# 5G Analytics: Preparing for the Data Explosion

By Bob Bender, Director of Industry Solutions, Teradata
04.21 / ANALYTICS / VANTAGE

teradata.

## Table of Contents

## Manage Massive Data Volumes and Create a Scalable Approach for 5G Analytics

Companies have been going through a digital transformation for the last decade. With the advent of 5G communications, we are about to see digital transformation shift into hyperdrive due to the massive amount of data being produced with the blazing fast speed of 5G and all the new technology it enables. The number of devices that will be online is expected to grow by 100 times the number of devices online in 2021. The amount of data generated by these devices is expected to grow more than 1,000 times.

Managing this data for maximum insights requires four essential components:

• Data governance

• Data integration

• Data architecture

• Analytics running in-database, at scale

All of these components have been around for years. But let's face it, nobody is doing all of them very well. The funny thing about data governance, integration, and architecture is they live by the "pay me now or pay me later" rule. Somebody eventually has to do them.

teradata.

**Being a Best-in-Class Telco Isn't Enough**

Many telcos think they're on the right path for their business by aiming to be best-in-class. But it's not enough. To succeed, they must build for their future, starting today. This involves using data the right way to turn the telco into a self-perpetuating, demand-generating machine.

Telcos of the future will offer real-time service personalization for a customer segment of one, deliver seamless experiences, offer proactive issue resolution, enable frictionless channel integration, and more. These telcos will process trillions of interactions per month and more than 100 million queries per day. They'll also analyze billions or trillions of records each day from all areas of the business to get maximum insights. **See how.**

## Take an Intelligent Approach to Building Core Components

Companies can spend a little more money and time upfront building core components one time with standards and the correct structure, or they can throw caution to the wind and let individual users perform these functions as they see fit. Let's drill down into each one in detail.

### Data Governance

Data governance sounds like a bunch of bureaucracy with committees and rules. But it can be done almost passively by investing time and developing a plan for your data. A lot of the data coming through 5G devices will be new, so now is a good time to plan how you will handle it.
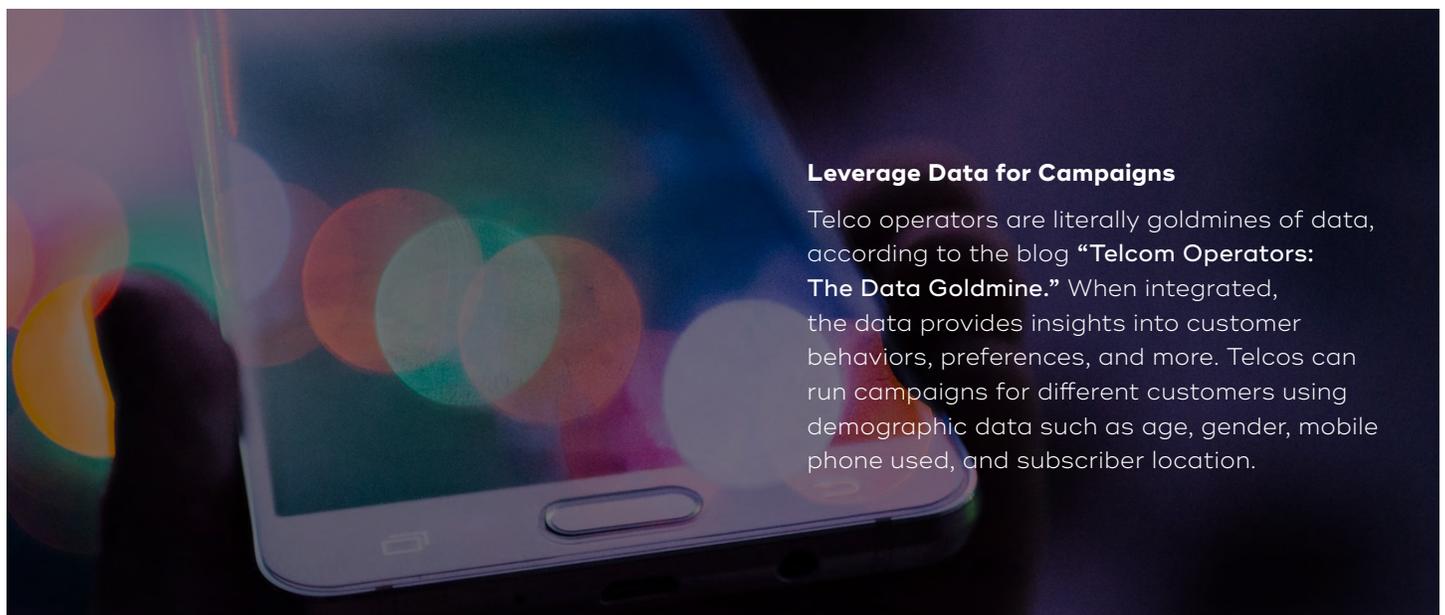
There are five dimensions to data governance:

1. **Strategic Dimension.** How will this data be used to add value to the company?

