

BIG DATA IN BANKING

Table of contents

What is Big Data?.....	3
How data science creates value in Banking.....	7
Best practices for Banking. Case studies.....	10
1. Fraud detection.....	10
2. Contact center efficiency optimization.....	11
3. Customer churn analysis.....	12
4. Risk management	13
5. Centralized marketing campaigns.....	14
About InData Labs.....	16

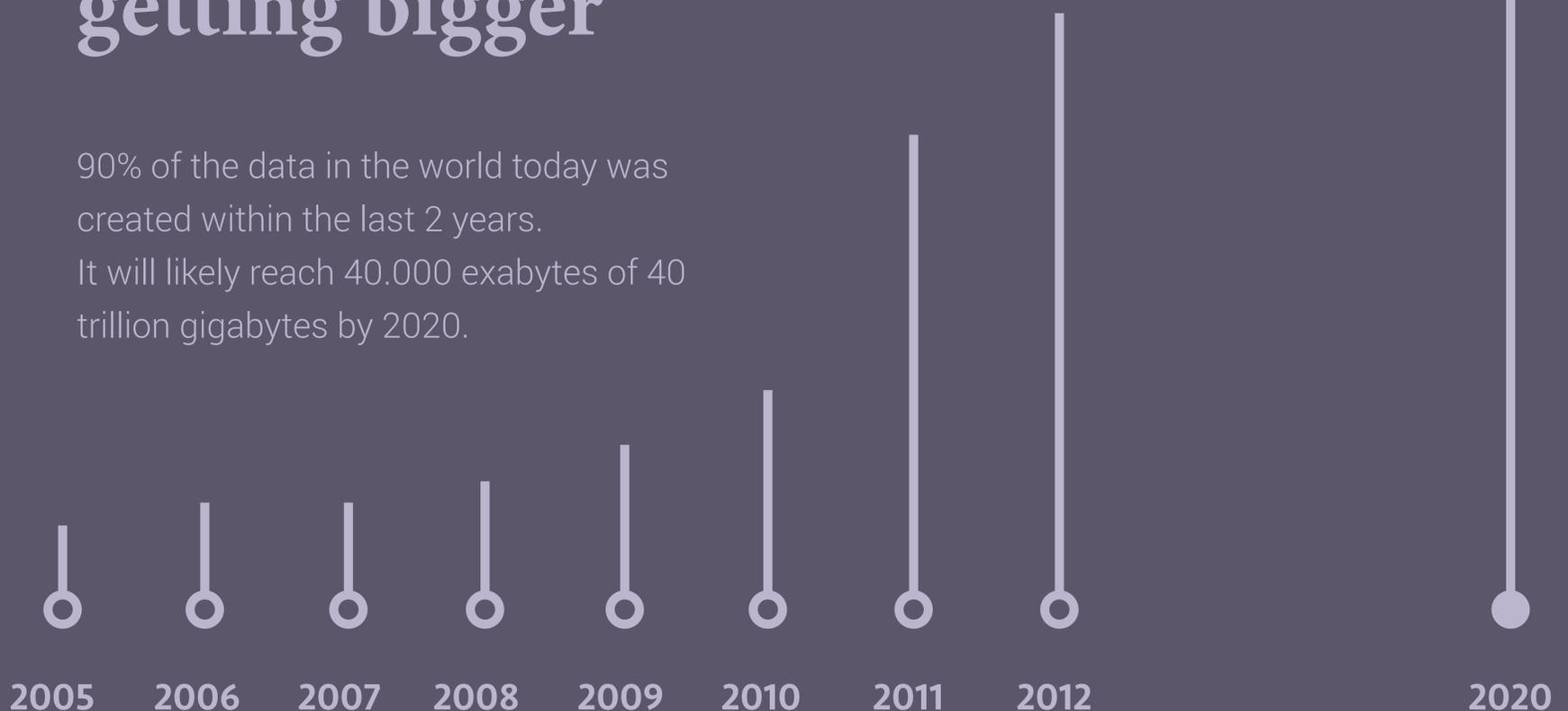
ABOUT BIG DATA

We create 2.5 quintillion bytes of data every day —

so much that 90% of the data in the world today has been created in the last two years alone. This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals to name a few. This massive, diverse and unstructured data which is impossible to process via standard software and data bases is called Big Data.

