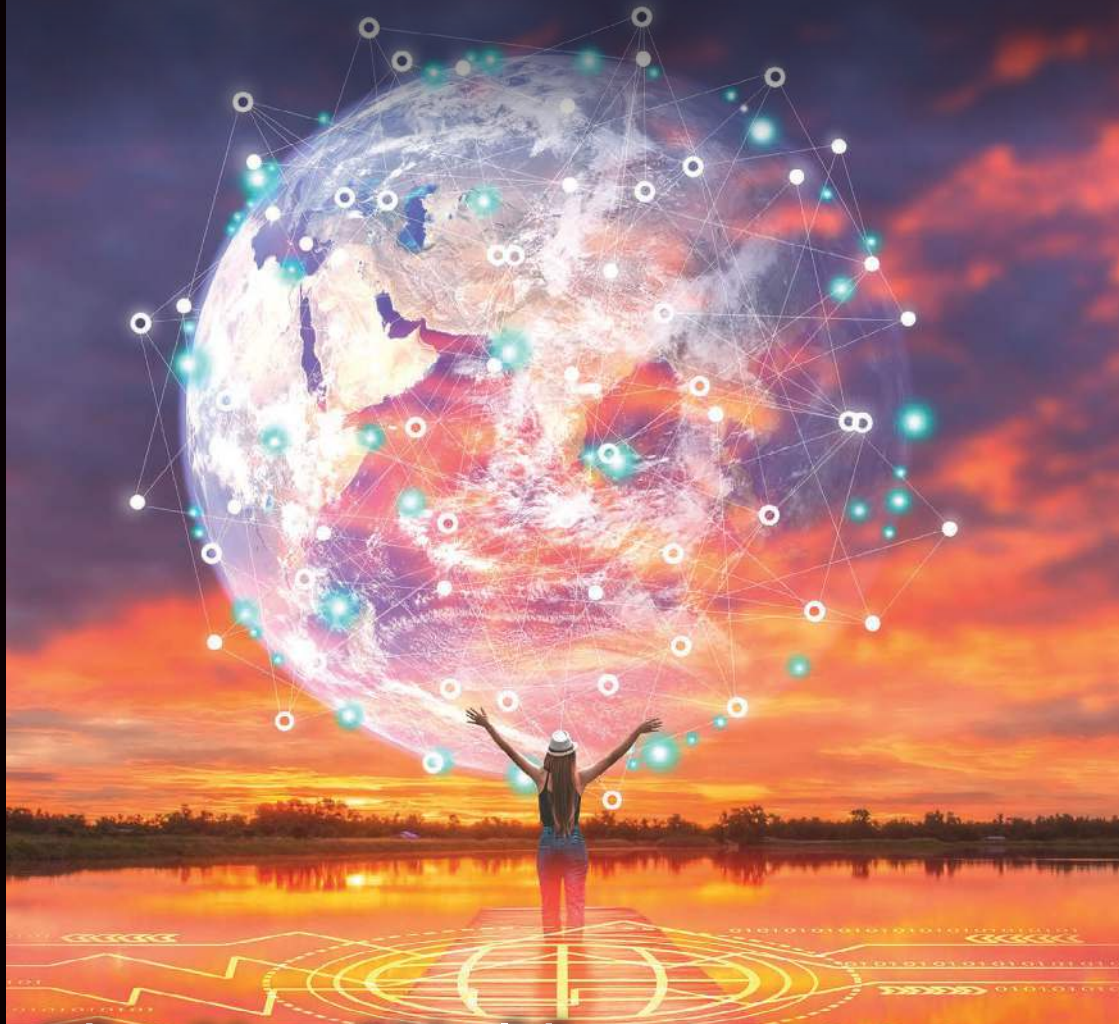
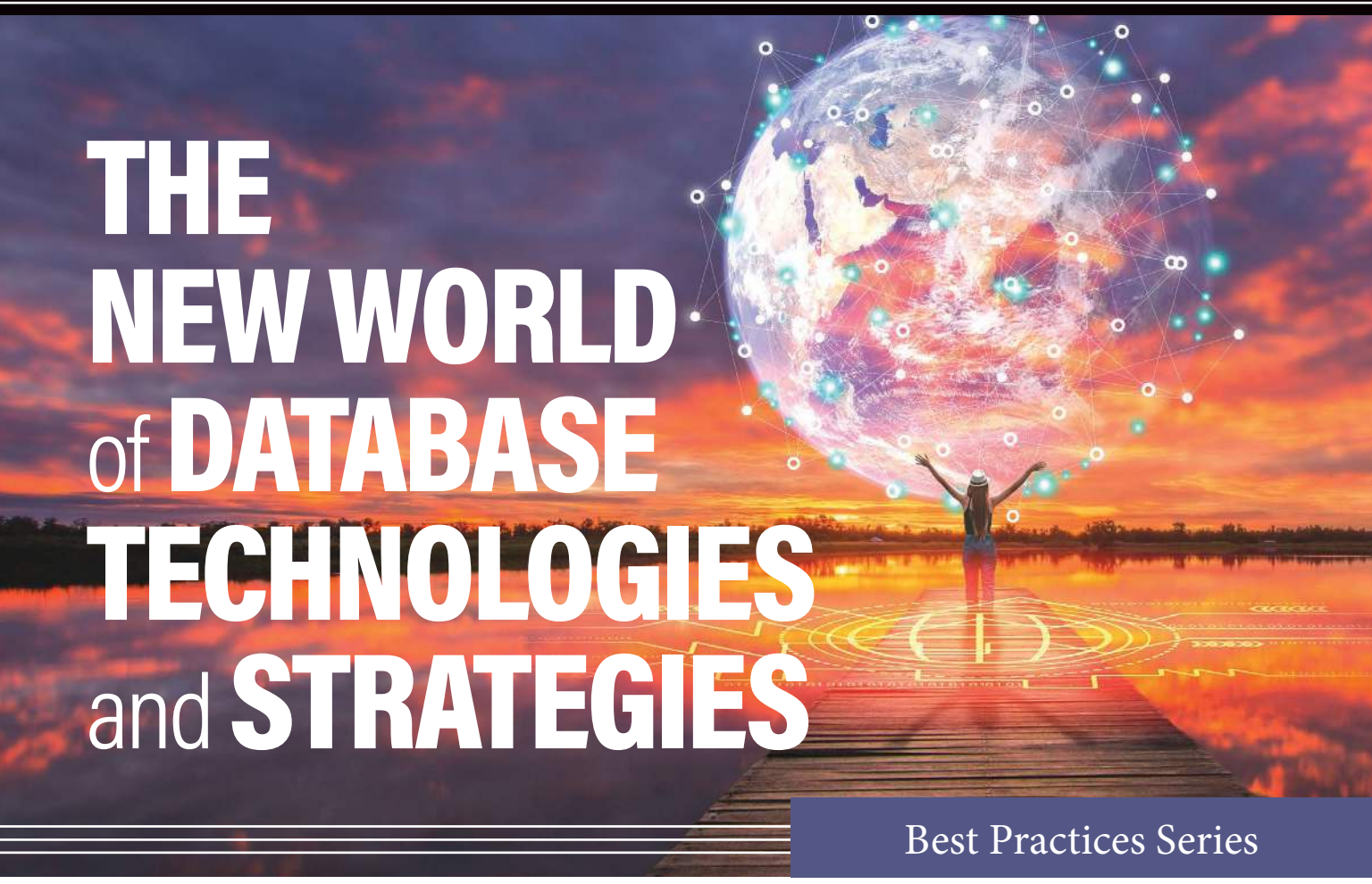
database
TRENDS AND APPLICATIONS
ONE COMPLETE MARKETING PROGRAM



