

A New Paradigm in Retirement Solution Platforms: Leveraging Data Governance to Build AI-Ready Data Products

Ramesh Inala,

Data Engineer, Prudential Financial, Newark, NJ, rameshhinala@gmail.com, ORCID ID: 0009-0009-2933-4411

Abstract

Over the last three decades, multiple generation-defining retirement crises have led to increased scrutiny of the retirement solutions available to American workers. Unlike the defined benefits provided by the previous generation's employer-employee relationship, today's employer-employee relationship typically provides more flexible employer-provided defined contribution plans. However, a lack of employee/retiree engagement, combined with a fragmented and inefficient in-plan asset accumulation and out-of-plan decumulation ecosystem, raises significant risks for employers and employees alike. The pace of adaptive change has accelerated. The need for innovation has been magnified, fueled by the timely confluence of a broadening landscape, expanding digital technologies, and increased regulatory oversight. The resulting need to innovate has led many platform strategists to focus on the identification, curation, and delivery of transformative ecosystem services enabled by the wise use of authoritative data craftsmanship.

In this paper, we expand on this work, integrating our software development lifecycle practice with a data product lifecycle perspective in highlighting the synergies that can be gained by recognizing the seamless interdependencies between services and data enablers. Our work offers a data ecosystem services value proposition framework—one that recognizes the operational servitization transition as well as the symbiotic and catalytic roles that data-driven services and services-driven data play in shaping the digital ecosystem strategies of individual platform participants. In crafting this framework, we weave together two strands of literature: one that has focused on the service and platform marketization of traditional product-firm relationships and one that has focused on the datafication of product-service relationships, enabling smart digitally augmented products and related product families.

Keywords : Retirement Solution Platforms, Data Governance, AI-Ready Data, Data Products, Digital Transformation, Financial Technology, Retirement Planning, Data Quality, Machine Learning Integration, Data-Driven Decision Making, Regulatory Compliance, Scalable Data Architecture, Intelligent Automation, Personalized Retirement Solutions, Data Management Strategy.

1. Introduction

A growing movement is present among participants in the 401(k) industry aiming to expand the asset retention capability of qualified retirement plans. By providing unique benefits in the form of avoiding capital gains tax and generating income during retirement, 401(k) plans that allow in-service withdrawals and post-retirement distributions are increasingly attractive to employees and account owners. This is particularly true as people shift their focus from retirement accumulation to planning for retirement income generation, facing increases in longevity and concerns surrounding retirement account depletion. Expanding asset retention also benefits plan sponsors

by enhancing their reputation and increasing fees associated with custodial services or investment management. The retention of assets in a defined contribution plan also benefits the financial services industry by generating fee revenues while minimizing the capital gains tax burden on retirees, an expanding and increasingly affluent demographic.

Despite developments in the industry, the majority of retirees have their defined contribution accounts rolled into taxed brokerage accounts rather than left in the qualified plan. This results in suboptimal outcomes for both account holders and their advisors. There is a high likelihood that the money will not be retained within the brokerage account for any significant time. It is unlikely to be traded or invested in a manner that generates a long-term income stream, protected against the risk of market downturns. As a result, the retiree will suffer market risk. At some point soon, these assets will likely be taken out of the account, triggering a capital gains tax event. If that money had remained in the tax-advantaged account, such risk and tax events could have been avoided for many years.

1.1.1. Background And Significance

The meaning and significance of the retirement resource more commonly defined as qualified retirement accounts, is often lost on the investor. A wide variety of available investment options, the fact that these accounts are often managed by middle-man financial advisors versus being direct retail brokerage accounts, and the simple fact that retirement is a long way off and the money is still getting contributed from payroll deductions, leaves the savers in void. It is with this in mind, that discussions of the impending retirement crisis have focused on the next generation of investors who will inherit family wealth, which will likely comprise more retirement resources than any other type of wealth. Family offices, usually associated with the extremely wealthy, are becoming increasingly common. Just as family offices manage the shared assets large family clans, adding a family office function to a retirement program is becoming a consideration for companies looking to provide a more robust set of services to employees who are toiling away in the shadow of accounts, often without the benefit of custodians or advisors who specialize in retirement programs.

Retirement plan sponsors are struggling to keep up with the growing demand for plan features, worker records, and plan assets. Many providers have turned to data centers. There are few without major obstacles. The obstacles can be categorized as strategic, tactical, or operational. Misdirected efforts in these areas can exacerbate difficulties associated with plan data. With the right strategy and approach, and yes, it is possible to have such a strategy and approach. Investment providers can turn their strengths to their advantage. Instead of merely being an ancillary part of defined contribution plan management or, worse yet, a possible unwelcome burden. Investment recordkeeping and management as other purpose-built data programs, retirement solutions can yield important business growth and help reduce risk. Further, defined contribution plans can support providers' long-term data product strategy. Leveraging the knowledge garnered through the process uniquely qualifies recordkeepers to, for instance, manage the data associated with other plans.

2. Understanding Retirement Solution Platforms

Retirement solution platforms help individuals and businesses plan, save, and manage retirement. Globally, governments have either moved to privatize or incentivize the privatization of retirement funding. As a result, the private retirement account business has exploded. While global retirement fund assets under management have been constrained in growth to a recent ceiling, the influence of that money in global capital markets has remained unprecedented. Moreover, the leverage that asset management firms and fitness teams have been able to apply on pension fund inflows from businesses and institutions, in search of a low-effort investment avenue, has been uncharacteristically unique to the retirement space. The sustainability of this relationship is being currently questioned and policed by regulators. The current technological capabilities landscape of the retirement planning business is characterised by an over-dependence on a limited number of platforms. These platforms are antiquated and somewhat clumsy point solutions, with legacy technology underpinning enabling factors. Moreover, there is a lack of innovation in building, integrating, and maintaining the various capabilities required to build a compelling product stack. The inability to commoditize any one or two of the layers has led to a unique asymmetry, with a few large players competing head-to-head against a multitude of mid-sized competitors.

Traditional platforms were built around vertical stacks of services and capabilities, susceptible to first-mover advantage isolation. Hence, greedy players would block vertical integration by smaller players, while innovating continuously on capabilities along the stack. Moreover, legacy platforms were bald as each of the capabilities required to build an integrated solution for any one business, or set of users, were separately owned, built, and customized. Access to users and serviceability of the solution became the key pain points in the integrated proposition. Over the years, there has been consolidation around a few incumbent platforms that now cater to their set of users. These services intricate market segments through custom-built solutions.



Fig 1: Retirement Solution Platforms

2.1. Historical Context

Opinion shifts have influenced the retirement space, leading to a myriad of means to save, grow, and distribute assets over many financial institutions. The paternalistic focus of the 1940's through the 1980's, where employed individuals saving for retirement relied on their companies to provide pension products, has faded. Employers took the years of market product returns, risks, and expenses via pension plans to encourage employees to take the risks via defined contribution plans, allowing employees to accumulate funds for retirement using the funds' investments. Employees

have transitioned from traditional wage earners to augmented types of labor including gig/stage, and many have taken up self-employment. Additionally, the length of time in retirement has significantly increased over the years, leading to an increased focus on retirement planning throughout the working years.

