

**TECH MAHINDRA  
DATA SERVICES  
HUMAN IN THE  
LOOP FOR  
BANKING AND  
FINANCE  
BANKING AND  
FINANCE**

---

WHITE PAPER

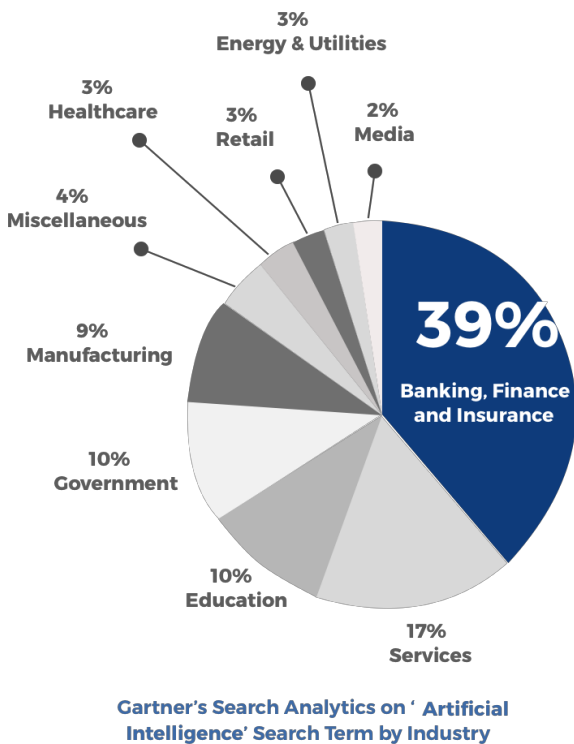


Connected World.  
Connected Experiences.

## Introduction

Artificial Intelligence is playing a visibly larger role in banking over the years. The industry has never shied away from trying new AI and ML offerings. A 2017 search analytics report by Gartner is evidence of the industry's acceptance of modern technology to solve some of the toughest challenges it faces. More queries related to AI and ML asked from Gartner are from the banking industry, more than most other industries.

Machine learning a reality in banking, a highly regulated and data-intensive sector. Banking is in a better position to reap the benefits of machine learning. A major reason for this is that banks churn out a copious amount of data every day in multiple forms. Banks can use this data to improve operations, increase revenues and enhance the customer experience.



## Where is Machine Learning Being Used in Banking?

Machine learning is gradually reshaping banking operations by revamping several processes and operations. It is making inroads into various aspects of banking with increasing use cases. Most machine learning use cases in banks fall under four categories.

- Customer-centric use
- Operation-focused cases
- Trading and portfolio management
- Compliance and supervision

The following are a few use cases that highlight how banks are leveraging machine learning to step up towards the future.

- Customer experience
- Digital assistance and chatbots
- Investment advice
- Fraud detection and risk management
- Regulatory compliance
- Equity Predictions/ Hedge fund management
- Wealth and portfolio management

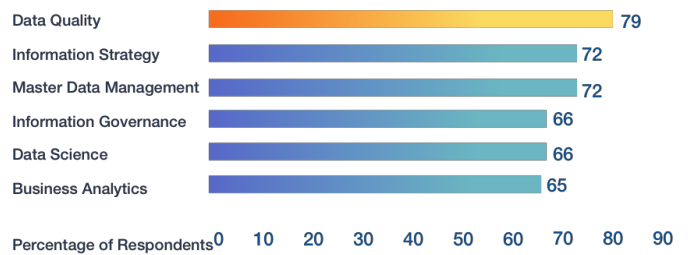
**“In God we trust, all others must bring data”**

- W. Edwards Deming

## Why is Data Quality important?

The efficiency of machine learning models is only as good as the data they process and get trained with. Any anomalies or inaccuracies in data result in suboptimal outcomes and lead to the model continuing to acquire inaccurate learning, affecting future outcomes.

Today, data quality is a colossal challenge that organizations are looking to address. A 2018 Gartner research highlighted that data quality is the primary aspect of the chief data officer role. (Figure below)



Contrary to common belief, machine learning goes beyond choosing and tuning algorithms. The model should be able to easily consume data and learn from it. It should also be predictable and scalable. However, for predictions and recommendation by the model to be dependable, it should be ensured that the data is correct. It should not have missing values, incorrect entries, duplication or redundancy.

