

**Moving beyond the
Artificial Intelligence (AI) hype**
to get strategic value from data science
in financial services





Executive Summary

Financial institutions can collect billions of data points from their technology infrastructure every day, with most of it going untouched. In fact, IDC Research has estimated a 61% growth in worldwide data anticipated to reach 175 zettabytes by 2025. Global businesses already have seen a huge opportunity from AI and automation and are heavily investing in AI projects, data, and the supporting infrastructure. This is where AI and Data Science can make a significant impact beyond automation to significantly change business models to match a new digital realm and new customer expectations, when it comes to services, accessibility, and technology capabilities.

By employing deep data analytic tools and cognitive machine learning to structured and unstructured data, businesses can have more insight than ever before and turn their data into meaningful assets to apply to problem solving, consumer insights and more. This paper looks into

how banks can gain competitive advantage by leveraging this data and data science and analytic tools to tap into customer behaviors and get ahead of the curve to offer improved value.

By exploring the next frontier of data science, banks, asset managers, and insurance companies have an opportunity capitalize on the latest trends in AI.

These include:

- Developing a strategy to source data given the boom in data creation but the challenge firms are facing related to data scarcity, which is pushing them to digitize existing data, generate synthetic data and tap into alternative data sources like social media data
- Exploring new approaches across the AI maturity curve which includes moving beyond simple business process management, automation robotic process automation to more intelligent automation approaches that bring in advanced Natural Language Processing (NLP) expertise powered by Data Science approaches that better understand content, context,

and intent specific to the banking and insurance lexicon

- Using AI to drive ahead business goals such as improved customer engagement across the bank and insurance lifecycle, touching on key trends relate to client acquisition, retention, engagement, product recommendations, and business intelligence across the digital ecosystem - and how to apply the right data science models / approaches to achieve those goals
- Advancing Data Science strategy, operations, and infrastructure including how to set up the right team structure to design and execute data science programs and address technology infrastructure best practices across the stack including cloud infrastructure considerations and regulatory and compliance concerns such as model auditability/interpretability and model lifecycle management

This whitepaper will explore how financial services firms and insurance companies can take advantage of each of these areas in detail.

Foreword:

Drivers for Data Science Application in Financial Services and Insurance

Data is growing at an unprecedented rate. IDC Research anticipates Financial Services data to reach 175 zettabytes within five years. A combination of data and advancements in processing power enable significant opportunities for the application of Artificial Intelligence (AI) to automation and insight extraction enabling significant improvements in cost, revenue, and profitability for Financial Services institutions. In fact, more than half of companies with innovation strategies are using AI to identify

opportunities in data that they would have missed otherwise.

According to Gartner, 37% of enterprises have adopted AI. Adoption is expected to grow – businesses are moving beyond Robotic Process Automation (RPA) to Artificial Intelligence and Intelligent Automation to enable new business strategies. 40% of digital transformation initiatives are slated to use AI by the end of 2019, generating an expected \$190.61 Billion in enterprise AI hardware, software and services investment by 2025.

Embracing the Next Digital Frontier

Businesses can generate significant insights by applying cognitive machine learning to structured and unstructured data. The capabilities enabled by cognitive machine learning are facilitating more personalized services and a unique digital experience for end users. Financial Services firms are using AI to reduce customer response times, automate repetitive customer service tasks, and personalize the customer experience by learning from customers’ past behaviors and the behaviors of similar clients.



At the same time, advancements in AI have enabled technology firms to compete head-to-head with Financial Services firms for the same clients and products by offering more sophisticated customer experiences and more targeted offers. As Financial Institutions morph into technology firms, technology firms are simultaneously moving into Financial Services by leveraging their comparative advantages in AI.

Recently, Apple launched a credit card that is linked to users’ Apple Pay and built right into its iPhone Wallet app; Facebook has added payments services, including its Libra cryptocurrency, and the list continues. The movement by Tech companies into fintech is undeniable, so what can banks learn? To remain competitive, banks must accelerate innovation. Google has a smaller IT department than HSBC but

has 100+ teams specifically focused on AI. Financial services firms are playing catchup to remain competitive with both FinTechs and established technology companies. Banks should combine their comparative advantage in Financial Services with new investments in AI and Data Science to compete effectively with technology firms moving into the Financial Services domain.

Challenges and Demands

Despite its promise, AI faces challenges to adoption. Talent and team structure are important. IBM recently predicted that 59% of all data science and analytics job demand is in Finance, Insurance, Professional Services and IT industries, and that demand for these roles will increase by nearly 40% in the next year.



Given this, we know firms have begun to put Data Science teams in place but are seeking guidance in how to design and execute on projects at scale. This is especially important because implementing data science requires new infrastructure and processes that banks are seeking help to architect and implement – for example, there is a migration from SAS programming language to less expensive, more flexible big data platforms that is challenged by a skill shortage required to migrate SAS models to Python and R. In terms of Intelligent Automation, some of the top challenges are that 75% of

organizations see information overload as a problem, and nearly half of knowledge workers find it challenging and time-consuming to find information, according to KM World. According to The HR Technologist, 85% see a failure to digitize, convert, and classify business inputs as a key transformation bottleneck – and yet, just 3% of organizations say they’ve scaled their RPA digital workforce. For Data Science programs, firms are realizing challenges associated with data quality and model confidence. Improving data quality, as well as leveraging synthetic data, is needed to power models. Another challenge is that Cloud native applications

still won’t have the same Cloud computing power needed to leverage the full extent of the potential of Data Science. As banks and insurance companies look to move beyond AI automation into more complex digital strategies leveraging Data Science, there is strong demand for Natural Language Processing (NLP) expertise, using AI to drive client behaviors including acquisition and client retention. This whitepaper will explore how to organize a NLP strategy, using AI for customer experience enhancement, and social AI around sales and marketing enablement to achieve business goals.



1

Data Strategies to Power AI and Data Science

Financial Institutions and Insurance Companies have a greater volume and variety of data today than at any point in history. It is expected that by next year, 1.7 Megabytes per second of data will be created per person. This unprecedented increase in data has generated massive opportunities for increased revenue and reduced costs - in 2018, banks saved \$41 billion through the use of AI. However, harnessing the potential of an unprecedented amount of

structured and unstructured data requires unique strategies and capabilities. Financial institutions have far from reached the tipping point for Data and Data Science strategies. As institutions look to better leverage available data as an essential part of powering their data science programs, firms also are showing an increased interest in incorporating different data sources and Social AI into their broader data science strategies.

The Data Challenge

A significant percentage of financial institutions' data is left untapped, essentially leaving opportunities and intelligence on the table. In fact, according to some sources, financial institutions use just 0.5% of available data. Furthermore, according McKinsey, only about 30% of banks it surveyed in 2018 had a data strategy in place. McKinsey mentioned that others had begun ambitious programs for enterprise data warehousing and other projects without a broader data strategy, resulting in disappointing results.