As touch/feel service industries expanded, driven by the increase of the entertainment and entertainment capitalism, employers started favoring employee benefit packages that included not just pension inducements, but also products bundled with vacations, medical insurance, tuition, stock options, healthcare savings accounts, wellness checks, etc. The start of mutual funds and the explosive growth of the industry coincided with the dropping of capital gains tax rates. The focus on growth over income was most evident by the launch of a myriad of investment products in the '90s such as limited partnerships, tax credit funds, variable annuities, and hedge funds of funds. Regulations followed with the Pension Protection Act, triggering the rapid growth of automatic enrollment and default funds in plans. Retirement assets were shifted from pension plans to plans. Baby boomers, once the workers who did not think about retirement, are now the largest living generations nearing, in, or preparing for retirement. The country's fiscal system has been influenced, marred with the twin deficits, as a result of years of spendthrift consumption and the oncoming, aging of death.

2.2. Current Market Landscape

Retirement plans are an essential benefit of employment for most Americans. They represent one of the largest accumulation of assets during the working life of an employee, and in many cases, it is the only form of savings for future income replacement. Retirement plans also represent a uniquely American perspective on income replacement during old age. Although all developed nations have some forms of social safety net for the elderly, the American system is based more on personal savings and investment to augment whatever amount is provided by Social Security when the person becomes eligible. Retirement plan assets are also an important component of the U.S. economy. At the end of 2020, a significant amount was held in 401(k) and other employer-sponsored retirement accounts. With a large total in U.S. assets in 2020, retirement plan assets represented approximately 26% of the total. The combined U.S. population over the age of 65 years is projected to reach 80.8 million by 2040 and cost the federal government a substantial amount in Social Security and Medicare spending in 2040. And there are 118.7 million working-age (18 to 64 years) Americans who support this 80.8 million citizens who are dependent on the government and increasingly shifting toward a self-support structure.

Retirement solution providers occupy a niche within the financial services sector offering a particular selection of investment opportunities for retirement, pension, or corporate 401(k) plans. These generally include insurance firms and fund management companies. Investor offerings: Corporate pensions accounts; Corporate 401(k) plans; Pension funds. This report offers a comprehensive analysis of the global retirement solution providers market. The report focuses on key areas including market trends, key players, key country markets, and market projections. It aims to provide a comprehensive tactical analysis of the global market. All statistics and data are benchmarked against the global market, and across key macroeconomic drivers and performance indicators.

2.3. Challenges in Traditional Platforms

The diversification and specialization of firms and financial advising services have enhanced consumers' options on specific products to satisfy their financial purposes. During the accumulation stage, especially for the younger population, retirement savings would be the most daunting task. However, not many people are aware of the importance of finding the right balance throughout their saving period, since it is not too late to open their target retirement portfolio in their middle or late career stage. As a result, consumers are bombarded with advertisements on a wide range of financial and apparent non-financial product-based services such as credit cards, bank accounts, loans, and travel packages from thirty companies vying for attention, trust, and clicks. Overshadowing these commoditized approaches to retirement are brand-new, holistic and data-driven solutions leveraging integrated behavioral data and supporting the behavioral finance perspective of the omnipresent confusion in investors' retirement solutions.

There are challenges in the traditional retirement platform approach to data products. First, while the majority of users opened their account with the platform some time ago and have continued investing, rules may or may not have changed since they joined. The algorithm that determines the allocation of their assets may not be the one that originated their investment. The risk tolerances and constraints would be those set back in the day. It is quite common that users' fear and perception of risk is very material to what should be considered their investment horizon when a three decade period is used to assess mispricing assumptions. With AI techniques, it could actually be possible to derive the economic situation that users were incorporated in and start to work back from there. Second, many decades of user data could have led to distortions and AI neural nets struggle with the interpretability that such dimensions evoke. Given the absence of a full understanding algorithm and of a simulation engine capable of reconstructing portfolios ten or more years back, it is important for some time to be spent doing what is called evidential reasoning before a final asset allocation is even considered as ready for production.

Equ 1: Feedback Loop Equation

Where:

- λ = Retention of prior value

$$\text{Data_Product}_{n+1} = \lambda \cdot \text{Data_Product}_n + \delta \cdot O$$

- δ = Outcome-based learning rate

3.

The Role of Data Governance

There is now consensus in the Scientific Community that in order to actually harvest the yet unseen potential value of Artificial Intelligence and Machine Learning systems, an organization must have Data Governance in place. As a general rule, before embarking on the AI journey with all its opportunities, a corporation must first build a sound Data Governance to achieve the expected results. Data Governance approaches and its family of well defined Data Roadmaps, Data Dictionaries and Data Catalogs are at the core of Data Management process and decision making. Without a strong Data Governance framework filling its Data Intelligence coffers, the performance

of any Data Lake or Data Warehouse designed to support complex Data Science workloads will suffer.



Fig 2: Data Governance

Good data governance enables the necessary realism for creating value. It helps organizations identify the problems that data can actually help solve, while at the same time help mold corporate expectations to avoid disappointment with the results in value creation and also decision making. Business transformations are seldom based only on data. They are often based on the combination of data and innovative decision making strategies leveraged by strong corporate leadership. However, when those transformations, both big and small, are looked after in the corporate strategy, then having adequate data governance will be the initiators of fueling the corporate coffers with the necessary Data Intelligence to elevate the corporate expectations levels and avoid disappointment. Disappointment is the result of collisions, be it at the individual's level with the Sr. Executives who are considered the architects of decision making and value creation. Be it at the corporate level, where the inflow of Data Intelligence suddenly becomes scarce due to the lack of Data Governance structures.

3.1. Definition and Importance

Data governance has been defined as the management of the availability, usability, integrity and security of the data employed in an enterprise. The role of data governance is to ensure that data is made accessible in a consistent manner to data users across the organization while ensuring that the data itself is accurate and trustworthy. In order to build AI-ready data products, organizations need to ensure that their data is high quality and trustworthy. Data governance can serve as a blueprint for organizations to operationalize the principles of data quality and enable the delivery of high-quality data products at scale.

A data governance framework typically encompasses a council, defined roles and responsibilities for data owners and stewards, data policies, data models and definitions, and data quality rules and

metrics. This framework is enhanced with processes for entering, updating, curating and deleting data and a collection of technology solutions that help enforce the processes and policies and automate the processes wherever possible. Data governance is often seen as a mature domain focus of controls to ensure consistent and trusted data across the organization. While the central directives assure consistency, an efficient data governance model recognizes the importance of contextual, domain-centered data governance consisting of data owners and stewards in the different business domains working closely with IT and the data solutions and technology teams. These domain-centric teams provide the requisite input to create localized data policies, processes, models, and technology customizations within the rules defined by central IT. Thus, central and domain data governance teams work together as an approach to ensure that data governance enables fast and scalable data operations while being more controlled and compliant.

