

THE RISE OF

The Disruptive Data Warehouse

TERADATA®

CONTENTS

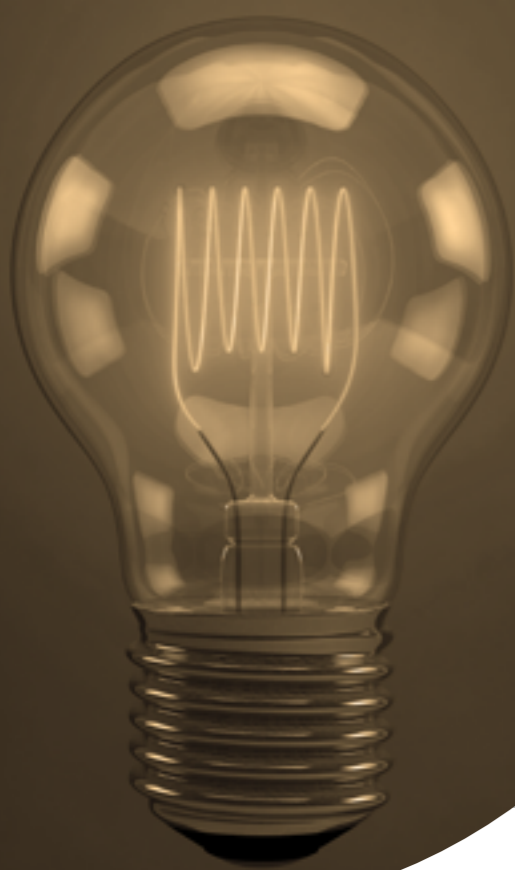
What Is the Disruptive Data Warehouse?	1
Old School	New School
Query a single database	Query multiple databases and platforms 2
The data warehouse is for business intelligence	The data warehouse is for everyone 3
The data warehouse is based on SQL	SQL? Absolutely. And Ruby, Python, Perl, R, Java and C++ 4
The data warehouse handles structured data	Yes—and JSON, XML, weblogs, and more 5
External analytics	Integrated and parallelized in-database analytics 6
Disk data storage	Hybrid storage with solid-state drives 7
Memory for standard caching	Intelligent memory: in-memory dynamic and automatic 8
Row-based data storage only	Hybrid row/column data storage 9
One choice: On-premises, dedicated system	Many choices: On-premises, in cloud, hybrid, hosted, mix 10
Always in the present	Time travel is possible 11
Large project, long lead times	Agile, self-service, untested data, incremental, phased delivery 12
Single Platform to Solve All Problems	Logical Data Warehouse 13
The Best of Both Worlds: The Disruptive Data Warehouse	14
Leading The Path To Technology Innovation	15

What Is the Disruptive Data Warehouse?

The data warehouse, as most people think of it, is passé. Associated with batch reporting, SQL only, long wait times—that's your father's data warehouse, and it shares little resemblance to today's data warehouse.

The data warehouse has been undergoing a continuous evolution. It integrates the best existing technologies and makes them easy to use, while embracing new technologies in an intelligent, cost-effective manner. It gives more business users access to your data. And it lowers the barrier to entry, putting innovation within any organization's reach. It is a disruptive data warehouse.

Join us as we take a tour through the disruptive data warehouse and show how it embraces the latest ideas, trends, and technologies in service of the business.





Old School: Query a single database

Not so long ago, organizations stored all of their data in a single database. This was the data warehouse that users and applications queried. It was okay because it handled their transaction data and processing needs.

New School: Query multiple databases and platforms

Today, organizations have different types of data sitting in various repositories and databases. The disruptive data warehouse has evolved to query these different repositories and perform various types of analytics. But it isn't just querying multiple systems and aggregating the results. The disruptive data warehouse automatically integrates processing across all the various engines. The processing is pushed down into the various repositories, and the results come back to one place where the processing is completed and the data is served to users at scale.

Bottom line: Data and the analytic power of multiple technologies can be applied to answer a business question, without the need to move data around. Multiple analytic engines can be invoked via the disruptive data warehouse—your mission control for analytical power.



Old School: The data warehouse is for business intelligence

The old school data warehouse had a single, primary use case: reporting-based business intelligence. It was passive, and it was static.

New School: The data warehouse is for everyone

Today's data warehouse is dynamic. It's interactive. It's for anyone who needs to analyze data—your BI team, data scientists, IT professionals, front-line workers, executives, business analysts, developers—everyone.