Given this, firms should focus on creating a streamlined data strategy with processes in place and the end goal of the value they look to gain from the data outlined to inform the strategy. However, the most important part of this is the data itself. And while firms create and have tons of data created per customer per day, obtaining the right data to power specific

data science models is a huge problem and data mining, digitizing paper-based data points, and acquiring data are all challenges that overall have contributed to data scarcity, or not having large enough datasets to feed the data science models institutions are wanting to use to overcome other business challenges.

Text Data

Data is the foundation for successful Data Science programs. And a lot of Financial Services data is in contracts, notes, emails, and otherwise in non-structured and sometimes even paper format.



One major initiative in Financial Services is the digitization of paper documents. Banks are tackling this problem using Optical Character Recognition (OCR) and Natural Language Processing (NLP) to digitize documents and extract specific elements. Such efforts are critical to the successful implementation of programs such as transitioning away from LIBOR to alternative benchmark rates and effectively using research reports, news articles, earnings releases, and other non-

structured data to inform research and investment management processes.

Another major initiative is the application of Named Entity Recognition to classify emails, call notes, and other unstructured documents so that customer processes can be automated across every major portion of a bank. At Synechron we've seen this use case in retail banking, wholesale banking, private wealth management, investment banking, and asset management.

Alternative Data

According to Greenwich Associates, 50% of institutional investors are planning to increase their use of alternative datasets in the near future. Greenwich Associates also reported that in the last year, alternative data budgets are up by 52%. And, with so much data publicly available, leveraging these new, alternative data sources will become a fundamental part of strategies from alpha generation, sales, marketing, and more.

In fact, FactSet research also shows 64% of asset managers believe that alternative data can help them beat an index benchmark and 59% of asset managers said that active managers can improve risk and returns profiles by leveraging alternative data.



Following this trend, over the past year, banks like JP Morgan Chase and Goldman Sachs have invested into alternative data, and Bloomberg and FactSet expanded their data offerings, with Bloomberg selling alternative datasets through its terminal, and FactSet's launch of its Data Marketplace, which includes alternative data as well.

Alternative data includes data outside of the typical data investors historically have used to evaluate a company or an investment, like financial statements, SEC filings, press releases, presentations, etc. Because of the huge amount of data available today for an entity, alternative data helps investors get information more quickly, accurately, and gleans more granular insights into a company than your average statements might and creates a more 360-view of the company, rather than a one-dimensional view of what looks good on paper. Alternative data sources can include geo-location data, research and industry data, social media, and additional sources like call detail

records, app usage, credit and debit card transactions, email and consumer receipts, web traffic, and more. In the European Central Bank Report, Twitter was found to be an "economically significant" predictor of international stock market sentiment.

One industry example of alternative data at work is Quandl, a data marketplace founded in Toronto in 2012, is focused on alternative data and was acquired by Nasdaq in 2018. Quandl promises to deliver actionable, profit-generating insights from alternative data. It sources, evaluates, and productizes undiscovered data assets to transform them in quantified, actionable intelligence for institutional clients.

However, this data is not without its own challenges – the number one challenge we've seen is that alternative data is not always the most accurate, because it will often have missing data. While there are good and reliable alternative data sources available, it has not reached its full maturity – yet. Given this, firms are hesitant to tap into these useful sources of

data because of their reliability. However, since one challenge of new data sources is data quality, this should remind firms to keep data quality at the forefront of their data strategies – regardless of if they are using alternative data or not. In fact, according to a McKinsey survey ranking typical challenges in data transformations at banks, Lack of front-office controls (e.g., poor quality of data entry at system of origin with no/limited validation) was the number one challenge listed.

There are ways technologies that can be applied to help streamline the use of "risky" data and significantly reduce said risks. For example, by applying machine learning firms can enhance data quality and better mitigate risk by catching gaps and data quality issues upfront during the ingestion process. This reduces risks that using poor quality or incomplete data might cause downstream. And, downstream, this might include, instead, applying Data Science to application review for automating test case analysis to ensure data quality and reliability.

Social AI

According to VentureBeat, Social AI is a form of collecting and sifting through customer history, user-generated content, and data from social media channels to generate more relevant content and has the ability to provide a better social experience overall.

When building a Social AI strategy for banking, firms must consider tapping new and existing data sources, synthetic data, alternative data, and social AI as parts of their larger data strategies.



Due to the broad nature of Social AI, data science programs that leverage social AI can include techniques like peer group clustering and analysis, which finds like-customers based on social factors like age, gender, location, marital status, financial standing, etc., to create specialized targeting toward those groups’ collective wants and needs, analyzing social media for sentiment analysis and other customer data, the social aspect of partner networks and open source communities working in collaboration for solutions, as well as alternative and synthetic data.

Simply due to the massive amount of data generated by both individuals and enterprises every day, data is one of the largest assets for any firm. While we have already outlined where data science and analytics can be applied to shed light on more customer information to predict behaviors, better engage, service, and communicate effectively with customers for better experiences overall, there is still the untapped areas within operational areas of the business. Social AI can also

be leveraged as AI for HR, new customer acquisition for targeting marketing, customer engagement and behavioral marketing, and more.

HR, as one example, captures a large amount of data on employees from their application process before they are even employed with the company, through their employment with the firm, and through to their exit from the firm. By analyzing this data, there will be valuable information that can translate into actionable insights to better understand employees, their pain points, as well as their level and timing of engagements. By leveraging this data, HR can help organizations improve employee engagement and relationships, finding patterns in engagement and performance levels, and help improve employee retention.

On the marketing side, data science can be leveraged to analyze online behavior plus historical data to more accurately target leads. Data science can further look

at identified leads and apply predictive analytics to determine the probability of conversion on a lead, and therefore recommend which leads require more nudging, and additional targeted offers and communication.

As previously outlined, data science is a crucial aspect of predicting and creating the right roadmap for customer engagement as well. By leveraging predictive analytics, firms can gain customer insights to predict how and when they will most likely engage. In one example, the AWS Machine Learning created a solution using Amazon SageMaker, a platform for building and deploying machine learning models, to create a system that makes predictions on purchases based on a customer’s past shopping behaviors. Based on these predictions, the solution then uses Amazon Pinpoint to send campaigns to customers based on these predictions to get them to engage and act on the prediction.

Data Sources

Alegio and Dimension Research have found that 96% of companies have run into problems with data quality, labelling and model confidence, stalling projects. Because of this, the use of social AI, synthetic data, and alternative data is on the rise.



For firms to leverage different types of alternative data, it is important to understand the sources and factors that can be analyzed as an alternative data source. A wealth of this data can be sourced from social channels – LinkedIn, Facebook, Twitter, Instagram, Pinterest, Yelp, Kickstarter, Glassdoor, Indeed.com,

and more. While investors have looked to social media for sentiment analysis for a while now, this data offers much more than that – geolocation services, image recognition, email data, receipts and billing statements, and more can all be applied to create a very detailed picture of a company or individual to predict activity,

financial worth, pinpoint problems, and more.

This information can be derived from factors including NLP trend analysis, financial modeling, image analysis, and more. So how can Social AI data feed use cases?