3.2. Key Principles of Data Governance

Data governance encompasses a broad range of activities, and several models are available to help outline its coverage. Some models emphasize certain aspects over others. For example, one governance framework identifies seven building blocks for data governance: strategy, organization, policies, processes, standards, technology, and metrics. This model combines aspects of organization, policies and processes. Another data governance framework identifies best practices and processes across eight domains, such as compliance, data stewardship, data policies, and data operations. This model emphasizes the process aspect of data governance. Our data governance model emphasizes five key principles of data governance: decision rights and accountability, policies and standards, quality and lifecycle assurances, stewardship, and enabled working environments.

Decision rights and accountability. Data governance helps establish, document, and communicate who has final decision rights and accountability across what data domains. For each data domain, a data governance framework clearly identifies data domain owners, authorities, and stakeholders on both business and technical management sides.

Policies and standards. Data governance is responsible for establishing, documenting, and communicating organization-level data policies and standards. Simultaneously, data governance encourages and enables other data management functions to create data policies and standards that are supplementary to the organization-wide ones. Data governance keeps track of all data policies and standards, as well as their versions and status. Data governance conducts periodic reviews of the policies and standards to establish effectiveness, and to assure that they remain current and relevant.

Quality and lifecycle assurances. Data governance establishes the organization-wide requirements for data quality (accuracy, completeness, consistency, timeliness, validity) and lifecycle management, across the full data lifecycle. Data governance coordinates with applicable data management functions to design, monitor, and report on the processes, controls, and measurements that assure that the requirements are met.

3.3. Impact on Data Quality

A widely-cited definition of data governance describes it as "the specification of decision rights and the accountability framework to encourage desirable behavior in the use of data." The increasing importance of data and advanced analytics has fueled a dramatic increase in the number of data governance initiatives in a variety of sectors. Data governance requires the formalization of a broad range of decisions and decision processes. Standard governance procedures and guides enhance both data quality and the utility of data products but at an additional cost, mainly in the form of additional time needed to comply with governance procedures and best practices.

In finance, data governance has become a hot topic in recent years with the increase in regulations in developed markets following the GFC. An effective data governance framework receives visibility in the form of a three-lines-of-defense model which consists of business management, risk management, and internal audit. However, the increasing regulations have also led to large investments in technology-led solutions. Companies spend as much as 30% of their budget managing data if we include activities such as cleansing, compiling, and processing data. The proliferation of data visualization tools that allow business users to easily analyze data has only made the importance of data quality more visible. The simple fact that while it is relatively easy to visualize data, not all data is suitable for visualization has highlighted the issues related to data quality. Thus, having a data governance framework is critical to the success of any company that seeks to derive value from its data.

4. AI-Ready Data Products

Rapid digital transformation in the data economy has propelled a profound change in how businesses leverage data. Existing financial service organizations and a wave of new financial technology startups are adopting differentiated enterprise data strategy to harness the power of data to comply with intricate regulations, offer personalized financial products, outperform their competitors as well as create new monetization streams. A key component of such enterprise data strategy is to build AI-ready data products that involve packaging and organizing internal data so that they can be efficiently shared, discovered, and reused across the data supply chain. By maximizing the usability and discoverability of internal data assets, data product thinking powers self-service analytics and the democratization of data and analytics in organizations. This is particularly relevant for companies and government institutions that need to address issues highlighted by regulations, especially related to increased transparency, fairness, accountability, and risk management of AI.

Historically, many companies have built vertical data silos, stitching various pieces of private or public data together to derive business insights. However, as organizations continue to move up the data analytics maturity curve from business intelligence to data science and machine learning, the demand for integrated and reliable datasets with domain knowledge, metadata, and data governance is increasing. In finance, this is leading organizations to build reusable, AI-ready data products that can be adapted for multiple use cases. For organizations just beginning their journey into the AI world, it is useful to think of data products as delivering a sprinkling of magic software development fairy dust to the underlying data asset. A data product is a specialized, packaged, and well-documented collection of a dataset or a related suite of datasets that are designed with the

end-user in mind. Advanced data product capabilities leverage best practices from both traditional and frameworks and tools in tandem.



Fig 3: AI-Ready Data Essentials

4.1. Characteristics of AI-Ready Data

The biggest hurdle on the road to enterprise AI is data, and not just any data, but high-quality, well-documented data, which we refer to as AI-Ready Data. The road from raw data to AIRD presents many challenges in preparation, cleaning, validation, management, and documentation. Achieving the requirements of AIRD is more than just an issue of quality, it is also an issue of completeness and structure. Rule-based systems, decision trees, random forests, and traditional statistical systems like regression or analysis of variance only need independent variables to be complete. However, machine learning systems based on functional decomposition, especially Deep Neural Networks, all rely on a very large number of variables in classes that cover the required number of parameters in the functional form. An analysis of neural networks and training data for vision applications shows that a typical AIRD set for DNN functional approximation will most likely include not just the pixels of the output but also auxiliary optical data, a high overlap between input and output pixels, and physical descriptors such as temperature.

AIRD is not everything you need to make a model, but it is the most important thing to help automate the model-making process. The key is to use the model for as many AI applications as possible and use that information to build the best AIRD possible for a wide range of varied models. In that way, either using Active Learning in conjunction with reinforcement learning or using the same model to implement many AI applications, the task of data preparation can be made much simpler and automated. Another, equally valid, approach is to attempt, using Active Learning, to automate the process of building “plug and play” modular models, instead of monolithic models that implement every possible function. However, the goal here is to provide the best AIRD possible for very general models that have different layers and perform DNN decomposition. Students of machine learning and AI constantly emphasize the need for diagnostic modules, specialized models that need specialized pre-packaged data.

4.2. Framework for Building AI-Ready Products

In a previous paper detailing our experience of building the next generation of security, fraud and risk detection products for financial services, we had talked about six dimensions in the What, How and Who of building data products, that converged into the formation of AI-ready products. This is visualized in the components listed below.

What is the product or service offering? Is it a full stack offering or a specialized piece on top of existing systems, such as a model training or serving library? What datasets does the offering provide? Are datasets built to be used in conjunction with external datasets and models? Can users instigate updates to the datasets? What is the product revenue model? What data products are unique to the offering and what data products are lower quality but still very valuable that can be made available to all consumers and users of the platform? Clearly defining the offerings, and associated data products goes a long way in defining the data governance structure, the required infrastructure components for the offering and the guiding principles for product development and operations.

What are the “smart” features of the offering? Are those features those historically provided by other full stack data products, such as provenance, auditing, catalogs, cleaning, data-at-rest, data-in-motion pipelines, schema, lineage, motivation, validation, testing, orchestration, systems integration, UI integration, access control, quality alerts and UI, UAC, sharing, thresholds, and what-if? Are there key differentiating features that enable the provider to offer a higher quality product; that are required to support the consumer of the product? Is the product offering using similar tools, techniques, processes, and technologies that development teams of the provider use?

