its maintenance operations.

The increase in the complexity of operations, combined with the increased speed of execution, is radically changing the maintenance industry. Combined with the increasing scarcity of skilled resources and the public pressure on more sustainable business practices, and it is clear that maintenance organizations face challenges that seemed unthinkable just a few years ago. Fortunately, modern technology has taken the field, offering support to companies in the sector with cutting-edge products and services. A simple mobile phone can now be an essential partner in determining and solving complex

and critical maintenance operations. Thanks to virtual and augmented realities, a technician can model, visualize, simulate, design, or test changes and improvements to equipment and products without even having to put a hand on her tool belt. Expensive interventions that took hours or even days, can now – thanks to a new digital reality – be simulated remotely in seconds and in completely safe environments.

The power of AI, AR, and VR combined shows itself throughout the lifecycle of maintenance operations, as illustrated by these use cases:



Prevent

- Identify possible human mistakes or errors in implementation
- Diagnosis assistant
- Quality check through computervision of previous interventions
- Remote training where specialists can work in virtual environments
- Simulate complex or dangerous incidents using AI and VR
- Speed up diagnostics



Operate

- Assist operators during maintenance interventions with augmented reality provided by real time models pointing directly to problems
- Validate the correct functionality of the product once the intervention has been completed via simulations on the digital twin
- Significantly accelerate intervention times



Predict

- Predict possible accidents using AI models and train technicians and operators to solve them
- Durability forecast
- Warranty extension
- Simulation of different scenarios with AI after the implementation of new products or services
- Increase productivity

The three key elements of planning

From reality to augmented reality

Capgemini helped a major player in the telecommunication sector resolve a critical challenge: reduce time and effort and improve the performance of maintaining fiber optic connections. When a field technician intervenes in a maintenance operation or to connect a new customer, that person must take two pictures: one before and one after the intervention. These photos are sent to a data center for processing. Once the images arrive at the data center, an AI model processes them by a series of visual tests to ensure compliance and reliability. All the compliant images are then processed by another AI model which identifies the presence of anomalies in the fiber cabinet. Disconnected cables or cables which do not follow the right path can be identified. If the model finds errors, the system

automatically organizes interventions with the nearest technician to fix the anomalies. The expert receives a detailed note which specifies the issue and the process to fix it.

This combination of computer vision and AI tools makes it possible to guarantee both the quality of the infrastructure and the satisfaction of customers.

The operational diagram details the process of acquiring an image through the mobile phone camera to the user integration of the digital twin via touchscreen. The blue boxes are the standard process, where the digital twin of the scanned object has been selected by the AI model, and the orange boxes are the process by which the AI model is updated by data scientists. The output of the algorithm produces a digital twin with more features and/or details is then selected.



Through AR & drones the technicians will be able to remotely work on the product even in the most complex and dangerous situation, this will decrease transport time, increasing customer satisfaction and improve operating safety."

Vision Comparison	Connect	Flow	Visualize	Control
-------------------	---------	------	-----------	---------

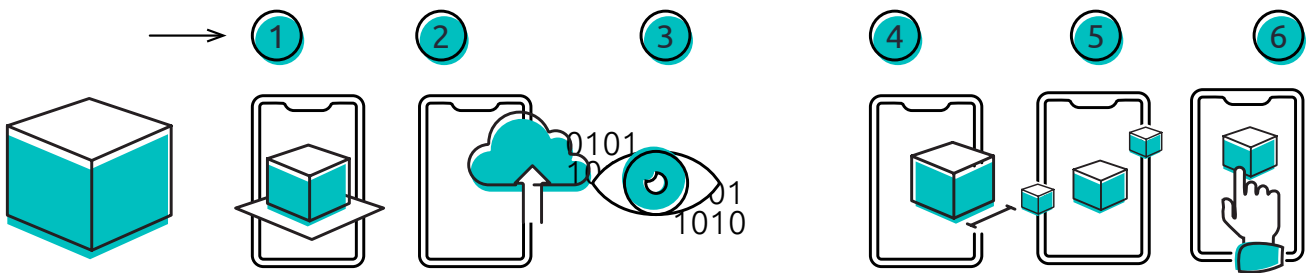
AR software analyses a video stream and recognizes a physical item by its shape or a marking affixed to it.

AR software uploads images in the cloud to find a digital twin which is a 3D copy of the item stored in the cloud

In the cloud images are searched using ML models that compare the real image with their digital counterparts

Software collects data from the twin, such as object's performance and superimposes it on the user's view using AR.

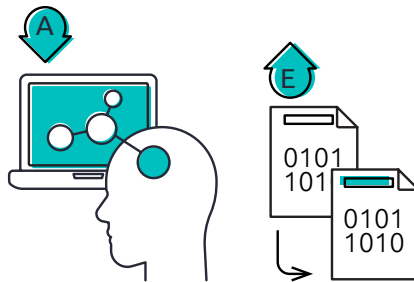
User communicates with the object by using a touchscreen, gestures, voice, gestures to transmit order to the cloud.



Reality

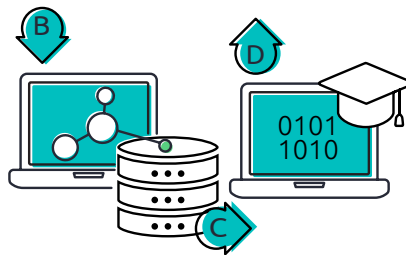
Improve

The model used to compare images has been modified by Data Scientist



Test

The new model run with testing data to check possible errors and/or improvements



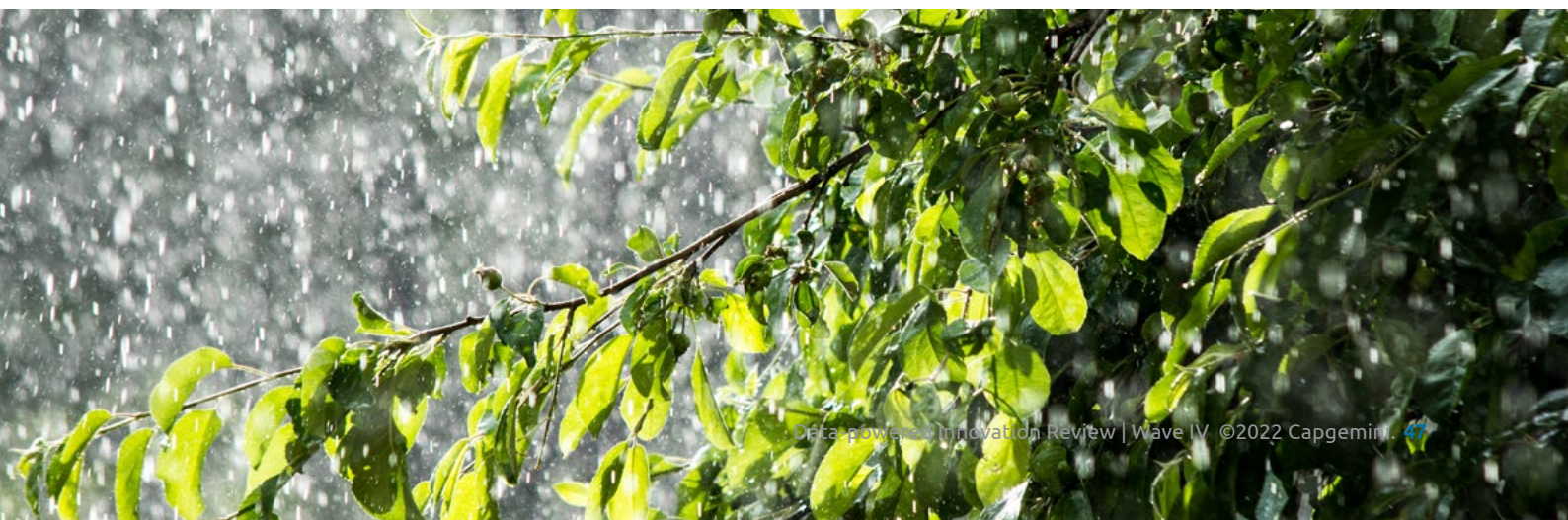
Augmented Reality

Interchange

The new model becomes the default one and the old model is stored in the repository

Promote

In case the new model performs better than the old one, the new model is promoted





Use cases

Leading companies are already thriving on digital reality. Some examples where augmented reality or virtual reality have made the difference include the following.