The disruptive data warehouse is built on the premise that there are dozens of usage models, and they all require an operationally robust means of serving data. In addition to serving BI reports, the disruptive data warehouse serves as a forum for executing diverse analytics, including via scripting languages, to interact with the data and discover nuggets of information that can help the business.

Bottom line: By allowing everyone in your organization to consume data, the disruptive data warehouse magnifies user driven innovation.





Old School: The data warehouse is based on SQL

Whether putting data in the data warehouse, taking it out or manipulating it, SQL is the one and only way to interact with data in an old school data warehouse.

New School: SQL? Absolutely. And Ruby, Python, Perl, R, Java and C++

SQL is a fundamental means of accessing data in a data warehouse and therefore a mainstay. But with an expanded user base, growing demand for analytic insights, and the ability to execute distributed functions across a variety of repositories, today's data warehouse includes additional tools that enable anyone to access data. A developer doesn't necessarily know SQL, and a business analyst who knows SQL doesn't necessarily know R or Python. But both users can get value out of a disruptive data warehouse. A function can be programmed in R, for example, and run inside the database where it pulls data out of tables, manipulates the data and writes back the results to the warehouse.

Bottom line: Users with specialized skills can use them to run their analytics against all the data.





Old School: The data warehouse handles structured data

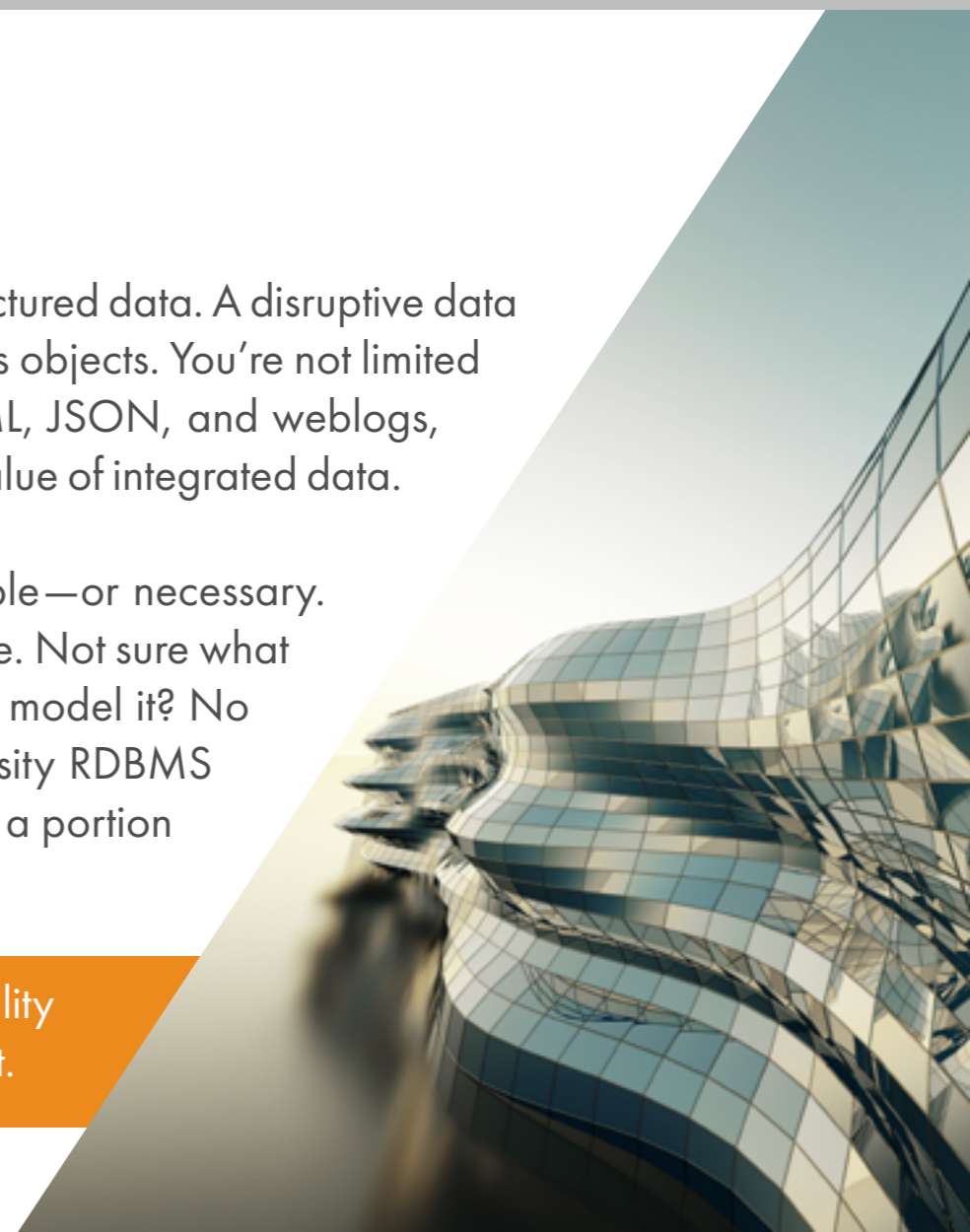
The old school data warehouse handles one type of data—structured data—in relational format, which must be modeled in advance.

