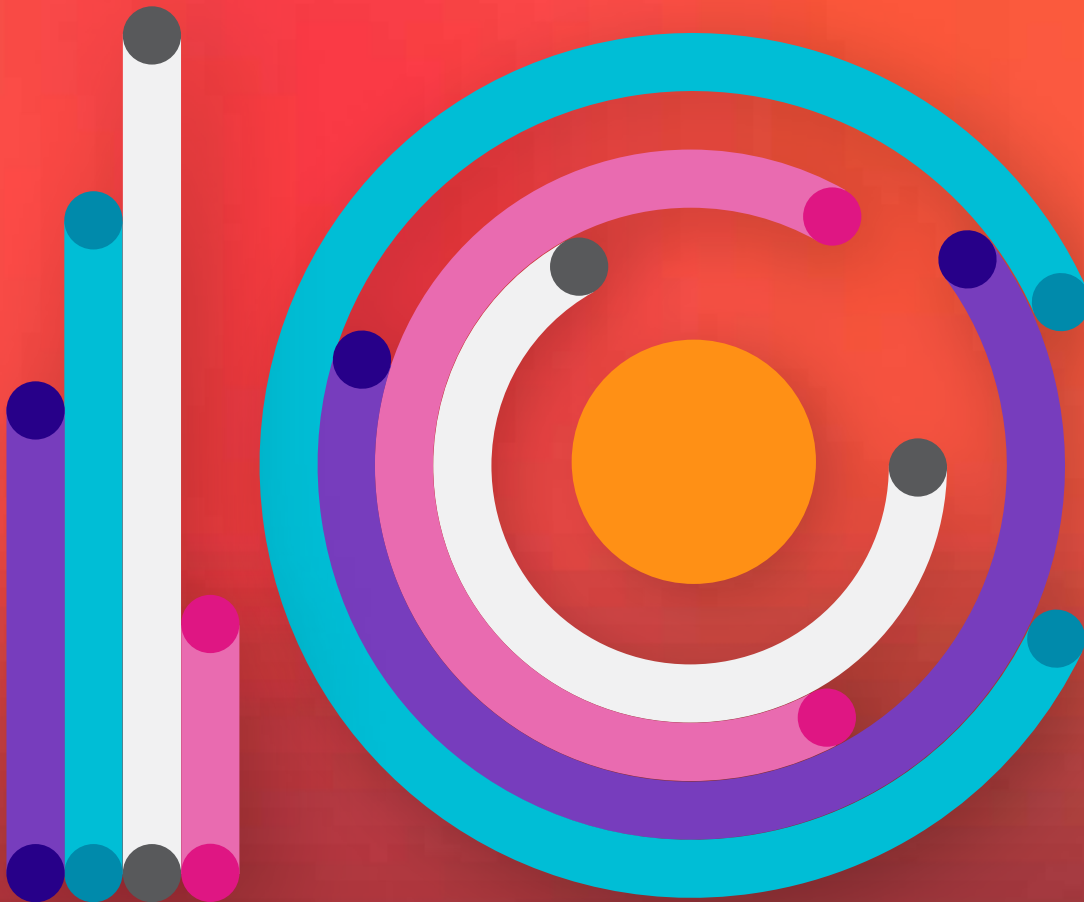# ThoughtSpot

Search & AI-Driven Analytics

TOP 10

EVALUATION CRITERIA

Introduction

It's been said that data is the new oil. An explosion of data sources and new technologies for capturing them are creating massive opportunities for companies. But in this new quest for insights, the last mile of data access remains the biggest obstacle.

Search has transformed our lives and along with it, our expectations for fast and easy access to information. In addition, smart content applications like Netflix or YouTube have been using the power of AI to automatically generate content recommendations that would be relevant to the end user.

These applications have become so fundamental, that it's hard to even think about what life was like before the power of search and AI, back when we were dependent on experts to get us access to the information we now get in seconds. Unfortunately, the BI industry today still feels a lot like what life used to be like in our personal lives back when we were dependent on experts. With search and AI-driven analytics, we believe it's possible for every human-being in your organization to access their data and get insights faster than ever before.

