



# Governance - Use Cases

do  
data  
GOVERNANCE



IBM

# Governance - Use Cases

A central graphic featuring a hand with the index finger pointing upwards, touching a glowing circular interface. The interface contains the text 'do data GOVERNANCE'. The word 'do' is in a large, dark grey font, 'data' is in a smaller, teal font above it, and 'GOVERNANCE' is in a smaller, dark grey font below it. The background of the interface consists of concentric circles and a hexagonal grid pattern.

do data  
GOVERNANCE

## Northern Trust gears up for GDPR through data governance solutions

Mention data governance to many corporate executives and you might get eye rolling and a quick change of subject. Business leaders are more focused on revenue rather than meeting regulatory demands. But the rise in enterprise data has also brought government-driven demands for privacy and protections that are forcing companies across the world to adopt tools to meet new requirements.

“Governance is a very abstract concept. Most people want to run away from anything close to governance,” said [Sanjay Saxena](#) (pictured), senior vice president of enterprise data governance at Northern Trust Corp.

Saxena visited theCUBE, SiliconANGLE’s mobile livestreaming studio, and answered questions from hosts Dave Vellante ([@dvellante](#)) and James Kobielus ([@jameskobielus](#)) during [IBM Fast Track Your Data](#) in Munich, Germany. They discussed the evolution of a data governance program within Saxena’s company and IBM’s involvement. (\*[Disclosure below](#).)

### Company had little data governance four years ago

Northern Trust, an international financial services company, did not have a significant data governance program four years ago. But market conditions and the expectations of government regulators led the company to develop a robust data control and management program that has slowly gained buy-in from the various department leaders.

“They are accepting the work we are doing, and they want to be a part of it,” Saxena said.

One of the motivations for developing data governance is the General Data Protection Regulation, which requires that companies protect the personal information of European citizens and be able to track where the data flows. The [laws](#) go into effect in 2018, and penalties for non-compliance are steep.

Northern Trust has moved to meet these new requirements by creating one central repository for all metadata and implementing a process to easily identify any sensitive information. Attention to data governance has actually improved company operations, according to Saxena.

“Customer names and addresses being wrong may not be much to a regulator, but it is really important to our business,” he said.

Saxena turned to IBM for help with data governance because it offered an end-to-end suite instead of segregated tools provided by other vendors.

“We’ve been able to make it a single integrated solution, and that’s really benefited us,” Saxena stated.



## Northern Trust gears up for GDPR through data governance solutions

Watch the complete video interview below, and be sure to check out more of SiliconANGLE's and theCUBE's independent editorial coverage of the [IBM Fast Track Your Data event](#). (\*

*Disclosure: TheCUBE is a paid media partner for IBM Fast Track Your Data. Neither IBM nor other sponsors have editorial control over content on theCUBE or SiliconANGLE.*)

### Since you're here ...

... We'd like to tell you about our mission and how you can help us fulfill it. SiliconANGLE Media Inc.'s business model is based on the intrinsic value of the content, not advertising. Unlike many online publications, we don't have a paywall or run banner advertising, because we want to keep our journalism open, without influence or the need to chase traffic.

The journalism, reporting and commentary on SiliconANGLE — along with live, unscripted video from our Silicon Valley studio and globe-trotting video teams at theCUBE — take a lot of hard work, time and money. Keeping the quality high requires the support of sponsors who are aligned with our vision of ad-free journalism content.

If you like the reporting, video interviews and other ad-free content here, please take a moment to check out a sample of the video content supported by our sponsors, tweet your support, and keep coming back to SiliconANGLE.

The Northern Trust Corp. has found regulations mandated by Europe's new General Data Protection Regulations to be valuable guidelines in the company's efforts to ensure clean, quality data, despite the law's potential as a limiting factor.

"We are treating GDPR as a foundation to our data governance program," said Sanjay Saxena (pictured), senior vice president of enterprise data governance at Northern Trust Corp. Saxena specifically recognizes GDPR laws as fundamental in the data quality strategy at his company.