New School: Yes—and JSON, XML, weblogs, and more

Today, in addition to structured data, organizations collect unstructured and semi-structured data. A disruptive data warehouse provides unprecedented flexibility, allowing you to store data as first class objects. You're not limited to rows and columns of structured data. You can store semi-structured data, like XML, JSON, and weblogs, natively. Methods within the database can operate on all data types, delivering the value of integrated data.

Given the variety of data, modeling every possible combination isn't always feasible—or necessary. With a disruptive data warehouse, modeling can be performed when it is appropriate. Not sure what the business value will be of a voluminous big data flow and can't justify the time to model it? No problem. Store it affordably in a data lake, such as Hadoop or a high storage density RDBMS that handles these data types, and create the schema at the time of access. Or move a portion of the data to the data warehouse to analyze it and determine its business value.

Bottom line: The disruptive data warehouse gives organizations the choice and flexibility to model when it makes sense, breaking away from the rigid modeling rules of the past.





Old School: External analytics

Data was extracted from the data warehouse, shipped over the network, and loaded into analytics tools prior to analysis. Data mining tools were relegated to a limited, representative subset of data for analysis, only to require extra time and effort when the sample was not representative.

New School: Integrated and parallelized in-database analytics

Time is of the essence. Business users need answers now, not next week or next month. The disruptive data warehouse eliminates the need to move data. Thousands of advanced math and statistical functions are integrated and parallelized in-database to analyze the entire set of data. Advanced analytics tools work with the warehouse to push the SQL equivalent of the analytic functions down into the data warehouse to run in parallel across all the data, while the disruptive data warehouse determines how to do this as efficiently as possible. As a result, business users get more accurate answers, faster. The network isn't getting bogged down, and the data isn't sitting exposed on a server.

Bottom line: Specialized analytics are run in parallel, in the database, and on all the data for the highest accuracy. No more waiting for extractions or moving data.





Old School: Disk data storage

The old school data warehouse relies solely on magnetic disk storage.

New School: Hybrid storage with solid-state drives

Storage technologies have evolved to deliver improved performance. However, this performance comes at a cost. Solid-state drives (SSDs) are pricey alternatives to disk. Instead of putting all data on SSDs, thereby driving up the cost, a disruptive data warehouse uses SSDs for only the most valuable, or frequently accessed data. The system *automatically* moves infrequently accessed data, also known as cold data, to slower magnetic spinning disks, and frequently accessed data, whether warm or hot data, to faster SSD storage.

Bottom line: The disruptive data warehouse intelligently and automatically uses storage to give you faster performance when you need it and cost efficiency when you don't.





Old School: Memory for standard caching

Old school data warehouses use standard caching methods for working memory, flushing the least recently used data out and releasing memory.

New School: Intelligent memory: in-memory dynamic and automatic

Some companies believe that the ultimate performance approach is to move all data into memory for pure in-memory solutions. The disruptive data warehouse leverages advanced engineering to apply in-memory processing where speed delivers real value, but uses standard caching for less frequently accessed data.

Here's how it works. The disruptive data warehouse automatically tracks and categorizes data as cold, warm, or hot. Cold data stays on disk storage where it is always accessible. Warm and hot data move to SSD storage. The very hottest data is also placed in-memory for cost effective, lightning fast performance.

Bottom line: The disruptive data warehouse knows what's hot and what's not. If cold data suddenly becomes frequently accessed, that data is automatically categorized as very hot and placed in-memory for blazing fast performance.