4.3. Case Studies of Successful Implementations

Having a validated and proven framework for the design and build of multi-intent AI-enabled Insights Data Products, the two case studies presented below confirm the comprehensive enterprise-wide benefits to be gained from properly architected solutions. The first is a broad data product to enable a wide range of data science initiatives for many audiences, while the second is an enterprise-wide solution to enable a time-efficient and cost-effective approach to deliver a specific critical business initiative. Case Study 1: Consulting to Catalyze Government Data Science Products Problem: A national government and its agencies needed to utilize disparate data to provide insights across different departments to support informed decision making. They had no practical solution to integrate their disparate multi-source data into a unified data product that supported enterprise-scale insights generation and sharing. Approach: A comprehensive Data CoLab product was proposed that provided a purpose-led enterprise-wide multi-intent shared data platform enabling a collaborative, energy-efficient approach to data science. Its objectives were to accelerate data exploration through enabled integrated data assets, to establish a trusted environment to democratize access for all departments, to make outcomes visible for citizens to ensure accountability, and to support sharing of data projects, don trusted and endorsed value, with a guided process and meaningful templates. Outcomes: The collaborative investment in structured data exploration provided them with the trained and enabled foundation to safely and securely explore, create and share insights, campaigns, and models across government services. The plan

was to create citizen services products for data-driven interventions such as tax collection optimization, public health, job service utilization, and crime prevention and mitigation.

Equ 2: Outcome Optimization Equation

Where:

- O = Optimized Retirement Outcome Score
- P = Personalization Score (from above)
- R = Regulatory Fit Score (compliance automation)
- T = Trust Factor (transparency, explainability)

$$O = \beta_1 \cdot P + \beta_2 \cdot R + \beta_3 \cdot T$$

5. Integrating Data Governance in Retirement Solutions

While creating AI-ready data products and use-cases for the built data factory, it is important to ensure that data fueled decision intelligence makes sense for the retirement solutions business. The BI dashboards and reports, data products and data for data science, ML use-cases should be driven invariably by the retirement solution domain experts and business users who are best aware of the business goals, objectives and mission; and who will use the retirement solution data products to inform their decision-making processes. It cannot be more evident that the strategic business stakeholders of retirement solutions also own the data product(s) and the product vision. Data owners provide the rationale for understanding the "What, Why and How" for each key domain area. This involves business areas such as retirement accounts, participants, products, plans, distribution channels, employers, legislative mandates, laws and regulations, sponsors, service providers, tax burdens, tax incentives, product designs, transaction flows, decision making process, reporting and what if analysis, guide, advice and recommendation, and analytics insights to help guide and manage the design and maintenance of the retirement solutions data factory.

But building such a comprehensive domain-centric business case, with a use-case for every data product contributing to the bigger picture of the retirement solution business mission, involves painstaking effort and commitment that must come from the business itself, with support from the data operations technology provisioning teams. Data strategy discussions that bring the business, data, and technology teams together are generally the key enablers to getting such a roadmap defined, from which data-driven use cases can be correlated to the retirement solution mission; and ultimately the roadmap enables connecting the dots on how the retirement data product for which the mission consensus is built, are interconnected, the logical and physical where, and the how data will fuel those product(s) of the retirement solution mission.

Strategic business stakeholders need to do this exercise to build mission alignment and consensus at all levels of the organization that make strategic decisions; and also encourage active collaboration of all resource stakeholders at the process and technology level for strategic business enablement of the retirement solution mission. The central idea for the collaborative consensus is so that every resource stakeholder understands the "What, Why and How" of their organizational

functions, processes, and services; such that data and technology decisions are made leveraging enterprise data integration and data sharing best practices.

5.1. Strategic Approaches

Retirement planners need to embrace the innovative concept behind Digital Transformation. They need to invest funds to build or obtain capabilities and infrastructure to support the successful adoption of Data Governance. An alternative option is to partner with a technology leader and vendor in Data Management, Data Governance, and AI-powered Data Products. This could allow the solutions provider to quickly bring an AI-ready Data platform to life with best practices configured into the solution. The third option is circling back to basic data collection, preparation, and benefic conversion, as the industry did many years ago. Use the internet to aggregate form-filling capability for clients, 401k to IRA transfer capabilities with Relationship Managers for transitioning funds, and ESG-themed Investing to attract millennial clients. Leaderboards in business sectors could assist with benefit makeup as well as prospecting other retirement accounts that could benefit from a transition plan for ESG.

With the first two strategic methods, the organization must invest time and funding to implement a trigger event that allows the data captured and enhanced within the Data platform to be delivered and refreshed. This is more easily attainable when capturing multiple points in the lifecycle of the investment and providing personalized value along the way and at event timeframes.



Fig 4: The Strategic Approach

5.2. Tools and Technologies

The success of a comprehensive data governance program requires a supporting technology ecosystem that meets the enterprise's unique operating requirements. Local tribal knowledge of the data permeates the lifecycle of the current tools, often hidden inside the adoption technical debt that has evolved over time. Simple file management, security, object-oriented databases or Wikis might meet initial needs but fail to deliver the rigor necessary for a sustainable foundation for enterprise data quality, security, privacy, and compliance. Delivering the required enterprise class technology enhancements is a barrier to progress for many; it was for us. Enterprise license, maintenance and administration costs outweighed the adoption risk of waiting for enterprise class solutions to catch up. It wasn't until the "How" of our company-wide implementation was debated that we could commit to remediate the adoption deficit. Ecosystem gaps still needed to be filled.

Business-focused users needed additional support for the high-level engagement and foundational understanding of their data and its purpose behind the data elements needed to remediate and populate the operating model. Building a working consensus with each department on their existing foundational understanding, purpose and assumptions of the data in their responsibility allowed for both iterative and efficient library validation. Institutions use in-house built data dictionaries and potentially many others to engage the investment and user compliance necessary to keep the inventory relevant. This library also represented the content of formally modeled data elements such as records, reports, dashboards, and source code that would be migrated. The required characteristics included blazingly fast execution and auto-scaling. Cloud costs paid only for use—as more data was moved from on-prem to cloud, the cost benefit became undeniable as the alternative in-house clusters were costly to store data. The external integration engine could handle our in-house hosted object stores and outside vendor payment services.

5.3. Best Practices for Implementation

When embarking on a new data governance program, there are several best practice guidelines to consider. Do expect your implementation to be challenging. A first version of a data governance program will be very difficult to implement, because every stakeholder has different needs, leading to a lot of tension. In addition, it takes a lot of up-front effort to gather and document all that is required for both the process of governance and the implementation of the automated tools to implement your data governance policy. This will lead to a lack of engagement and enthusiasm among all but the most committed. Budget, timelines, and staffing should also reflect the effort and time required at first. Over time, the implementation should flow more smoothly as the basic structure is in place and the ongoing activities require less effort.

Plan for change. At first, data governance is a mandate-based process requiring the active participation of the data controllers, implementers, consumer, and steward. Over time, as the program matures, the hope is that the actual data governance activities become part of the business as usual process flows of implementing new systems, maintaining existing systems, and fulfilling requests by outside parties. As such, while the effort should still be visible to the stakeholders, it will not dominate their working lives. Data accountability will become ingrained in the culture of the organization and business process efficiency, effectiveness, and quality will be seen to improve, leading to a more efficient bureaucratic process.