The New World of DATABASE TECHNOLOGIES and STRATEGIES in 2023

Best Practices Series

database
TRENDS AND APPLICATIONS



THE NEW WORLD of DATABASE TECHNOLOGIES and STRATEGIES

Best Practices Series

The world is changing, and the database world is changing just as fast. And, to a large degree, the changes we see in the world are courtesy of databases. AI, advanced analytics, and on-demand digital services all depend on the viability and robustness of the databases that support them. But building, maintaining, and growing the data environments required for this new world call for new approaches as well as technologies. This means rethinking the way data environments are built, optimized, and managed.

Of course, the database world has not idly sat by while these needs have emerged. In recent years, we have seen the development of new generations of databases, from distributed databases that allocate data stores across clustered networks to graph databases that show visual displays of the relationships between data. In-memory databases allow for processing of massive amounts of data in near real time, and cloud databases push administration, scalability, and storage out to third-party providers. The rise of data lakes and, subsequently, data lakehouses, means the rise of repositories in which vast varieties and volumes of data can be maintained for later analysis. On top of all this, security and privacy concerns dominate corporate agendas in the face of relentless hacking and breach attempts.

recent years is the rise of data management and analysis skills for mainstream business users, especially with the rise of generative AI tools and platforms. Database management has become an enterprise-wide venture that requires the ideas and insights of all constituents.

A data-driven enterprise needs to leverage real-time insights, as well as be able to provide those elements, without so much as a hiccup, any time in the day or night.

These developments mean greater agility, capacity, and capabilities for data environments. At the same time, there's a need for data teams to evolve and adapt to this more data-centric world.

The following are ways to achieve greater advancement:

- **Educate the business user, educate the database team.** One significant change in
- **Keep the business front and center.** Every investment in data technology—as well as strategic initiative—needs to have a use case behind it. Ultimately, it's the business that foots the bill for data environments, and it's important that business leaders identify their top priorities. An essential part of business-first database management is data governance. It's important that

The rapid pace of change and continued adoption of advanced technologies mean enterprises will need the skills to leverage database technologies.

governance practices and policies not only assure data quality, compliance, and security, but also align with what the business needs.

- **Make data architecture and data modeling matter—more than ever.** As database technology is considered, the enterprise's data environment needs to be mapped out and aligned with current and future requirements of the business. This is important, as significant investments will continue to be made in data technology, and these investments will need to be as focused as possible. Data and business managers need to have a sense of what their organizations' data requirements and capacity will be today, but also 1, 2, or 3 years out. What kind of growth is anticipated? As the architecture is rolled out, data modeling comes into play, which is essential for top-performing database design.
- **Jazz up queries.** Typically, queries have been within the domain of data professionals who prepare the coding and reporting at the requests of users. This has slowed analytics delivery, and the impetus is now on enabling end users to design their own queries as they need them. The challenge is that, for years, data team members and their business counterparts have been frustrated by the complexities involved in building SQL commands. While NoSQL databases have risen in prominence in recent years, it's still important that staff members have a fundamental understanding of SQL to make the most of the query language. In addition, new tools and concepts arising in the market—even those employing AI—can ease the use of this important query tool.
- **Make sure data is all backed up, available, and secure.** A data-driven enterprise needs to leverage real-time insights, as well as be able to provide those elements, without so much as a hiccup, any time in the day or night. Backup and recovery always need to be on and as close to real-time as possible. Database replication has come a long way in recent years, and cloud services now can provide backup for datasets of any size. This ensures one degree of data security. In addition, while it's not news that data is highly vulnerable to hackers, malicious code, and inside threats, it is news that fast-growing attack surfaces are occurring in data-driven enterprises, particularly those that rely on digital services, cloud, and ecosystems or partners. From the beginning, security needs to be built into processes involving sensitive or personally identifiable information, incorporating the latest encryption and access controls.
- **Keep tabs on performance and potential glitches.** "Observability" is the new watchword for data environments, both in terms of performance as well as tracking lineage. Data teams need to employ data analytics and monitoring tools to not only catch but also prevent system slowdowns, glitches, and potential outages.
- **Keep on automating.** The trend in the database world—as well as across the information technology space—is toward increasing automation. However, much of this is occurring in fits and starts, with many data processes still requiring manual control and scripting. Many low-level database functions—from backups to schema changes—can readily be automated, enabling data professionals to focus on serving the requirements of the business.
- **Keep everyone's skills refreshed.** The rapid pace of change and continued adoption of advanced technologies mean enterprises will need the skills to leverage database technologies. There is an insatiable demand for data engineering, database administration, data science, and data analysis skills. Stepped-up recruiting efforts may help, but the competition for talent is steep. The best strategy is to provide opportunities to existing data staffs to pursue continuous learning through outside educational programs, internal training, certifications, and coaching.
- **Look to DataOps, DevOps, and agile methodologies.** Continuous improvement, working hand in hand with the business, is the name of the game. Data pipelines deliver quality data on a constant basis, and collaborative methodologies such as DataOps assure that everyone is working on the same page, assisted by automation.