Big Data is only getting bigger

90% of the data in the world today was created within the last 2 years. It will likely reach 40.000 exabytes of 40 trillion gigabytes by 2020.



2.2 million terabytes of new data is created every day

Nowadays data management is becoming a critical differentiator that separates market leaders from all others. Most enterprises face Big Data, which is so large that it is impossible to process it using traditional software tools. Forward-thinking companies actively crunch their high-volume unstructured data to get competitive advantage and find new business opportunities.

Today's high-end technologies make it possible to realize value of Big Data. For example, retailers and banks can analyze behavioral trends and social media activity of each customer and provide personalized product offerings. Medical organizations and Google can detect and track the emergence of disease outbreaks via social media and web search trends.

Oil and gas companies can take the output of sensors in their drilling equipment to make more efficient and safer drilling decisions.

Big Data is a trend across business and IT, which refers to new technologies that can analyze high-volume, diverse data from traditional and digital sources inside and outside the company. Leveraging Big Data analytics leads to more confident decision making, which means greater operational efficiencies, cost and risk reductions.

Big Data relates to data creation, storage, retrieval and analysis that is remarkable in terms of volume, velocity, and variety:

Volume.

Massive volume of data is contributed by many sources of constantly updated data containing financial, environmental, location and other information: transactions, social media, use of smartphones and internet of things. For example: Facebook ingests 500 terabytes of new data every day; a Boeing 737 generates 240 terabytes of flight data during a single flight.

Variety.

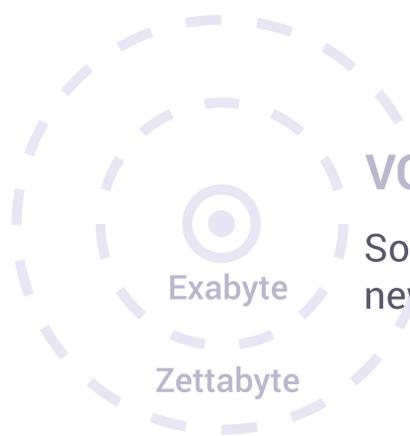
Data today comes in different formats: geospatial data, 3D data, audio and video, and unstructured text, including log files and social media. Managing, merging and analyzing different varieties of data is a challenge for many organizations.

Velocity.

Data is streaming in at exceptional speed and should be timely processed. Clickstreams and ad impressions capture user behavior at millions of events per second; high-frequency stock trading algorithms reflect market changes within microseconds; machine to machine processes exchange data between billions of devices.

What is big data?

Three key differences
between analytics and big data



VOLUME

So big, we need
new metrics



VARIETY

An abundance
of sources



VELOCITY

Real time
data processing

What does this all mean?

It means that globally, companies are turning to big data strategies in order to gain an edge over their competition. They realize that good business decisions are now data driven, and not intuitive-like. They analyze data to better understand and reach their customers, develop new revenue streams and improve operational efficiencies.

Big data adoption grows at different rates in each vertical industry. Such markets as retail, financial services, telecommunications, and media are making considerable investments to effectively use their data to drive value.

The reason behind these verticals being the forerunners is that they have a lot of customers generating plenty of data, and a continuous need to keep customers happy so as not to lose them.

For example, the widespread use of increasingly granular customer data can enable retailers to improve the effectiveness of their marketing and merchandising.

Data analytics applied to supply chains and operations will continue to reduce costs and create value and new competitive advantages for growing retailers' revenue.

BIG DATA IN BANKING

90% of financial institutions in North America think that successful big data initiatives will define them as winners in the future.¹

According to Gartner, big data in the banking industry has the highest level of opportunity due to the high volume and velocity of data available. Globally, Financial Services and Banking is taking the lead in applying progressive big data technologies and data science techniques, followed by Telecommunications and Retail.

There are multiple internal data sources in banks - including relational databases, XML data, Data Warehouses, enterprise applications such as ERP and CRM. Banks also have a large amount of external data about their customers in the form of website visits, tweets, Facebook wall posts, searches, streams, videos, etc. This huge amount of data needs to be stored, processed and analyzed in order to let banks solve real business problems that bank face nowadays.