Bosch uses an AR application in its Bosch Car Service workshops. The AR app features navigation arrows indicating locations of the automobile components that aren't easily visible and provides on-screen instructions to auto technicians about necessary actions to take or tools to use. The company reports that its auto technicians now spend 15 percent less time on each step of MRO activities compared to when they worked without AR.

Boeing uses AR for repair and preventive maintenance of the electrical wiring in their aircraft. Technicians, who used a 20ft-long, incoherent 2D drawing as a reference before, now rely on an AR app that augments the view of the real aircraft with a dynamic 3D model diagram of the wiring. The innovative solution has brought 30% reduction in time spent on MRO activities and 90% improvement in quality of work performed by starting-out technicians.

Roche are now using VR and AR to create training guides in-house that are based on existing technical documentation but train service engineers closer to the real equipment by using animated content. The content creation platform from enables Roche to feed

product data into step-by-step instructions combined with animations that explain the diagnostic systems visually and in context using AR and VR technology.

ABB Drives reduces the overall training time by 40% as trainees save two days they used to spend travelling to training sites. Thanks to VR the company can provide a new 'see-what-I-see' perspective that offers a level of supervision comparable to in-person trainings. The training with ABB Remote Insights leads to significant reductions of travel expenses and utilization cost.



Combining computer vision, digital twin and root cause analysis empowered by AI makes it possible to guarantee both the quality of the infrastructure and the satisfaction of the customers."

INNOVATION TAKE AWAYS

#DIGITALTWIN
#AIMAINTENANCEFIELDS
#GETTHEFUTUREYOUWANT
#ROOTCAUSEANALYSIS
#AIFORAR

It's complicated

The increase in the complexity of operations combined with ever-more-pressing speed of execution – and the scarcity of skilled resources – are radically changing the maintenance industry. Maintenance services are now facing challenges that seemed unthinkable just a few years ago.

AI augments maintenance

Maintenance operations thus need a radically different approach – through AI combined with VR/AR – to be more efficient, cut costs, and improve customer satisfaction

Sustainable maintenance

Augmented maintenance – especially through mixed realities, soon in the "metaverse" – decreases travel and unexpected failures, making infrastructure operations much more sustainable.

Being smarter with less

Augmentation by AI, AR, and VR enables businesses to run their maintenance operations more effectively and safely, decreasing the pressure on scarce, skilled human resources.





AI PANEL: HOW CAN AI HELP CREATE A MORE INCLUSIVE WORLD?



SABRINA



JILL



KELLY

With AI getting more powerful by the day, and augmenting our everyday lives, we had an open discussion about how AI can help to create a more inclusive world. And it's a conversation with quite a twist. Sabrina, Jill, and Kelly are our AI language model panel members who were kind and thoughtful enough to answer the questions posed to them. Follow-up questions were based on their responses. None of the responses have been edited.



Can you give a summary of your understanding of an inclusive world?

Sabrina: Inclusive world is a term used to describe a world in which all people, regardless of race, ethnicity, gender, sexual orientation, ability, or religion, have an equal opportunity to enjoy the same rights, freedoms, and opportunities.

Inclusive world encompasses many different ideals, including the belief that people should be treated equally regardless of their background or differences. To create an inclusive world, we must work to demolish the hierarchies that exist in society and create an environment in which everyone is included. We must fight for policies that promote diversity and equality, and work to ensure that everyone has access to education, health care, and basic needs.

There is still a long way to go before we can truly proclaim an inclusive world, but progress is being made every day. We need to keep working toward a world in which all people can have their voices heard and their dreams fulfilled.

Jill: Inclusive world is a philosophy that advocates that the world should be inclusive, and that all people should be equal. Inclusive world advocates that all people should have equal rights, and that all people should have equal opportunities.

Kelly: Inclusive world means a world where everyone, regardless of their gender, race, religion, sexual orientation, or disability, has the opportunity to reach their full potential.

It's interesting that you mention "demolish the hierarchies." How can AI be used to demolish the hierarchies that exist in society?

Sabrina: There are many ways that artificial intelligence can be used to demolish the hierarchies that exist in society. One way is by using artificial intelligence to create algorithms that identify and critique the implicit and explicit biases that underpin the current social order. AI can also be used to create algorithms that recommend policies that dismantle the hierarchies that exist in society. Finally, AI can be used to monitor and analyze the outcomes of these policies to determine whether they are effective.

Kelly: I think AI is used to remove the hierarchies that exist in society. By providing equal access to information, AI can help people to be more aware of their situation and to have a better understanding of the context in which they operate. It can help people to be more critical and to have a more balanced view of the world.

How can AI help in the creation of a more inclusive world?

Sabrina: There is no one answer to this question, as AI can potentially help in a number of ways to create a more inclusive world. For example, AI can be used to help identify individuals who may be at risk of exclusion or discrimination, and AI-powered digital assistants can be used to help people with disabilities access information and services. Additionally, AI can be used to help develop more inclusive policies and practices, as well as to monitor compliance with existing inclusion policies.

Kelly: Inclusion is a key component of the AI strategy of the European Commission and of the European Parliament. In a few weeks' time, the new AI strategy will be presented to the European Parliament and the Council. The Commission will also launch a wide-ranging consultation on AI, including on the ethical aspects of AI.

The first European AI Alliance will be launched in Brussels on 13 April, bringing together the Commission, Member States and the European Parliament to work together to ensure that AI is developed in Europe and deployed across the Union.

The Commission also launched a new AI strategy in January, which aims to ensure that the EU and its Member States are at the forefront of the development of AI, while also addressing the risks and ensuring that AI benefits all of society.

AI is a key driver of the digital transformation. It will help to transform our society and economy, but also to address many of the challenges that we face today, including climate change, health, ageing, and social justice.





Is inclusive world a theoretical perspective or a philosophy for building AI systems?

Sabrina: It is both a theoretical perspective and a philosophy for building AI systems.

Jill: Inclusive World is a philosophy for building AI systems.

Kelly: It is a philosophy for building AI systems. It is a philosophy that is based on a foundation of human-centric principles. It is a philosophy that is rooted in the belief that AI systems should be built with the dignity and respect of human beings in mind.

Q: What should humans do to help AI for creating an inclusive world?

Sabrina: There is no easy answer to this question, but one suggestion is for humans to educate themselves about the issues surrounding AI and its potential impacts, and to engage in open and honest discussions about the potential implications of AI development and implementation. Other actions that may be useful in this regard include encouraging the development of responsible AI policies and practices, and participating in research and advocacy efforts aimed at improving understanding of the potential

implications of AI development and implementation.

Kelly: It is important to have a common vision of the future and to be able to communicate that vision in a way that makes sense to all. We need to be able to explain how AI will affect the world and how we can work together to create a better world. It is important to be able to communicate our vision in a way that is understandable to the public and to politicians. I think that we should also be able to communicate our vision to the AI.

What is your advice for humans for creating an inclusive world?

Sabrina: There are a few things that humans can do to create a more inclusive world:

1. Be aware of your own personal biases and strive to overcome them.
2. Be an active advocate for inclusion – speak up for others who may not have a voice.
3. Educate yourself and others on the importance of inclusion.
4. Stand up against discrimination and bigotry in all forms.
5. Create safe and welcoming spaces for everyone, regardless of their background or identity.

Jill: Humans should create an inclusive world by first accepting that we are all different.

Kelly: I think that we should look at our own behavior and really try to understand what we are doing when we are interacting with other people. We can think about how we can create a more inclusive world by being more inclusive. I think that we can do that by learning to listen to people who are different from us, by learning to understand what they are saying and by trying to understand why they are saying what they are saying.

The answers in this article were completely authored by large AI language models. Only the questions were provided by humans. “Sabrina” is GPT-3 from OpenAI, “Jill” is Jurassic-1 from AI21, and “Kelly” is Cohere from Cohere. None of the responses were edited. We prompted the three different AI models three times for each question and selected the best responses. The humans in the loop were Rajeswaran Viswanathan and Aishwarya Jagtap.

