



**Deloitte.**

*Together makes progress*

Hyperpersonalization in the age  
of High-Performance Computing

How financial institutions can differentiate  
to beat the value compression

# Introduction

Hyperpersonalization in financial services is the artificial intelligence (AI)-powered, real-time delivery of individualized products, services, insights, or experiences that contextualize transactional, behavioral, and emotional data across the customer's physical and digital footprint. As financial products become increasingly commoditized, differentiation now lies not in what banks offer, but in how, when, and through which channels they deliver value. Institutions that control the point of distribution—owning the digital and physical touchpoints of engagement—capture a disproportionate share of margins by shaping the customer's financial journey end to end.

This white paper explores how hyperpersonalization, powered by high-performance computing (HPC) and AI/machine learning (ML), is redefining the customer experience in the financial services sector. It presents a compelling argument for moving beyond traditional segmentation models and embracing a "market of one" approach, in which banks tailor products, advice, and interactions based on an individual's real-time behavior, preferences, and financial lifecycle stage.



## Introduction

Banking value compression

Hyperpersonalization

High performance computing (HPC)

Looking ahead

Endnotes

Authors

# Banking value compression

Core banking products like deposits, payments, cards and loans have become commoditized with very little differentiation. Customers are shopping around for the best deal and able to switch easily. As a result, banks are facing intense competition – and not just from other banks but from tech startups, which are able to offer services at lower cost and leveraging revolutionary new technologies. Banks are also burdened with regulatory changes and aging technology systems, which are a significant drain on their capital.<sup>1</sup>

Banks are therefore experiencing “Value Compression”— a phenomenon in which the margin i.e. the difference between the price banks can charge for services and the cost of delivering those services reduces over time. As a result, banks are able to capture less value per transaction.

Countering value compression requires creating new revenue streams, enhancing customer loyalty and continuously finding ways to differentiate offerings. Hyperpersonalization could be one strategy that may help.



Introduction

**Banking value compression**

Hyperpersonalization

High performance computing (HPC)