Typical banking sources of **big data** include:



Customer bank visits



Call logs



Web interactions



Credit card histories



Social media



Transaction type



Banking volumes

Happily, advances in technology - such as processing power, data warehouse storage - as well their reliability make it easier for banks to apply them for solving high-impact business problems. Moreover, it is now data scientists who play a crucial role in applying big data tools and mathematical algorithms to each specific business problem. And it is no doubt, that banks of all sizes, shapes and forms need to incorporate data science into their operating models.

What are bank business problems that data science can actually solve? They are multiple. Application of big data in banking varies from marketing (marketing campaigns efficiency, next best offer, personalized messages), through operational efficiency (ATM cash optimization, capacity planning) to human resources management (predictive recruitment modelling).

¹ Microsoft and Celent, "How Big is Big Data: Big Data Usage and Attitudes among North American Financial Services Firms", March 2013.

Application of data science in banking



At the same time, more than half of big data efforts in financial service companies are focused on achieving a customer-centric view. For example, analysis of big data can help banks better target the right product to the right customer at the right time. By correlating the social activities of a customer with a spending pattern, banks can customize and optimize the timing of their product offerings.

Also, analysis of social media helps banks predict customer churn. As the study by Ernst and Young shows, **63% of customers** in the United States trust online personal networks and communities on choosing various banking products.

Moreover, 45% of customers comment in social media channels on the quality of service they received.

The ability to monitor customer sentiment gives banks early signals - and allows banks to be proactive in improving the customer experience, their engagement with the brand, thus saving costs and preventing revenues loss.

Analytics techniques can also play a significant role in fraud detection - allowing organizations to extract, analyze, and interpret business data to increase the probability of fraud and implement effective fraud detection systems.

All in all, big data opens huge opportunities for banks. According to Ernst & Young predictions, by 2030, banks will deepen their personal connections with customers via data analysis techniques that might seem fantastic by today's standards.²

² E&Y, Building the bank of 2030: top eight global trends.

Case study #1 Fraud detection

BUSINESS PROBLEM

A leading Eastern European bank needed to improve its credit cards fraud detection capabilities. The existing process was detecting and preventing some fraud, but as the size and type of fraud varied and changed correspondingly, the bank required a new, more sophisticated system to provide efficient protection for the bank's customers.

SOLUTION PROVIDED

InData Labs offered to build an intelligent system that would detect **fraud** in real-time. The system applies the data analysis approach to create patterns on each client's behavior based on all of their previous transactions live. Then it enriches this information with third-party data (e.g. social media data, geo-location) - such as geo data taken from the bank's mobile application. Transactions that do not fit into the cardholder's profile are marked as suspicious. Cardholders' profiles are updated with every single new transaction - the system uses self-learning techniques in order to constantly adapt to any changes in cardholders' behavior.

BENEFITS

- Fraudulent transactions are successfully distinguished from legitimate transactions
- Fraud prevented or minimized
- Decreased operating costs thanks to an automated approach

Case study #2 Contact center efficiency optimization

BUSINESS PROBLEM

A mid-size European bank faced a problem of rapidly rising call center operating costs and declining customer satisfaction. The bank's agents often had to make multiple transfers and switches of calls, with sometimes no result for the person who was calling. It became evident that the bank needed to improve the contact center operations.

SOLUTION PROVIDED

In order to solve these problems, InData Labs has proposed to tackle two major issues:

- To reduce overall volume of calls and the number of transfers
- To turn call centers into sales centres - implementing a customer-centric approach

InData Labs has developed a solution that solves both of these issues simultaneously - a system that would actually understand why people call in order to offer the best service in the shortest time possible.

The system provides call center analytics that merges together real-time and historical data to anticipate customer needs ahead of time. With this system the bank's agents are able to learn in advance the reason behind customer calling and solve the problem before the customer even expects it to be solved.

For example, if the client encounters that his card or online banking passcode has been blocked, the system detects it in advance. It predicts that the client will call the contact center very shortly and ask to unblock it - so in real-time the system notifies the contact center operator about the problem thus giving some time to handle it. Thus, early problem identification allows to shorten client's waiting time and enhance client's user experience in addition to providing bank systems with more time to deal with the problem in advance.