The hype is behind us. It's now time to evaluate today's search and AI-driven analytics vendors on what matters most to creating new insights: ease of use, data volumes, user scale, and whether you will need an army of consultants to integrate these new technologies into your existing BI and analytics environment.

In this book, we present ten different criteria that you can use to evaluate search and AI-driven analytics products - everything from search intelligence, to automated insights, data modeling, and total cost of ownership.

Table of Contents

1

Training Time

Despite $69B spent annually on BI software and services, there's only 22% adoption in the enterprise.

Traditional BI products require you to take multi-day classes or get certifications before you can use them.

Meanwhile, over a billion people use Google every day. Do you remember going to your first Google training class?

3 Days

Average duration of a beginner BI training class

The Less Training Needed, the More Adoption Grows

Most BI products are designed for business analysts who need to go to a week-long training class to become productive. Even IT teams need training to support these products effectively.

This training requirement and the continuous need to stay on top of technical skills is why the BI industry is plagued by such a terrible adoption problem (22%).

In contrast, today's most popular consumer tech services that are driven by a search interface don't require any training. Google, Yelp, Uber, Mint, Amazon, and many others rely on search to drive their user experience. If you had to go to a training class to use those products their adoption would be terrible, too.

This is the reason consumer companies measure their adoption in millions, while enterprise technologies measure in thousands.
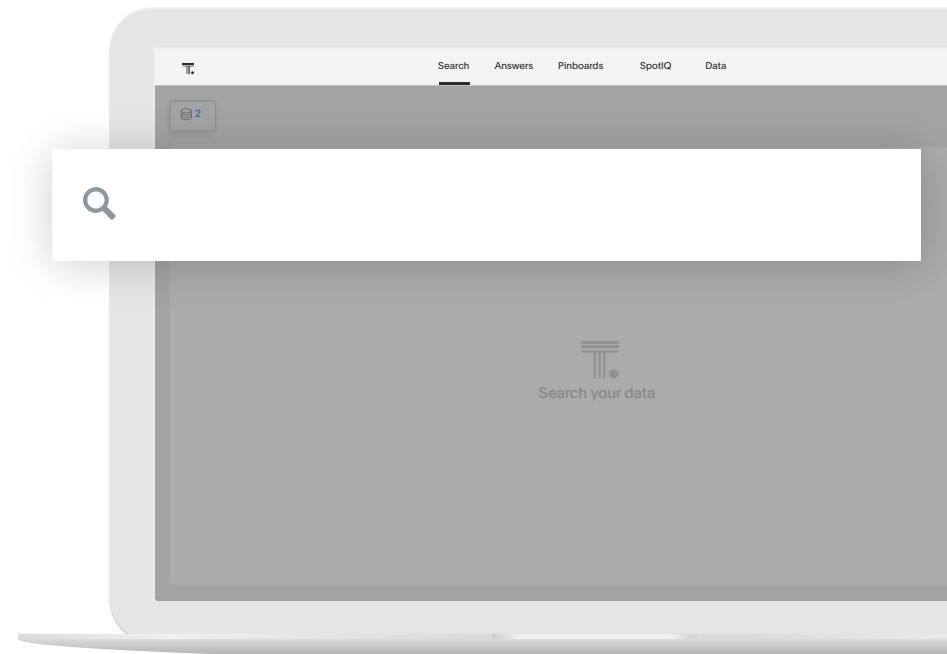
Ask vendors for the length of a typical training session for non-technical users, business analysts, and IT and BI teams.