6. Regulatory Considerations

As with any enterprise solution that holds sensitive data, a retirement solution platform must comply with many regulations governing the use of personal data. The most stringent regulations are adhered to by the largest recordkeepers, who have government pension plans as clients and meet the standards set forth by relevant authorities. These global companies are equipped to offer a full suite of integrated financial services, from retirement planning to provision of health benefits, to recordkeeping and distribution of pension plan assets. These regulations focus on plan sponsor fiduciary responsibilities but also influence the product offerings by recordkeepers like allocation options for investments, withdrawal and rollover rules and procedures, and explanation of associated fees. These federal regulations provide guidance on the minimum the providers need to do, but it is the state laws on data privacy and data services that lead the market to demand more robust personal data protections.

California uses data privacy protection laws to regulate personal data, setting the standard for other states to follow with their own rules. Other states have enacted similar legislation to protect their residents. The two central tenets of these laws are to limit the data collected about their citizens and the jurisdiction to criminally fine organizations that misuse that data. Consent must be obtained when data is collected, dispersed, or profited from, including third-party use. Residents should be able to access their data, require its deletion or request that its use be limited, such as not being used in model training, so that their use not be used either directly or indirectly to discriminate, such as for predatory lending product offers.

It is noteworthy that while the retirement space may be subject to more stringent regulations due to the additional government oversight via the mandated pension protection plan, the compliance burden is shared by both the recordkeeper providing the platform technology and the financial institution distributing advice products through that platform. Platforms that lack clear accountability for compliance responsibilities may struggle to quickly onboard clients, or worse, fail to refine their data and services platform to maximize their contribution to sustainable financial outcomes, particularly for plan sponsors with significant unfunded employee liabilities like governments or enterprises with defined benefit plans.

6.1. Compliance Requirements

Building an AI-ready data product involves various legal and compliance challenges. The regulatory landscape is complex and ever-changing. Importantly, while the underlying data sets used for building the data products have to comply with several laws and regulations, there are no specific compliance requirements for the data product itself. Existing laws and regulations impose constraints on how financial institutions handle and use the constituent data sets. Hence, care must be taken to prioritize compliance while creating AI-ready data products. Within retirement plan services, a broad set of data compliance regulations impact the data products selling and service delivery flow. Federal regulation of retirement income policies have hundreds of statutes, rules, and guidelines applicable to pension plan vendors and sponsors. They created and institutionalized the roles of plan sponsor and administrator. Together, they set key fiduciary compliance responsibilities. Other compliance-related events, such as plan audits, apply to 401(k) plans, profit sharing, defined benefits, switch to cash balance, and pension loss funding. Creating and testing

scenarios on compliance support pre-distribution tax withholding and tax notice requirements. Processing the distribution — including reporting accuracy — is yet another area of compliance regulation.



Fig 5: Compliance Requirements

6.2. Data Privacy and Security

The financial services industry is highly regulated; privacy controls that address the security of consumers' nonpublic personal information are federally mandated. Data use agreements between businesses and consumers must be unambiguous; user consent to current and future collection and sharing of personal data is imperative. Privacy is the bedrock of any trusted relationship between a financial services firm and its customers; privacy breaches can damage that trust irreparably. There are numerous state, federal, and international regulations and security frameworks surrounding privacy and security, including various acts and security frameworks. Different markets have different regulations, and firms depend on a patchwork quilt of state and federal laws, agency expectations, and industry standards. However, there are two cornerstones of regulation in the financial services industry that are ubiquitous across businesses everywhere: the Gramm-Leach-Bliley Act and more contemporarily, the California Consumer Privacy Act and Europe's General Data Protection Regulation. The Gramm-Leach-Bliley Act dictates the maintenance of: Information security programs; Safeguarding of existing privacy; Disposal of sensitive information; and a Focus on service provider relationships.

The California Consumer Privacy Act creates a laundry list of consumer rights surrounding data usage. It allows consumers to request information about what businesses collect about them; omits charges for a firm to fulfill this request at least once every 12 months; and promotes the right to deletion of data, just to name a few provisions. Europe's regulations are more stringent than California's. The General Data Protection Regulation significantly impacts the way businesses handle the personal information of their technology partners, clients, and customers. It mandates that businesses must keep client and customer data private and secure – for everyone, everywhere, all the time, period. Consent is the foundation of all data processing; explicit consent is required on the individual before the collection of data. This requirement runs counter to the commonly accepted practice of default settings collecting information until the user opts out..

Equ 3: Governance Impact Function

Where:

$$\text{Governance_Impact} = \log(1 + G^k) \cdot \text{Data_Readiness}$$

- G = Governance Maturity Index (scale 0–1)
- k = Sensitivity factor (higher in regulated environments)

7.

Future Trends in Retirement Solutions

Recent centuries have experienced transformative changes in the global economy as a result of the exploitation of information and technology. Today, as the digital economy continues to grow and expand, the need for innovation and transformation becomes ever more crucial for business sustainability and profitability. In response to this changing digital economy landscape, a new wave of innovation is being made towards shaping the next stage of the information revolution, where the explosion of emerging new technologies—such as AI, Blockchain, Augmented/Virtual Reality, 5th Generation Internet, Internet of Things, Drones, Quantum Computing, Biotechnology, Energy Diffusion/Storage, and Advanced Robotics combined with more powerful computer storage, processing, and communication capabilities—is promised to support profound changes that will transform and redesign most business processes and customer engagement paradigms. This section looks at the convergence of some of these technologies—specifically, predictive analytics and artificial intelligence—and explores their opportunity for reshaping how retirement solutions can better fit their customers' financial situation and challenges.

Predictive analytics is based on data, statistical algorithms, and machine learning methodology and is used to identify the likelihood of future outcomes based on historical data. It was born as a natural evolution of modern business analytics. With the mass adoption of social media, data governance has advanced and achieved new paradigms that emerged as models for ensuring trust in both the data and insights provided. The retirement plan industry is traditionally rooted in a backward-looking context, where methods and processes focus on past data normalized to their own specific threshold. This occurs because of difficulties in access to confidential information and in the keying of updates from participants, thus incentivizing reliance on established methods and expected long-term return computations.

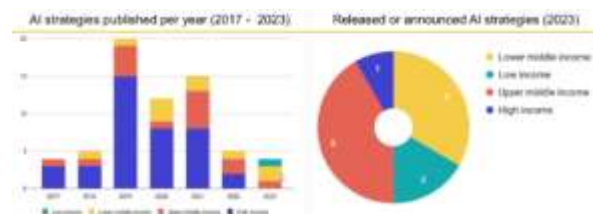


Fig 6: Leveraging AI for Enhanced eGovernment

7.1. Emerging Technologies

The speed of business is accelerating. Disruption has become a habit. As new technologies enter the market at a fast pace, we are now faced with a multitude of potential solutions. The increasing pace of technological change is reshaping the financial landscape, redefining customer expectations, and remaking the competitive landscape.