BENEFITS

- Reduced operating costs
- Improved cross-sales and up-sales success rates
- Increased customer loyalty

Case study #3 Customer churn analysis

BUSINESS PROBLEM

A large European bank confronted a serious issue - it was losing customers in favor of a competing bank. So its management needed to answer the question - to identify the customers that were most likely to churn and offer them a superb customer service in advance to keep them happy and loyal.

SOLUTION PROVIDED

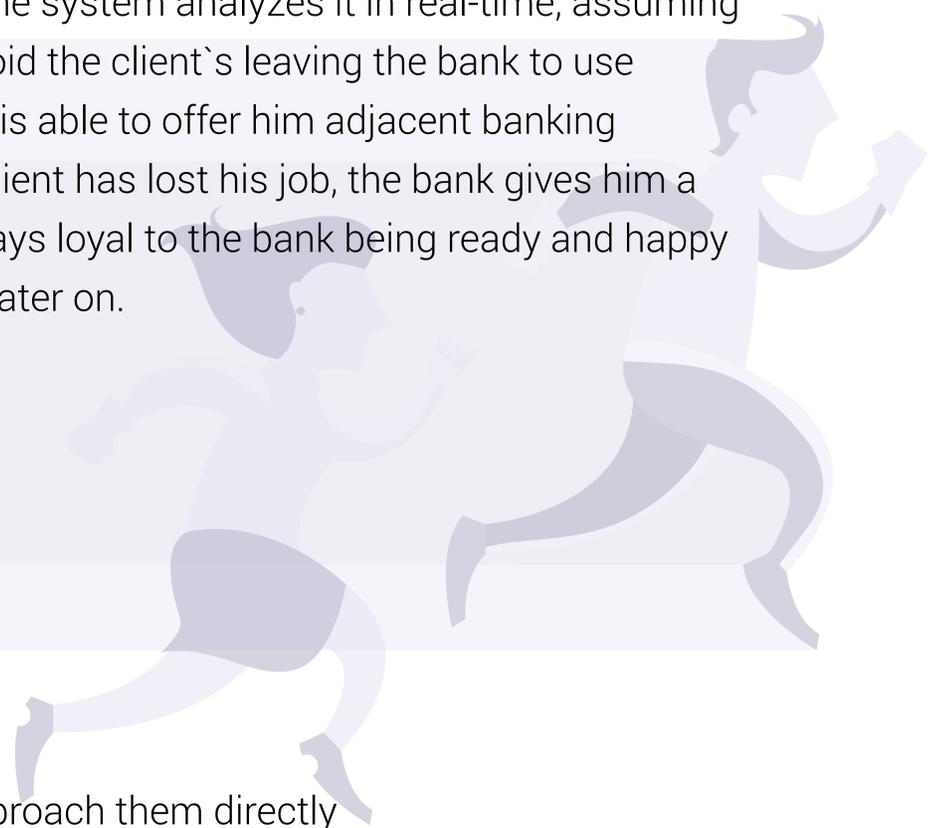
InData Labs has developed a customer churn prevention system delivering analytics on each client on a real-time basis.

In this decision-based system, mathematical algorithms and machine learning techniques are applied on massive amount of historic data available on lost customers. On the basis of this data a churn model is developed. The system identifies very accurate behavioural patterns and apply them successfully to existing customers. Each customer is given a score which measures potential attrition on the basis of as many factors as possible. Moreover, the system models every situation and gives recommendations on the next best action for the customer to prevent churn (e.g., providing competing offers such as higher rates, fewer fees, or even a gift).