Failure to ignore quality may lead to the introduction of biases and inaccuracies in the machine learning model which will continue to increase over time. Thus, bad data can start to impact your future decisions from today. This particularly can be detrimental to a bank as these biases may result in substandard customer services, higher operational costs and increased fraud risks.

In addition to this, banking regulations are more stringent than in other sectors and industries. If your model learns based on incorrect representation, it may even lead to non-compliance, attracting severe penalties from regulatory bodies. Therefore, ensuring high-quality data is not an option but a mandate for banks.

Tuning hyperparameters, building an ensemble and performing feature engineering are some of the recommended ways to improve data quality. However, it needs more than these approaches in banking as they focus on refining data quality without the need for more data and are only effective to a certain extent. Banking, a highly customer-centric sector, requires a significant amount of contextualized and annotated data.

## What Does Data Contextualization Do?

Even a huge volume of current and historical data may be unable to deliver if it fails to generate insights that banks look for to make strategic decisions across various departments. To do that, there is a need for precise classification of parameters based on a thorough understanding of data and patterns hidden in it. This means it requires massive and organized data which is feature-rich and correctly labeled.

Contextualization makes it possible to derive valuable and actionable insights from every piece of data at various stages and refines outcomes. It ensures that machine learning models are trained to deliver enhanced output with the desired accuracy and relevance.

It is also important to note that banks generate data in different formats and on different platforms. This includes text, images, voice, videos, social media interactions among others. Hence a linear approach to decipher the collective data is more futile than inefficient.

As an example, only analyzing the number of telephonic interactions of customers or their frequency of visiting a branch in a specific period may be of little avail if you are looking to know the reasons for attrition or reduced transactions.

Now, add other parameters to this such as social media activity, visited outlets, customer profiles (technophile, saving-focused, shopaholic etc.). This enables you to create machine learning models that only tell you the reasons for inefficient customer service but can also predict any such instances—quickly and accurately. For instance, an angry tweet from a dissatisfied bank customer can show the gaps in services delivery and help you fill it.

## Use Cases of Our Contextualization Services

Here are a couple of examples of how you can leverage data contextualization in banking.

### Customer Payment Data

Annotation derives valuable insights from customer payment data, such as identifying spending patterns and channels of transactions, and customer categorization. While the insights it can deliver are endless, it requires to context understanding to be able to fully leverage this data.

Generally, customer payment data involves number, nature, time and location of transactions and exists in an unprocessed manner. We synthesize the payment data with other data sources like personal information of customers and their needs and preferences. We then add other key parameters—for example, customer intent, influences, triggers and so on.