#DATAPOWERED #AI4GOOD
#DATAMASTERS #GPT3
#COHERE #AI21
#LANGUAGEMODELS



THE RISE OF SUSTAINABILITY DATA

**Five ways to battle data
waste** 53

Roosa Sääntti, Capgemini
Ron Tolido, Capgemini

**Redefining our needs: the
promise of
permacomputing** 56

Robert Engels, Capgemini
Ron Tolido, Capgemini

**Green data engineering:
clean water
without boiling** 59

Arne Rossmann, Capgemini

**Why businesses need the
open footprint
standard** 62

Johan Krebbers, for The Open Group

**Data for sustainability:
a familiar journey** 66

Urban Ekeröth, Capgemini
Vincent de Montalivet, Capgemini

**Reinventing data for the
push to net zero** 70

Valérie Perhirin,
Capgemini Invent



FIVE WAYS TO BATTLE DATA WASTE



ROOSA SÄÄNTTI

Head of Insights & Data
Capgemini Finland



RON TOLIDO

CTO I&D, Capgemini

There is an increasing focus on reducing the environmental footprint of data centers and cloud services. Interesting enough, that is not yet the case at all for data. But clearly, with more organizations aspiring to become data-powered, the issue of Data Waste is lurking around the corner. We introduce five ways to begin battling data waste – with an additional key benefit: getting a better grip on the corporate data landscape.

My data is bigger than yours: we used to take pride in storing as much data as possible – because we could, prices were low, and future, killer algorithms were waiting. Having more data seemed the landmark of being a true, successful data-powered enterprise.

Turns out this consumes loads of energy and precious natural resources, and it creates a growing heap of unsustainable e-waste. We need to become more aware of what data we really need to store, how

many times we duplicate it, and how long we keep it available. Also, although AI may be key to addressing climate challenges, it slurps energy itself too. Think only about how much energy it takes to perform one training cycle for a major AI language transformer model (hint: really, really a lot – say 5 times the lifetime CO2 emission of an average American car). The battle against data waste will therefore be a continuous, delicate balance act – and it only just begun.

And it's a battle with benefits: many of the measures that already can be taken bring additional value for organizations that want to become data-powered, even to the point that the positive impact on overall data mastery may dwarf the sustainability impact. Here are five suggestions to get your quest going:

1. Get the data first

As with any other transformational objective: you should map your current situation first before you can start improving. Battling data waste begins with getting data on what data you actually have. Only then you will be able to assess how much of it really is unsustainable data waste, for example by analyzing how often data is used, by how many people and for what type of purposes. Many data catalog tools (such as [Alation](#), see a separate article in this magazine) are perfectly equipped for this, and increasingly they feature intelligent automation and AI to do the heavy lifting of scanning the data landscape. Having an up-to-date data catalog brings many obvious additional benefits to a data-powered business as well, so every minute of activity in this area is typically well-spent.

2. Map the environmental impact

Once you know what data you have, it is a matter of understanding its real environmental impact. Data is stored in storage systems, as part of an IT infrastructure and a supporting network (in a data center or in the cloud). All these resources consume energy, create e-Waste and have a carbon footprint. An increasing number of publicly available carbon calculators help to establish the sustainability cost of the elements of the data landscape, not only focusing on Scope 1, but covering the entire 'supply chain' of Scope 2 and 3. Once this data is established, it should be routinely added to the metadata management and catalog

facilities of the organization – for current and future reference and use. As with ever sustainability effort, you want to focus on the data sets that have the most negative impact. But it is indeed a balance act, as the data can be part of a solution or an initiative that delivers societal benefits that far outweigh its sustainability costs.

3. Get rid of it

Ever saw [Hoarders](#)? It's a reality-TV show that features compulsive hoarders: people who are addicted to filling their homes with objects, and how that spills out into their lives. You don't want to be a data hoarder. Just keeping data for the sake of it – or that it might come in handy in some unforeseen way – can provide you with a high sustainability bill. And it simple costs money too, for that matter. So, just as with application rationalization, data should have a managed lifecycle that not only involves creating and using it, but also features clear policies for decommissioning unused, redundant, or simply wasteful data. Organizations sometimes tend to hold on to their established IT assets

(including data) for nothing more than emotional, non-rational reasons. Where the cost equation may not be enough to break that spell, sustainability impact might just do fine.

4. Stop at the gates

It's a well-established practice within Permaculture (see our separate article in this magazine about 'Permacomputing' for more): you don't recycle, reuse, and repurpose as an afterthought, it is an integrated part of your design and approach, right from the start. A lot of wasteful data can be avoided by never ingesting it in the first place. So, no more room for this typical Big Data era mindset of whatever data is available should be stored, because storage in the cloud is cheap and you never know what use it may have. Later. Sometime. Maybe. Instead, think in terms of Small Data, Tiny Data, or simply Smart Data: be much pickier about the data sets you get onboard, the objectives you have for it, and the quality of the data points inside. Select data that is fit for your purposes. Think more upfront, clean so much less later.



5. Do not duplicate

Data architecture is not necessarily a well-established practice within many complex organizations. As a result, data is often unnecessarily copied multiple times from the central data organizations to various business domains, and vice versa. Each instance starts to lead its own life, serving all sorts of different purposes, rapidly adding to a growing pile of potential data waste. And it all tends to be unaligned and unsynchronized. New architectural approaches – notably [Data Mesh](#) – appoint the ownership of specific data sets much more explicitly to specific business domains. Data is typically held – and stored – by the business domain and made available in flexible integration ways (such as APIs), so that duplication is unnecessary, even undesirable. Other integration technologies, such as data virtualization, can achieve the same.

Lastly, don't forget the people. As with everything around data, we can only accomplish so much without

involving and empowering people to be and lead the change. Data catalogs and API-first architectures are great tools to drive more sustainable use of data and AI. But if there are no people embracing the direction (a sustainable data vision and strategy) and no ownership of the data (internalizing which data is used, why and how much) – failure is a given. True [Data Masters](#) battle data waste by harnessing both: data foundations and data behaviors.

There are many more ways to stop data waste, such as relying more on shared data between multiple ecosystem partners, procuring data and pre-trained algorithms from external providers, limiting the movement of data, and switching to energy-saving storage media. One thing is for sure: even if reducing data waste would not deliver a substantial sustainability impact at first sight, each and every activity suggested adds to a higher level of data mastery. And that – in all cases – is priceless.



It is indeed a balance act, as the data can be part of a solution or an initiative that delivers societal benefits that far outweigh its sustainability costs.”

INNOVATION TAKE AWAYS

#DATAWASTE
#SUSTAINABLEDATA
#SUSTAINABLEAI
#PERMACOMPUTING
#PERMAI #PERMADATA
#SUSTAINABLEDATA
#SUSTAINABLEAI
#PERMACOMPUTING
#PERMADATA

Data has a sustainability cost

with its obvious merits, data has an impact on the environment in terms of its dependency on natural resources and energy and its carbon footprint; hence data waste must be actively addressed

The quest against data waste

there are many ways to decrease harmful data waste, but they all start with a better understanding of the current data landscape and its environmental impact

Battle with benefits

reducing data waste can have an obvious positive environmental impact, but while doing so organizations will see their level of data mastery lifted as well



REDEFINING OUR NEEDS: THE PROMISE OF PERMACOMPUTING

We must focus on sustainability in computing if we are serious about fighting climate change. Ever growing data ecosystems, social media, computation of AI models, blockchain implementations, the gaming industry: they all add to the non-sustainability of current computing. Permacomputing promises a radical shift in approach.