Leading and supporting databases in the 2020s call for analytic and predictive capabilities that help enterprises keep ahead of intense global competition. All aspects of corporate management and service or product delivery rely on greater intelligence within supply chains, inventories, delivery networks, and customer channels. The technology that will bring this greater intelligence—from AI to robotic drones—needs reliable datasets. This amplifies the roles of data teams, who need to produce value from the rising amounts of data pulsing through enterprises. ■

—Joe McKendrick

How to Know When to Use a NoSQL Database

If you asked me what database to use back in 2008, the answer for pretty much every use case was SQL Server, Oracle, or MySQL. Fast forward to today and there are 100s of options for relational and NoSQL database technologies. So how do you choose when to use a NoSQL database?

PROS AND CONS OF NOSQL DATABASES

Understanding the strengths and weaknesses of NoSQL as a whole, along with the different types available, can guide you to the right decision.

PROS OF NOSQL

Schema-less Flexibility: You have a variety of data models to choose from, including document stores, key-value pairs, time-series, graph, wide column, and many more.

Horizontal Scalability: NoSQL's flexible structure also lends itself to better horizontal scalability and distributed processing.

High Performance: NoSQL doesn't go through data normalization, so your queries don't need to use joins, which drives high performance on large data sets.

Purpose-built: Relational databases are often used as a one-size-fits-all option for every use case, which can impact functionality and performance. NoSQL databases focus on specific use cases, so you can get a fit-for-purpose solution.

CONS OF NOSQL

Eventual Consistency: Many NoSQL databases have eventual consistency, so if you have workloads that require ACID compliance, that could be a poor fit.

Querying Learning Curve: If you or your DBA team only have an SQL background, then it can take some time to adjust to querying in NoSQL.

Specialized Skill Sets: While NoSQL databases are popular today, relational databases still dominate the market and that's reflected in database administrator (DBA) skill sets.

Database Sprawl Potential: The flip side of having many purpose-built databases is that you can end up with a database sprawl issue leading to an overly complex environment.

WHEN TO USE NOSQL DATABASES

Here are several common types of NoSQL databases and their use cases:

- **Key-value:** Sometimes all you need out of a database is a simple key-value structure that provides fast reads and writes due to the lack of overhead.
- **Document:** The JSON support makes this a favorite of development teams.
- **Wide Column:** When you're dealing with massive data sets, and you want distributed processing that allows you to read and write every node for high availability and eliminating single points of failure.
- **Graph:** The relationship between objects is the most important data and needs to be treated as first citizen.
- **Time-series:** Your data is primarily time-based, such as security camera events or financial transactions.
- **In-memory:** You need to speed up your primary database performance through in-memory caches.

5 LEADING NOSQL DATA PLATFORMS

- **MongoDB:** A source-available document database, released over 14 years ago, is the most popular NoSQL option due to its developer-friendliness, its maturity, and the breadth of API support it has.
- **Apache Cassandra:** An open-source wide-column database designed for massive, distributed data sets.
- **Redis:** The most popular in-memory database, often complementing another database technology to speed up its performance.
- **Amazon DynamoDB:** An AWS cloud key-value store designed to be cloud-native, with fast performance.
- **Azure CosmosDB:** Microsoft's NoSQL PaaS offering is multi-model, acting like a database Swiss Army knife that's compatible with MongoDB, Cassandra, Gremlin, and PostgreSQL APIs, as well as Azure Table.

"When choosing the best NoSQL database, consider your data's unique needs, scalability requirements, and desired flexibility. Seek a balance between performance, ease of use, and community support to pave the way for a successful and efficient data-driven journey."

WHAT TO DO WHEN YOUR DATA SET BECOMES LARGE AND COMPLEX

NoSQL databases are excellent for working with large data sets but scaling them up properly requires specialized expertise. Without the right architecture and tuning, it's easy to run into a situation where your costs grow exponentially, and your performance suffers. Working with a managed service provider that has a specialized NoSQL practice, such as Datavail, ensures that you're getting the most out of your NoSQL investments through optimization, performance tuning, on-going support, and site reliability engineering services. ■

DATAVAIL
www.datavail.com