Old School: Row-based data storage only

In the old school data warehouse, data is stored in rows, the way it has always been. Row processing is efficient—but only some of the time.

New School: Hybrid row/column data storage

Row processing has an important place in the data warehouse, but so does columnar processing. The disruptive data warehouse allows you to use the best approach for your data. Store tables in rows, columns, or *both* rows and columns. The result is the ultimate in flexibility to deliver the best performance for every situation.

Bottom line: The disruptive data warehouse delivers the best of both worlds—row and columnar data storage—allowing you to apply the right partitioning model to the right table and the right data.





Old School: One choice: On-premises, dedicated system

The old school data warehouse is a dedicated system deployed on premises. The expense makes it a major commitment.

New School: Many choices: On-premises, in cloud, hybrid, hosted, mix

Today, organizations have choices about deployment models. The cloud lowers the barrier to entry, converts capex costs to opex, and enables organizations to try a disruptive data warehouse without making a significant financial commitment or taking up space in the data center. Alternatively, you may find that a hybrid approach that melds the best of both worlds is optimal for your organization.

Bottom line: With a disruptive data warehouse, organizations have choice and flexibility in deployment and expense categorization.





Old School: Always in the present

The old school data warehouse is inherently always in the present and therefore doesn't reflect changes to the business as they happened over time.

New School: Time travel is possible

The disruptive data warehouse features temporal capabilities that deliver automated point-in-time analysis. Like a time machine for analytics, users can run a query against the warehouse exactly as the data looked in the past. For example, if a user transfers from department 123 to 124 and runs a personal expenditure report for last year, the old school database will look for spending in the current department. It will not see any spending for department 124 because all the spending was in department 123. The disruptive data warehouse allows the user to run a report for last year and get the correct results for department 123.

Bottom line: The disruptive data warehouse reflects changes to the business and delivers reporting based on how the business existed at specific points in time.





Old School: Large project, long lead times

Adding new data to an old school data warehouse is a lengthy, tedious process that involves multiple parties and results in long lead times. Users have to wait months to access and analyze their new data, by which time the need for analysis may have changed.

New School: Agile, self-service, untested data, incremental, phased delivery

Self-service capabilities in the disruptive data warehouse enable users to load, test, query, and merge their new uncleaned, untested data with the rest of the data in the warehouse. All of this is done in a low-risk manner in a Data Lab, which is sectioned off from the rest of the database and ensures visibility and adherence to IT governance. The Data Lab can be run against the production environment, enabling the user to sandbox within the production environment and test additional data.

Bottom line: The BI team is no longer a bottleneck in the analytics process. The disruptive data warehouse enables rapid exploration and experimentation, allowing a wide range of users to uncover business value efficiently while staying within IT governance processes.





Old School: Single Platform to Solve All Problems

Until recently, the data warehouse consisted of a single platform and database. The performance, scalability, economics, and analytic feature set of the data warehouse were tied to the capabilities of the one underlying technology.

New School: Logical Data Warehouse

The modern data warehouse consists of multiple platforms and technologies working together seamlessly to solve business problems. The Teradata® Unified Data Architecture™ is a logical data warehouse that presents users with a single interface to all the resources they need in an integrated analytical ecosystem.

Bottom Line: The capabilities of the data warehouse are no longer tied to a single physical platform and database. Instead, the data warehouse is a logical system consisting of best-of-breed technologies that deliver powerful, agile business analytics.



The Best of Both Worlds: The Disruptive Data Warehouse

The hype cycle, startups, open source vendors, and the tech press often give the mistaken impression that the data warehouse has been standing still and is bound for immediate replacement by the new, new thing.

The reality is that the disruptive data warehouse embraces all that is great in new technology while providing all the stability, reliability, recoverability, security, and availability enterprises require to run their business.

The bread and butter of your business, transactional data, belongs in rows most of the time. All data is not equal (think about data on your own laptop); not everything deserves to live in-memory. Data is coming in fast and furious, and once you find business value in it, you'll want to model it so that more people can access it.

A lot has changed in the past five to ten years. The disruptive data warehouse has kept pace with these changes, integrating new technologies in an intelligent and cost-effective manner while paying heed to the changing needs of the business. New tools and self-service capabilities expand the data warehouse user base, allowing the business to uncover valuable insights faster than ever before.