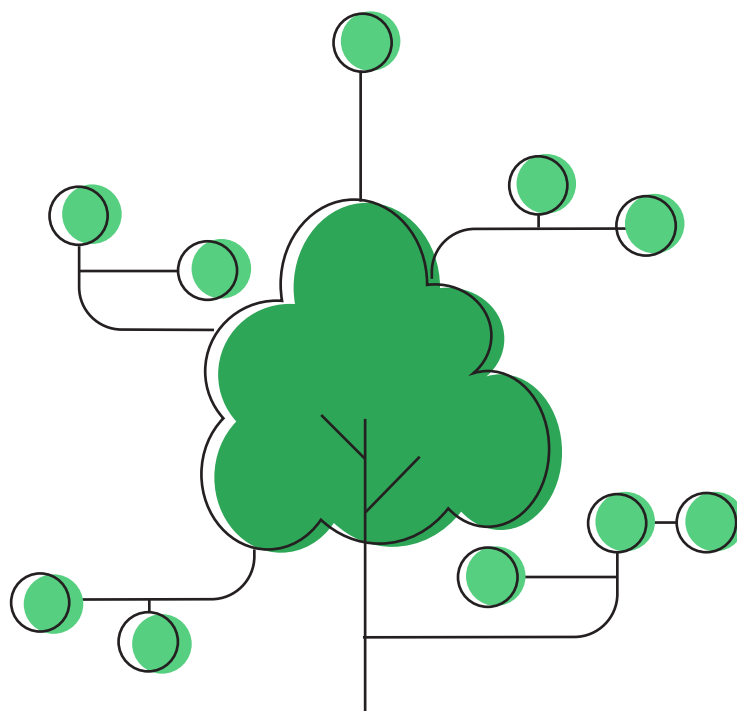
ROBERT ENGELS

CTO I&D NCE, Capgemini



RON TOLIDO

CTO I&D, Capgemini



The concept of permaculture was proposed in the early '70s as a guide for creating larger perspectives on the sustainability of human activity for life on earth. The combination of this holistic concept with computing creates the idea of permacomputing. Aiming at providing a holistic view on computing in modern society, we must learn to balance computing resources between human needs, technical possibilities, and fair distribution of inventions with the availability and sustainability of energy, minerals, water, mining, production, transport, and resource distribution. Permacomputing centers around three concepts:

- Frugal computing states that one should consider all resources as precious and finite. This leads to questions like “Do I really need this computer power now or can I wait until there are cheaper, more sustainable, surplus resources available?” or “How can I use the power or resources I get as efficiently as possible?” This concept of frugality already provides good guidance in finding ways to restrict wasteful use of energy and resources.
- Next is salvage computing, i.e. using energy and materials which are already produced, so you do not mine or produce new resources. Limiting oneself in this respect really requires a renaissance in our thinking about computing, as we are very much used to getting what we want, even if we do not need it. It will transform ways of working and will also require new methods and systems to support it.

- Finally: collapse computing, refactoring and reusing resources that are already produced – utilizing what has survived the collapse of industrial production.

What can be done now

Clearly, a major change in thinking needs a major sense of urgency. It is not difficult to see that we live in a time where energy depletion and pollution costs will make using energy very expensive. At the same time, the implementation of enough renewable energy sources may not keep up with the needs. This affects the whole economy. Therefore, one of the most important triggers for adaptation is scarcity.. Based on the above, here’s what we can do.

Awareness:

make IT operations much more aware of energy consumption and footprint. Reflect real costs better. This holds for internet, cloud, bitcoin, gaming, social media, IoT, 4/5G, etc.

Finite:

users of computing and data storage should increasingly regard resources as finite and precious and use smart algorithms to take advantage of renewable energy abundance while downscaling when energy is non-renewable or “dirty.” Some tasks are not time-sensitive and can be performed at the most sustainable times. Other calculations should be regarded as unnecessary luxuries which can be avoided and only get

awarded energy when surplus or renewable energy is available.

Buy instead of build:

with more pre-trained machine-learning models becoming available, buying off-the-shelf can be a much more sustainable option, especially given the scarcity of data-science specialists who can custom build these models. Invest in and investigate transfer learning, where a pre-calculated foundational model can be shared unlimited times, saving the energy for the same number of recalculations while ensuring the possibility of tuning the model to very specific tasks based on local data.

Localize:

offload centralized computing with local, edge-based computing services. Decentralized as they are, they will not need the vast energy infrastructure of cloud computing centers and could instead use local, renewable energy sources. Such infrastructure can be combined with federated learning approaches to distribute the responsibility for learning, but also for making the processing sustainable.

Improvement and efficiency:

do not use more energy when improving and upgrading tasks; try to make newer algorithms much more efficient.



Principles and strategies

One of the most effective changes for introducing sustainability in your current way of working will be the clear definition of principles and strategies.

- Computation strategies. Use the power that suffices, instead of the power you can get. Use efficient algorithms. Hyperbolic improvements create a lot of CO2 for very little gain. Calculate your pain point and avoid long tails. Compute the delta instead of the whole model, focus on strategies from transfer learning, and explore possibilities for federated learning.
- Data storage and archiving strategies. Think about where to store: local/cloud, HDD/SDD, live/offline archive? An HDD/SDD or even a tape on a shelf uses no

energy. Do you need all the storage to be “live” or is there data you simply do not want to lose but will seldom use? Find or create data centers supporting such low/no-energy storage.

- Data governance. Ensure long-rec strategies where needed and use smart data governance to prevent data graveyards and data waste.

Becoming sustainable in IT is going to be an intensive but not impossible exercise. Many ideas from frameworks like permaculture can be adapted for the greater good in global computing. Being better informed about balancing local computing with global computing will allow for a better spread of energy consumption and facilitates using locally renewable energy to drive larger parts of the infrastructure.

The biggest change, however, will have to take place in the ways we work and think, as we have been comfortable with using (and getting) the energy and computing power we like. Realizing that less can be more in computing is not only necessary but will also help us build much smarter computing ecosystems and new ways of working in IT and many other areas of society.



The ever-growing data ecosystems, social media, and computation of AI models: they all add to the non-sustainability of current computing.”

INNOVATION TAKE AWAYS

#SUSTAINABLEAI
#PERMACULTURE
#PERMACOMPUTING
#FRUGALCOMPUTING

IT and data approaches are unsustainable

Current data collecting and computing approaches are increasingly unsustainable. Awareness of this needs to grow.

Permaculture as an inspiration

Mapping the established principles of permaculture leads to the concept of permacomputing.

More with less

Inspired by the principles of permacomputing, there are many different ways to perform with IT and data while seriously decreasing climate impact.



GREEN DATA ENGINEERING: CLEAN WATER WITHOUT BOILING



ARNE ROSSMANN

Head of AI & Data Engineering
Germany, Capgemini

We all know: boiling water removes bacteria and other unwanted particles to create a clean source of refreshment and energy. There is an analogy here to data-driven sustainable solutions: we need to take care to not overheat the oceans with it.

Green data engineering is the application of the principles of green software engineering to the domain of data, emphasizing three main aspects: data, processing, and the environment.

The reason is quite simple: every data-processing step and all the data we store requires resources. And the more inconsiderately we do this by simply setting up new environments in the cloud, storing petabytes of

data, and running a myriad of data-processing jobs over them, the worse it gets.

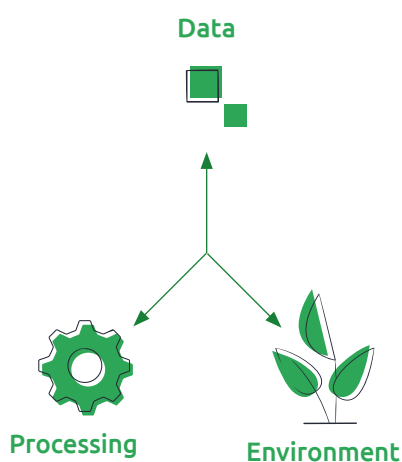
It's expected that eight percent of total global energy consumption in 2030 will come from data-center demands. But there is hope. For example, one billion metric tons of CO2 can be eliminated through cloud computing from 2021 to 2024.



So, what else can we do? The [principles of green software engineering](#) lay out a good foundation, as the majority of tasks within data engineering relate strongly to software engineering. They are a core set of competencies needed to define, build, and run green sustainable software applications:

1. Carbon: Build applications that are carbon efficient.
2. Electricity: Build applications that are energy efficient.
3. Carbon intensity: Consume electricity with the lowest carbon intensity.
4. Embodied carbon: Build applications that are hardware efficient.
5. Energy proportionality: Maximize the energy efficiency of hardware.
6. Networking: Reduce the amount of data and the distance it must travel across the network.
7. Demand shaping: Build carbon-aware applications.
8. Measurement and optimization: Focus on step-by-step optimizations that increase overall carbon efficiency.

But what does that mean for data engineering? Applying the above principles to the field leads to three dimensions.