Saxena spoke with Rebecca Knight (@knightrm) and Dave Vellante (@dvellante), co-hosts of theCUBE, SiliconANGLE's mobile livestreaming studio, during the [IBM Chief Data Officer Summit](#) event in Boston, Massachusetts. They discussed Saxena's tools and methods for maintaining data under GDPR, as well as the benefits of highly regulated information for customers and businesses alike. (\* *Disclosure below.*)

As data volume increases and creates issues around discoverability for companies, Saxena is taking an approach rooted in the highest level of protection to provide the best customer experience. "There's a huge amount of synergy between GDPR and information security. ... We are in a systematic fashion trying to figure out where all our sensitive data is and whether it is controlled, protected, etc. ... All these things come together very nicely from a GDPR perspective," he said.

### Clean data for stronger customer relationships

GDPR provides a guideline for better serving customers when the volume and categories of data are virtually limitless, according to Saxena. “What we are doing in GDPR is mapping out sensitive data across hundreds of applications and creating that baseline for the whole company so that any time a regulator comes in and wants to know where a particular person’s information is, we should be able to tell them in no uncertain terms,” he said.

In addition to developing client relationships, regulated data enables more effective marketing for businesses. “Now you have that information tagged, it’s all nicely calibrated in repositories ... you can use that for your analytics, you can use that for your topline growth, or even see what your internal processes are that can make you more effective from an operations perspective,” Saxena said.

With an amount of structured and unstructured data that feels incomprehensible, companies like IBM are proving valuable by offering tools to help organize information, according to Saxena. “There are ... tools available in the marketplace, including IBM’s tools, which help you map the data. ... There are other tools that help you develop a metadata repository,” he stated.

As data becomes increasingly valuable, and quality harder to come by, Northern Trust Corp. is turning its attention to recruitment and talent development to maintain productivity. “It’s hard to recruit in these areas. ... We get interns ... who have the technology knowledge, and we couple them up with business experts to work in collaborative teams,” Saxena said.

These efforts coupled with a commitment to strong customer relationships has Saxena confident in his company’s ability to keep pace with data innovations. “It’s still evolving from my perspective. ... Have we attained a state of perfection? No. We are getting there in terms of more optimization, more emphasis, and more money and financials being put on data quality,” Saxena concluded.

Watch the complete video interview below, and be sure to check out more of SiliconANGLE’s and theCUBE’s coverage of [IBM Chief Data Officer Summit](#). (*\* Disclosure: TheCUBE is a paid media partner for the IBM Chief Data Officer Summit. Neither IBM, the event sponsor, nor other sponsors have editorial control over content on theCUBE or SiliconANGLE.*)

### Since you’re here ...

... We’d like to tell you about our mission and how you can help us fulfill it. SiliconANGLE Media Inc.’s business model is based on the intrinsic value of the content, not advertising. Unlike many online publications, we don’t have a paywall or run banner advertising, because we want to keep our journalism open, without influence or the need to chase traffic.

The journalism, reporting and commentary on SiliconANGLE — along with live, unscripted video from our Silicon Valley studio and globe-trotting video teams at theCUBE — take a lot of hard work, time and money. Keeping the quality high requires the support of sponsors who are aligned with our vision of ad-free journalism content.

If you like the reporting, video interviews and other ad-free content here, please take a moment to check out a sample of the video content supported by our sponsors, tweet your support, and keep coming back to SiliconANGLE.

### Business challenge story

As the Internet of Things (IoT) becomes an increasingly fundamental technology for effective government, cities are embedding sensors in everything from buses and garbage trucks to water systems and public buildings. By analyzing data from those sensors, modern cities will be able to design smarter public services, make wiser policy decisions, and manage day-to-day operations much more efficiently.

Yet even without any investment in sensor networks, today's cities already contain millions of the most intelligent and versatile "sensors" that have ever existed: human beings. A public-spirited citizen with a smartphone is an incredibly valuable source of data for government agencies, because they will provide accurate feedback on the status of the city's systems in real time.

The only problem is collecting and analyzing the data fast enough. In Jakarta, a district of 10 million people is divided into five cities, 44 sub-districts, and 267 villages. The city government receives an average of 1,400 messages per day via its custom-built Qlue mobile app, which allows users to submit feedback about public services. On top of this, citizens send an average of 130 SMS messages per day to the governor's mobile phone, and many more via other channels such as email and Twitter.