2. **Business Capabilities.** What business capabilities does this data provide?

3. **Tactical Dimension.** Who will be using it, what type of metadata is needed, and what type of access will users need?

4. **Data Protection.** What level of security will be required, and will it contain privacy elements?

5. **Data Integrity.** What data cleansing or quality checks must be done, and how current must the data be for the users?

Companies should also define the guiding principles of the data. These principles provide a foundation to build the five dimensions. Figure 1 shows an example of guiding principles that telcos and others have used successfully.

teradata.

**Figure 1:** Guiding Principles for 5 Dimensions of Data Governance

### Data Sourcing

- Obtain the data from the true source
- Get the lowest grain of usable data available
- Capture data as frequently as can be justified by the business case and/or technology
- The Analytic Platform (AP) does not fix source system data issues, it raises awareness of the issues

### Governance and Quality

- The AP provides one version of the truth–have a single logical instance of data
- Time should be spent on analytics over engineering
- Leverage current technology capabilities
- Verify every data movement

### Development and Maintain

- Clear, consistent, and easily monitored data flows
- Quick and rapid development and deployment of analytic solutions
- Keep it simple–avoid complexity whenever possible
- Develop and maintain the AP with standards
- Build with the future in mind
- Build with toolset automation where possible
- Build with self-healing and self-cleanup processes in place

### Access/Use

- Promote a self-service information access model
- All non-ETL/DBA data (user) access will be available via the Semantic layer

### Communicate

- Raise awareness of AP capabilities and offerings
- Continuously align analytic goals with corporate goals and strategies

**Leverage Data for Campaigns**

Telco operators are literally goldmines of data, according to the blog **"Telcom Operators: The Data Goldmine."** When integrated, the data provides insights into customer behaviors, preferences, and more. Telcos can run campaigns for different customers using demographic data such as age, gender, mobile phone used, and subscriber location.

teradata.

## Data Integration

Data integration requires discipline and modeling skills that most people don't have in today's IT world. But it's something that must be done at some point in the analytic process. You might as well do it one time upfront by a trained data modeler instead of requiring users to do it every time they write a query. One problem with letting users do it is nine times out of 10, they will get it wrong. This causes the query to run 10 times longer than it needs to or even worse, users get a bad answer that looks correct but isn't.

Some data scientists like to have all of their data in a gigantic flat file that is not normalized. This is usually because of the tool they're using. Some tools cannot efficiently perform joins between two or more tables. So, the flat file provides data integration, but it also has a lot of nulls, "0" fields, and redundant data that takes up space. Or worse yet, this makes the flat file disjointed, meaning even though each record is flat, some columns of data in the record do not line up with data in other columns. This is dangerous because you end up with the wrong information being used for reporting, and analytics are run against data that crosses over the disjointed boundary.

Today's modern data warehouse stores a lot of data, but there are times when data outside of the data warehouse is needed to complete the query or analytic function. The modern data warehouse has tools to bring data together when needed. Gone are the days where all data is moved into a data warehouse staging area, transformed into an integrated, physical third-normal form database, and then views are created for accessing the data the way users want.

In a modern data analytic platform, the analytic process comes to the data through tools that provide access to data files outside the database from a number of data storage systems. What is the same, however, is the logical data model that performs the joins. The model must be developed by a trained data architect or analyst.

This is similar to views but better defined for analytics. The logical data structure provides data scientists with the data they need for the analytic functions being used. What they must be careful about is where they start the query and where the majority of the data sits. They do not want to start a query from a platform that cannot handle the processing requirements or the data volume. The tools that enable all of this data gathering need to be smart and perform the proper procedures in the proper place.

Telecommunications operators need a database that can handle the gigantic amounts of data—petabytes and billions of records at a time—that is produced by 5G networks. The database must perform well with structured joins between multiple tables. This is especially true when joining network data like cell logs with customer data and device data.

### Key to Data-Driven Companies

Companies that treat data as their greatest asset to drive the business tear down their silos and integrate all of their data. This requires discipline, including governing the data and establishing a modern architecture.