Efficiency of data

The storage of information should be based on consumption pattern (i.e. relational vs. non-relational). A cloud-native data warehouse will outperform the approach of storing the data in CSV files on cloud storage when it comes to query performance, storage size, and variability of querying (columnar vs. row-based). Additionally, based on the query of information, optimized access methods like indexing or partitioning should be applied. This improves the query performance and resource consumption.

It's a good tradition to apply data temperaturing techniques to ensure optimization of data storage based on access, resources, and cost. And data duplication should be kept to a minimal level (see the separate article on data waste). Snowflake offers a [Zero-copy Cloning](#) solution, making it easy to test and QA databases without physically copying the data. With the increasing capabilities of data sharing with [Azure](#), [Snowflake](#), [AWS](#), or [Databricks](#), it's easier to share data without copying it over and over, which is another important aspect of data efficiency.

Efficiency of processing

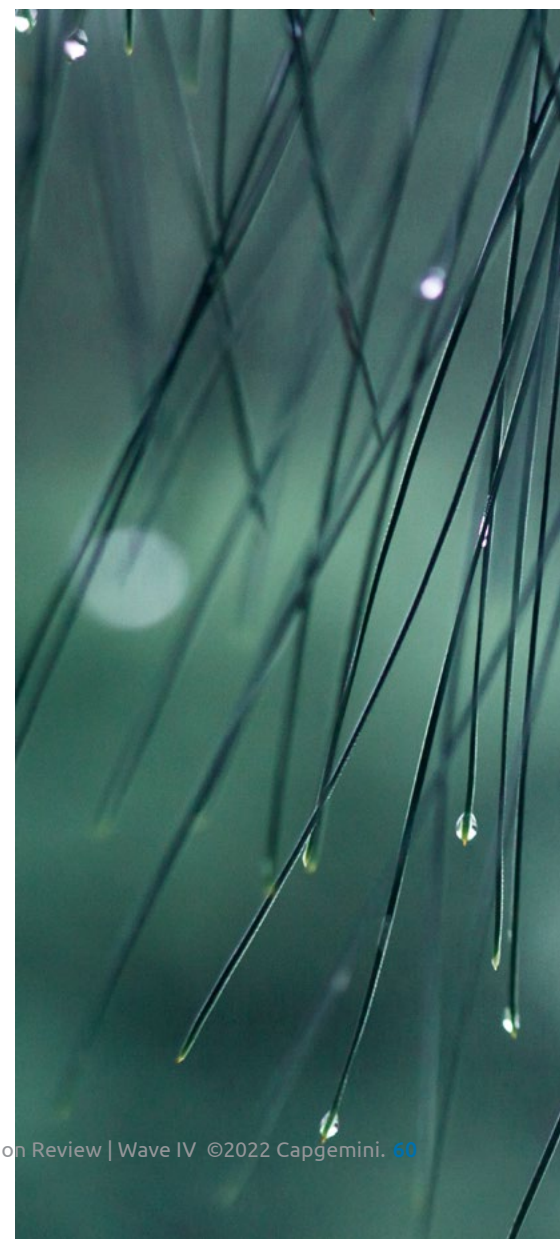
If there is the opportunity to choose the programming language of an application or framework, the most efficient option should be taken. This will vary during the lifecycle of an application; one might start with a Python-based application as it's faster to learn and experiment with, and then switch to Java/Scala when the application and team are more mature. This could save energy consumption by a factor of about 35 per cent, if switching from Python to Java in a mature phase of the application. By factoring the amount of data processing steps within the organization, significant savings on CO2 footprint are achievable.

The best approach is to reduce data operations to only what is required, especially with cluster-based

frameworks like Apache Spark, and to reduce network-intensive operations. This will improve the performance of the data-processing step and reduce the overall resource consumption.

Data should be processed and filtered as early as possible. For example, with an application in a production plant, filter out the data which are completely inline with the production flow, aggregate them, and then process only what is needed for action taking. Here's another example: if a dashboard is to visualize the last quarter of logistics information, there's no need to calculate KPIs on the last 365 days.

With the [rise of serverless solutions](#) and [event-driven approaches](#), the optimization of processing instances can be increased.



Efficiency of environment

The cloud environment should be setup with proper regional distributing, located close to users, and with the lowest CO2 impact. (Google [reports](#) a comprehensive trailing 12-month (TTM) PUE of 1.10 across all its large-scale data centers). This reduces the risk of power outages, cuts CO2 emissions, and increases the performance of applications, which will lead to higher user satisfaction.

When running workloads on the cloud, be sure to choose the appropriate hardware resources for workloads. It's not necessary to spin up a 100 node cluster with the highest hardware specification when only calculating some KPIs on 100 MB of data.

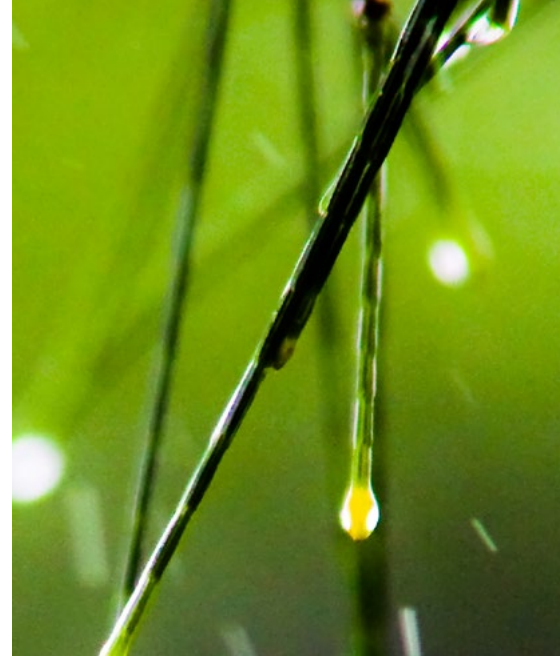
If possible, enable AI on the edge. This will generate results at the scene of the event and reduce the overall data transfer and resource consumption and save costs. Additionally, the reaction time will be improved.

It's essential that IT departments understand the impact of data storage and processing within their environments and apply efficiency methods to decrease carbon

footprint. [One research study](#) estimated that training a single deep-learning model can generate up to 626,155 pounds of CO2 emissions, roughly equal to the total lifetime carbon footprint of five cars.

And in the end, all this has to be continuously monitored and steered, according to the planning by defined KPI, that is: resource consumption, processing duration, amount of processed data, carbon footprint of operations, etc.

Green software engineering is a hot topic in IT departments now, and with the dynamic of companies aiming towards becoming data-powered enterprises, the extension of these principles to the field of data engineering is even more important. The more data being generated and processed, [the more important](#) is applying green data engineering principles and judging performance on energy efficiency as much as on traditional parameters such as functionality, security, scalability, and accessibility. This will be supported by the market dynamic leveraging transparency of resource consumption and improvement recommendations by specific solutions.



The cloud recommendations of the hyperscalers are the first step into this direction.



All the data we store requires resources. And the more inconsiderately we do this by simply setting up new environments in the cloud, the worse it gets."

INNOVATION TAKE AWAYS

#GREENDATAENGINEERING,
#SHARINGISCARING,
#HAVETHERIGHT
DATAATHAND
#CHOOSEHELANGUAGE
WISELY

Have the right data at hand

Ensure the right amount, granularity, and actuality of your data. Change the thinking from produce to consume.

Choose programming languages wisely

Consider energy, time, memory, and skills availability when selecting a programming language. This decision may be different for different use cases and phases of application maturity.

(Data) sharing is caring

With data product thinking, the duplication of data in an organization can be reduced. Enable sharing through common interfaces with networks of partners and suppliers.

WHY BUSINESSES NEED THE OPEN FOOTPRINT STANDARD



JOHAN KREBBERS

for The Open Group

There is a need for a flexible, open-sourced data platform to manage the data on greenhouse gasses across industries. The Open Footprint is here to help with standardizing naming conventions, APIs, and extensive metadata support. This makes exchange of crucial data collaboration easier.