Here are a few examples in different financial services use cases on how this data can be used:

In investment banking, customer sentiment can be taken from social media data on companies, which in turn can impact stock prices.

In venture capital, where data derived from crowdfunding platforms like Kickstarter and Indiegogo would inform sentiment around a company or product, popularity, demand, etc. to help VC firms make informed decisions before providing funding.

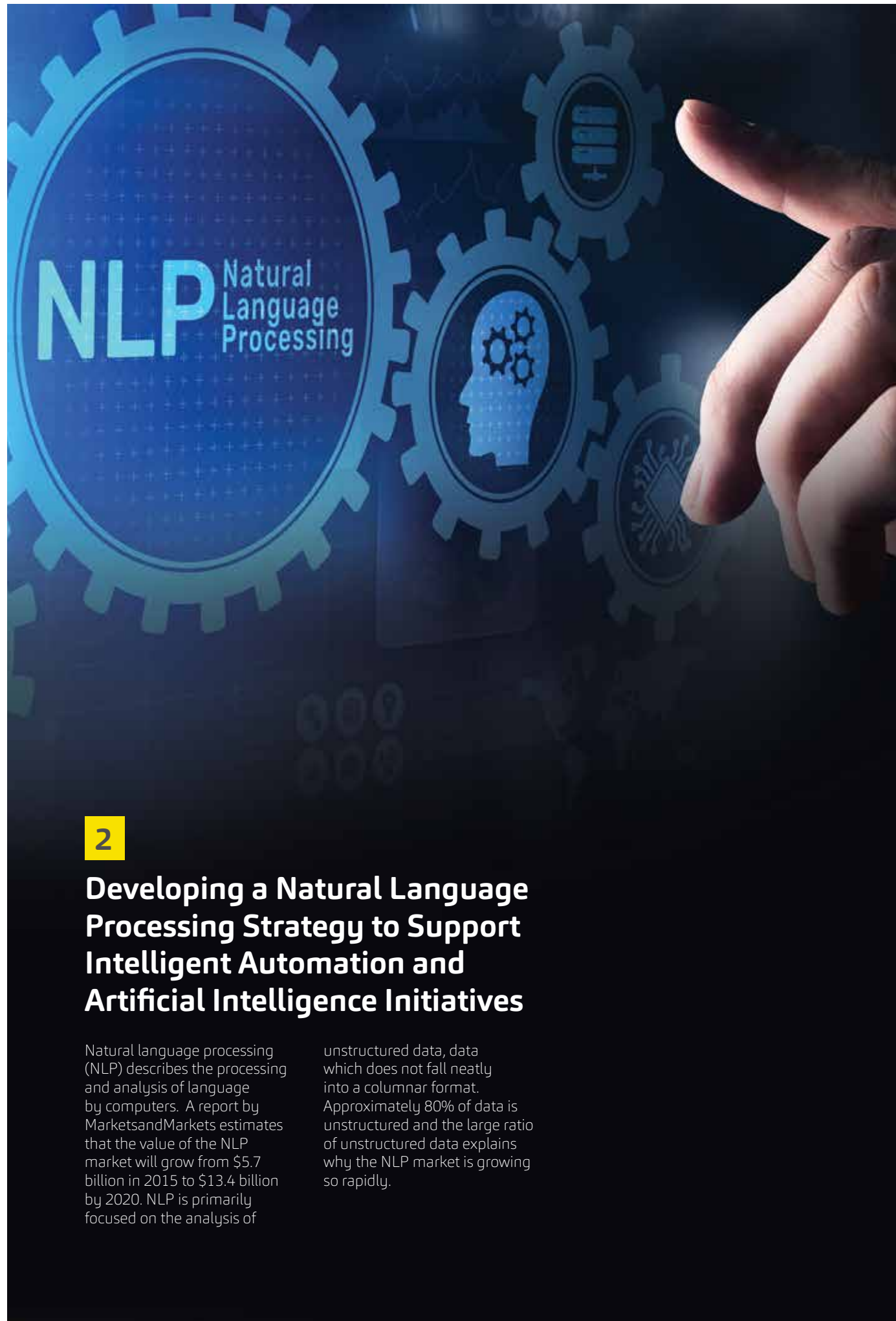
For hedge funds, a version of satellite data with image analysis could glean information that would help predict sales data by looking at the number of cars in shopping center parking lots.

In insurance, Synechron’s Visual Risk tool looks at alternative data sources, like google map image analysis and analysis of reviews on sites like Yelp to inform underwriting models by applying the data analysis against potential risk violations against these factors. For example, a form requiring the distance of a building from a fire hydrant could automatically pull that information from google maps imagery. If a restaurant doesn’t have a liquor license and claims to not be subject to risks related to serving alcohol, but a review on Yelp mentions drink service, or an image posted at the restaurant features a beer bottle, a violation would be cited.

In today’s digital ecosystem, the sheer amount of data created by both businesses and individuals every day is astounding – and means that, on some level, every company is becoming a data company. By thinking about data as this valuable asset, data becomes a powerful tool and can lend its value to power any data science program.

By leveraging data science in the aforementioned areas, firms will see benefits in cost-reductions, as the models will be able to create actionable insight that they would otherwise spend on campaigns that may or may not hit correctly, and likely overspend on these campaigns and engagement opportunities that they could have gotten from their own data. It also can more quickly and

accurately predict of individual customers and/or employees better than any human could. To harness the full potential of data, firms must look to three prominent data trends as part of solidifying a data strategy: synthetic data, alternative data, and Social AI to use both internal and external social data sources to garner powerful analyses.



