

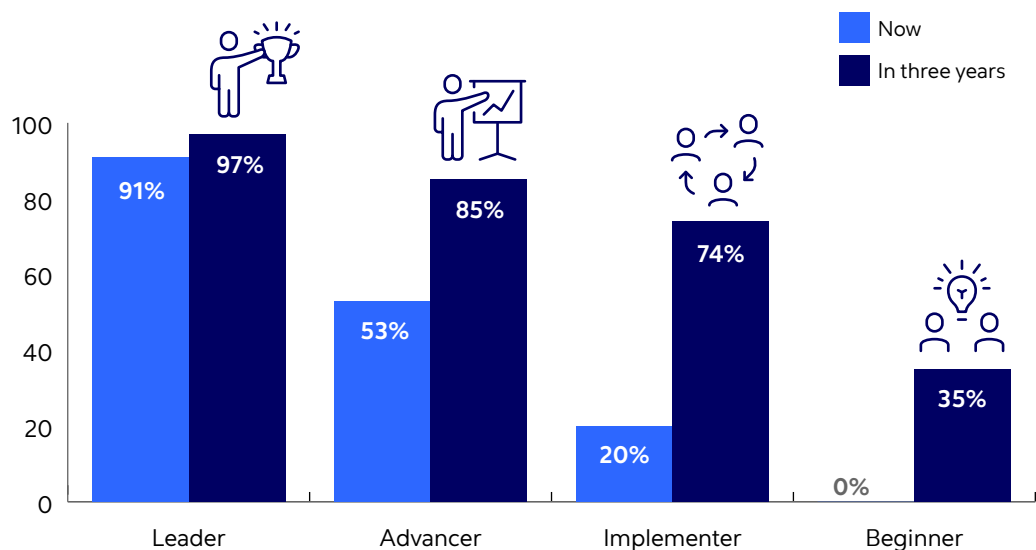
It's All About the Data (But Not Just Any Data)



Modernizing data = AI maturity

Nearly all AI leaders have mature data modernization practices

Percent of respondents who say they're maturing or advanced in data management.



Response base: 1,200
Source: ESI ThoughtLab/
Cognizant Figure 4

It's clear from our study that there's an inextricable link between AI maturity and data management (what we call "data modernization"): the work involved with ensuring the accessibility, reliability and timeliness of data for AI and analytics. Nine out of 10 AI leaders say they're in the maturing or advanced stages of data management, while literally none of the AI beginners rate themselves that way (see Figure 4). At the same time, having a proper IT architecture and data modernization processes in place was the most important lesson learned for beginners (60%).

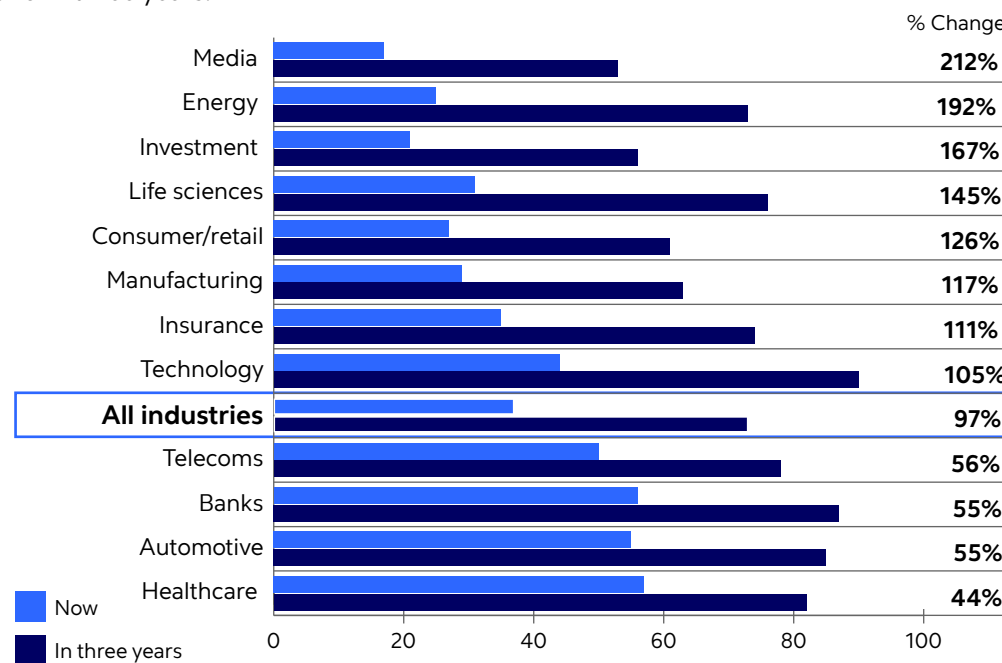
What's more, more than half of healthcare, banking and auto manufacturing businesses – the industries with the highest percentage of leaders in them – have already made significant progress in modernizing their data, and an overwhelming majority expect to by 2023.