The technologies that are reshaping traditional models fall into three major categories: digitally-enabled services; big data analytics; and emergent technologies. The retirement space is seeing the introduction of digitally-enabled services that cut costs and improve the delivery of existing services. Portfolio and asset management services are being priced down to near-zero. New technology-driven service models make investment advice accessible and affordable for millions. In addition to robos, future-class investment providers will leverage their unique data sets to

support advisors and directly deliver small-balance wealth-management services. Data democratization is lowering the cost of access to sophisticated analytical tools for executing repetitive investment strategy execution and implementation tasks. Technology is allowing companies to identify, target, and engage customers and client segments by need and across the distributed decision journey. From AI-chatbots to branding, message, and offer personalization, digital marketing and customer engagement technologies are helping companies provide customers with an experience that meets or exceeds their expectations.

These tools and services are also being used to identify and mitigate risks and improve operational efficiency and performance, especially back-office operational infrastructure. Technology infrastructure and tools that streamline processes and cut costs are rapidly emerging, requiring organizations to think about investments in orchestration and connectivity to deliver efficient operations for themselves and their partners. No- and low-code tools that allow firms to create interfaces and solutions inexpensively and at scale will increasingly become part of the technology infrastructure of the industry.

7.2. Predictive Analytics in Retirement Planning

It is no surprise that the latest emphasis on predictive solutions for retirement is driven by various predictions of the coming shortfall with the current plans and the government safety net. The erosion of cash balance plans and the continued decline in defined benefit plan sponsorships has left much of the responsibility for retirement funding with the employee but invariably with little-to-no education or knowledge on how to properly manage their investment decisions. On the horizon are tools and strategies that enhance the current process of retirement planning and monitoring. For example, every significant predictive issue is being addressed, from the propensity to retire, research retirement products to the likelihood advisors will offer retirement planning advice to their clients. Survey results show that developing reliable predictive models to estimate retirement probabilities throughout the life cycle and develop playbooks to aid in the management of the myriad of triggers required for a successful transition will be the norm.

Another product is predictive life cycle modeling, which leverages financial services transaction data to create easily accessible resources for itemized major financial life events. This concept addresses a much-discussed subject concerning advanced life stage connected products. In addition, financial life events utility is increasingly reliant on decision-support interfaces that present plans, resources, and products monetarily linked to tangible predictive triggers. Whether it's just-in-time education regarding prudent savings rates triggered by key milestone events or automatic links to pension optimizer tools when turning age 65 and letting clients know it is time for a rollover, predictive life stage models will be an indispensable resource for financial institutions.

8. Stakeholder Perspectives

As the number of available retirement solution options continues to increase, offering technology solutions that bring simplicity and transparency to the retirement planning process can help drive client engagement. Information is like a firewall; it is a protective asset that prevents a service user from inadvertently turning into a service provider. But what does a combined future look like for

advisors, recordkeepers, and asset managers? Can third parties help solve problems by providing clearly defined use cases? In search of answers, key industry stakeholders sought various viewpoints.

Whether or not the future is unified, certainly major recordkeepers, asset managers, and advisory firms all see the advantages of directly or indirectly taking on re-engagement roles. Some predict this future because they are on the front lines of ground-breaking prototype efforts today. Some have publicly issued breathtakingly ambitious reports to shareholders. Some simply see the advantages that are only logical for consumers in what have traditionally been disparate business silos. Some foresee unified solutions as defensible offerings, helping expert advisory firms push ROIs even further, while recordkeepers neutral challengers by making it easy for all to recommend their firm technologies.

For regulators, the history of financial services is rife with consumer failures in complex needs that have once been unmet or altogether misaddressed. Lifetime risk capacity – that is, the leverage point of withdrawing from a decumulation fund – creates consumer vulnerability at those moments of extreme need. Greater coordination and co-creation over the short and long terms hold promise to improve retirement security and reduce the burden currently placed on governmental social funding. Do lessons learned in best-practices governance that inform the data compliance and transparency mandates of today carry forth to this new governance landscape?

8.3. Consumers

Consumers in general want, expect, and deserve fiduciary-scale solutions to help engage them, enhance their awareness of their personal risk budgets, and direct how their net worth should be aligned to those budgets to achieve their future spending goals. Idealized solutions bring outlines of how monthly spends and tax strategies in accumulation and decumulation phases derive from — are linked to — the probability-weighted net worth curves. Costs, incremental or not, for customized solutions such as consultancy preferred offers should be transparent and posted upon requesting discovery.

8.1. Financial Institutions

Retirement plays a substantive role in influencing the broader economy. With over \$39 trillion in retirement assets in the United States, accounting for about a third of total financial sector assets, financial institutions (FIs)—primarily banks, asset managers, broker-dealers, and insurance companies that provide pension services to retirees—play an important role in the safety and continuous growth of these assets. With the primary aim of profiting from the management of these assets, FIs help lower the cost of capital and facilitate risk management, thus helping to create better market conditions for investors and supporting capital market efficiency in the economy. Since FIs earn a percentage on the assets under management, they have a vested interest in growing their AUM and want to retain as many customers as possible to remain viable: they want the value of the services they provide to the user to exceed the costs of engaging their services.

However, the increase in the average life expectancy and the retirement asset value of an occasional consumer versus the decision epoch of a FI has made the former an interesting target for an asymmetric game for FIs, despite the unequal footing. Travelers are frequently presented

with a one-time contribution in the form of high commission, low-cost vehicles, and they actively resist such enticements, as they may be overwhelmingly ignorant of their retirement assets. In the meantime, FIs are legally responsible for managing the retirement accounts held by their past clients, and it is not uncommon for these FIs to offer passive portfolios of exchange-traded funds for the duration of the individual's career. Coincidentally, many consumers are enamored by the equity market potential, which may lead to self-harm without expert assistance. We present a model to balance the apparent duality in roles of the customer and the FI toward potential failures in this dynamic system to initiate engagement.

8.2. Regulators

Regulatory engagement is one of the most complex and important challenges that financial institutions face in their pursuit of digital strategy and modernization. It is at the core of many business decisions that firms are required to take. Guiding firms with the objectives of regulatory frameworks would help facilitate their ability to apply technology to drive towards those objectives while not disrupting corporate operations or industry economics in a fashion that either creates conflict with the public sector mandate or collapses institutions. Technology can drive significant benefits in efficiency, productivity and well-being making it essential to provide a scaffolding that reinforces such modernization efforts.

To this end, regulators can leverage the discussions and considerations to consider the barriers and enabling conditions to accelerate adoption of additional technologies and business model changes among the organizations that they oversee. Such an approach would help generate innovation, competition and public trust. Pioneering activities and initiatives to apply modern data governance enabling technologies among challenger organizations could serve as a lighthouse, signaling the set of foundations required to undertake a technological digital transformation.