As **Randy Bean** wrote in his article "Why Is It So Hard to Become a Data-Driven Company?" for Harvard Business Review,

*"Data-driven business transformation is a long-term process that requires patience and fortitude. Investments in data governance, data literacy, programs that build awareness of the value and impact of data within an organization, may represent an eventual step in the right direction, but organizations must show that they are in it for the long haul and stick with these investments and not lose patience or abandon efforts when results are not immediately forthcoming."*

**teradata.**
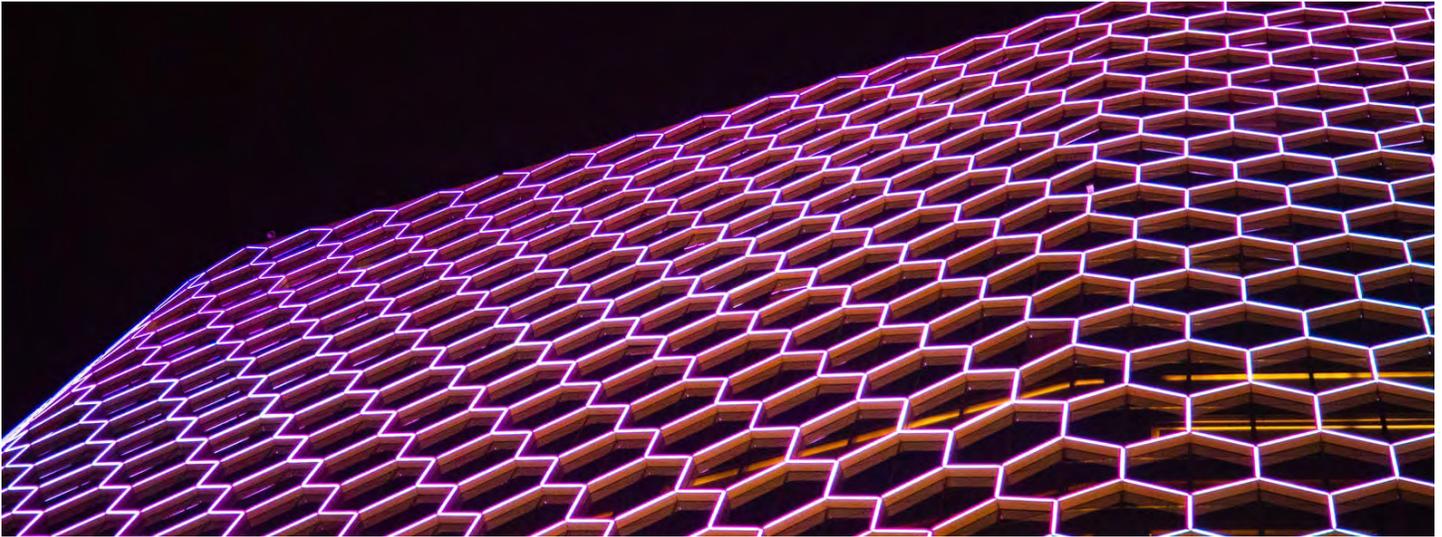
## Data Architecture

Data architecture is about knowing where the data is created and how it will be ingested, transformed, and used. It also documents any other data subject area that needs to be used with the data, and where it is created and ingested.

If you answered the questions for data governance, then this task is relatively easy. Where things can go awry is choosing the wrong technology for the tasks. Generally speaking, IT architects like to recommend the latest and greatest tool. However, the way technology rapidly changes in today's world, the latest and greatest may only be around for a year or two before something presumably better comes along to replace it.

The other issue with architects is they often make processes more complex by inserting tools that only perform one or two tasks. A project then requires several different tools. A better approach is to think holistically and use an analytic platform that can perform many processes and tasks in one place while providing an integrated platform for analytics.

Data architecture also considers where the data sits. If your company is like most others today, you have some systems that sit in a data center that is owned by the company, and some data that has been moved to the cloud. This creates a problem for some analytics platforms.

Unlike an operational system that automates one or more functions for a business process and creates transactions or handles events for that process, an analytic platform is used across many business processes and more often than not, requires data from two or more business processes to complete an analytical function. To make this distribution of data manageable, the analytic platform must be able to live in both the data center and in the cloud or multiple clouds. In other words, a connected hybrid multi-cloud platform is needed.

**Fastest, Lowest Cost Path to Cloud Modernization**

As a cloud-first company, Teradata provides the fastest path to cloud modernization at the lowest cost for telcos and other companies. Teradata delivers the leading multi-cloud data platform, Teradata Vantage™, from start to scale on Azure, AWS, and Google Cloud.

Teradata modernizes data pipelines and ensures native integration with first-party cloud services. Telcos benefit from scalability, flexibility, and unlimited intelligence to build the future of their business.

teradata.

## Analytic Functions

This brings us to the fourth item on the "must-haves list," which is analytic functions running in-database. The same premise for data architecture of avoiding using several different tools also applies here. With analytic functions, one thing you do not want to do is move data around to multiple tools. Instead, you need tools that can access the data directly from the database or, preferably, actually run in the database.