Further, the same industries expecting particularly high growth in AI maturity are the same ones that expect the greatest gains in data modernization (see Figure 5). In all, businesses spend about 35% of their AI budgets on data modernization, or about \$13.3 million per company. Beginners spend even more, or 44%.

Industries expecting to accelerate their AI maturity will also surge in data modernization

Percent of respondents who say they're maturing or advanced in data management now and in three years.



Response base: 1,200

Source: ESI ThoughtLab/Cognizant

Figure 5

Data that matters

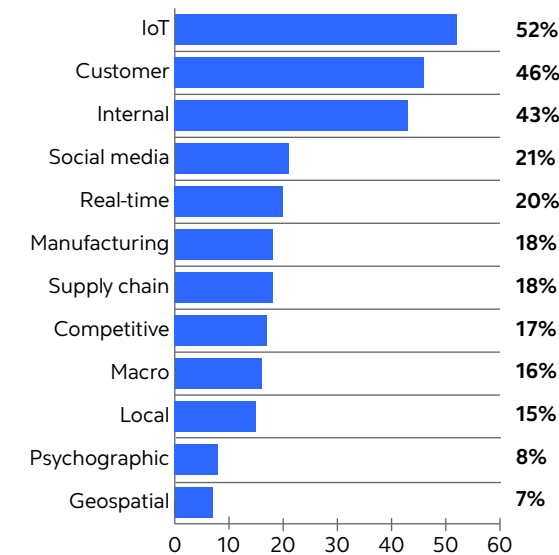
Increasingly, businesses are finding the most easily accessible data sets aren't enough to make the most intelligent decisions. By 2023, we'll see businesses pulling from wider and more diverse data sets for AI-driven insights (see Figure 6).

Today, IoT, customer and internal information are the main types of data integrated into AI applications. In many cases, this is simply because of the sheer volume of accessible data generated by sensors and customer interactions. But other forms of data are where the greatest insights often lie, particularly when such data is combined.

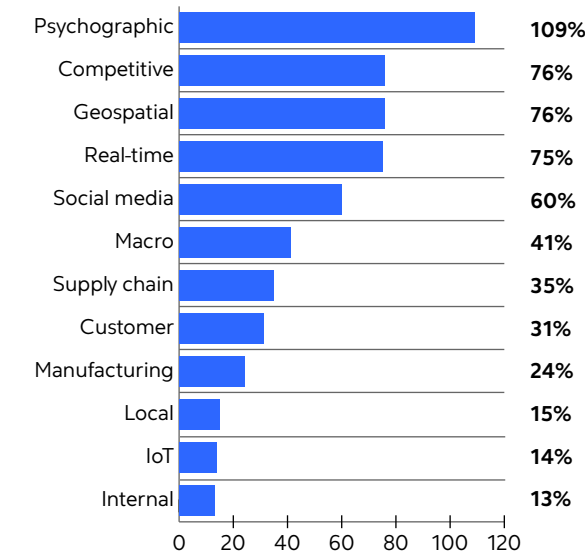
For example, we worked with a home goods store that wanted to know where people went when they left the store. By combining geospatial data with other data, we discovered that 30% of people go to McDonald's. That insight became very valuable for cross promotions.

Data sources will expand greatly in three years

Data used for AI today
Percent of respondents integrating each type of data into their AI applications.



Fastest growing over three years
Percent growth in the use of each type of data.



Response base: 1,200
Source: ESI ThoughtLab/Cognizant
Figure 6

Further, during the pandemic, global brands will see different regions of the country and the world open for business at different paces. For one of our clients, that has meant continually adjusting its product mix, product placement and product sourcing as supply and demand conditions change unpredictably, region by region.

We're using machine learning to refine our analytic models to predict the effects of everything from ongoing infection rates to regional weather conditions on future sales and demand trends. This data includes medical information from leading healthcare providers, as well as historical internal data, such as same-store sales. At each step, we're using our agile analytics methods to make sure we deliver the analytics the business needs most as the recovery from the pandemic unfolds.