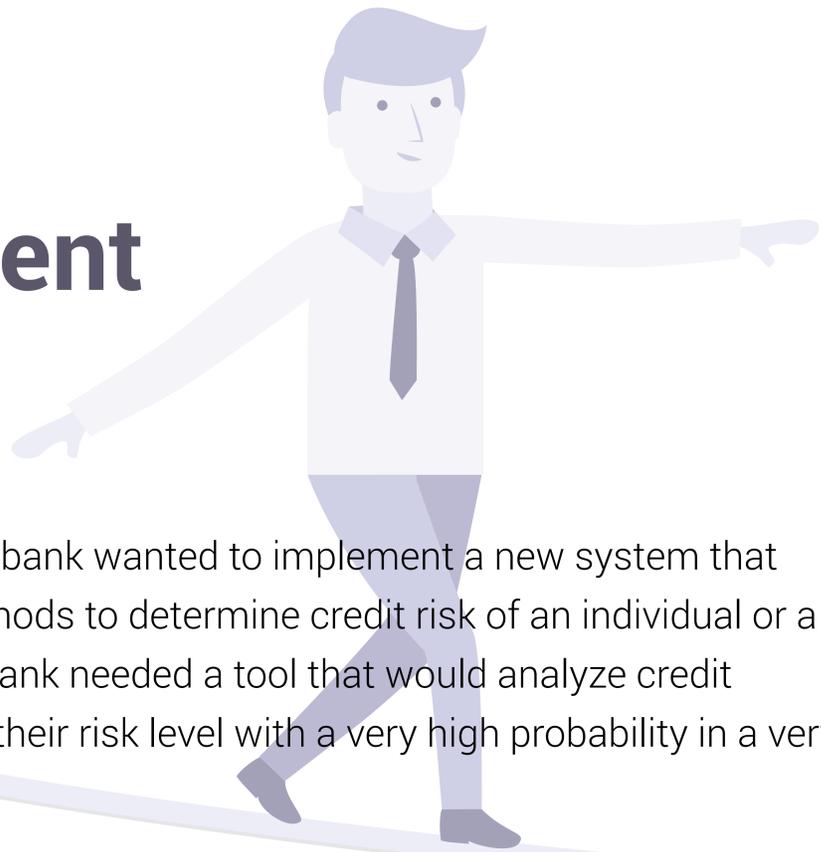
For example, the bank's client used to get salary on a bank account at the certain date of each month, however these transactions stop. The system analyzes it in real-time, assuming that the client has changed or lost the job. To avoid the client's leaving the bank to use his/her new employer's preferred bank, the bank is able to offer him adjacent banking products (e.g. savings account). To verify if the client has lost his job, the bank gives him a courtesy phone call. As a result, the customer stays loyal to the bank being ready and happy to keep using this bank's products and services later on.

BENEFITS

- Upgraded customer service
- Increased customer loyalty
- Reduced attrition rates
- Identification of profitable customers to approach them directly



Case study #4 Risk management



BUSINESS PROBLEM

A large Western European bank wanted to implement a new system that would apply accurate methods to determine credit risk of an individual or a legal entity. Basically the bank needed a tool that would analyze credit applicants and determine their risk level with a very high probability in a very short period of time.