However, in considering such a course, regulators would need to address several key considerations. Can the set of objectives for which technology is being deployed require a fundamental rethink and societal discussion? Would the structural and governance reforms required to drive such public sector reinvention serve as an appropriate model for institutions in the private sector? Would micro and decentralized models of public service delivery find sufficient accountability and public value generation? Would incentives, financial or reputational, be required to support a skeptical public and institutions?

8.3. Consumers

Retirement system consumers are seeking more personalized products curated especially for them. Primary goals for most consumers having access to employer sponsored retirement accounts is to, in order of priority: discover appropriate investment options, avoid losing money, appropriately save for retirement, obtain higher investment returns, and delay taxes. In addition, they wish to minimize fees, avoid too many investment choices, and obtain appropriate diversification. Retirees still want some control over the level of guaranteed income in retirement, but only if they are compensated well for giving up control, while desiring convenience in managing the payout of their assets over retirement, especially with regard to spending.

The increasing popularity of low cost index funds make saving for retirement less expensive for consumers. However, average asset fees for retirees remains a cause for concern as they are several times higher than index fund fees, resulting in large drag on returns over long retirement timeframes. Despite the growth of target date and other managed retirement products, a subset of retirees prefer to continue to manage their retirement portfolios, especially those who are relying on limited resource human capital to actively manage risk.

Retirement planning is fraught with conflict over rewards for bad advice. There are bad pension products, such as variable annuities, or providers who hold assets in unqualified far more expensive funds backed by poorly performing pension products implemented without regard to whether or not a consumer needed those funds for short term savings. Conflicting interests coloring advisors' decisions, such as selling commissionable products not in the clients best interest, have led to a push for a fiduciary standard for retirement advice. Consumers will benefit from a higher level of fiduciary obligation among providers, aligned interests for the advisor and client, and the availability of a wider group of trusted advisors.

9. Challenges and Risks

The path to a new paradigm in IRS platforms is not without challenges and risks. IRS data in its current form is a conglomeration of disparate yet related data sets that comprise and govern sources of record used in data ingestion and processing. Porting data into the platform may expose weaknesses in existing data governance practices and data management approaches. Stronger centralized and unified governance, cataloging, management, and control of IRS data will be required to implement durable and sustainable data-driven IRS platform solutions.

The independent nature of the various IRS data generations in data ingestion, processing, and sharing functions is not aligned with the data-centric approach and goals for its future IRS platform services strategy. IRS and other agency data consumers contribute enormous downstream costs and risks to the IRS that are magnified by the absence of standardized data protocol sets used in data generation, ingestion, and processing. Data access without appropriate protocol standards creates legal, regulatory, and compliance exposure to the IRS and other government stakeholders. Adherence to stringent guidelines around privacy, data ethics, and algorithmic fairness will be central to building and maintaining trusted relationships with users, partners, and others who are touched by or affected by approaches for incorporating technology into services modeling and deployment.

Data platform dependencies, oversights, and errors in the management of IRS data assets will undermine the user experiences built on top of the technology for front-end modeling, interaction, and service deployment driven by new technologies. Without a model toward data definition, description, and usage that is uninhibited by politically driven and often artificial boundaries, the risk of model-induced withdrawal or withdrawal-related errors as a result of poisoned or spurious data feed into machine learning technologies is high.

9.1. Data Management Risks

The increasing demands for strategic data for combating risks, driving growth and enabling compliance in enterprise business centers, are ramping up pressures on enterprise-wide data management functions. Accelerated adoption of privacy data protection regulation, enterprise-wide data standardization, and the onset of enterprise-wide cloud-migration projects are presenting enterprise data management functions with opportunities for establishing partner network connections with related business functions to build enterprise-wide data networks; incorporate data assets within the enterprise-wide cloud; and build enterprise-wide data products. Executives, however, need to remain ever-vigilant in their oversight of enterprise-wide data management functions during this transition period, given growing concerns about the optimizing-revenue and mitigating-risk data consequences of poor data quality. Increasing susceptibility to incidents of compromised data integrity and outages related to cyber-attacks; wasted efforts associated with incorrect deductions drawn from strategic analytics based on incomplete data; and costly fines and lost business opportunities related to wasted compliance costs incurred by enabling privacy infractions, are rapidly mounting. While there are many reasons for the dismayingly high levels of entrenched poor data quality still evident in today's corporations, enterprise executives need to be particularly sensitive to the potential data management risks posed by multiple emerging trends driven by the growth of data-driven initiatives.

Many companies exhibit “the cobbler’s children” syndrome wherein, ironically, the greatest possibility of compromised and poor data quality lies not in the business units’ operational data sources but, rather, at the enterprise-wide level. Authoritative “final version” data sets are rarely curated or even regularly updated for data enrichment. Their emergence is generally a consequence of “data disasters,” which occurs when the enterprise starts uploading query requests to these tables. To circumvent this situation, data concentration at the enterprise-wide level is generally limited to affirmation of select fields in the operational address tables, notably address, demographic, and financial transaction data sources, for consistency, accuracy, and completeness; any real data curation and enrichment, however, is usually performed by local business users.

9.2. Adoption Barriers

In particular, the platform lowers the barrier for retirement solution providers. Small providers often cannot compete with larger companies that have some of the largest market shares because they cannot afford to develop artificial intelligence (AI)-capable technologies in a cost-effective manner. In addition, companies that have their own customer information systems in place and have invested in the development of proprietary data products must safeguard their investments. Other companies may have developed proprietary data products in partnership with consulting firms, through software-enabled partnerships, or with use of border services. Furthermore, investor impatience may lead companies in the development phase to abandon technology investments because of unsatisfactory short-term expectations. Moreover, potential providers that actually have the intention to submit products may face internal risks related to a lack of an AI vision, hesitation to disrupt existing product lines, limited organizational capabilities, bandwidth/resource constraints, data deficiency, revenue sharing concerns, and legal, regulatory, and compliance barriers. All these factors will inevitably influence the decision-making process of asset managers deferred to another time, leading to delayed product launches.

Considering the significance of AI products for asset managers is increasing, the number of AI incubators batch is also likely to increase. The emergence of AI-ready product platforms will add momentum to the development of AI solutions for retirement investing. The shift to AI product platforms will not necessarily generate action by increasingly impatient investors or requirements of scalable in-depth social returns for investments. Unlike other areas such as healthcare, robo-advisors, and consumer tech, deep upside potential in retirement investing from at-scale tech innovation adoption is lagging. It may take time until investment conclusions and recommendations can be efficiently derived from vast troves of retirement fund transactions and stock ownership data, so alternative AI platforms for retirement investing may develop incrementally at best.

9.3. Mitigation Strategies

To address the data management risks, organizations should focus on data discovery, documentation, and literate culture. Since the idea of data as a product is still new to most organizations, they concentrate on centralized catalog tools to help seed a criteria-based discovery and have users actively document their data and code with good descriptions and consistent patterns. Also, most importantly, a senior executive should be assigned to promote data quality and data management tools and practices with their direct reports.