"At the same time, we knew that this was just one of many projects that will involve big data analytics. So, we decided to build a general-purpose big data platform that we could use to capture, store and analyze huge volumes of data for any use-case."

### Transformation story

JSC knew that it needed a platform that could handle every aspect of big data collection, management and analytics. It needs to be able to ingest a huge variety of different types of data—everything from unstructured text on social media to JSON data from the web and relational data from traditional database systems. It also needs to deal with cultural change, as some government departments are only just beginning to move from paper-based processes to digital workflows.

"Data governance is very important," says Diory Paulus. "The danger with any analytics project is that if you put garbage in, you will get garbage out. And this is especially difficult to guard against when the datasets are so large."

Diory Paulus, Head of Data & Analytics at Jakarta Smart City (JSC), a management unit under the Communication, Informatics and Statistics division of the Jakarta Provincial Government, explains: "It's important to listen and be transparent when citizens send feedback about public services—but in a city as large as Jakarta, the sheer number of messages makes it impossible to respond fast enough if you are handling every message manually. We wanted to find a way to process feedback more quickly and analyze and prioritize the most important issues."

He continues: “We looked at the Gartner Magic Quadrant and asked several leading companies to show us their products. IBM was one of the few vendors that could answer all our questions and had a truly comprehensive vision, embracing big data collection, data warehousing, advanced analytics and governance.”

IBM® Global Business Services® helped the JSC team design a solution that would provide a big data hub for integrating information from the citizen feedback app and social networks, as well as government services such as transportation, healthcare, water distribution and other departments.

The solution uses IBM InfoSphere® DataStage® to extract, transform and load data from all these sources into a central data lake, built on IBM BigInsights®, and a powerful data warehouse, which runs on IBM PureData® System for Analytics. IBM InfoSphere Information Governance Catalog provides a robust data governance framework, making it easier to align the technical systems with the business requirements and processes. As analysis tools, the solution provides IBM Cognos® Analytics for reporting and dashboarding, IBM SPSS® Modeler for advanced analytics and predictive modeling, and IBM Text Analytics for categorization and sentiment analysis of unstructured text.

“We are taking our first steps with many of these technologies, and the guidance and training we have received from IBM has been very valuable,” says Diory Paulus. “We are very keen to learn and share our knowledge with other government organizations, both within Indonesia and internationally.”

## Results story

JSC has already started using the big data platform to streamline the way it handles citizens’ feedback and provide new insight into the most important topics.

Diory Paulus says: “There is a real opportunity to use data to make smarter policy decisions. For example, we can analyze all the feedback we receive, and identify patterns. Recently we looked at the top ten villages, ranked by the number of complaints. We saw that one of our villages had the highest number of complaints for two consecutive months, and when we looked into it, we found that most of the problems were related to garbage collection.

“We plotted all the incidents on a heatmap and overlaid the route of the garbage trucks, which showed us that the most complaints came from some areas where the garbage trucks don’t drive through. We were then able to work with the village council and the Jakarta waste management department to improve the routing and scheduling of the garbage collections—and sure enough, the number of complaints is now decreasing again.



“This is just one example of the power of big data, once you can collect and analyze it effectively. Traffic management is another key area—we can integrate traffic complaints data with data from Waze and the GPS systems in our buses, and get new insight into where the bottlenecks are.”

JSC is also investigating ways to use analytics to combat fraud by looking into utilization of the smart cards that citizens use for subsidized school fees and other services. The city is in the process of launching a new “Jakarta One” card, which will integrate a much wider range of payment services. JSC expects that rapid fraud detection capabilities will be increasingly important to ensure that fraudsters cannot take advantage of the system.

Diory Paulus concludes: “With the IBM platform, we are now able to say ‘yes’ whenever a government department asks us for help with data and analytics. We know we can collect and store however much data we need, and we have the analytical power to process it and deliver results quickly.

“We are excited about the next steps on this journey—for example, using Text Analytics to make it easier to categorize and prioritize the messages we receive by email and SMS, or on Twitter. This should save many hours of manual processing, and help us ensure that we always deal with the most important issues as quickly as possible.