2

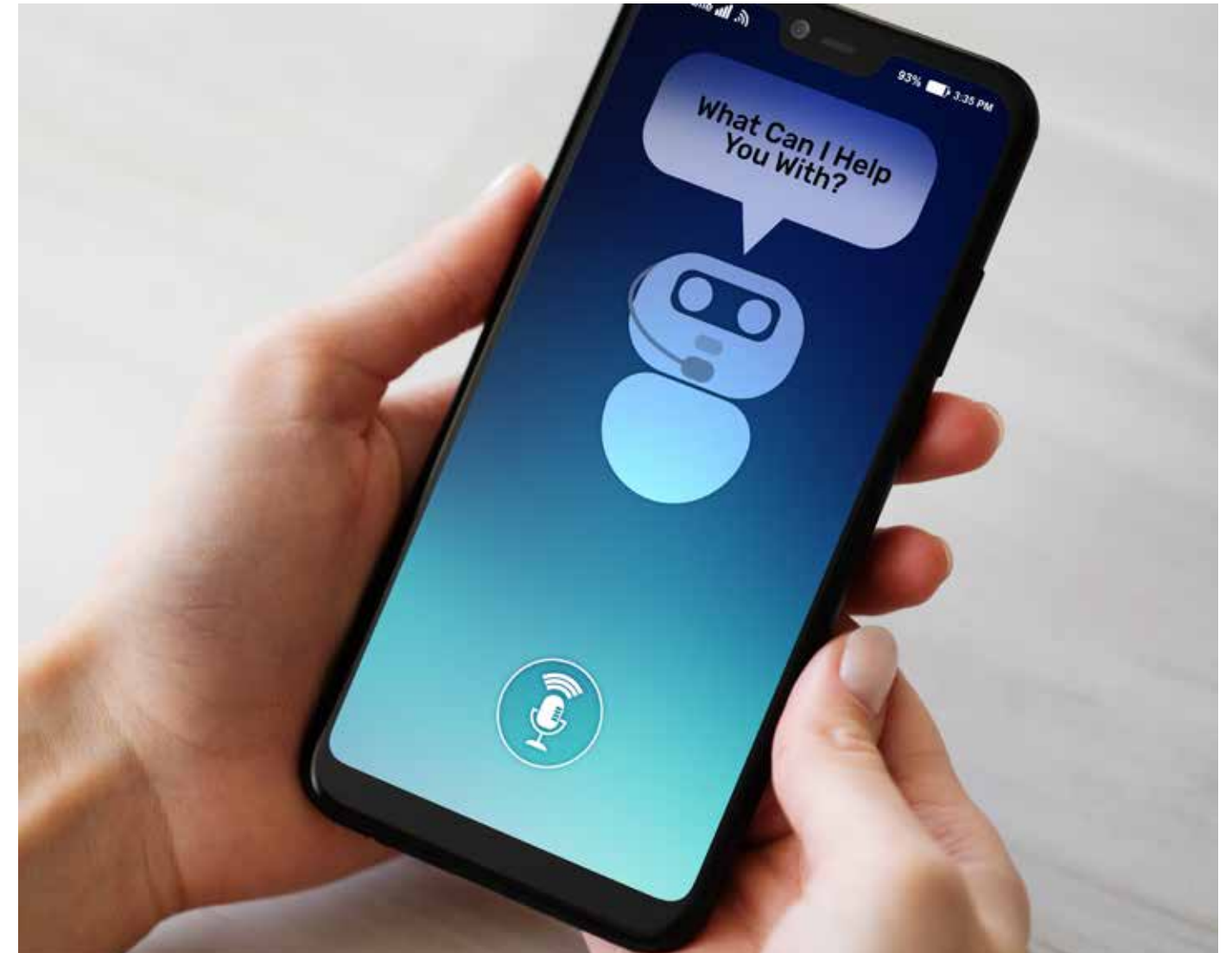
Developing a Natural Language Processing Strategy to Support Intelligent Automation and Artificial Intelligence Initiatives

Natural language processing (NLP) describes the processing and analysis of language by computers. A report by MarketsandMarkets estimates that the value of the NLP market will grow from \$5.7 billion in 2015 to \$13.4 billion by 2020. NLP is primarily focused on the analysis of

unstructured data, data which does not fall neatly into a columnar format. Approximately 80% of data is unstructured and the large ratio of unstructured data explains why the NLP market is growing so rapidly.

Robotic Process Automation (RPA) and Bot Automation

By 2020, at least 20% of people living in developed nations will use AI assistants to help with everyday operational tasks. Additionally, 40% of virtual assistant users are expected to interact primarily with new applications that support conversational user interfaces with AI. In fact, Gartner recently found that RPA is the fastest growing market in enterprise software, increasing at over 63% in the prior year.



Chatbots and NLP will save companies \$8 billion per year in customer support costs alone, primarily through the reduction of cost to hire and train call center assistants. Using AI to analyze customer interactions can drive higher customer acquisition rates and enhance the customer experience (CX). Customers are more likely to recommend companies that consistently and effectively implement AI solutions.

To date, NLP has primarily supported RPA strategies for workflow automation as well as supported customer service functions

including automated voice response technology and chatbots. Banks are currently looking at two key strategies for advancing their NLP strategies:

1. Leveraging Conversational AI to tap into new, customer-driven, conversational data,
2. Intelligent Automation (IA) to refine NLP analyses of both structured and unstructured data with more intent and analysis infused into these models.

By collecting a new class of unstructured data from conversational AI, banks can use analytics and visualization to learn more about their customers and create even more value-added services and platforms for additional points of engagement and a better understanding of pain points and desires.

NLP to Power Intelligent Automation

For use within financial services, IA requires an ability to understand industry specific language (a financial services lexicon) to refine data understanding and create more intelligent automation models.



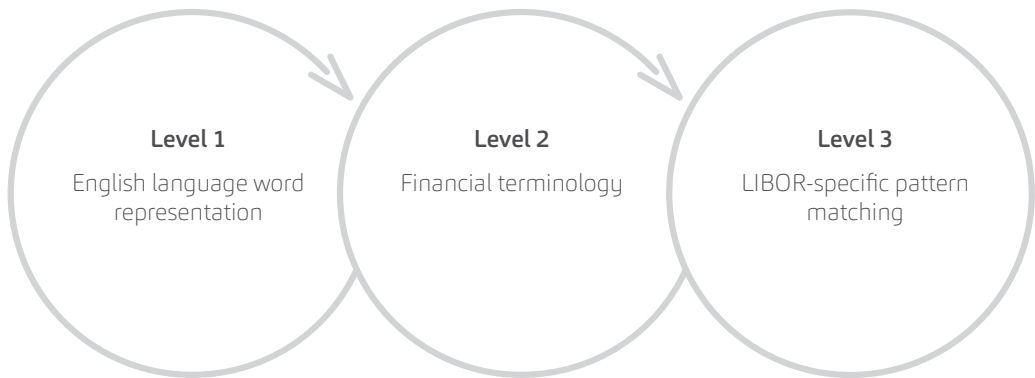
Case Study

Synechron recently took on this challenge in creating its Data Science Accelerator for LIBOR Impact Analysis, which enables financial institutions to identify and quantify their LIBOR exposure at a contract- or institution-wide level, by identifying LIBOR-related terms in the banks' systems. It does so by leveraging a combination of Optical Character

Recognition (OCR) and NLP to identify any reference to specific benchmarks (e.g. LIBOR) and data attributes contained in a contract and to return those results in a structured form.

Importantly, the NLP model analysis goes beyond simple keyword recognition to understand the intent, content, and materiality of that reference gaining meaning from a complex, proprietary financial services terminologies developed. It achieves this by using three levels of

Inbuilt NLP Models to understand language and patterns and uses reinforced learning to remove outliers over time:



Industry Use Cases

Bank of America's virtual assistants have surpassed more than 3.6 million users and assisted 12 million client requests to date. Orange Bank's Chatbot handles 24,000 conversations a week, with 20% of those queries occurring outside of regular working hours, and Commonwealth Bank has a Chatbot that understands more than 500,000 ways of customers asking for 500 different banking activities. Additionally, JPMorgan Chase streamlined its back-office operations and saved >360,000 hours of labor using Chatbots.



Additionally, SAS offers a software called SAS Platform, which it claims helps banks improve customer experiences and analyze customer feedback using NLP. The firm also offers SAS Visual Text Analytics, which leverages NLP to scale the human act of reading, organizing and extracting useful information from huge volumes of textual data.