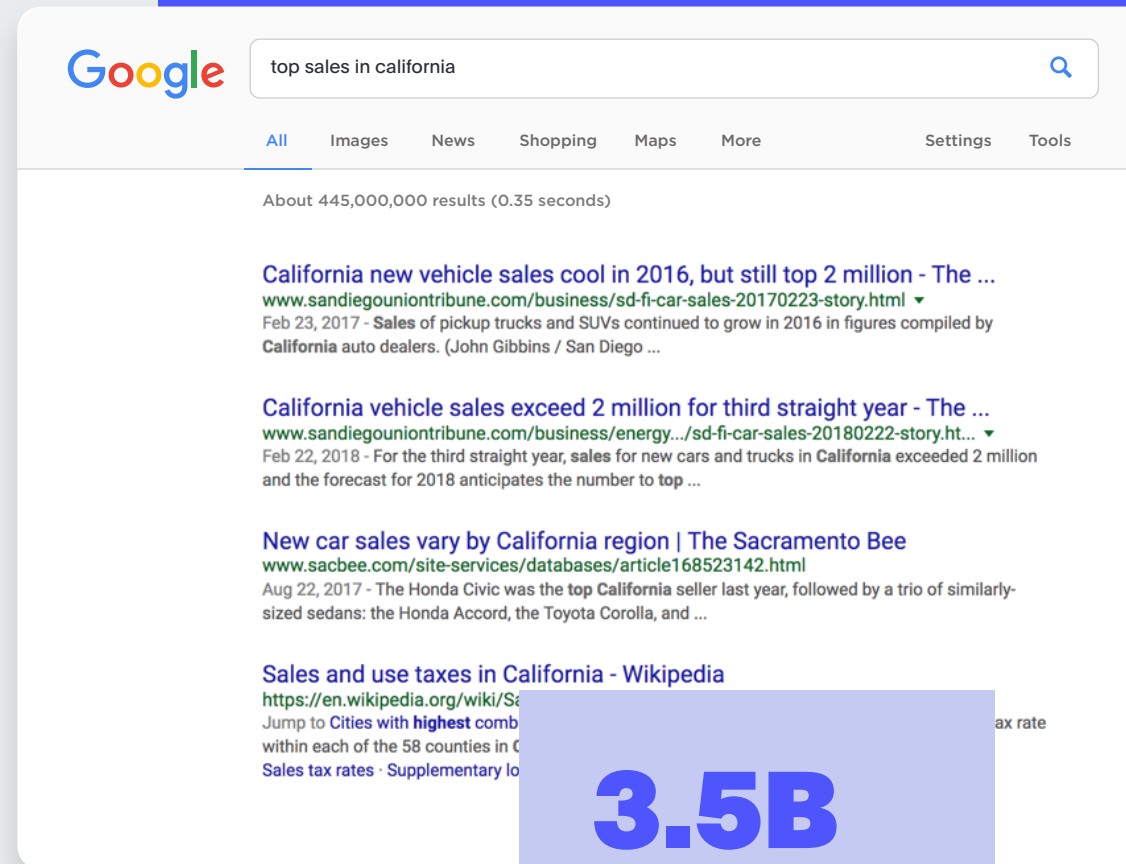*64% of business users are confused by legacy BI interfaces.*

2

Search Experience

You use search every day on consumer websites such as Google, Amazon and Facebook. All three are similar, but work slightly differently. Google returns lists of web pages, Amazon lists of products, and Facebook lists of friends and events.

Most BI products have search boxes designed similarly to return ranked lists of pre-built reports of dashboards.

But for search to reach the next level in BI, a fundamentally different approach is required. If you type "revenue last year in California", you don't want a list of ranked reports and dashboards. You want a single number. This requires a new kind of search experience designed for numbers that is very different from the search engines powering the consumer web.

Google   top sales in california

All   Images   News   Shopping   Maps   More   Settings   Tools

About 445,000,000 results (0.35 seconds)

California new vehicle sales cool in 2016, but still top 2 million - The ...
www.sandiegouniontribune.com/business/sd-fi-car-sales-20170223-story.html ▾
Feb 23, 2017 - Sales of pickup trucks and SUVs continued to grow in 2016 in figures compiled by California auto dealers. (John Gibbins / San Diego ...

California vehicle sales exceed 2 million for third straight year - The ...
www.sandiegouniontribune.com/business/energy.../sd-fi-car-sales-20180222-story.ht... ▾
Feb 22, 2018 - For the third straight year, sales for new cars and trucks in California exceeded 2 million and the forecast for 2018 anticipates the number to top ...