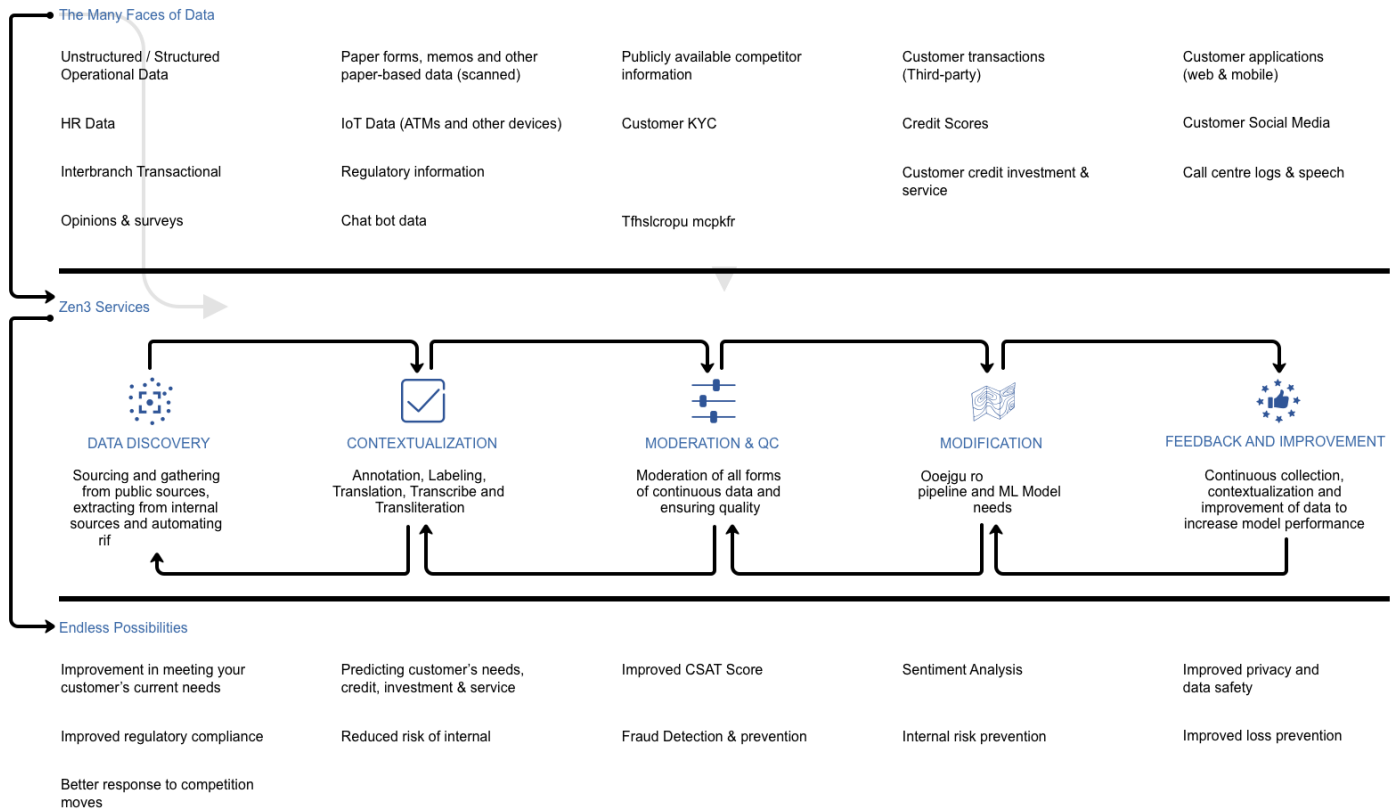
In the final stage, we transform the synthesized data by enhancing it with the use of AI, IoT, big data, machine learning and other data sources. Thus, we help you optimize the outcomes to make better strategic decisions.

### Fraud and Risk Management and Prevention

Fraud prevention is a priority area for banks. Majority of CIOs in the banking sector are reported to be most concerned about transactional data protection. It is followed by customer segmentation, call center virtual assistants, sentiment analysis among other fraud risk concerns.

Our annotation experts use their expertise in machine learning, NLP, social media listening, and analytics to identify risk indicators, anticipate frauds, verify customers, and build models to minimize risk.

# Data in Its Many Forms Needs Context!

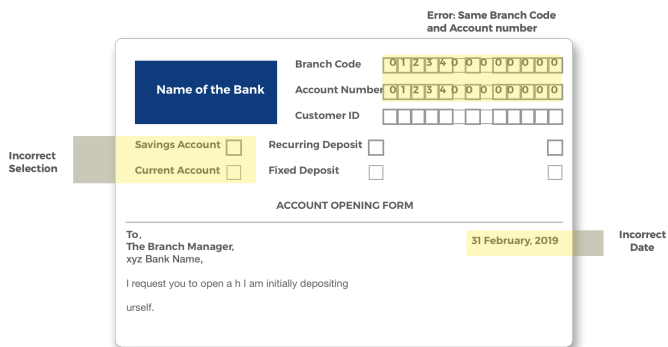


## How Do We Contextualize Data?

Following are some of the ways we contextualize your data from all existing formats and annotate it to enhance the capabilities of your machine learning model.

### Image Labeling

Following are some of the ways we contextualize your data from all existing formats and annotate it to enhance the capabilities of your machine learning model.



We identify and label all transactional information including correct store name, types of good sold at the store and by the merchant along with the correct industry codes. Additionally, we also label purchase location (GIS long-lat) globally through the web or by calling the store directly.