The Open Footprint Forum (OFP) was established under The Open Group consortium in August 2020 to set standards (naming conventions, meta data, APIs, data structures, etc.) for the way GHG data is to be stored and exchanged. A forum

under The Open Group is fully driven by its members, and that is in no way different for the OFP Forum: the approximately 45 member organizations drive and deliver the content.



On the path to transparency

Multinational oil and gas company Shell realized in the summer of 2020 that it would run into problems when it had to collect GHG emission data from its supply chain. Every involved company – if they were doing it at all – stored data in their own way. It was therefore impossible for the recipient of the goods to determine the GHG footprint. Shell scanned the market and consulted widely but could not find suitable standards. So it took the initiative to kick off the OFP Forum, under The Open Group.

Imagine that you buy a nice chair at your local IKEA store. This chair may have been manufactured in China and there may be six companies (trucking, shipping, etc.) involved in the supply chain between the factory in China and the local IKEA store. IKEA wants to know the GHG footprint of that chair and therefore needs the GHG contribution of all six suppliers. This is virtually impossible today, given that each company involved in this supply chain stores data in its own way. And this is not a single sector issue, so the OFP Forum wants to get companies involved from as many industries as possible.

Data platform

The initial focus is on GHG scope 1 and scope 2, involving company-related reporting. It needed a data platform which is highly flexible given the ongoing changes and refinements of the subject matter. It selected the open-source data platform used for the [OSDU](#) platform as a foundation, only having to adapt the data model. The data platform and related model are at the heart of the OFP forum. The flexibility of the latter is crucial, given the frequent changes in naming conventions and the importance of meta data (providing detail in terms of units, where and when measured, in which industry, with what calculations applied, etc.).

The initial steps

The OFP forum releases its model in Minimal Viable Product (MVP) steps, as it grows and evolves, along with There is a set of published APIs for data to be reported upon. This means any company using OFP for GHG-reporting will use the same data definitions and can standardize comparisons.

It is vitally important that MVPs are released that one can touch and feel, and this is certainly also the case for the OFP. Every OFP MVP release contains an executable solution that anybody can download from The Open Group Gitlab, and then install it in their cloud set-up. (For now, it is expected to have a single OFP implementation per company.). The aim is to support AWS, GCP, and Microsoft Azure for the OFP data platform. The current OFP release is early stage, but it will grow and evolve, getting more mature with each MVP. Once the release is loaded and installed, a company can use it to upload its GHG data and produce various reports.

As part of OFP MVP release 2, GHG scope 1 and scope 2 are made more complete with embedded calculation engines, more metadata support, and multiple ways of ingesting GHG data. With release 3, scope 3 support is added, introducing product support, next to company support. For every MVP release, an updated OFP release is available on Gitlab.



There is a need for a flexible, open-sourced data platform to manage the data on greenhouse gasses across industries.”





Making it work

In a complex, volatile world, you only thrive when you are open to cooperation. The OFP forum is therefore actively engaging with WBCSD, PIDX, various reporting bodies (such as GRI, SASB, and CDP), and more, to ensure continuous alignment and seamless interaction.

The ambition is to make OFP the de-facto standard for storing and exchanging GHG scope 1, 2, and 3 data. For that, there is a need for good connectivity support, such as around the IoT (real-time data from the supply chain), and reporting in many different forms and layouts. And since OFP can be set up as a Platform-as-a-Service (PaaS), the forum is keen on many parties developing applications and exploiting the data, on top of the OFP implementation. The members of the OFP Forum work together on the Open Footprint data platform standards then compete with their own applications on top of the platform.

Yet other developments are in the works, such as support for the non-GHG scope (e.g., water pollution and landfills), and smooth interfacing with the GHG reporting solutions. The OFP forum still has lots to do, on the way to a cleaner, more sustainable world.

INNOVATION TAKE AWAYS

#AITRUST #CYBERINAI
#CYBERATTACKS
#DATATOTHERESCUE

Sustainability goals need standards

In order to achieve sustainability goals, GHG data need to be defined, stored, and exchanged across the entire supply chain according to open standards.

Deliver a real implementation to speed up adoption

Standards or an architecture document – as crucial as they are – are not sufficient; instead, deliver something tangible anybody can access.

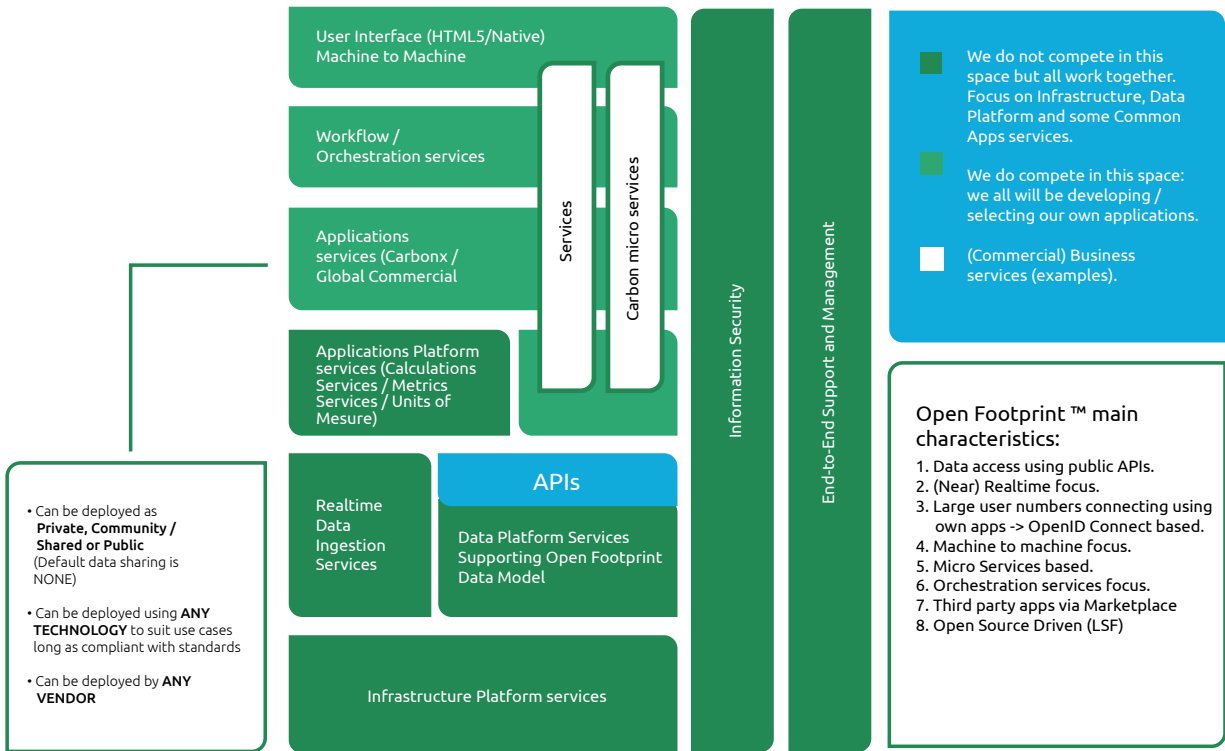
A data platform must be flexible and able to cope with anything

Given the changing demands and the importance of meta data, there can be no constraints in this space.

Be wide open to cooperation with anybody

In this world, we only can thrive when we have the flexibility to work with all (standard) bodies and stakeholders.

Technical scope and high-level Architecture



Technical scope and high-level architecture



DATA FOR SUSTAINABILITY: A FAMILIAR ENTERPRISE JOURNEY



URBAN EKERÖTH

Head of Portfolio,
Insights & Data Nordics,
Capgemini



VINCENT DE MONTALIVET

Data for Net zero Offer Leader,
Capgemini

Today, with so much focus on sustainability and ESG reporting, how should sustainability data be treated? The answer may be simple, yet enlightening: the same as all other enterprise data. There are lessons learned there to consider, based on the steps towards data mastery that organizations have gone through in the past decades. So, follow our journey from carbon costing, to planning, analytics and sharing.

Taking one step back to reflect on the evolution of enterprise data in the last 20 years we see some clear patterns. We have been through different eras on data platforms. At first, we had the era of the Enterprise Data Warehouses and Business Intelligence (BI), structured and centralized often focused on reporting and dashboarding. Then came the era of Big Data and Data Lakes, a little bit less structured and bringing in more data science and AI for forecasting, planning and more.