Over the next three years, companies will double their use of psychographic data and ratchet up their reliance on competitive, geospatial and real-time data by about 75%. No AI program is complete without "voice of the customer data" gleaned from sources such as social media and call center analytics. While much of that data is generally lost or overly summarized, it's essential to extract insights from that data and get them to relevant business teams.

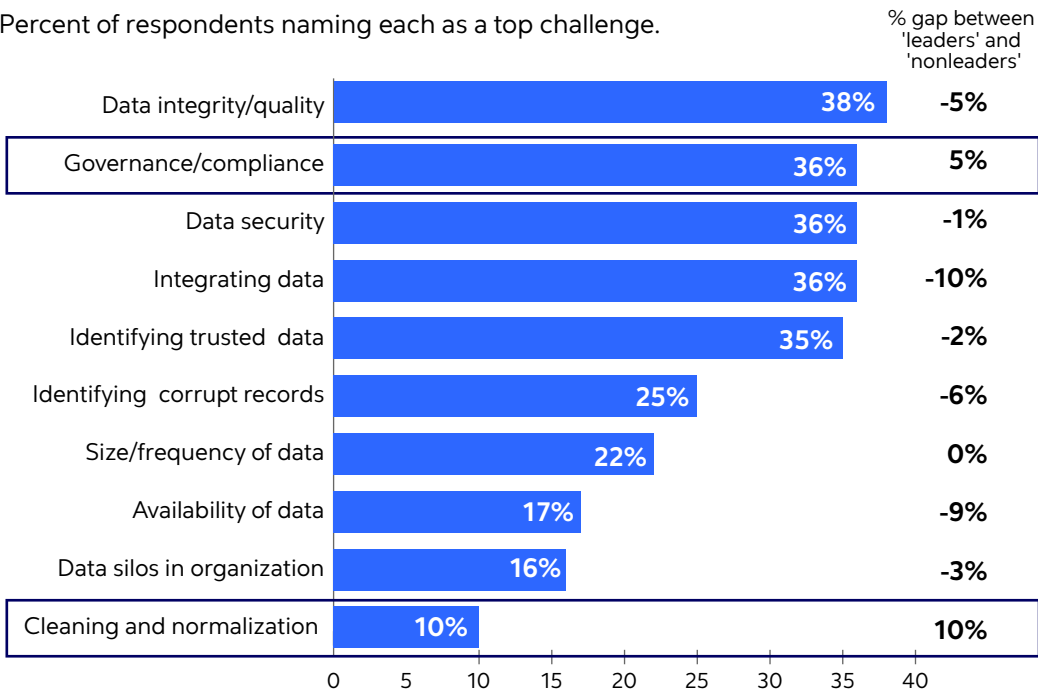


Perishability and other data challenges

With data modernization comes challenges – most of which don’t ease up to any great degree as maturity increases (see Figure 7). This makes sense, given that advanced AI maturity also means more scaling of pilot solutions, more use of diverse data sets and a greater shift toward modernizing data (i.e., through the use of data lakes and the cloud) vs. simply managing data (i.e., with databases and storage).

Compliance, for example, becomes more difficult as organizations scale their AI solutions around the world, and cleaning and normalizing data becomes twice as difficult for leaders compared with non-leaders as they leverage richer data sets. Identifying trusted data continues to be a challenge as businesses turn to third-parties for external data, as does ensuring data integrity as businesses work to ensure ethical algorithmic decisions. In short, as AI gets used for more powerful business outcomes, the responsibility grows to meet ever higher standards.

Top 10 data challenges



Note: Multiple responses permitted.
Response base: 1,200
Source: ESI ThoughtLab/Cognizant
Figure 7

As businesses embrace a wider selection of data types, a particular challenge is managing data perishability – ensuring the data used for insights is current, accurate and relevant. Unlike ERP and other structured data, more dynamic data types – such as IoT, social, real-time, geospatial and psychographic – come with a shorter expiration period. By using machine learning, businesses can continually assess the timeliness, accuracy and relevance of their data and analytic models, testing millions of new models against real-world data to continuously provide new scenarios and ranges of forecasts for changing conditions.



As businesses embrace a wider selection of data types, a particular challenge is managing data perishability – ensuring the data used for insights is current, accurate and relevant.