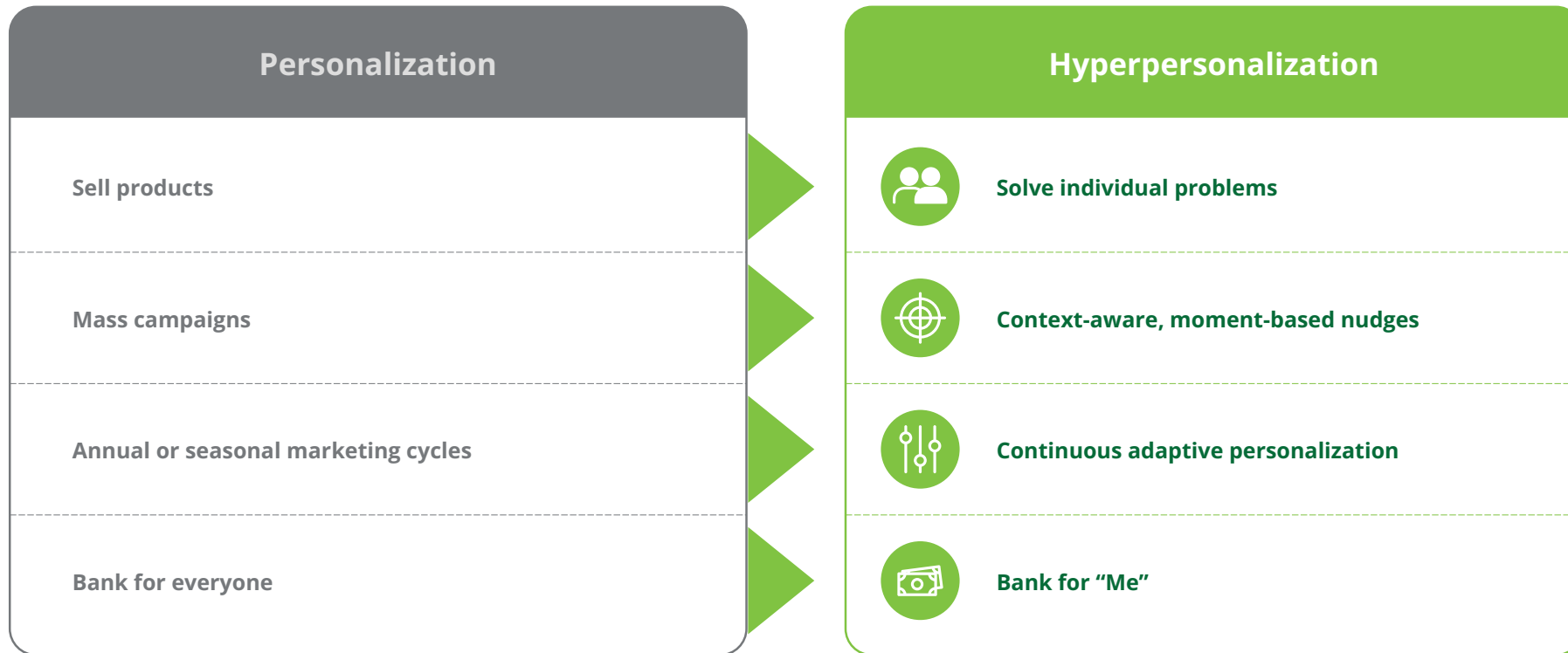
Looking ahead

Endnotes

Authors

# Hyperpersonalization

Hyperpersonalization is about creating highly tailored products, services, and experiences for each customer. This goes beyond simple personalization, which says, "Here's a product you might like." Hyperpersonalization would say, "Here's the solution you need right now, based on what you're doing, feeling, and aiming for."



Hyperpersonalization includes analysis of spending behaviors, life events, channel preferences, and context to proactively anticipate individuals' needs and then respond to those needs by tailoring a product or service in real time that addresses those needs. Fulfilling such needs accurately and immediately will allow banks to capture higher value per transaction.

Introduction

Banking value compression

**Hyperpersonalization**

High performance computing (HPC)