Synechron also worked with a major consumer bank to build on its customer-first business strategy and add next-level data science models to its existing Pega-based client comment management and action automation platform to

enhance client services. The existing system automated the processing of structured data, but unstructured data required human review. It spanned complex document reviews across phone/branch/fax/email and auto-retrieved customer data, and routed feedback across almost 1,000 case types, 8 geographies and 14 fulfillment groups. Synechron was asked to assess the system's technological capabilities, enhance decision processes and infuse next-level business intelligence by using data science models to enhance labeling, grammatical understanding and intelligent decision-making across the system for bot processing. By using

state-of-the-art NLP models to enhance explainability, testing extraction methods including regular expressions, deep learning and machine learning models to make client comments structured and comprehensible to a bot, Deep learning models for Named Entity Recognition (NER) and end-to-end classification using cutting-edge, word-embedding models such as ELMo and Google BERT and pre-trained NER models for common labels, Synechron was able to implement scalable models to easily adapt to new use cases and changes in business rules and improved the client servicing system processing times.

Conversational AI

Conversational AI is gaining attention among banks and insurance companies with large call center voice data, text/messenger chat data, and other interpersonal touchpoints. Conversational AI deciphers everyday interactions – voice interactions in smartphones, spam filters in emails, customer support chatbots, and online foreign language translation.

In fact, it is estimated that 50% of analytical queries will be generated via search, NLP or voice, or will be automatically generated and that 40% of virtual assistant users are expected to interact primarily with new applications that support conversational user interfaces supported by AI. Financial services professionals estimate that by 2027, 66% of customer interactions will be completed by chatbots, generating a growing source of data for analysis and business



Within banking, who is leading the charge and how can banks capitalize? A survey conducted by American Express found that more than six in ten U.S. consumers say that their go-to channel for simple inquiries is a digital self-serve tool such as a website (24%), mobile app (14%), voice response system (13%) or online chat (12%). It would be impossible to cater to these needs without leveraging conversational technology which provides businesses with the opportunity to interact with customers/clients on human terms without using machine-driven interfaces. They can be chatbots, voice assistants, call bots, and research assistants.

To tap into this trend, bank leaders have done the following:

- Rabobank leverages a Google Assistant serving 1,500 – 2,000 clients per day for requesting account balances and setting specific spending limits.
- Capital One extended skills to tracking spending and making payments to all Alexa-enabled devices, including Echo Show.
- HSBC is rolling out Pepper the robot to its branches to “enrich the value and depth of customer engagements

and the overall branch experience” by offering product information, self-service banking capabilities, and more.

It’s not just customers driving the need for Conversational AI, NLP and Machine Learning – it is estimated that Chatbots and NLP will save companies about \$8 billion per year in customer support costs alone through a reduction in the cost to hire and train call center assistants.

Industry examples

Some of the biggest players in financial services have begun implementing NLP, and the industry will continue to mature as it becomes clearer how to create transparency within NLP and AI, and how to communicate more complex, industry-specific information.



Getting deeper into NLP usage in the industry, there are fintechs cropping up that focus specifically on these use cases, and will help mature NLP for banking use cases:

- Personetics is a London-based company founded in 2010. The company offers a chatbot called Assist, which they claim can help banks and financial institutions give their customers personalized banking

- services and aid them with product discovery using NLP. Personetics is currently working with six of the top 12 banks in North America and the European Union, serving over 60 million customers across the globe and servicing over 6+ billion personalized customer interactions to date.
- Kasisto, a NYC-based company, offers a conversational AI platform,

KAI, which can be used by banks to build their own chatbots and virtual assistants. Because it’s rooted in AI reasoning and natural-language understanding and generation, it can handle complex, industry-specific questions about finance management. Kasisto has so far been the platform powering DBS Bank, Standard Chartered Bank and TD Bank’s AI Assistants.

3

Transforming Customer Experience using Data Science to Achieve Broader Business Goals

As banks look to take their data science strategies to the next level, Applied AI needs to be a key focus. This means using AI to drive business goals such as improved customer engagement across the bank and insurance lifecycle, risk management around critical business processes and more. To dig deeper into how applied AI works, we'll focus on applying AI to financial services sales & marketing business goals, covering key areas related to Customer Experience enhancement including client acquisition, retention, engagement, product recommendations, and business intelligence across the digital ecosystem. We'll review how to apply the right data science models / approaches to achieve these goals.

One of the key areas where financial services firms are looking to add value is in customer experience (CX). CX is one of the major drivers for business-driven use cases, using data science to build and personalize customer expectations. According to a report by Efma and Capgemini, only half of global retail banking customers are happy with their services. Findings suggest that

retail banks need to collaborate and personalize customer experience with satisfaction notably higher among customers who had been offered personalized digital experiences proactively (49.1%) than those who had not (39.5%).

Given the financial sector is going through a radical transformation, largely driven by customer-centricity, emerging technologies have the power to redefine user experience through data-driven models. Significantly, a report from research firm Kantar noted that those financial institutions that lead in customer experience (CX) have a higher share of deposits and a better chance of cross-selling other products and services to their current customers. So, this is the value data science can offer to attain it as firms embark on their innovation strategies, and it is crucial that part of this focus on collecting and preparing data to feed into more analytics that will lead to more positive customer experiences, and the application of data science models, approaches, and techniques to leverage applied-AI to the right business cases with specific business goals in mind for success.

Client Acquisition and Retention

Predictive modelling is a huge focus for personalization, a top priority for banks seeking to acquire and retain the right clients. Companies that get predictive analytics right can greatly improve customer experience. User-centered insights enable banks to understand problems from a customer's point of view.



Real-time analytics enables customer experiences to be tailored at the same pace as data is collected and analyzed. Spotify demonstrates the value of real-time analytics by using algorithms of customer behavior to tailor playlists on the fly, leading to higher retention and improved customer experiences.

As firms ramp up digital transformation initiatives, data is becoming increasingly important, especially as data science applications come into play. Gartner's top 10 data and analytics technology trends for 2019 highlighted continuous intelligence (CI) as a major trend this year and beyond, reporting that 50% of significant new business systems will incorporate "Continuous Intelligence" – using real-time context data to improve decisions. According to Gartner, CI is similar to real-time data but builds on that to combine with business operations, processing current and historical data to prescribe actions in response to events.

To achieve the level of analytics previously

described to create meaningful and valuable customer insights and predict their behaviors, firms must leverage their data quickly and correctly – and have the right infrastructure in place to manage the amount of data and glean the amount of analytics and meanings from all that customer data.

Generally, predictive analytics and modelling are thought to be front-office techniques to engage with the customer. However, there is value in using models for the middle- and back-office. Even processes that are not customer facing have a significant impact on the customer. Fraud detection is one example. Synechron partnered with Quantexa to develop a data science driven solution focused on anti-money laundering (AML) and fraud detection. The tool uses peer group analysis to identify anomalies in trading activity, resulting in fewer red flags and false positives, as well as an enhanced customer experience by reducing actions like card freezes on suspicious activity due to a false positive alarm.