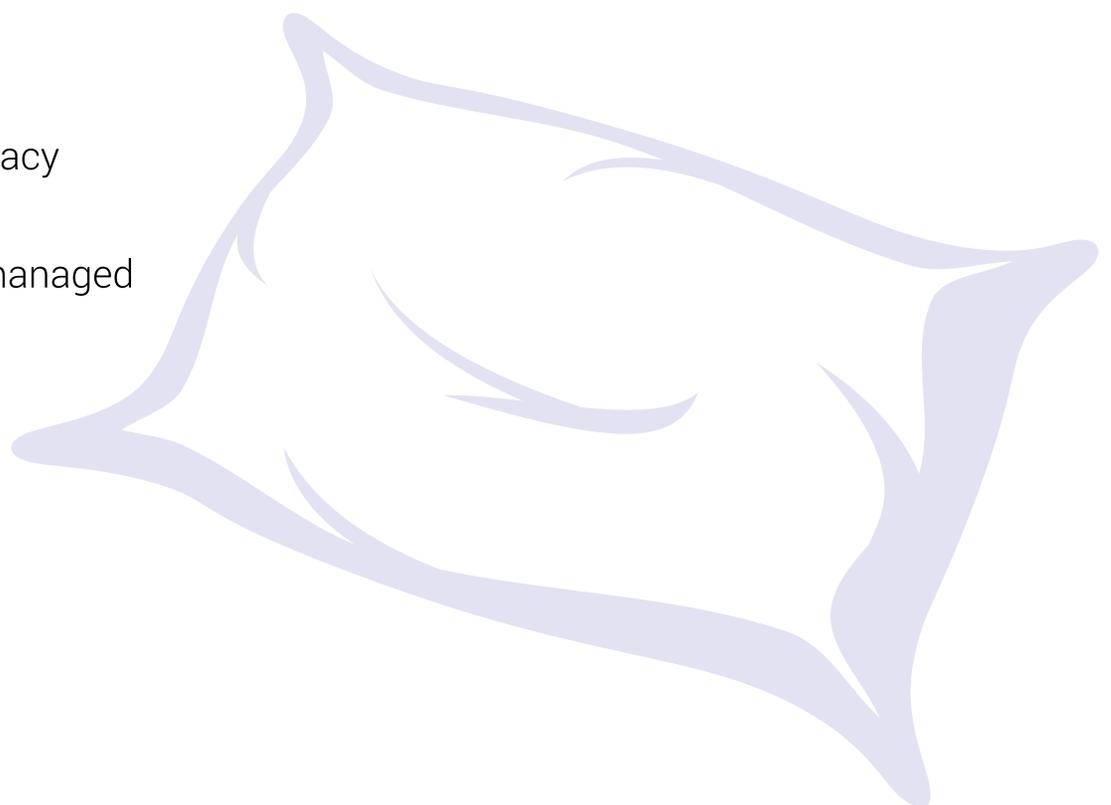
SOLUTION PROVIDED

To help the bank, InData Labs has developed a sophisticated analytical tool that is able to predict loan defaults with a very high probability. The high probability is reached thanks to a high-end predictive model developed by InData Labs' data scientists on the basis of enriched analysis of various data sources mix. In terms of historical data, the system gathers and analyzes internal sources of information in the bank, such as credit reports and applications, repayment rates of credit applicants and any information on default and recovery for borrowers. It also analyzes data from emails, website usage and call centers. This sophisticated analysis is then enriched with data analytics from local credit bureaus and a behavioural information from social media activities and other online sources of information (blogs, Google search, etc.).

As a result, the solution is able to provide the credit score for each applicant almost in real-time and with a very high predictability of risk.

BENEFITS

- Higher credit scoring accuracy
- Improved credit decisions
- Credit risk controlled and managed
- Healthy credit portfolio



Case study #5 Centralized marketing campaigns

BUSINESS PROBLEM

A European bank was losing its market share. It became evident that the bank was no longer meeting customer needs. The bank identified that the major problem was hidden in its ineffective marketing efforts. A seasoned solution was required to solve the following burning problems:

- Ineffective marketing campaigns
- Fragmented marketing strategy across various channels
- Long evaluation of campaigns performance
- Challenges to present marketing campaigns results to senior management in real-time