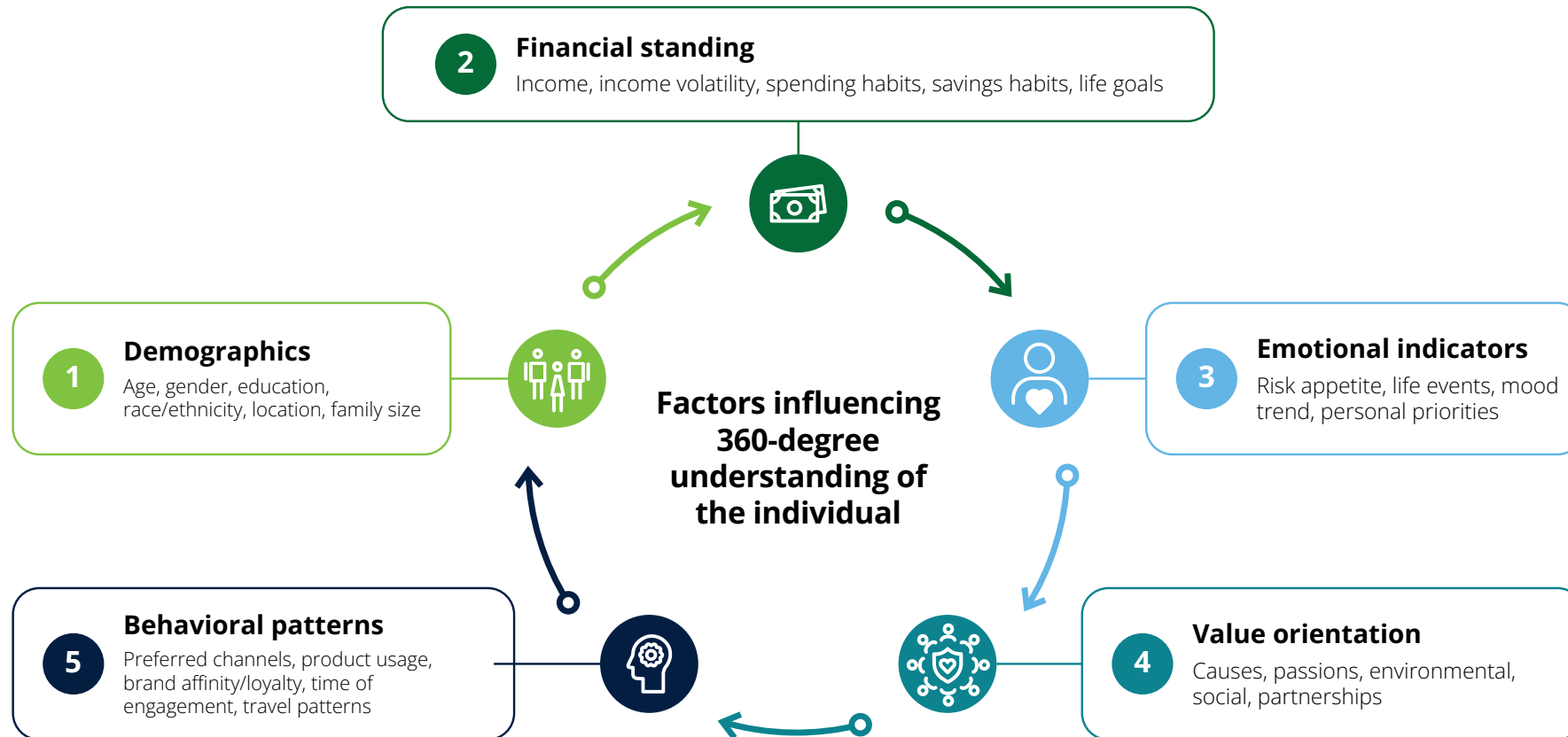
Looking ahead

Endnotes

Authors

## 1. A 360-degree understanding of the individual

Hyperpersonalization requires institutions to shift from addressing segments to engaging each customer as a market of one. This means going beyond demographic or transactional data to build a 360-degree understanding of the individual, as described in the following diagram:



Introduction

Banking value compression

**Hyperpersonalization**

High performance computing (HPC)

Looking ahead

Endnotes

Authors

## 2. Contextual intelligence across the journey

After understanding who the customer is, banks will need to know the context of the customer (i.e., where they are) and what they need to understand:

- The customer's current goal or pain point (e.g., saving for a house, looking for emergency liquidity)
- Where they stand in their financial life cycle or decision-making process
- Their most recent interactions and behaviors, across physical and digital channels

This contextual intelligence will enable banks to design experiences that will naturally integrate into the customer's life:

- Anticipate the next best action (e.g., recommending an investment product to someone who just received a large bonus)
- Avoid irrelevant or untimely offers (e.g., suggesting a credit card balance transfer offer to a transactor)
- Deliver seamless omnichannel continuity (e.g., picking up from where the last mobile or branch interaction left off)

## 3. Timeliness and micro-moment activation

Timing is everything. The most sophisticated personalization strategy falls flat if the response comes after the moment has passed. Hyperpersonalization ensures that actions are taken not just intelligently, but also instantly; with latency ranging from milliseconds to minutes, depending on the use case. Examples include:

- Sending a spending limit alert before a customer exceeds their monthly budget
- Triggering a personalized investment offer right after a market dip that matches their portfolio
- Customizing and delivering services via a push notification based on the customer's real-time activity (e.g., customer walking into an airport lounge receives a push notification on the customized offers)

Importantly, hyperpersonalization does not always mean in real time. It means timely within the context of the customer's need—delivered at the most relevant point in their decision-making window.

72% of customers desire immediate service, and 70% expect any representative they interact with to have full context of their interactions<sup>2</sup>