Banks also use predictive modelling in the back-office to predict system outages, forecast customer losses, estimate customer lifetime value, and optimally segment a firm's client base. For example, Sprint is able to identify a segment of customers ten times more likely to cancel than other customers. Firms able to pinpoint potential losses at the right time can prevent defections and increase profitability.


Recommendation engines represent another strong use case of applied AI. In one use case, Synechron's Relationship Management Dashboard Accelerator leverages a combination of advanced data visualization, data analytics, NLP and NLG, RPA, and more to predict what the next best action is for wealth management clients and prioritize client calls with auto-generated actions, predict client attrition risk and auto-deploy remediation measures, and also includes an AI-powered Investment Recommendations Bot based on historical data.


Identifying patterns for selling opportunities across different customer segments


Peer group analysis and cluster analysis is an essential part of data mining for data science, machine learning, pattern recognition, and more. It entails grouping a subset of data in such a way that members of a cluster have defined similarities to each other that make them markedly different than other sets. According to McKinsey, analytics-based customer segmentation is vital in shifting away from traditional processes in which banks have segmented customers to create more innovative models with advanced analytics and machine-learning algorithms allowing for a deeper, more nuanced understanding of customers.


Exhibit 1


Advanced analytics and machine learning can classify customers into microsegments for more targeted interventions.











Customer type	True low-risk	Absent-minded	Dialer-based	True high-touch	Unable to cure
Targeted intervention	Use least experienced agents provided with set scripts.	Ignore or use interactive voice message (segment will probably self-cure).	Matching agents to customers and live prompts to agents to modify scripts.	Focus on customers able to pay and at high risk of not paying.	Offer debt restructuring settlements early for those truly underwater.
Impact	Agent-client conversation guided by onscreen prompts based on probability of breaking promises.	10% time savings allows agents to be reassigned to more difficult customers and specific campaigns.	Can lead to increased "connection" and higher likelihood of paying.	Added focus addresses higher probability of default rates in this segment.	Significant increase in restructuring and settlements enhances chance of collecting at least part of debt.

Source: McKinsey analysis

McKinsey on Payments | August 2018

An example of successful cluster analysis is Synechron's Wealth Tech Accelerator for Client Prospecting. Wealth managers use

the attributes of existing clients to identify the most profitable potential prospects and their likelihood of becoming clients.

Applying Data Science to Customer Experience Driving Marketing and Sales

Mercator Advisory Group's recent Insight Summary Report, Omnichannel and Branch Banking: The Current U.S. Consumer Banking Environment, found that only 1/3 consumers want to be contacted about new products and services by their financial institution. Given this, financial institutions need to be very strategic in how they cross- and up-sell to existing customers and need to be smart about analyzing what their customers actually want, so when they do provide new services, or find a new channel for engagement, the customer is receptive.



Understanding this in creating data science models is vital in this digital and data-heavy climate because CI uses AI and ML to continuously interpret data, find the patterns, and identify the value – no matter how much data, how unstructured it might be, or how complex the patterns are. This allows firms to mine all the data they are collecting from customers and their engagements as they happen, and act accordingly. Over time, this will help firms build narratives with this data, and this would be true on a customer-wide basis, as well as an individual-customer basis, finding trends in what many customers are doing/feeling/

acting, and targeting individual customers at milestones or during low points to strategically re-engage. By understanding this, firms can maximize their opportunities to cross-sell and up-sell to existing clients in addition to just retaining them – creating a symbiotic relationship where the bank gets more business from the customer by creating more value-added services to help the customer through more intelligent engagements. One example for this new type of opportunity that can be mined from looking at the full customer lifecycle and predicting the right engagement points

around milestones and other critical points is around big life moments. Take for example the home-buying process. Traditionally, the bank's involvement ends with closing on the home and signing the mortgage. However, by adding offerings like value-added, direct to customer services critical to home owners ranging from banking and insurance services to smart device management, utilities and home renovations allow banks to be more fully engaged through milestones associated with home ownership for future lending and refinancing opportunities.

Applying Data Science to Customer Experience Driving Marketing and Sales

Mercator Advisory Group's recent Insight Summary Report, Omnichannel and Branch Banking: The Current U.S. Consumer Banking Environment, found that only 1/3 consumers want to be contacted about new products and services by their financial institution. Given this, financial institutions need to be very strategic in how they cross- and up-sell to existing customers and need to be smart about analyzing what their customers actually want, so when they do provide new services, or find a new channel for engagement, the customer is receptive.



Understanding this in creating data science models is vital in this digital and data-heavy climate because CI uses AI and ML to continuously interpret data, find the patterns, and identify the value – no matter how much data, how unstructured it might be, or how complex the patterns are. This allows firms to mine all the data they are collecting from customers and their engagements as they happen, and act accordingly. Over time, this will help firms build narratives with this data, and this would be true on a customer-wide basis, as well as an

individual-customer basis, finding trends in what many customers are doing/feeling/acting, and targeting individual customers at milestones or during low points to strategically re-engage. By understanding this, firms can maximize their opportunities to cross-sell and up-sell to existing clients in addition to just retaining them – creating a symbiotic relationship where the bank gets more business from the customer by creating more value-added services to help the customer through more intelligent engagements.

Pairing high-touch, personalized human touchpoints with digital innovation for improved CX – Industry Use Cases

In one industry example, USAA, a credit union that serves current and former members of the armed forces and their families, communicates with members via telephone and online, relying on employees (many of which are also former armed forces members) to connect and empathize with member needs. This, paired with focusing its digital technology to uncover customer needs, allows USAA to provide customers with the most relevant products and services, creating positive customer experiences. Due to this, USAA ranked at the top of research firm Katar's CX rankings for FS.



Russian bank Sberbank offers an AI-based tool called Tips that helps customers manage their finances more intelligently. The tool analyzes individual customers’ banking behaviors to provide personalized estimates for what will happen in their future, help them set goals, and connect them to the products that will be most useful for those specific financial goals.

McKinsey also reported trends in realizing the value of analytics in banking, sharing the case study of a top consumer bank in Asia. While the bank had a large market share, it ranked below competitors in products per customer. By leveraging advanced analytics to look at multiple big data sets focused on customer demographics and key characteristics, products held, credit-card statements, transaction and point-of-sale data, online and mobile transfers and payments, and credit-bureau data, the bank was able to reveal unsuspected similarities that allowed it to define 15,000 microsegments in its customer base. This, in turn, gave them context to build a next-product-to-buy model that increased the likelihood to buy three times over, and therefore

improve their customer to products ratio.

Synechron also worked with a major bank to create a system to predict customer cash flows. Knowing that customers desire self-service options to do what they want, when they want (usually outside of normal bank hours), a major bank sought to enable its customers to predict cash inflows and outflows based on their historical spending patterns. Using the bank’s mobile app, customers could proactively manage account balances. The tool resulted in an improved online banking experience for thousands of customers, and the bank saw a reduction in customer overdrafts as well as an increase in deposits, which were attributed to improved transparency and increased customer trust. The tool also led to increased levels of customer acquisition and retention.