[Side Box - Apple debuted a version of merchant identification with their Apple Card to provide both the correct purchase type and easy visual reminder of the purchase to help in budgeting.]  
Design Notes - We will show labels inside labels here.

Image labeling helps the bank in determining:

- The propensity of your customer in buying at the store
- The propensity of customers converting that transaction to an EMI
- The type of purchases your customers usually make
- Providing better information to the customers in their purchases

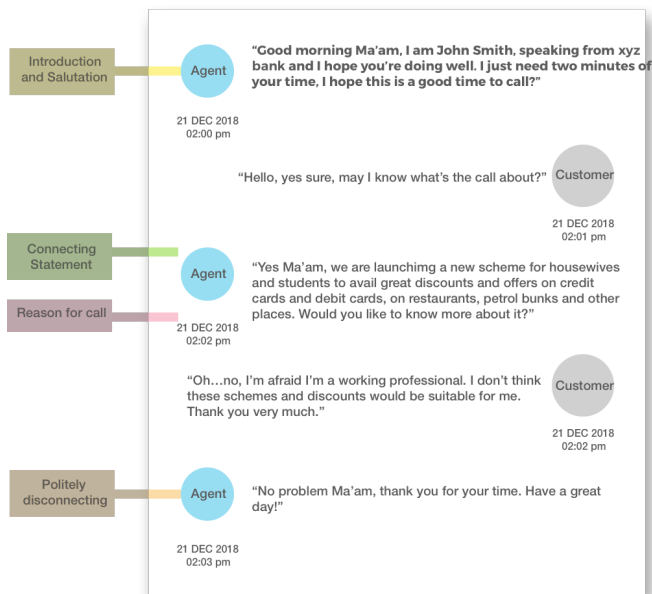
Here is another example for image labeling.

Design notes - find a better image. We will show labeling and an extra layer of information here

We label all such information as given in this image and ensure that your transcription team uses correct spelling and remaining details are accurate. We also use the email addresses and social media profiles of your customers to add other required pieces of information. We further tag them to facilitate more targeted campaigns

## NER/ PoS Tagging

Following are some of the ways we contextualize your data from all existing formats and annotate it to enhance the capabilities of your machine learning model.



Design notes - please find a better example of a call center transcript, take it from the Sayint team

We collect all speech data for your customers from your call center and convert it to text. We then use our sales and customer service expertise to analyze the text and label it.

[side box content - We are experts in Indian languages, including Indian English]

Zen3 assures a better accuracy in it than most of its competitors

## Text Classification

Design Notes - Insert image of a chat conversation, an email conversation and call center script. Label sentiment or request type.

We perform text classification of the notes for all your customer interactions available in the text form. It helps you build a corpus of data to identify the intent behind the communication through your machine learning model.

Using text classification, you can improve chatbot performance or build automation programs to respond to customers via either email or voice call.

## Text Summarization

Design notes - insert image of a financial regulatory filing sheet.

We are quick and efficient in summarizing data from regulatory notices and filings. Besides, we scan large data sets online and share the summary as well as social media sentiment analysis.

## Human-Reviewed OCR

Design Notes - Insert image of a paper form or a handwritten request letter

We annotate and tag the entire data available in the paper form throughout the journey of a customer. Text-tagging with correct annotation improves machine learning models and helps summarize the data. Using this, you make sure your bank takes the right and timely action.

## Social Media Information Tagging

Even with limited information on application forms of customers, we build their online and social media profiles and generate insights into all their activities online by tagging and summarizing their profiles.

## Why Zen 3

Zen3 has a proven expertise in aggregating and labeling your customer data using analytics, segmentation and metrics. We clean your data by removing any anomalies and annotate it to make it a high-octane fuel for your machine learning models.

Here are a few reasons why you can consider Zen3 a dependable AI partner.