Introduction

Banking value compression

**Hyperpersonalization**

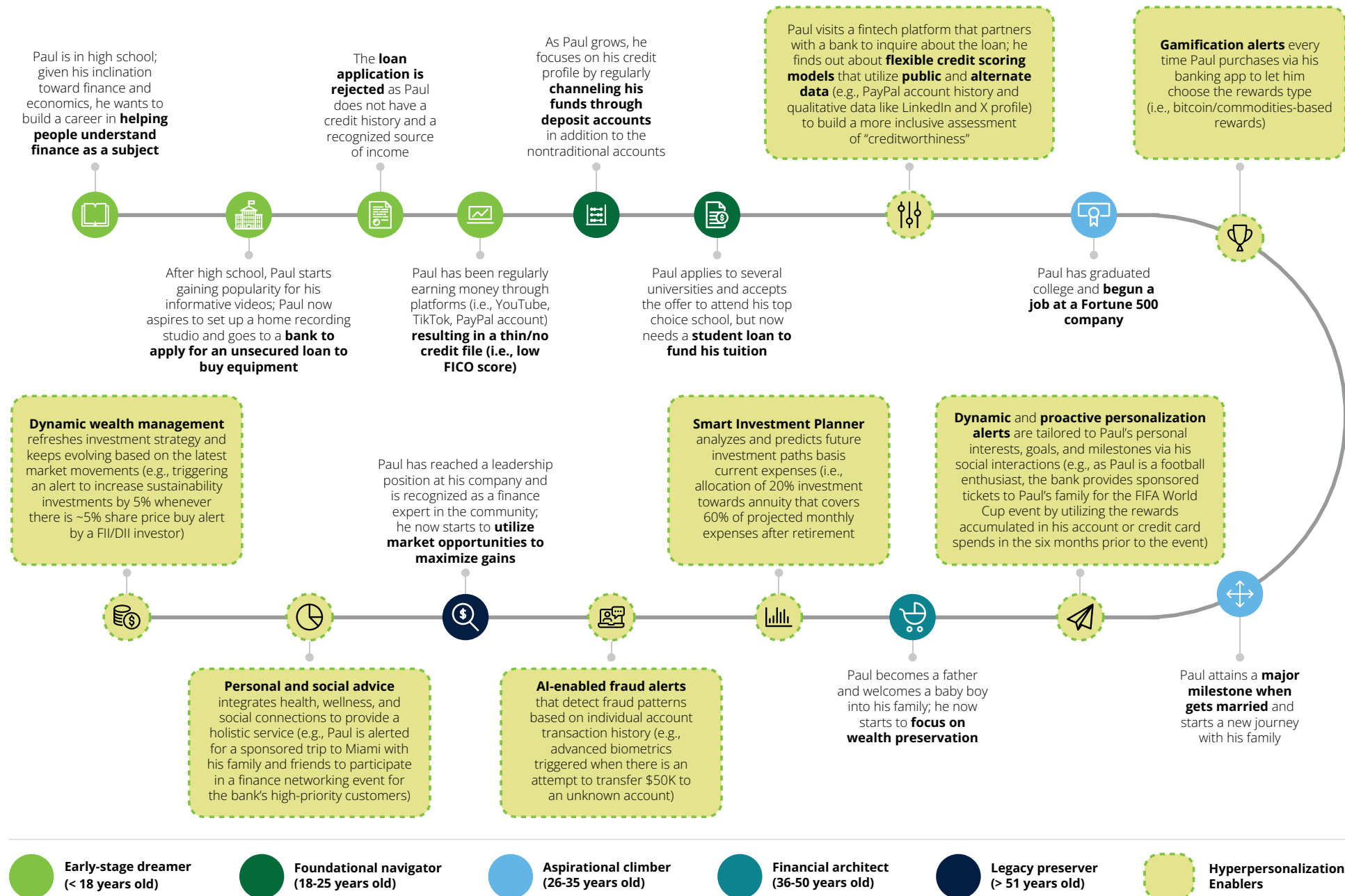
High performance computing (HPC)

Looking ahead

Endnotes

Authors

# Childhood to retirement: hyperpersonalized banking along life's journey



- Introduction
- Banking value compression
- Hyperpersonalization**
- High performance computing (HPC)
- Looking ahead
- Endnotes
- Authors

# When context needs scale

Leading banks are increasingly combining rich customer data and advanced AI operations to deliver real-time, tailored experiences at scale. The following use cases highlight where global banks are driving measurable impact:

Solution	Description	Capabilities that make it real
<b>Automated savings and spend insights</b>	<ul style="list-style-type: none"> <li>Always-on AI analyzes monthly cash flows, categorizes spends</li> <li>Provides budget tips based on spending habits</li> </ul>	<ul style="list-style-type: none"> <li>Deep customer data</li> <li>Real-time event streams</li> <li>First-party AI platforms</li> <li>Model ops at scale</li> </ul>
<b>Personalized market data research and analytics</b>	<ul style="list-style-type: none"> <li>Integrates market data, research, portfolio analytics, and client behavior</li> <li>Identifies what's most relevant to each client, in real time</li> </ul>	<ul style="list-style-type: none"> <li>Internal data products</li> <li>Quant libraries</li> <li>Robust entitlements</li> <li>Cloud-native delivery</li> </ul>
<b>AI-driven cash flow and working capital analytics</b>	<ul style="list-style-type: none"> <li>AI algorithm analyzes cash flows and creates custom enterprise-grade forecast in seconds</li> <li>Determines "next best action" for receivables, payables, and borrowing requirements</li> </ul>	<ul style="list-style-type: none"> <li>Embedded enterprise resource planning (ERP) connectors</li> <li>Data science capabilities</li> <li>Treasury data network integrations</li> </ul>

In contrast, many regional banks face significant compute constraints, fragmented data, batch-based cores, and tighter budgets, leaving them dependent on campaign-style personalization that falls short of the millisecond responsiveness needed for true in-the-moment engagement. So how do regional banks gain necessary scale without headcount or skill parity with large banks?

**Financial Institutions implementing hyperpersonalization generate 40% more revenue compared to competitors<sup>3</sup>**

Introduction

Banking value compression

**Hyperpersonalization**

High performance computing (HPC)

Looking ahead

Endnotes

Authors

## Digital twins

After understanding who the customer is, banks will One key capability that could make hyperpersonalization possible is digital twins.<sup>4</sup> A digital twin is a virtual replica of the real user, that is continuously updated with the customer's data, behaviors, preferences, and risk tolerances. This information is used to continuously predict the customer's actions.

A bank would need to maintain one such digital twin for each of its customers, which could range from hundreds of thousands to millions. The bank would need capabilities like prediction models, risk and pricing simulations, and behavior clustering. All of these will require massive parallel computing that traditional cloud computing infrastructure is not optimized to economically run. The speed, volume, and complexity require high-performance computing.



### Speed

Response time in milliseconds, critical for moment-based engagement



### Volume

Petabytes of customer interactions, financial data, and context for millions of users



### Complexity

From credit scoring to fraud detection to customer intent modeling

Introduction

Banking value compression

**Hyperpersonalization**

High performance computing (HPC)