“Most important of all, analytics helps us show our citizens that their feedback really makes a difference. By showing the people of Jakarta that we are listening, we can keep them engaged in helping us build a city that is better for everyone.”

## Business challenge story

For the Dutch health insurance company CZ, a fast time-to-market is crucial when testing and launching new software applications for use in its business—CZ not only needs to be able to respond to customer requirements swiftly, but must also meet regulatory deadlines.

For example, each year, insurance companies in the Netherlands need to update their systems to align with changes in the country's healthcare system. The deadline for completing these changes is "Prinsjesdag"—the day when the reigning Dutch monarch announces the main features of government policy for the coming parliamentary session—and there are significant penalties for any insurer who fails to complete their changes on time.

"This means our testing environment must work as smoothly and automatically as possible to ensure we can launch applications on schedule. For instance, copying and loading data from our production environment into our testing environment needs to be quick and easy, and it's also important that the test data is referentially intact, to ensure proper application testing."

However, to comply with data protection regulations, CZ must also ensure that it protects clients' privacy as it moves data between its production and non-production environments. These requirements will soon become even more strict, when the upcoming European General Data Protection Regulation (GDPR) comes into force.

Ed de Cock, Test Coordinator at CZ, takes up the story: "The pressure is high. Sometimes a new application has to be operational in just a month, while updates of existing applications need to be released every four weeks or even every other day.

Han van der Vinden, Test Manager at CZ, explains: "The regulations state that we must not use real customer data for development and testing purposes, to ensure that our clients' privacy isn't threatened in the case of lost or hacked data containing details such as addresses or bank details. It's not just about complying with regulations—we also need to protect the company's reputation and ensure that our customers trust us as stewards of their personal information.

"However, we still need to use realistic data to test our applications fully. A solution is to mask our production data—that is, replace any sensitive personal information in the dataset with 'dummy' information that isn't related to any real customers—before we move it over to the test environment."

CZ had previously developed its own tool for masking data, but this no longer met privacy legislation standards.

Ed de Cock adds: “Our homegrown tool didn’t anonymize data sufficiently to meet new regulations, and also left room for errors. For example, if you added a new table to your production database, you would also need to remember to add a new table to your testing environment—but because this wasn’t automated, there was a chance of it being forgotten. That could lead to errors and inconsistencies in the testing process.

“What’s more, indicating which dataset you wanted to load into the testing environment was very time-consuming. For instance, if you wanted to select people with certain characteristics, you first needed to run a query to extract that data from the production database, and then load it manually in the testing environment.”

CZ knew it was time to find a new automated approach to test data management that was secure, accurate, and would enable software engineering teams to work at a faster pace.

## Results story

With IBM InfoSphere® Optim™ Test Data Management in place, CZ can easily protect its clients’ data, comply with privacy regulations, and meet deadlines for changing regulations—while cutting costs and boosting efficiency.

“Optim’s subsetting feature is a huge advantage, because it allows us to reduce the amount of data that is copied to our test environments,” says Han van der Vinden. “This reduces the time taken to create copies by around 50 percent, which contributes to faster development cycles overall. It also makes it easier to spot changes in the database structure and keep the test database consistent with the development and production environments.”

Additionally, the solution’s automation features minimize errors and boost efficiency, as Ed de Cock describes: “Optim helps us select and load data more quickly, apply changes consistently across the testing environment, and ensure that all relevant items are copied when creating test datasets—so we can rely on our test results. The next step is to be able to perform automated tests quickly and accurately. We aim to achieve a 10 percent reduction in overall development cycle time within a year, helping us get new releases into production more quickly.

“By speeding up our software development processes, the solution makes it easier for us to respond swiftly to customer requirements, and meet regulatory deadlines for implementing changes in healthcare policy.”

At the same time as boosting efficiency and accelerating software development, the solution also cuts expenditure. Specifically, by enabling the use of smaller datasets in the test environment, Optim reduces total storage requirements and costs.

Flexibility is another key benefit: whenever new masking requirements arise, CZ can write the specific routines it needs and integrate them seamlessly into its testing processes. This may prove vital in helping CZ comply with new data privacy regulations such as the GDPR.