AI, Deep Learning

A/B Testing, Experimentation, Simple ML Algorithms

Analytics, Metrics, Segments, Aggregates, Features, Training Data

Cleaning, Anomaly Detection, Prep

Reliable Data Flow, Infrastructure, Pipelines, ETL, Structured & Unstructured Data Storage

Instrumentation, Logging, Sensors, External Data User Generated Content

## Proven Quality

With over 15 Million judgments done for clients in the past 3 years, we've maintained a robust quality assurance process that drives analysis accuracy of >95%. Training ensures consistent analytical standards, bringing individual disagreement rate below 5% across projects.

## Talent Advantage

Human-powered data pipelines require committed personnel, and we've maintained a 0% attrition rate for HIL middle management over the last 3 years, all while scaling up hiring in support of data pipelines based across the US, Europe, and Asia. Our dedicated talent acquisition teams based out of Seattle, Singapore, Hyderabad, and Delhi ensure continued access to premium talent.

## Focus on Value

Our HIL projects center on proven cost-management practices, with prices 30-50% below leading competitors. An absolute dedication to process-improvements has led to a consistent long-term throughput acceleration across all data annotation efforts. We have even proactively identified new efficiencies for large scale labeling efforts by suggesting workflow improvements to our clients.

## Deep Expertise

With 130+ editorial resources with rich digital & print media experience currently engaged, we have the expertise needed to support even the most complex, media-rich data pipelines. This subject matter expertise is coupled with deep knowledge of the content management tools needed to administer these efforts efficiently.

## Staffing Capabilities

With 130+ editorial resources with rich digital & print media experience currently engaged, we have the expertise needed to support even the most complex, media-rich data pipelines. This subject matter expertise is coupled with deep knowledge of the content management tools needed to administer these efforts efficiently.

## Background

While Banks are trying to leverage everything they have at their disposal to reduce customer churn, regulatory bodies are aggressively pushing for demanding compliance norms.

Conversational Analytics or Speech Analytics refers to a platform to acquire, transcribe and analyse 100% of all customer conversations across channels including phone, email, chat and social. These insights help satisfy contact center efficiency benchmarks, agent performance and regulatory standards simultaneously, thereby increasing revenues.

The Banking and Financial Services industry includes Banks, Lenders, Mortgage Services, Collection Agencies, Brokerages, Insurance and other financial services.

## Opportunity

These institutions have 3 major challenges:

1. Customer Churn because of dissatisfaction which directly impacts bottom-line
2. Rising complaints, investigations and penalties from different regulatory bodies
3. Losses due to agent attrition, productivity and training challenges.

Learning from customer conversations is the best action that any financial services institution can take to address these problems. Conversational Analytics is the most powerful platform for achieving these goals.

“With Zen3, we understand our customers better and we have been able to relay this information to our marketing teams which has helped us significantly in better targeting translating to improved revenues.” - Vice President, Customer Success

## Solution

Automated QA and Root Cause Analysis of 100% of all customer conversations will help banks and other financial services agencies produce significant cost savings. While this helps in reducing average call handling time and improves productivity, it also has a direct correlation with customer satisfaction levels. Analyzing conversations also helps understand knowledge or training gaps in each agent which helps the organization tailor training programs and improve First Call Resolution.

“We did less than 3% audit earlier but with Zen3, we’ve achieved 100% audit. We have also saved tremendous amount of time using the search functionality on Zen3’s dashboard to identify keywords, sentiments, topics and contexts directly from conversations.”

- Compliance Manager, A leading retail company

Banks are consistently sharing success reports on how improving agent performance and onboarding time has significantly reduced churn. Speech Analytics helps identify exact gaps in training and makes training programs more focused.

Our Speech Analytics solution helps you uncover valuable insights to improve agent performance. We analyze both real-time and historical communications

# Tech Mahindra



[www.youtube.com/user/techmahindra09](http://www.youtube.com/user/techmahindra09) [www.facebook.com/techmahindra](http://www.facebook.com/techmahindra)  
[www.twitter.com/tech\\_mahindra](http://www.twitter.com/tech_mahindra) [www.linkedin.com/company/tech-mahindra](http://www.linkedin.com/company/tech-mahindra)  
[www.techmahindra.com](http://www.techmahindra.com)

Copyright © Tech Mahindra 2020. All Rights Reserved.  
Disclaimer. Brand names, logos and trademarks used herein remain the  
property of their respective owners.