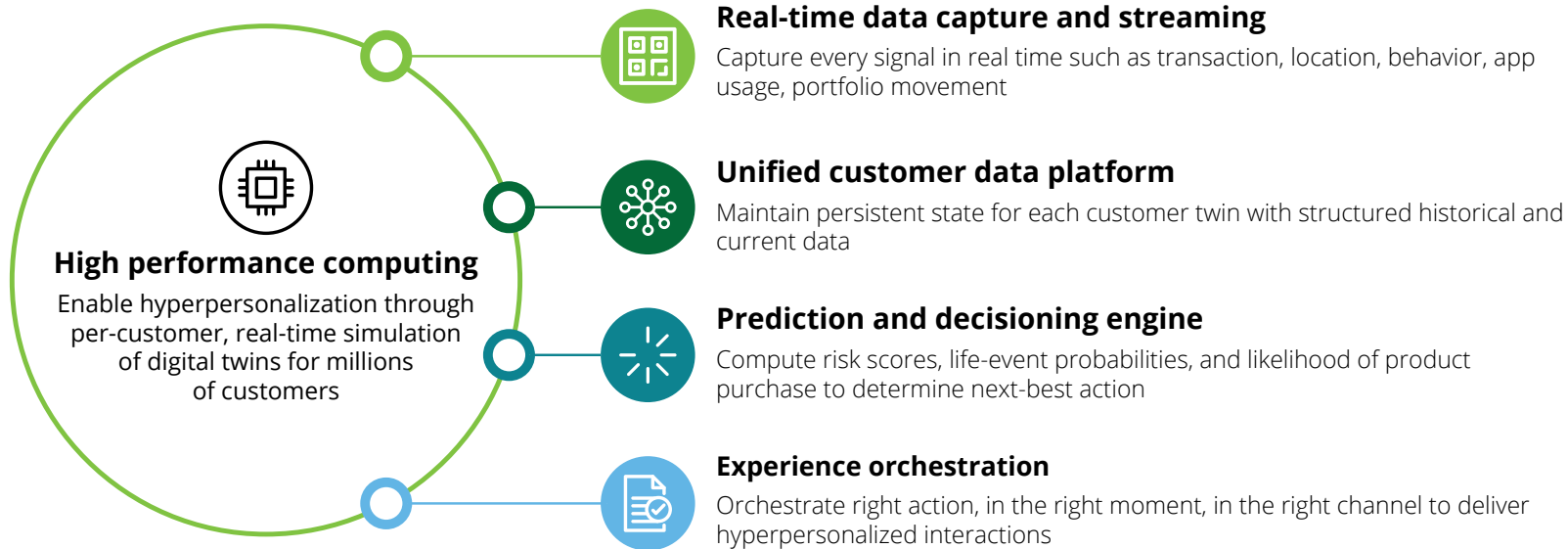
Looking ahead

Endnotes

Authors

# High-performance computing (HPC)

High-performance computing refers to the use of many compute nodes working in parallel to solve complex problems at scale that are otherwise difficult to solve by normal servers.<sup>5</sup> HPC is different from traditional cloud auto-scaling capabilities, which include scale-up (more power to one machine) or scale-out (more machines to handle separate tasks). HPC uses many machines that are tightly synchronized to solve the same problem at the same time. HPC should not be confused with supercomputers, which are a purpose-built, extremely powerful subset of high performing computers. HPC, in the context of hyperpersonalization, can be achieved with commodity hardware running on-premises or in the cloud that provides necessary compute required for real-time digital twin simulations per customer.



Modern HPC solutions empower banks to analyze millions of transactions and data points in real time, with sub-second latency, enabling truly responsive and intelligent decision-making at scale<sup>6</sup>

Introduction

Banking value compression

Hyperpersonalization

**High performance computing (HPC)**

Looking ahead

Endnotes

Authors

# Augmenting existing banking systems

A bank has several systems that could be classified into two categories:

- Systems of record: Include customer relationship management, core banking, loan servicing, card processors, and data warehouses. These systems are product-centric and optimized for transaction processing.
- Systems of engagement: Include user-facing channels like online portals, mobile apps, chatbots, social media, and wearables. These systems are experience-centric and optimized for multichannel orchestration and low-friction journeys.

Neither of these systems are, however, designed for simulations and predictions. HPC is not a replacement for either of these systems. Think of it as a new “systems of intelligence” layer that rightly sits between these systems of record and engagement.