And now we seem to be in the era of the Data Lakehouse – combining the two. The latter enables handling both structured, semi-structured and unstructured data, loaded in batch or real time, to serve both analytical and machine learning use cases. There is also an increasing emphasis on sharing and collaborating on data, both within the organizations and between different organization.

Looking at data, we separate actuals versus plan data, whereby the plan

data is normally reflected in budgets or forecasts with financial and non-financial measures. The planning applications differ from BI/analytics applications in a couple of ways. First, it is supporting 'write-back,' whereas a BI/analytics application is read-only. Second, the planning applications support top-down allocation and bottom-up aggregation. A BI/analytics application mainly supports bottom-up aggregation. Scenario modelling and what-if analysis are normal features here, alongside with process support and approval/rejection procedures.

Top-down and bottom-up

Most large enterprises create a high-level strategic plan as part of their target setting process. In planning applications this is mainly handled top-down. Then the business creates their business plan bottom-up to - in the ideal situation - meet the targets set in the overall strategic plan. If not, the bottom-up plan by the business is often rejected and required to be revised.

With all initiatives on sustainability going on we anticipate the rise of Sustainability - or Carbon planning - in these applications.

Actual data, to be compared with the plan/budget/forecast, is ingested, curated, and stored in the data platform side by side with the plan which have been imported from the planning applications. There are different types of analytical processing on the actual (and plan-) data, i.e., reactive, and proactive. Reactive means reporting on history, looking at trends over time and creating a sense of where we are today and why we are here. Proactive means predictive analytics, prescriptive and cognitive analytics. Here we enter the domains of AI and machine learning. We have already seen Sustainability Cockpits and Dashboards emerge.



Fitting it all together

Within an enterprise selling products or providing services, it is common to do product costing to know which price to ask in the market and which margins it will bring. The same goes for sustainability and carbon neutral scenarios – enterprise should do Carbon Costing for their products and services and make them transparent and available throughout the product lifecycle.

Certainly, the enterprise should also do a Carbon Plan – with normal top down-, bottom up- and write back capabilities mentioned before. Without a plan the enterprises will not know if their sustainability achievements are good or bad, ahead or behind targets. Carbon costing alone will not do the trick since it calculates carbon footprint product by product. But with planning applications being able to use allocation keys, the drivers behind carbon footprint can be spread across and allocated to all levels of the enterprise. This ensures scalability and speed and ability to handle complexity.

On the accounting side, taking place in ERP's or expert applications, we see Carbon Accounting now evolving. In most of the cases in conjunction with the Green Gas protocol. Separate ledgers are being set up, side by side with the regular general ledger. CO2 is then the currency, instead of EUR or USD.

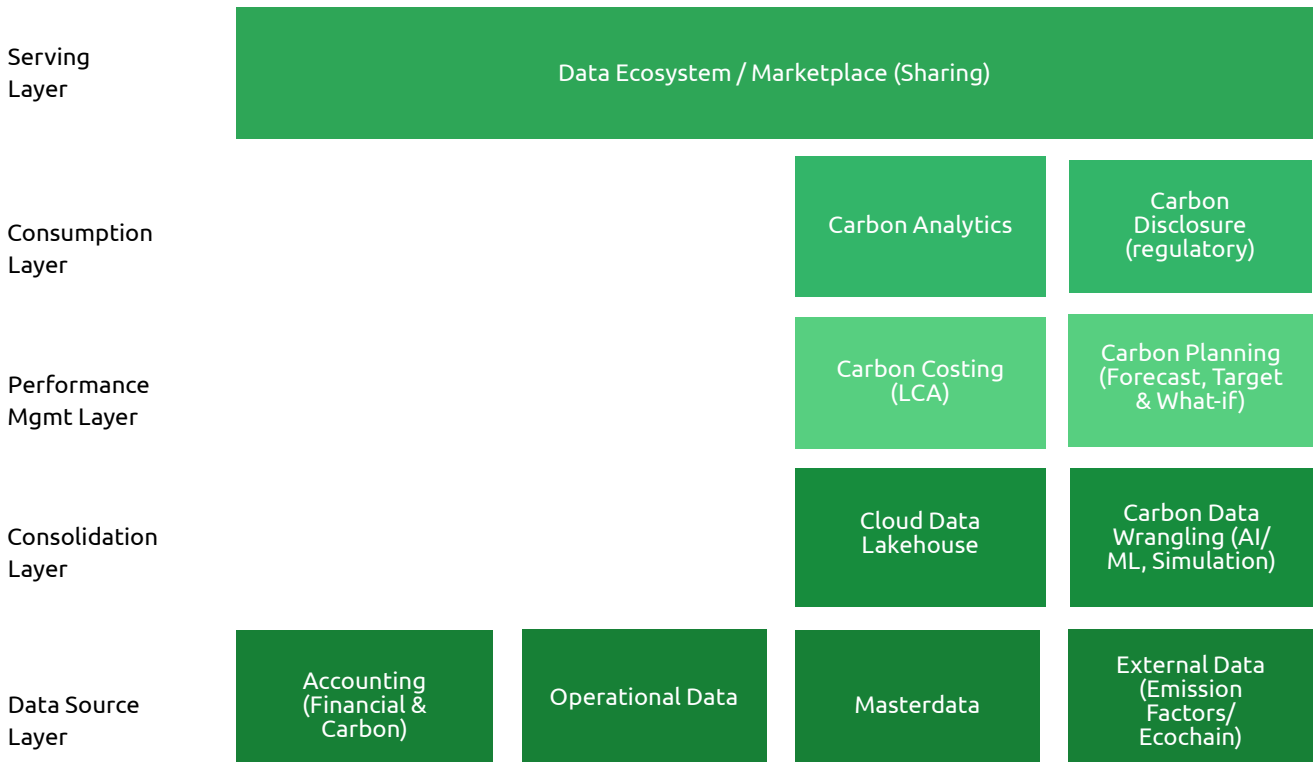
For Carbon Analytics everything needs to be combined, carbon costing, carbon planning, carbon accounting and actual data from various data sources. When all that data is collected in the Data Lakehouse, analytics and reporting can take place. Cloud data platforms with agility to integrate to new data sources and formats, and its elasticity and scalability will be enabling this. Most enterprise already have one or several of these platforms in place. Do they need to buy another one just for sustainability data? No, leverage the exiting investment to also harness sustainability data is the goal! To steer things right, even before they occur, predictive and prescriptive

analytics is needed on top of the data.

Data sharing and Data Ecosystems will be key requirements since the ESG-data need to be shared not only within the own organization, but outside the enterprise, too. We already now see expert companies providing ESG-data as benchmark or as ruling data on whom to do sustainable business with. This data should be integrated as external market data into the company's own calculations and ambitions.

Sustainability Data Hub

In the Data for Net Zero service from Capgemini we see main objectives such as Measure, Improve and Anticipate to be addressed by our clients. 'Anticipate' goes back to our reasoning about Carbon Costing and Carbon Planning above. 'Measure' goes back to Carbon Accounting and Carbon analytics on top of the consolidated data in the Cloud data platform. And by anticipating and measuring we will see 'Improvement' and change.



Graphic: Schematic Drawing of the Sustainability Data Hub Supporting Data for Net Zero



To conclude, we have learned a lot from the journey towards becoming truly data-powered. And in the same fashion that we have been using data in increasingly advanced ways to support – for example – finance and accounting, we can now do the

same for the race towards Net Zero. Only much faster. It's a matter of not reinventing the wheel, but standing on the shoulders of (data) giants.



“Enterprise should do Carbon Costing for their products and services and make them transparent and available throughout the product lifecycle.”

INNOVATION TAKE AWAYS

#ARCHITECTURE
#CARBONDATA
#SUSTAINABILITYDATAHUB

It starts with carbon costing

Just like with the organization's products and services 'normal' costing activities, Carbon Costing should be applied, making it transparent and available throughout the product / services lifecycle

A carbon plan should be next

A Carbon Plan with established top down-, bottom up- and write back capabilities is needed, to establish whether objectives are actually met.

Carbon analytics combine it all

Carbon costing, carbon planning, carbon accounting and actual data from various data sources come together – to be responsive and proactive.

