2022

(OCIENT)™

Beyond Big Data

THE RISE OF HYPERSCALE

CONTENTS

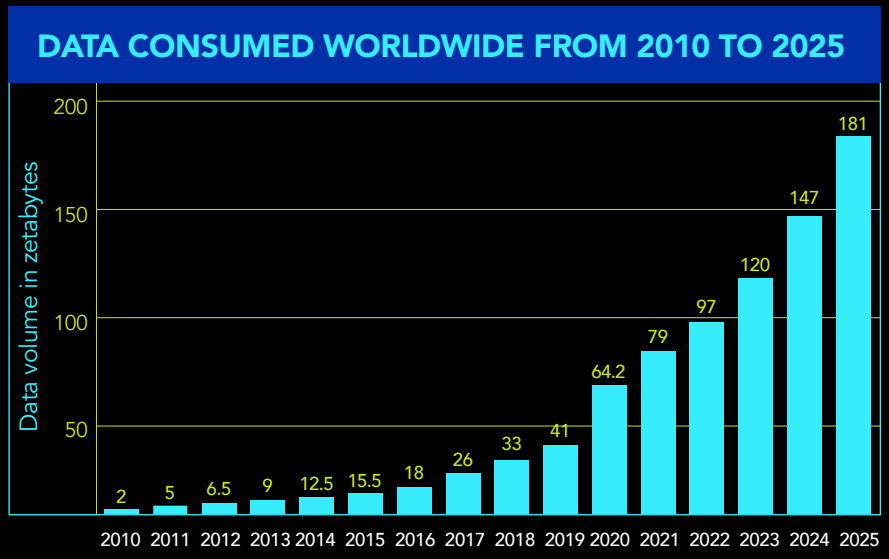
INTRODUCTION

It's no secret that data is growing at extraordinary rates. In 2020, people were generating 1.7 MB of data every second, and that already nearly unfathomable number continues to increase steadily. In fact, current growth rates point to a total data volume of over 200 zettabytes – 200 billion terabytes – by 2025. That's more than 11.4 billion years of HD cat videos.

But it's not just data volume that's increasing. The velocity at which data is being generated and the variety of what's being collected are also ticking steadily upward. Data is coming in from more sources, from more locations, and for a wider variety of use cases than ever before. For example, each of today's connected cars produces about 25 MB of data per hour. By 2025, there will be about 400 million connected cars on the road, constantly gathering information about location, engine health, road conditions, and more. That means that every single day, cars alone will generate 240,000 terabytes of data.

Put simply, the world is moving beyond big data. Not only will tomorrow's biggest businesses be collecting and storing trillions of rows of data – they will need to perform continuous, complex analysis on those hyperscale data sets, using technology to run thousands or even millions of queries every hour, around the clock. This will keep their businesses running smoothly, enable the development of new products and services, bolster customer satisfaction and help them find the proverbial needles in the haystacks: the business insights that will lead to continued prosperity and future growth.

To better understand this shift, Ocient surveyed 500 technology and data professionals across a variety of industries who are already managing active workloads of 150 terabytes or more. With the goal of understanding the top opportunities, challenges and concerns associated with the rise of hyperscale, we've used this report to identify key trends, compiling them together in this, our very first Beyond Big Data report. Here, you'll learn more about the importance of hyperscale data analytics, the challenges it presents, and the technical capabilities tomorrow's businesses need today.



https://www.statista.com/statistics/871513/worldwide-data-created/
https://www.statista.com/topics/1918/connected-cars/#topicHeader__wrapper/

HYPERSCALE DATA ANALYTICS

hy•per•scale da•ta an•a•lyt•ics

/ˈhīpər skāl/ˈdadə,ˈdādə/ /ˌanəˈlidiks/

noun

The complex computational analysis of ultra-large data sets — those comprised of trillions or more records — in interactive time (seconds to minutes versus hours to days)

EVERY 2 DAYS we create as much information as we did from the beginning of time until **2003**.

OVER 90% of all the data in the world was created in the past **2 years**.

THE TOTAL AMOUNT OF DATA being captured and stored by industry **doubles every 1.2 years**.

Assuming the average row of data in a database is 100 Bytes of data, the 181 Zetabytes of data projected to be consumed annually by 2025 will consist of 1.81 sextillion rows of data. That's 1.81 billion times 1 trillion rows of data. A printout of just one trillion rows of data would circle the earth 73 times. Scrolling through it on a screen would take 600 years.