Introduction

Banking value compression

Hyperpersonalization

**High performance computing (HPC)**

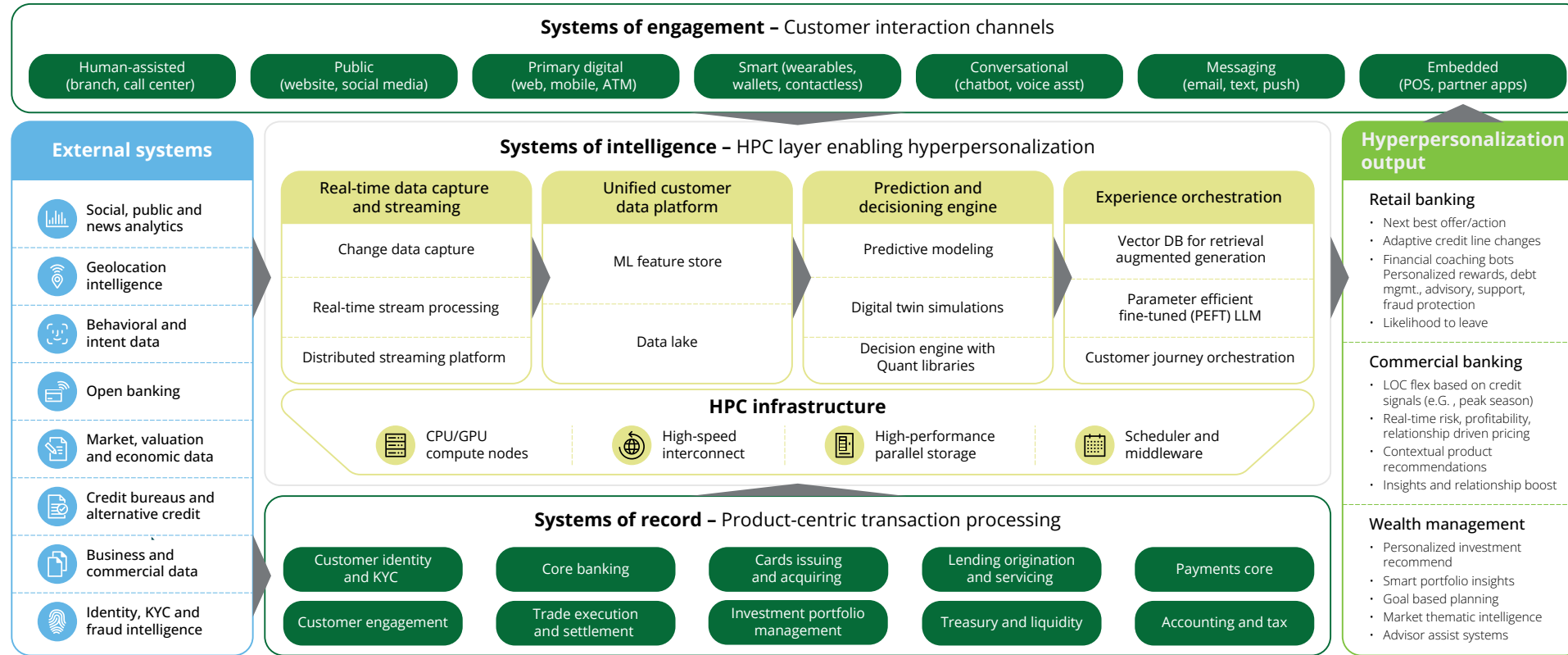
Looking ahead

Endnotes

Authors

## Business architecture for hyperpersonalization

The business architecture presented here lays out how financial institutions can harness the power of high-performance computing to source and process a wide variety of inputs; run predictive analytics and digital twin simulations; and deliver hyperpersonalized insights in the right language and channel, and at the right moment.



This will involve sourcing data from a wide variety of systems ranging from internal systems of record to external business and market data, and also behavioral and intent signals coming from social media and devices. HPC infrastructure will enable real-time capturing and processing of such data into a unified data platform that is then used for large-scale and continuous simulations and predictions. The platform will have to run through the bank's quant libraries (mathematical and statistical tools necessary for quantitative finance) for pricing and risk calculations, which will then be factored into custom offer decisions. An experience orchestration layer will then distribute these offers in the applicable language, context, and interaction channel.

Introduction

Banking value compression

Hyperpersonalization

High performance computing (HPC)

Looking ahead

Endnotes

Authors

## Implementation considerations

Implementing HPC for hyperpersonalization requires more than just compute power. It will require strict controls on consent, data usage, and security. Banks should consider built-in customer transparency and choice throughout so these systems of intelligence act as a secure, privacy-protecting layer without duplicating or leaking sensitive data, which can help maintain customer trust.



### Customer consent

Trust is essential to the effective delivery of banking services. As hyperpersonalization includes processing significant customer data, banks need to make sure that they have obtained explicit opt-in consent from their customers. Customers should further have the ability to granularly choose what they share and easily turn off data-sharing as needed. Compliance with data protection laws and regulations, such as General Data Protection Regulation (GDPR) and Central Consumer Protection Authority (CCPA), is non-negotiable.



### Governance

Regardless of the consent, banks should carefully select what minimum data is required for the purpose of specific use cases and only collect that information. This will not only reduce data sourcing and storage costs but also strengthen regulatory compliance. Also, banks should establish clear boundaries for invasive inferences such as pregnancy, divorce, or medical conditions. Hyperpersonalization is not surveillance.



### Data quality and security

A wrong prediction isn't neutral—it will irritate the customer and damage trust. Quality of the underlying data is therefore key. Unfortunately, a lot of this data will come from legacy systems of record that typically have data quality issues. Therefore, consider investing in establishing data lineage, identifying trusted sources of data, and considering zero-duplication architecture to avoid outdated copies. Finally, a lot of this data is confidential and sensitive. Therefore, banks should apply data at-rest and in-transit encryption as necessary.