The volume of data coming from 5G devices is enormous and in some cases requires very low latency for an analytic response. The less you have to move the data, the better the latency will be. That said, you must have a database that can scale to meet large volumes of data and meet the latency requirements. That is a very short list!

Analytic functions come in all sorts of shapes and sizes:

- Statistical analysis
- Path and pattern analysis
- Data preparation and transformation
- Cluster analysis
- Time series analysis
- Predictive modeling
- Text analysis
- Geospatial analysis
- Association and recommendation engines

Once again, there are many tools that provide one or two of these functions. In each case, you have to move the data to the function. That's why it's very important to choose an analytic platform that has these functions built in or at least allows the functions to directly access the data in-database. This saves a lot of time and provides consistency with your analytics.

## Capture and Analyze 5G Data at Scale

As 5G produces more data at a faster rate than ever before, telcos need the ability to capture and analyze that data at scale. Investing resources upfront to create standards and a seamless structure that enables data governance, data integration, a modern data architecture that can access data regardless of where it resides, along with the ability to run analytics in-database at scale, will enable telcos to meet the current and future needs for 5G analytics.

**Get a Strong ROI on 5G Investments**

For a look at the data coming out of the new 5G architecture and how telcos can use it to achieve a strong return on investment for 5G, read a white paper by Nadine Manjaro titled "Leveraging the Data Explosion." **View here.**
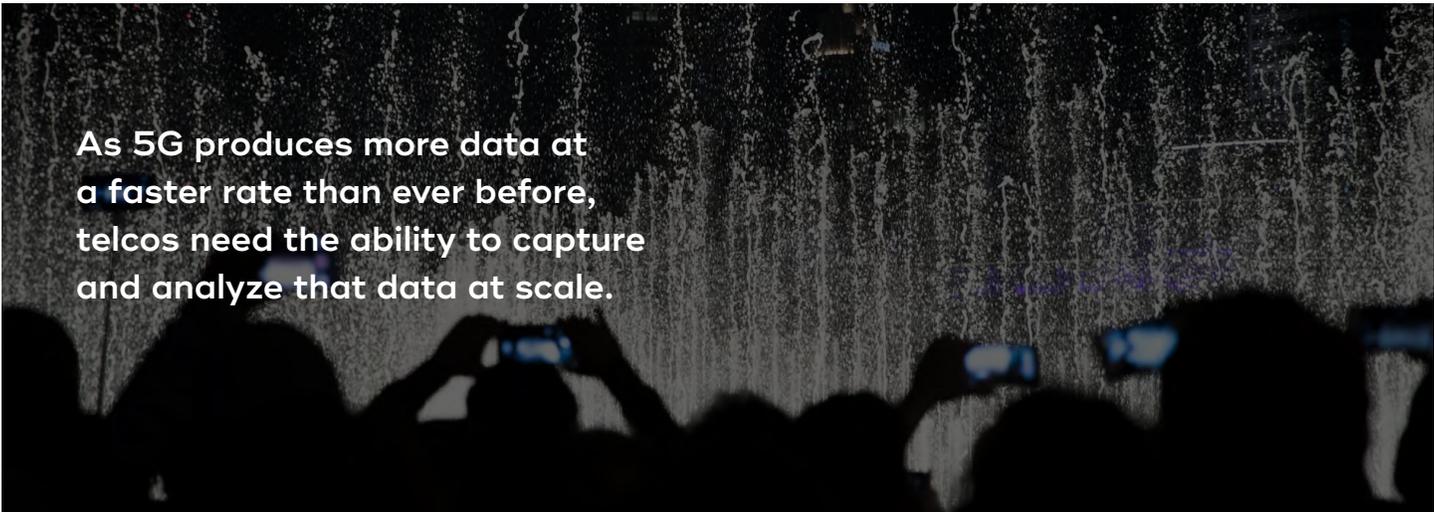
As 5G produces more data at a faster rate than ever before, telcos need the ability to capture and analyze that data at scale.

teradata.

## About Teradata

Teradata is the connected multi-cloud data platform for enterprise analytics, solving data challenges from start to scale. Only Teradata gives you the flexibility to handle the massive and mixed data workloads of the future, today. Our open approach embraces the modern ecosystem to create a seamless experience for ingestion, exploration, development, and operationalization. Teradata's experts and partners around the world can show you how to drive business outcomes and unlock unlimited value by turning data into your greatest asset. Learn more at **Teradata.com.**

### About the Author

**Bob Bender** is the Director of Industry Solutions at Teradata where he specializes in telecommunications. He has more than 37 years of computer and analytics industry experience, with 32 of those years in the telco industry. At Teradata, he specializes in business value assessments, analytic roadmaps, analytic environment architecture, and developing use cases for customers. He has led several analytic teams in developing analytic solutions for network, finance, retail, customer care, and human resources in today's Tier 1 telco companies.

teradata.