User-centric insights also create the unique ability to be able to empathize with customers, understand what they care about, points of satisfaction and delights, dissatisfaction and annoyance. This information can be analyzed using data science to uncover latent things that customers may not even be aware they

do or do not want or prefer based on that analysis of behaviors. This creates a unique opportunity for data science teams to use advanced models and advanced analytics to find out those needs, and thus provide really valuable opportunities for new customer solutions that would otherwise be unmined.

Financial institutions can take a data-science led user-centric approach to analyze data about their customers, and, in turn, better understand their wants, needs, and actions. Research shows that 77% of high-performing customer service teams rate their ability to leverage artificial intelligence as excellent or above average. Given this, banks are using machine learning and data science increasingly as a whole on CX in interacting with customers. Whether it’s to predict at what point do clients drop off in the process, how far a chatbot can be used before a human needs to step in and intervene, and finding more sophisticated, accurate ways to gauge emotion using technology beyond sentiment analysis. One way to do this is to analyze sentiment coupled with intent and context.

Amelia

IPsoft’s Virtual Agent, Amelia, is coined “the most human AI” is one example of this. Forrester Research stated that Amelia’s use of Natural Language Generation (NLG) is used to engage with real-tome conversations with humans, making it a digital colleague, a customer service agent, an IT support staffer, and an operations expert. Everest Group named Amelia the most capable intelligent virtual agent due to its ability to make contextualized conversations in more than forty languages, and her ability to make personalized decisions based on real-time data analytics.

- This level of engagement and connectivity is made possible by:
- 1. Semantic Understanding and Memory – to achieve meaning and concept
 - 2. Episodic Memory - to build business knowledge
 - 3. Contextual Comprehension – to understand the context of “it” is a home, a car, etc.
 - 4. Emotionally Responsive – to create the perception of empathy for customer experience – for example, being able to tell a customer is getting irritated, lonely and just wants to talk to someone, etc. and escalating to a human based on emotional cues (tone, etc.) , and by sounding emotionally human in the tone of its own responses.



4

Completing the pre-requisites for a data science ecosystem

Data science is only applicable if developed within an infrastructure that enables scalable data acquisition and cleansing as well as fast, flexible, and scalable model development, testing, and deployment. Legacy systems pose a challenge to banks trying to adopt modern AI practices

Banks are in the process of modernizing their technology stacks to enable flexible, scalable, and auditable machine

learning model development. While some banks are choosing to partner with third-party providers, many are building their own technology stacks consisting of Spark, Hadoop and technology for Model Lifecycle and Data Lineage management.

How to set up teams and be successful

Creating a data-driven organization with a strong data science practice requires a diverse mix of skills. A data scientist is a combination of statistician, programmer, and mathematician with deep domain knowledge and skills in product management.



This skill set is difficult to find in a single hire, especially given the complexity of the field and the years required to obtain mastery of areas including linguistics, statistics, software development, product management, and mathematics combined with deep expertise in one or more domains within a field. An optimal team will often consist of individuals with one or more complementary skills including machine learning engineers, software developers with knowledge of machine learning algorithms, statisticians, and experts in fields including artificial intelligence, machine learning,

- **Business Consultants** – to bring financial services and insurance domain expertise as well as best practices in data governance, target

- state operating model creation, innovation management, and more.
- **Data Engineers** – to create consistent and easily replicable data pipelines for data scientists. These experts bring the organization industry-leading skills in areas including data governance, architecture, and ETL as well as Computer Science skills.
- **Data Scientists** – to produce mathematical models for prediction/recommendation during on strong skills in Statistics, Math, Computer Science, and more.
- **Machine Learning Engineers** – to deploy mathematical models on a high-compute, cloud-enabled technology architecture, drawing on
- a multi-functional skill set across data science and software engineering
- **Application Development Supporting team in UI/UX/DevOps**

Employing the Right Data Science Infrastructure

While data science can be pivotal in banks’ business strategies to solve their most complex business problems and create transformative new innovations and approaches, industry domain knowledge is crucial.



Beyond the need to adhere to a specific and complex financial services lexicon, data science programs need to be able to operate in line with the highly-regulated financial services environment. And with most bank clients operating on a global-scale, both region-specific and global requirements, which is constantly changing, must be followed. Additionally, existing data science teams need guidance on how to design, execute, and build data science projects, products, and infrastructures. This will require flexible big data platforms, and SAS for models and infrastructure using the Cloud.

An important part of achieving this is having in place the infrastructure for data ingestion, data modelling. To do this, firms

- should put consider their Data Science Technology stack, which can help outline how firms:
- **Acquire data sources** - including streaming/in motion, staging/at rest, operational systems and distributed, centralized and monolithic, and more
 - **Organize** – through processing and data preparation, data modeling, and logical data warehouses stakeholders can organize their data for effective processing, training, and virtualization
 - **Analyze** – once the data is organized, it can be analyzed by line-of-business applications covering execution and operationalization with key areas of focus including traditional reporting, smart data recovery, visual exploration, storytelling, and advanced analytics
 - **Deliver** – the analysis can be delivered to relevant devices, applications or data stores by system integration, automation or APIs.

Data Science Operations

One key consideration is having the operations in place to deal with highly-sensitive financial data in the highly-regulated financial services environment.

This is why creating a detailed data governance plan is critical to developing a data science strategy and using any data within models. This is important to keep data safe, and to deter data-associated risks like data provenance, validity, privacy risk for personal and financial data particularly. A governance plan will safeguard highly sensitive financial services data and makes sure data is of high quality to be useable and focuses on the ways data is used and changed through different processes and how that can be recorded.



Another limitation can sometimes be the team itself – if only a handful of people know how the solution works, what the bugs are, and how to fix it, no one else can fix it if a production issue arises, and therefore the risk is too great to be scaled. These challenges can be addressed to move from having data scientists working in siloes to the institution working at scale on production-grade machine learning products, the correct documentation needs to be in place, including DevOps, Model Lifecycle Management (MLM) and more.

One way to address these challenges is to create data ontologies, which organize correlated data, searchable terms for data on specific subsets of clients can be found without having to compromise the data privacy and right to be forgotten required by global data regulations. This can act as a library for the banks’ data, where it can be categorized and analyzed so firms can understand how data is related, different types of data, and why categories are important not just for organization but to create data trails and comply with regulatory requirements, better allowing banks to understand all of the data in their business and be able to explain it ontologically.

Another way to achieve documentation is annotation against the models to achieve enterprise-scale data science done

effectively. Traditionally, many legacy systems are rules-based and want to move to predictive. By adding annotations to the models to work at scale, business analysts can understand what is going on for regulatory visibility and auditability, and so that they can understand the process behind the models to be production-grade.