### Infrastructure modernization strategy

While HPC can be built from commodity hardware, banks must balance the scale vis-à-vis cost to a select location for such clusters between clouds or within their own data centers. Alternatively, a hybrid model can be considered also to bring the best of both worlds. Again, various considerations such as latency, data sensitivity, and capacity will play into the decision.



### Talent and operating model

Hyperpersonalization requires multidisciplinary teams—data engineers, AI scientists, HPC specialists, and domain experts—working in agile, cross-functional pods. Regional banks must define the right build vs. partner strategy, invest in upskilling or cosourcing talent for AI and HPC capabilities, and establish an operating model that seamlessly connects business, technology, and compliance functions.

Introduction

Banking value compression

Hyperpersonalization

**High performance computing (HPC)**

Looking ahead

Endnotes

Authors

# Looking ahead

Value compression in banking is real. The banks that win the hyperpersonalization race will be the ones that shift from processing transactions to predicting customer needs. HPC and AI capabilities will allow building real-time digital twins, personalize every interaction, and turn data into actionable and monetizable insights. The future of banking may not be defined purely by the biggest balance sheet but by the smartest decision engine.

Deloitte's deep banking experience combined with our virtually unmatched AI and data engineering depth makes us someone you can trust in this journey. We have the necessary capabilities and alliances with industry-leading hardware and software vendors to plan, build, and operate such HPC clusters.



Introduction

Banking value compression

Hyperpersonalization

High performance computing (HPC)

**Looking ahead**

Endnotes

Authors

# Endnotes

1. Katherine P, "[The true cost of legacy systems: A deeper dive into banking IT modernization](#)," Digital Bank Expert, August 11, 2025.
2. Ram Prabhakar, "[Top 10 ways to improve customer experience in banking 2025](#)," Xerago, last updated July 11, 2025.
3. Brent Walker, "[Hyper-personalization ROI for credit unions & banks](#)," Psympl, September 4, 2025.
4. Frances Yu, Brian Campbell, and Timothy Murphy, "[From manufacturing to medicine: How digital twins can unlock new industry advantages](#)," Deloitte Insights, June 5, 2025.
5. IBM, "[What is high performance computing \(HPC\)?](#)," accessed December 2025.
6. Stuart Midgley, "[HPC networking: A ballad in four parts. \(Part 1: Latency\)](#)," The Daily Dug, December 9, 2020.

Introduction

Banking value compression

Hyperpersonalization

High performance computing (HPC)

Looking ahead

**Endnotes**

Authors

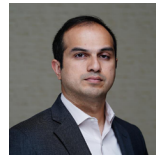
# Authors



**Tushar Daru**

Principal

Deloitte Consulting LLP  
[tdaru@deloitte.com](mailto:tdaru@deloitte.com)



**Vaibhav Sathe**

Managing Director

Deloitte Consulting LLP  
[vasathe@deloitte.com](mailto:vasathe@deloitte.com)



**Aditya Kumar Jha**

Senior Consultant

Deloitte Consulting India Pvt. Ltd.  
[adkumarjha@deloitte.com](mailto:adkumarjha@deloitte.com)



**Abinash Roy**

Senior Consultant

Deloitte Consulting India Pvt. Ltd.  
[abinroy@deloitte.com](mailto:abinroy@deloitte.com)

We thank **Saurabh Sanjay Soparkar** and **Jackson Atassi** for their valuable contributions.

Introduction

Banking value compression

Hyperpersonalization

High performance computing (HPC)

Looking ahead

Endnotes

**Authors**

# Deloitte.

*Together makes progress*

This publication contains general information only and Deloitte is not, by any means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor.

Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

#### **About Deloitte**

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee (“DTTL”), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as “Deloitte Global”) does not provide services to clients. In the United States, Deloitte refers to one or more of the US member firms of DTTL, their related entities that operate using the “Deloitte” name in the United States and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting. Please see [www.deloitte.com/about](http://www.deloitte.com/about) to learn more about our global network of member firms.

Copyright © 2026 Deloitte Development LLC. All rights reserved.