Choice and flexibility mean that the disruptive data warehouse meets users where they are right now. And intelligent automation ensures that new technology is leveraged the right way, at the right time, with minimum technical debt. Now and in the future, the disruptive data warehouse is poised to provide the tools and capabilities businesses need to take advantage of all of the data.

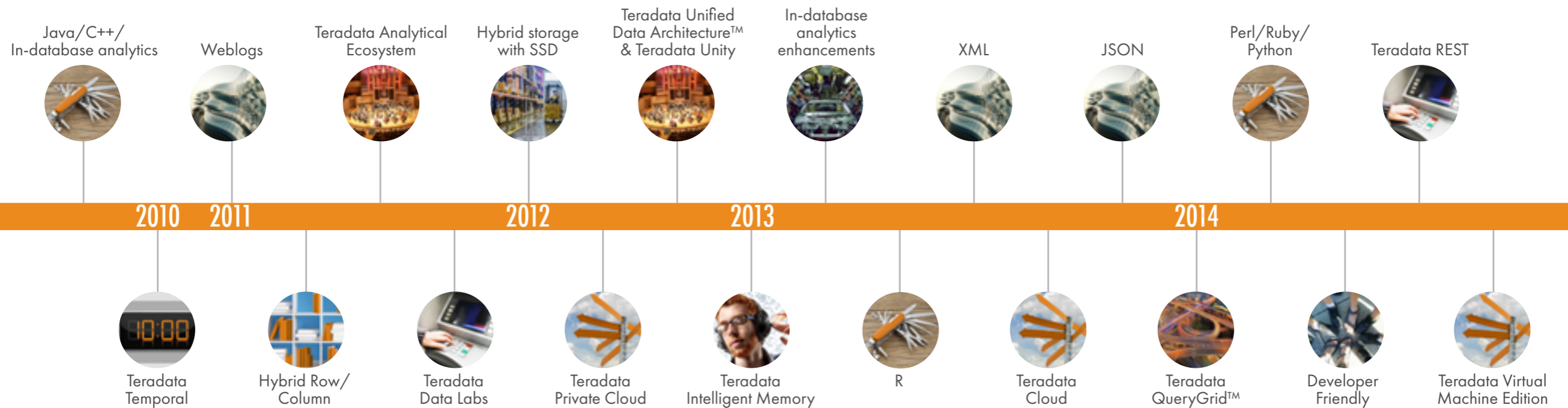
In short, the disruptive data warehouse is:

- **Smart:** Optimizes processing and use of resources, dynamically managing hot, cold, and warm data
- **Inclusive:** Of many types of users, many types of analytics, many data types
- **Strategic:** Fast answers, smart use of resources
- **Flexible:** All deployment models, whether on-premises, cloud, hosted, or hybrid
- **Not your father's data warehouse**

Leading The Path To Technology Innovation

Teradata® has been at the forefront of data warehouse technology, pushing the limits and in many cases pioneering the disruptions themselves.

Teradata believes the data warehouse is a strategic business tool that is delivering more business value today than ever in history. The Teradata integrated data warehouse continues to develop technology disruptions enabling companies to achieve the business agility and market differentiation needed for competitive advantage.



THE RISE OF THE DISRUPTIVE DATA WAREHOUSE

New School Thinking	Teradata Innovations
Agile, self-service, untested data, incremental, phased delivery	Teradata Data Labs, Teradata REST
Hybrid storage with solid-state drives	Hybrid Storage with SSD
Hybrid row/column data storage	Hybrid Row/Column
Integrated and parallelized in-database analytics	In-Database Analytics Enhancements
Intelligent memory: in-memory dynamic and automatic	Teradata Intelligent Memory
Logical Data Warehouse	Teradata Unified Data Architecture™, Teradata Unity
Many choices: On-premises, in cloud, hybrid, hosted, mix	Teradata Private Cloud, Teradata Cloud, Teradata Virtual Machine Edition
Query multiple databases and platforms	Teradata QueryGrid™
SQL? Absolutely. And Ruby, Python, Perl, R, Java and C++	R, Ruby, Python, Perl, Java, C++
The data warehouse is for everyone	Developer Friendly
Time travel is possible	Teradata Temporal
Yes-and JSON, XML, weblogs, and more	XML, JSON, Weblogs

Unified Data Architecture and QueryGrid are trademarks and Teradata is a registered trademark of Teradata Corporation and/or its affiliates in the U.S. and worldwide.