In one case study example, Synechron worked with a tier-one bank to use data science for predictive AML suspicious activity reporting. The large, global bank wanted to move beyond rules-based transaction monitoring toward predictive suspicious activity monitoring of bank money transfers, foreign exchange, mortgage transactions and other transactions where it served as the originator or recipient. Synechron built advanced Data Science and Machine Learning models to identify and predict potentially suspicious transactions using Logic Regression, Random Forest and Support Vector Machine algorithms and designed the solution to deliver continuous testing/model training to reduce false positives & annotate the rationale for analyst review.

To be the most effective, banks should look to implement a single unified platform for Model Lifecycle Management (MLM) that is hosted on the cloud and is technology agnostic to work in all

environments. This would serve as a source control for models and algorithms. In this platform, Automation via DevOps could enable continuous integration and delivery, as well as automated data and model versioning, thorough data and model lineage, and integrated development, testing, validation, and production. A single platform will give banks the tools needed to create one place so the workflow is shared from data scientist a model can go to an engineer, so passing of the baton is automated. In this system, source control is controlled for governance matters, and everything is backed up, and anything everyone has done is backed up, and everything is versioned.

Financial institutions want to use AI to drive client behavior around key topics such as acquisition and retention, but existing data science teams need guidance on how to design, execute, and build data science projects, products, and infrastructures. This will require flexible big data platforms, and SAS for models and infrastructure using the Cloud, and leveraging future-proofing technologies.



Conclusion

For financial institutions and insurance companies today, digital transformation and innovation strategies are top-line agenda items. To this end, data science is critical to any innovation program. As firms delve into their data science agendas, there are several key themes to consider in crafting a successful strategy.

To do this, as outlined in this paper, firms need to ensure they have the right infrastructure in place, and critically implementing a data strategy to source and prepare data to power AI and data science. Once these critical steps are complete, firms can make all of this actionable via NLP strategies, and re-thinking their business models to take a CX-led approach using data.

Accelerating Digital for Banks, Asset Managers, and Insurance companies.

Synechron is a leading Digital Consulting firm and is working to Accelerate Digital initiatives for banks, asset managers, and insurance companies around the world. We achieve this by providing our clients with innovative solutions that solve their most complex business challenges and combining Synechron's unique, end-to-end Digital, Business Consulting, and Technology services. Based in New York, the company has 18 offices around the globe, with over 8,000 employees producing over \$500M+ in annual revenue.

18

Global
Locations

8000+

Team Members
Globally

Vertical

Focused only on Financial
Services

\$500M+

Privately Held and Self-
funded

150+

Marquee
Clients

Power of 3

Digital, Consulting and
Technology services

Our Value proposition

A unique approach to market differentiation in the financial services domain

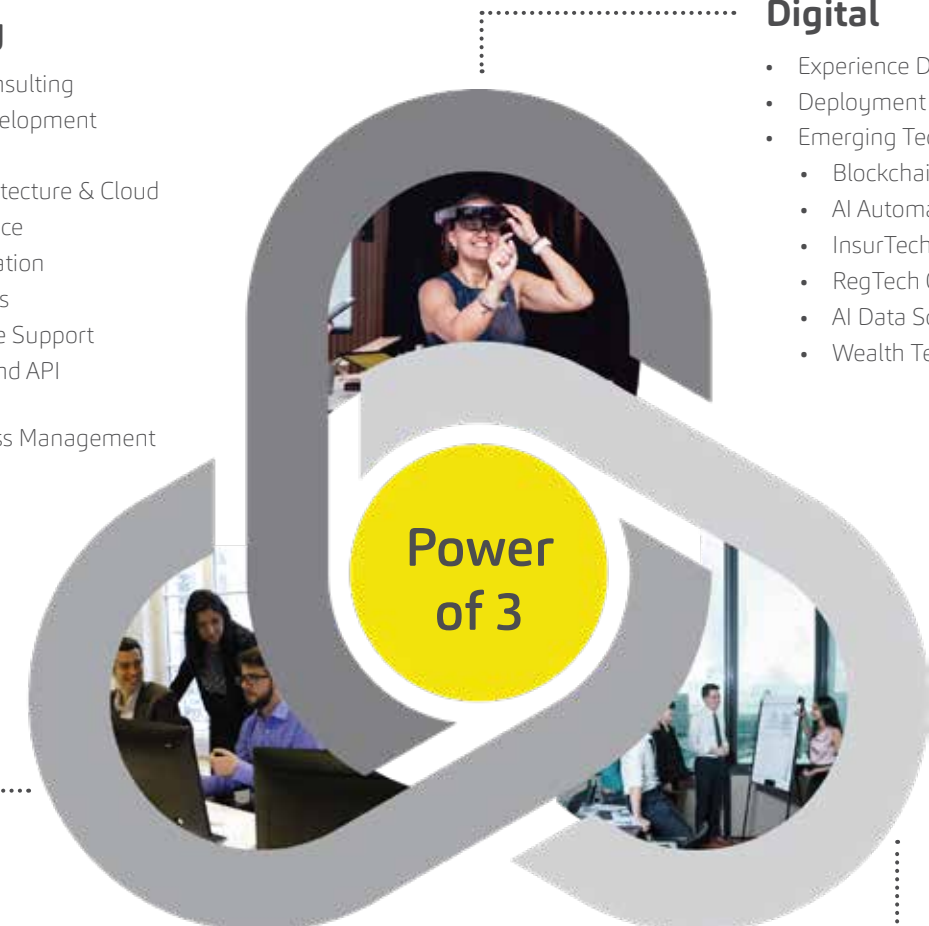
Unlike other firms, Synechron's "Power of 3" approach and financial services expertise gives us a competitive edge to tackle our clients' problems from any vantage point with great depth. Synechron combines the "Power of 3" - business process knowledge, digital design and core technology delivery excellence - to drive transformative solutions. We have the unique ability to provide an end-to-end approach, from business consulting through technical development to digital enhancement. This empowers us to deliver solutions to some of the toughest business challenges.

Technology

- Technology Consulting
- Application Development
- Automation
- Enterprise Architecture & Cloud
- Quality Assurance
- Systems Integration
- Data & Analytics
- IT and Database Support
- Microservices and API Development
- Business Process Management

Digital

- Experience Design
- Deployment and DevOps
- Emerging Technology Frameworks
 - Blockchain COE
 - AI Automation COE
 - InsurTech COE
 - RegTech COE
 - AI Data Science COE
 - Wealth Tech COE



Business Consulting

- Enterprise Strategy, Architecture and Transformation
- Client Experience and Lifecycle Management
- Data Science, Engineering and Governance
- Regulatory Change and Compliance
- Finance and Risk Transformation
- Innovation Management
- Program Management Services

Global Footprint



Synechron
Digital / Business Consulting / Technology

www.synechron.com | Email: info@synechron.com

Proprietary material

"This material and information is the sole property of Synechron and is intended exclusively for general information purposes. Any rights not expressly granted here are reserved by Synechron. Please note that copying, modification, disclosure of data, distribution or transmission of this material without prior permission of Synechron is strictly prohibited."