“The IBM solution empowers us to keep our clients’ personal data safe, protecting the company’s reputation and preserving our customers’ trust,” concludes Ed de Cock. “It also cuts costs and accelerates software development processes, boosting efficiency and enabling us to meet regulatory deadlines.”

## Business challenge story



The Food Supplies and Consumer Welfare Department of the Government of Odisha offers a lifeline for some of the state’s most vulnerable citizens; its subsidies help to ensure that people living below the poverty line have access to adequate food and other household necessities.

With hundreds of thousands of cases to manage every year, the Department found it challenging to ensure that its benefits were being directed to citizens who were truly in need.

A spokesperson from the Government of Odisha explains: “Most of our benefits claimants truly need our support to make ends meet, but unfortunately there are always individuals who will try to abuse the system. For instance, some people would under-report their income and claim benefits that they were not really eligible for, while others would use multiple ID cards to claim benefits more than once.

“These kinds of fraud divert valuable government resources away from the citizens who needed them the most—and we wanted to put a stop to it. The problem was that we lacked an efficient way of confirming claimants’ eligibility for welfare assistance.

“In the past, our best option was to hire external investigators to visit claimants and check on their circumstances, but this was too expensive and time-consuming to apply on a larger scale. We realized that if we wanted to stamp out benefits fraud, we needed to find a smarter way to work. And that’s where IBM came in.”

## Transformation story

In a first-of-its-kind project in India’s government sector, the Food Supplies and Consumer Welfare Department used a suite of IBM Information Management solutions to unify citizen data from multiple systems of record, and analyze it to determine which citizens were making wrongful claims.

Over the course of three years, the project team gathered approximately 45 million records from dozens of government offices, including files such as driver’s licenses, electricity bills, income tax returns and pension statements. IBM® InfoSphere® DataStage was used to build reliable, repeatable processes to extract, transform and load this data from the source systems, while IBM InfoSphere QualityStage helped to identify and remediate data quality issues with individual records.



Once all the data has been brought together and stored in a central IBM DB2® database, the solution uses IBM InfoSphere Master Data Management Standard Edition to compare the records from each system, use a probabilistic matching engine to identify records that belong to the same person, and link them to a single “master record” for each citizen.

“The IBM solution provides sophisticated matching algorithms that help us create a single view of each citizen’s income and property, and compare this against our welfare claimant data to identify potential cases of fraud,” recalls a spokesperson. “For instance, if the records show that an individual owns a car or a house, receives a large pension or pays at least a modest amount of income tax, then it means they aren’t living below the poverty line and are not eligible to claim benefits.”

Looking beyond the welfare system, better information management has the potential to enhance citizen services on a much broader level. With a standardized, centralized store of citizen data, other government departments stand to gain a much more accurate and comprehensive view of citizens, which they can use to drive more efficient and responsive citizen-centric services.

## Results story

With the new solutions in place, the Food Supplies and Consumer Welfare Department has a much faster, accurate and effective way of identifying fraudulent welfare claims.

“IBM Analytics have completely changed the way that we work,” says a spokesperson. “Today, we can catch fraudsters almost immediately. With the new system, 60 percent of records that previously would have required investigation are flagged up automatically, which means there’s much less need for us to hire investigators to go home-to-home to check up on claimants. This is delivering big administrative cost savings, which we can re-direct to other initiatives.”

The solution also helps to improve ongoing fraud detection, as a spokesperson explains: “When someone attempts to register for this welfare program, the system automatically checks their records and, if they do not meet the eligibility criteria, their application will be rejected. By taking a proactive approach, and stopping fraudsters before they even get into the system, we are saving significant effort and resources.”

Most importantly, the new approach is helping to ensure that the right benefits go to the right people.

A spokesperson concludes: “Since introducing IBM Analytics solutions, we have detected around 500,000 wrongful welfare claimants and have removed them from our program. Now, the benefits that they were unjustly receiving can be used to support genuinely needy people instead. We are ensuring that deserving citizens get the help they need, while keeping tighter control over welfare costs and making better use of taxpayer money—it’s a win-win situation for everyone.

“This project is the first of its kind in the Indian government sector, and we are proud to be leading the way and showing what is possible with a modern approach to master data management. We look forward to helping our colleagues in other departments and from other states learn how they can benefit from these technologies too.”