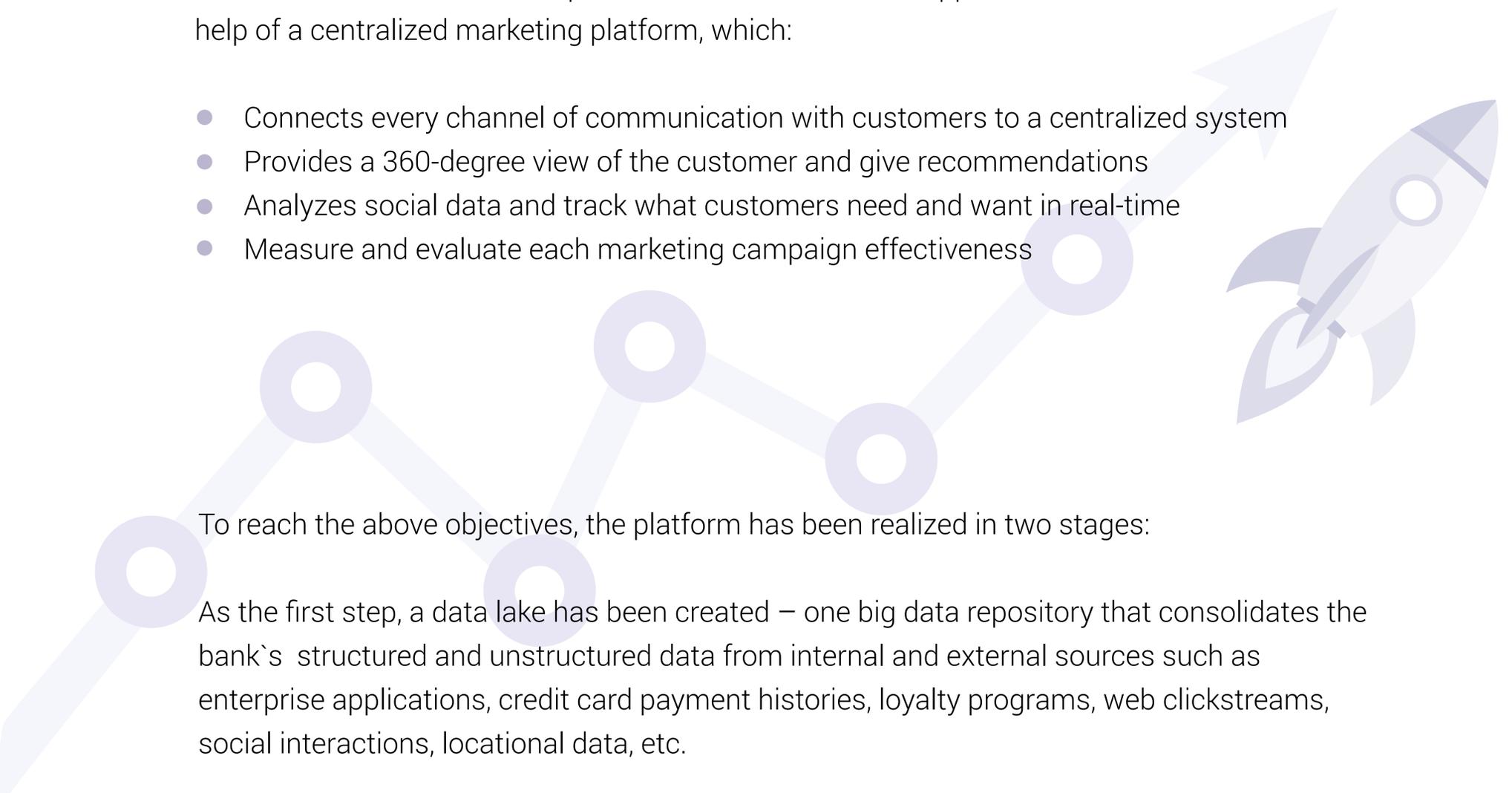
SOLUTION PROVIDED

InData Labs has offered to implement a customer-centric approach within the bank with the help of a centralized marketing platform, which:

- Connects every channel of communication with customers to a centralized system
- Provides a 360-degree view of the customer and give recommendations
- Analyzes social data and track what customers need and want in real-time
- Measure and evaluate each marketing campaign effectiveness

To reach the above objectives, the platform has been realized in two stages:

As the first step, a data lake has been created – one big data repository that consolidates the bank's structured and unstructured data from internal and external sources such as enterprise applications, credit card payment histories, loyalty programs, web clickstreams, social interactions, locational data, etc.



At the second phase, InData Labs has built a consumer intelligence software which applies mathematical algorithms and machine learning techniques to match all data available for one particular customer. Thus, one single enriched profile for every customer is created. This profile gives clear vision on each consumer's preferences to buy certain types of products. The system is also able to provide recommendations on which products to whom at what time might be offered. Now the bank's employees conduct marketing campaigns via all or selected channels and measure effectiveness shortly with the easy-to-make and user-friendly visualized reports.

BENEFITS

- Substantial increase in average campaign response rate
- Centralized marketing strategy across various departments
- Quick and easy reporting to senior management

ABOUT INDATA LABS

Our team delivers high-end engineering services & intelligent data analysis to achieve increased profitability of your business through constant innovation, insightful & data-driven management.

Leveraging the latest big data technologies with a highly professional & talented team of data engineers, statisticians & mathematicians, we help our clients solve high impact business problems in Marketing, Merchandising, Supply Chain, Fraud Detection, Risk Analytics to name just a few areas. Our core industry competencies are FMCG & Retail, Finance & Telecom.



Big Data Strategy Consulting

- Use cases definition & prioritization
- Architecture design
- Road Map elaboration and Strategy report delivery



Engineering

- End-to-end deployment and management of Big Data platform
- Big Data integration services
- System engineering & technical support services



Data Science

- Customer analytics with our internal flagman products - MoneyGraph & Snipe
- Mathematical models & algorithms developed by our R&D Lab
- Predictive analytics solutions

More information about InData Labs services is available on the Web at www.indatalabs.com