The only way to make this data usable? The widespread deployment of databases that can rapidly collect, store and analyze trillions or more rows of active data.

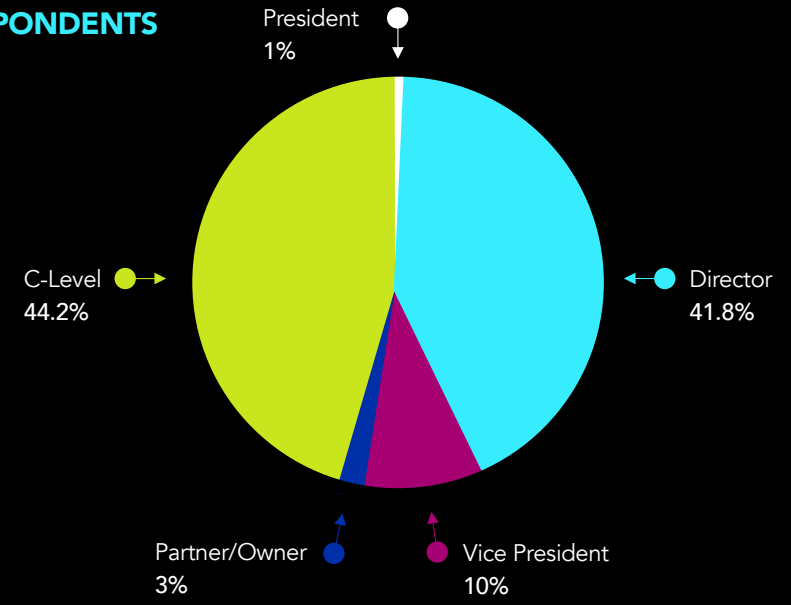
https://cloudnine.com/ediscoverydaily/electronic-discovery/date-fun-facts-big-data-ediscovery-trends/
https://www.idc.com/getdoc.jsp?containerId=US46410421

RESEARCH METHODOLOGY

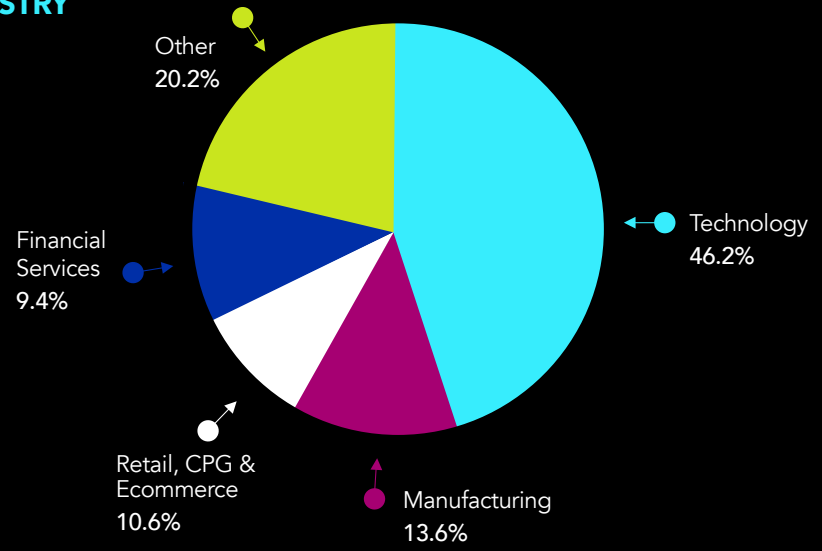
In May 2022, we partnered with Propeller Insights to survey 500 high-level technology savvy business leaders who manage data workloads of 150 terabytes or greater.

Their companies' annual revenue ranged from US$50 million to US$5 billion. Approximately 50% of respondents represent companies with annual revenue greater than $500M.

500 RESPONDENTS BY ROLE



500 RESPONDENTS BY INDUSTRY



KEY FINDINGS

1 DATA ANALYSIS IS TIED TO FINANCIAL SUCCESS.

Rapid analysis of hyperscale datasets is now imperative. Key decision makers recognize that real-time data analysis helps businesses succeed by driving efficient decision making, generating critical business insights, and predicting future outcomes.