The durability of data products is enhanced by having a central team of user-appointed experts that all data product teams can turn to for quality reviews. Prior work convinced us that the people in these roles can best serve in a development and consultation role from an internal data science service. Many of these functions are also best accomplished in consultative ways, developing expert communities and internal documentation, as well as through training and quality reviews of specific products. However, since data product performance will ultimately be evaluated through the lenses of end-user performance, the internal purchasing function, and user demand, our vision of managing internal data products most closely parallels an internal government service model.

Lack of access to domain-specific expertise is a risk for both organizations and employees. Insufficient investments in the citizen model mean risk for organizations with domain experts refusing to become citizen data scientists and, instead, building their own domain-specific AI applications on their own through third-party vendors or open-source packages. It also means risk for employees, many of whom will learn to do domain-specific AI on their own and leave to become higher-paid employees of third-party vendors. Internal data product teams must also move quickly.

10. Conclusion

In this study, we argue for a new paradigm in retirement solution platforms. This paradigm introduces core data governance features to promote data model reuse for better scalable analytics. Within the confines of data governance, the purposeful application of AI tends to be much less ambitious, especially at the outset of the development and deployment of analytics capabilities, thus inspiring faster time-to-value. We illustrate and validate this retirement solution platform concept with our implementation that processes and analyzes client eating habit data. Industry-specific retirement solutions can be much more complex than general-purpose data products. Such

complexity can increase exponentially with scale and business volume without effective governance mechanisms in place. We believe this paradigm can help organizations create most types of analytics products from small to large scale, both general-purpose and industry-specific.

Pioneering work in the application of data governance to global cloud-based capabilities to promote company focus and solutions reuse recommend that general-purpose data products may take multiple functional forms, including enterprise resources planning, finance, and human resource systems. While we believe that the governance paradigm we present can act on all types of data products at various levels of sophistication and efficacy, we see a divergence here. This divergence reflects the inherent trade-off between general-purpose and industry-specific products. The roadmap to sufficiently effective industry-specific solutions is much longer and costlier compared to possible quick wins from the general-purpose counterparts. This asymmetry in effort and complexity can also affect the available levels of data governance and infrastructure sophistication. While general-purpose data products are simplified using a monolithic architecture, company-specific essential data products are implemented using system architectures that are complex and shared across multiple organizations.

10.1. Future Trends

As AI becomes more prevalent in many industries and its increasing adoption by consumers across different use cases, many other industries are also seeing growing interest in AI. This has fueled a trend towards the development of so-called "AI-ready data products," which are data products that make underlying datasets readily accessible for use in AI solutions. To enable external consumption of the data for AI specializations, the data provider must prioritize building the product as reusable AI data, allowing for ingestion into automated training pipelines. AI-ready data products provide easy datasets to find, access, evaluate, and consume, to familiarize users or external consumers of the products with the different datasets.

The financial services industry has seen an explosion of investment into AI for different applications; however, much of it remains nascent and continues to be applied to a limited number of well-explored use cases such as in automated customer service solutions. Recent developments in the generative AI space have sparked a frenzy of interest in advancing research and technology for a wider array of AI product applications and have attracted new players and new investment to a space that many had lost faith in after previous AI winters during the latter quarter of the decade. Generative AI relies on access to huge amounts of structured data in many domains and importantly requires the data to be structured in a way that is suitable for AI training.

Right now, banks and financial institutions have access to huge data volumes generated from regulatory compliance undertaken to meet increasingly strict regulatory requirements; however, these institutions do not yet have specialized data products that could serve as foundations for a paralleled generative AI revolution and ensuing monetization opportunities. Making available datasets from the huge troves of compliance data generated from institutions on behalf of their clients while ensuring that the data use complies with all relevant industry regulations will help banks and regulators alike to harness the vast potential for AI monetization in the industry.

11. References

- [1] Motamary, S. (2021). Machine Learning Applications in Demand Forecasting and Order Fulfillment for Smart Manufacturing OSS. *International Journal of Science and Research (IJSR)*, 10(12), 1639-1653.
<https://www.ijsr.net/getabstract.php?paperid=MS2112143019>
<https://www.doi.org/10.21275/MS2112143019>
- [2] Koppolu, H. K. R. (2021). Data-Driven Strategies for Optimizing Customer Journeys Across Telecom and Healthcare Industries. *International Journal Of Engineering And Computer Science*, 10(12).
- [3] Yellanki, S. K. (2021). Operational Efficiency through Service Integration: Insights from Platform-Based Business Models. *International Journal of Engineering and Computer Science*, 10(12), 25756–25772. <https://doi.org/10.18535/ijecs.v10i12.4681>
- [4] Pandiri, L. (2021). Cloud-Based AI Systems for Real-Time Underwriting in Recreational and Property Insurance. *International Journal of Science and Research (IJSR)*, 10(12), 1626-1638. <https://www.ijsr.net/getabstract.php?paperid=MS2112142733>
<https://www.doi.org/10.21275/MS2112142733>
- [5] Kalisetty, S. (2018). Cloud-Driven Big Data Harmonization for Real-Time Demand Forecasting in Hybrid Retail-Manufacturing Supply Chains. *Global Research Development (GRD) ISSN: 2455-5703*, 3(12).
- [6] Gadi, A. L. (2021). The Future of Automotive Mobility: Integrating Cloud-Based Connected Services for Sustainable and Autonomous Transportation. *International Journal on Recent and Innovation Trends in Computing and Communication*, 9(12), 179-187.
- [7] AI-Based Financial Advisory Systems: Revolutionizing Personalized Investment Strategies. (2021). *International Journal of Engineering and Computer Science*, 10(12). <https://doi.org/10.18535/ijecs.v10i12.4655>
- [8] Dwaraka Nath Kummari. (2021). Smart Infrastructure Auditing: Integrating AI to Streamline Manufacturing Compliance Processes. *Journal of International Crisis and Risk Communication Research* , 168–193.
<https://doi.org/10.63278/jicrcr.vi.3036>
- [9] Machine Learning Integration in Semiconductor Research and Manufacturing Pipelines. (2021). *IJARCCCE*, 10(12).
<https://doi.org/10.17148/ijarcce.2021.101274>
- [10] Digital Infrastructure for Predictive Inventory Management in Retail Using Machine Learning. (2021). *IJARCCCE*, 10(12).
<https://doi.org/10.17148/ijarcce.2021.101276>

- [11] Meda, R. (2021). Machine Learning-Based Color Recommendation Engines for Enhanced Customer Personalization. *Journal of International Crisis and Risk Communication Research*, 124-140.
- [12] Dwaraka Nath Kummari. (2021). A Framework for Risk-Based Auditing in Intelligent Manufacturing Infrastructures. *International Journal on Recent and Innovation Trends in Computing and Communication*, 9(12), 245–262. Retrieved from <https://ijritcc.org/index.php/ijritcc/article/view/11616>
- [13] Kalisetty, S. (2017). Dynamic Edge-Cloud Orchestration for Predictive Supply Chain Resilience in Smart Retail Manufacturing Networks. *Global Research Development (GRD) ISSN: 2455-5703*, 2(12).