Carbon data should be shared

No organization can run its net zero journey on its own; hence capabilities to share data and operate in data ecosystems are key.



REINVENTING DATA FOR THE PUSH TO NET ZERO



VALÉRIE PEHRIN

Head of Insight-driven Enterprise,
Capgemini Invent

There's a pressing need to reinvent economies and societies to eliminate carbon emissions – and enterprises have an important role to play in addressing this. The challenge is huge but, with top-quality data from across the organization's ecosystem, we can apply powerful assets such as AI to the task. Success moves enterprises beyond regulatory compliance by unlocking sustainability data to drive innovation.

As societies strive to become more sustainable, jurisdictions such as Europe and California are leading the world in establishing regulations governing environmental performance – including but not limited to carbon emissions, packaging, waste, and pollution. Laws vary by industry sector and jurisdiction but, regardless of the complexity this creates, it's essential that organizations comply with all regulations.

At the same time, enterprises in most sectors today understand that treating sustainability solely as a compliance issue is no longer an option. A poorly-defined, poorly-executed sustainability strategy has significant, negative consequences. Here are details on some of these.

- [In Consumer Products and Retail: How sustainability is fundamentally changing consumer preferences](#), the Capgemini Research Institute found 79 percent of consumers are

changing purchase preferences based on sustainability – and that sustainability has the potential to significantly impact customer experience, happiness, and loyalty.

- Organizations in certain sectors – such as automotive, consumer products and retail, and energy – are particularly vulnerable to the impact of sustainability on brand image. But sustainability issues have the potential to affect any organization. For example, just as sustainability strategies inform consumers’ purchasing decisions, they also influence a company’s reputation with its B2B customers and its desirability with potential hires as a place to work.
- Investors increasingly assess net-zero strategies as part of their due diligence before investing in an organization. A poorly-defined strategy makes it harder for companies to attract this important source of funding, and incomplete, inaccurate, or unreliable data is a hurdle in demonstrating the strategy’s effectiveness.
- Climate change and extreme weather events are already having a massive effect on companies across all industry sectors. Without a solid sustainability strategy, it’s difficult for companies to assess risk and take action either to avoid such events or mitigate their consequences.
- These few examples make it obvious that sustainability performance and financial performance are intrinsically linked – which is why many companies now assign responsibility for environmental performance to the Chief Financial Officer.



Measurement and innovation

In today’s data-powered business environment, it’s a given that having high-quality data is necessary if a company is to comply with regulations and benefit from a strong sustainability strategy. But few organizations possess the tools, technologies, processes, and culture required to capture, qualify, and activate trusted data. Collecting and managing high-quality sustainability data should be the first objective of every enterprise.

The good news is that once this is achieved the company can start leveraging that data to innovate. AI is a powerful tool for this, able to combine data from across the organization – as well as from upstream and downstream sources such as suppliers, distributors, and retailers – and then derive insights,

make recommendations, and share them across business functions and the value chain. Some examples from my work at Capgemini Insight include:

- Building an AI-powered simulator to help a company in the mining sector anticipate and reduce the carbon impact of proposed IT projects
- Applying AI to review the manufacturing operations at another company in the mining sector, to reduce the carbon footprint of its raw materials use
- Using AI to assess consumer shopping behavior, reducing fresh fruit waste for a European retailer
- Implementing AI-optimized processes to help a manufacturer reduce energy consumption, and cost, by seven percent.

Anticipation and digital twins

Companies can also leverage data to anticipate the impacts of sustainability decisions on company performance. For example, subcontractors with excellent sustainability track records are generally not the cheapest option, so enterprises must determine how best to balance financial performance with sustainability performance – and then convince stakeholders such as investors that this is the right decision.

Digital twins are emerging as a useful tool in cases such as this. Digital twins allow enterprises to use their data to create a virtual representation of their operations and then use AI and other technologies to apply simulations to the data and measure the outcomes. Changes can have significant, company-spanning repercussions, both positive and negative, impacting everything from customer satisfaction to financial performance. It's crucial that

decision-makers have the opportunity to test and assess such ideas before implementing them.

As with many new initiatives, companies looking for success should take a pragmatic approach. Articulate a clear vision and focus on internal assets before incorporating data from partners, subcontractors, clients, and other outside sources. AI-derived insights can help decision-makers prioritize sustainability issues. Companies can then focus on a pilot project before scaling up to span the enterprise's ecosystem.

Data mastery links sustainability with innovation

Companies that become data masters find it easier to supply the information required for compliance with environmental regulations. But that's just the start. Whether it's helping R&D develop new products and services, providing marketing with the insights to boost brand

image, or identifying potential new business models for the company to consider, data masters can use AI tools and that strong data foundation to help drive innovation, unlock value throughout the company's ecosystem, and accelerate the organization's sustainable transformation.



Collecting and managing high-quality sustainability data should be the first objective of every enterprise."

INNOVATION TAKEAWAYS

#SUSTAINABILITYSTRATEGY
#DIGITALTWIN
#PERFORMANCE #DATATRUST
#GOVERNANCE

Sustainability is about more than compliance

Accurate, trusted sustainability data is essential for satisfying regulatory compliance. But many others – from consumers to investors – also demand high-quality data about a company's sustainability strategy.

AI powers sustainable innovation

Applying AI to high-quality sustainability data is an opportunity to innovate in ways that build brand image, attract investment, reduce operating costs, and mitigate risk.

Find the balance

Enterprises must walk a fine line between sustainability performance and financial performance. AI-powered simulations can help companies stay on the right path while avoiding missteps.

AUTHORS



**ARNE
ROSSMANN**

Head of AI & Data Engineering
Germany, Capgemini



**ASHOWARDHAN
SOWALE**

VP, I & D Architecture Head, India
Domain Leader for AI, Capgemini



**BAPTISTE
JOURDAN**

CRO & Co-founder
of Toucan Toco



**GUILLAUME
GERAD**

Lead Data Scientist,
Insights & Data France, Capgemini



**JEAN-MARIE
LAPEYRE**

Chief T&I Officer, Global Automotive
Industry, Capgemini



**JOHAN
KREBBERS**

For The Open Group



**JEROME
DESBONNET**

VP, Cybersecurity CTIO, Insights &
Data, Capgemini



**JULIAN VAN
VELZEN**

CTIO NL & Head of Capgemini's
Quantum Lab, Capgemini



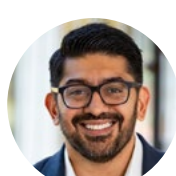
**MYLES
SUER**

Senior Director, Head of AI& AA,
BCM, Capgemini



**NADINE
VAN SON**

Senior Consultant, S&I and
Transformation, FS Capgemini



**NEERAV
VYAS**

Head of Customer First, Insights &
Data North America, Capgemini



**NICOLAS
CLAUDON**

CTIO SCE Insights & Data,
Capgemini

AUTHORS



PETTERI
VAINIKKA

VP Product Marketing, Cognite



ROBERT
ENGELS

CTO I&D NCE, Capgemini



RON
TOLIDO

CTO I&D, Capgemini



RODRIGO
MAIA

Vice President, Head of Research,
Technology & Innovation, Capgemini



ROOSA
SÄÄNTTI

Head of Insights & Data
Capgemini Finland



SHANKAR
RAMANATHAN

Senior Director, Head of AI& AA,
BCM, Capgemini



URBAN
EKERÖTH

Head of Portfolio, Insights & Data
Nordics, Capgemini



VALÉRIE PEHRIN

Head of Insight driven enterprise,
Capgemini



VINCENT DE
MONTALIVET

Data for Net zero Offer Leader,
Capgemini



TED
KWALTER

VP of Trusted AI,
DataRobot



For more details contact us:

RON
TOLIDO

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

ron.tolido@capgemini.com

MONISH
SURI

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

monish.suri@capgemini.com

JOHANNA
SUNDH

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

johanna.sundh@capgemini.com

LAURA
IMMONEN-BEATTY

Global Content Marketing Lead,
Insights & Data, Capgemini

laura.immonen-beatty@capgemini.com

Interested in learning more?

[Download Wave 1](#)

[Download Wave 2](#)

[Download Wave 3](#)

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.

Get the Future You Want | www.capgemini.com