2 DATA WORKLOADS ARE GETTING BIGGER, FASTER, AND MORE COMPLEX.

This can be particularly difficult for companies undergoing digital transformations.

3 LEGACY SYSTEMS AREN'T BUILT TO HANDLE HYPERSCALE DATA ANALYSIS

but upgrading old systems can be costly, complicated, and risky.

4 SECURITY, COMPLEXITY, AND BUDGET ARE THE TOP CONCERNS

for teams handling hyperscale data workloads.

5 STAFFING IS A CHALLENGE.

Data professionals are critical to their organizations' success, but there is a dearth of people with the right talent, skills, and experience to fill these roles.

THE MAJORITY OF EXECUTIVES SEE FASTER DATA ANALYTICS AS KEY TO THEIR ORGANIZATIONS' SUCCESS.

As the number and type of hyperscale datasets increases, so does the potential impact of their analysis.

The ability to harness data at hyperscale in a cost-effective way will enable enterprises and government organizations to better understand customer and citizen experiences, including how products and services are used and deployed in the real world in near real time. This level of analysis will be a critical force in enabling **greater agility and flexibility** among increasingly automated supply chains and business operations.

With hyperscale data analytics paving the way for greater intelligence, growth, and innovation in virtually every industry today and in the future, organizations will be able to leverage smarter predictive models to better protect and prevent potential negative events or outages from occurring, thereby **saving time and resources** on disaster recovery as well.

78% of respondents agree that faster data analytics are tied to revenue growth. The more quickly companies can analyze data, the better they can respond to changes in the marketplace. The result: improved sales, higher customer satisfaction, and a healthier bottom line.

64% use insights from hyperscale datasets to make business decisions.

62% say data analysis insights are essential for business planning and company strategy.

HARNESSING HYPERSCALE DATA IS NOT JUST AN OPERATIONAL EXPENSE.

IT IS AN INVESTMENT CRITICAL TO DRIVING INCOME.

> FOR A TYPICAL FORTUNE 1000 COMPANY, JUST A 10% INCREASE IN DATA ACCESSIBILITY WILL RESULT IN MORE THAN $65 MILLION ADDITIONAL NET INCOME.

Richard Joyce, Senior Analyst at Forrester

https://www.forbes.com/sites/brentdykes/2019/03/28/the-four-key-pillars-to-fostering-a-data-driven-culture/?sh=747c50477d90

EXECUTIVES ALL AGREE: VOLUME, VARIETY, AND VELOCITY ARE ALL GROWING EXPONENTIALLY.

Clearly, data and data analysis merit special attention. But as companies increase their digital connectivity, bringing in a wider variety of data types from more disparate sources than ever, the challenges just keep getting bigger and more complex.

Nothing can be done to stem the tide of data. Companies must find a way to keep up, or they will fall behind.

97% of survey participants expect their data to grow fast or very fast in the next 1-5 years.

98% agree that it's somewhat or very important to increase the amount of data analyzed by their organizations in the next 1-3 years.

FOR LEGACY SYSTEMS, THE RUNWAY IS COMING TO AN END.

For years, companies have been finding ways to keep their legacy systems alive, with varying degrees of success. But as the amount of data and our need to analyze it increases, the ability of legacy systems to scale and flex sufficiently is increasingly threatened.

IBM Db2 (61%), **Cloudera** (49%), and **Teradata** (41%) are at the top of the list of legacy systems that key decisions makers are looking to replace. In many cases, these and other legacy products are reaching end of life and/or end of support.

59% of our survey respondents are actively looking to switch data warehouse providers.

The top three reasons they cite:

- **The technology is legacy and I'm looking to modernize (40%)**
- **The system(s) isn't/aren't comprehensive enough to meet my needs (42%)**
- **Not flexible enough (36%)**

With so many legacy systems on their way to obsolescence, decision makers are putting a lot of thought into what comes next. Speed, agility, and integration are top of mind, but replacing legacy systems – some of which have been in place for decades – can be an overwhelming proposition.

When the positives of a new data warehouse outweigh the negatives associated with sunsetting familiar technology, **it's time to make a move**. Our respondents showed that in addition to speed and performance upgrades, having a flexible and agile system and improving data access and integration across the enterprise are key.

The top three features motivating decision makers to consider new technology:

- **Improved speed and performance (57%)**
- **Flexibility and agility (55%)**
- **Improved access and integration (54%)**

THE NUMBER ONE CONCERN AMONG SURVEY RESPONDENTS IS MAINTAINING SECURITY AND COMPLIANCE.

63% of survey respondents are concerned about the security and compliance of their hyperscale data sets. Whether that means bolstering cybersecurity protections or adhering to GDPR and CCPA, **keeping data safe is paramount** among the professionals we surveyed. That means protecting data with a variety of tools including SSO applications, encryption and role-based access.

The top three challenges our respondents are facing:

- Maintaining **security and compliance** as data volume and needs grow **(63%)**
- Scaling data management and analysis in a **cost-effective manner (49%)**
- **Streamlining** the number of systems and ecosystems under management **(48%)**

THE SOLUTION NEEDS TO SCALE, NOT THE COST.

> Organizations clearly want and need to be able to harness massive amounts of data to stay competitive and drive better decisions, but IT's budgets aren't as scalable as public cloud resources.

Doug Henschen, Vice President and Principal Analyst, Constellation Research

65% of our survey respondents are concerned or very concerned about their ability to hire enough experienced and specialized talent to achieve their data strategy needs.

The top three most sought-after roles:

- **Chief Data Officer (64%)**
- **Data Analyst (60%)**
- **Data Scientist (41%)**

STAFFING IS A CHALLENGE, ESPECIALLY AS DEMAND FOR EXPERIENCED, SPECIALIZED TALENT OUTPACES SUPPLY.

Chief Data Officers are relative newcomers to the C-Suite. A 2021 PwC study showed that only 21% of the top 2,500 publicly traded companies in the world have CDOs in place, and almost half of those were appointed in or after 2019. Top businesses are just beginning to recognize the **value and importance** of hiring data professionals.

However, because data science is a relatively new field that requires specialized skills, education and experience, these jobs are difficult to fill from the C-Suite down. In fact, according to Quanthub, the number of job postings for data scientists in 2020 was **three times the number** of searches (or demand) for the same role.

https://www.strategyand.pwc.com/de/en/unique-solutions/digital/cdo-2021/strategyand-cdo-study-2021.pdf

https://quanthub.com/data-scientist-shortage-2020/

CONCLUSION

Data never stops. It's being generated everywhere: by our smartwatches, our phones, our TVs, our homes, our cars, the airplanes we fly in, the places we visit...and the list goes on. The most successful companies aren't just collecting this data: they're analyzing it.

As decision makers prepare for their organizations to collect unprecedented amounts of information, they're looking to upgrade and modernize their legacy systems. They're prioritizing security, while also working to reduce complexity and cost. They're putting together plans to skill up quickly with teams who can remove barriers to data intelligence so they can seize opportunities to leverage data in better, more valuable ways. And they're preparing to take the next step in their data journeys – a step that will bring about the kind of innovation that powers new products, experiences, services and business models.

IT'S TIME TO LOOK BEYOND BIG DATA AND PREPARE FOR THE RISE OF HYPERSCALE.

LET OCIENT TACKLE YOUR TOUGHEST DATA CHALLENGES.

Ocient is more than just a hyperscale data warehouse company. We provide next-level technology alongside an expert team with a proven track record of building trillion-row plus data storage and analysis systems, from databases and data warehouses to mainframes and supercomputers.

- With over 300 person years of experience in database, data storage, and data warehousing systems, Ocient's team of industry experts deeply understand the problems customers have analyzing and managing data at hyperscale.
- We engineered the Ocient Hyperscale Data Warehouse to optimize data analysis at hyperscale.
- Our solutions-first approach takes on the heavy lifting of designing, configuring, and deploying proven end-to-end solutions enabling customers to accelerate time to market.
- Ocient removes the talent and resource constraints many organizations face when challenged to innovate at hyperscale.
- We form close partnerships with customers to solve complex business and technical challenges from initial discussion through to successful production deployments.
- Ocient's cost-effective implementations relieve customers of the time-intensive burden of building an end-to-end solution, while enabling them to launch new products, services, and revenue streams powered by hyperscale data to grow their business.
- Ocient offers ultimate flexibility of deployment, on-prem or in the cloud, with solutions tailored to customers' business requirements and use cases.

(OCIENT)™

Don't miss your chance to get ahead of the rise of hyperscale.
Visit us at **Ocient.com** to learn more and book a demo.