New car sales vary by California region | The Sacramento Bee
www.sacbee.com/site-services/databases/article168523142.html
Aug 22, 2017 - The Honda Civic was the top California seller last year, followed by a trio of similarly-sized sedans: the Honda Accord, the Toyota Corolla, and ...

Sales and use taxes in California - Wikipedia
https://en.wikipedia.org/wiki/Sa
Jump to Cities with highest comb   ax rate
within each of the 58 counties in C
Sales tax rates · Supplementary lo

3.5B

### searches per day on Google

Not All Search is Created Equal

Many BI products advertise a search box. It is important to understand how each of them work. Does it only search pre-built reports and dashboards? Does it only look at metadata? Does it merely return a list of matches? Does it use any guesswork in estimating results? Or does it provide a single answer?

Some approaches rely on programmable algorithms that interpret what the user is asking and provide error-prone estimates for answers. Others modeled after web search return a long list of ranked search results of pre-built reports that the user has to wade through.

Meanwhile, the newest breed of search-driven analytics engines search through all the underlying raw data, compute results, and then present charts and numbers based on those real-time calculations

"Search" has many flavors - document, metadata, dashboards, or numbers. Determine which best meets your needs.
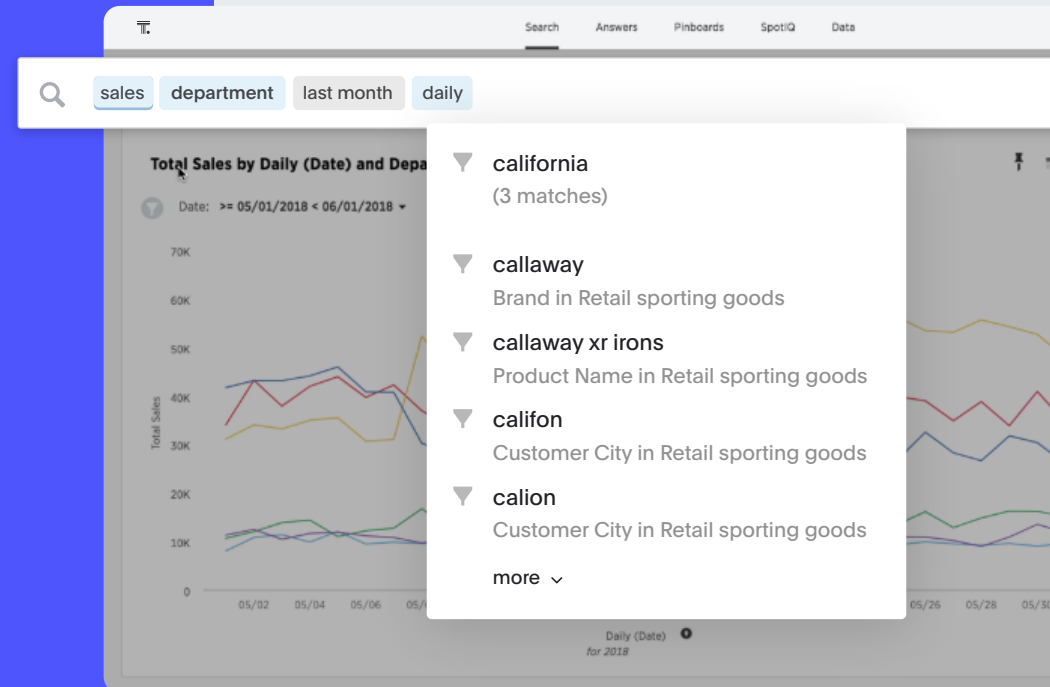
> **By 2020, 50% of analytical queries will be generated via search, natural language processing or voice, or automatically generated.**
>
> *Source: Gartner*

3

Search Intelligence

Google changed consumer search forever when it invented the PageRank algorithm that ranked pages by how many other pages link to them. This was different from how Facebook grew using graph search for social networks, or how Amazon's faceted search made it easy to browse large catalogs.

Search technologies in the BI world today mostly equate to a BI analyst either setting up a database of pre-defined search terms and answers for a business user to "discover", or providing search-based access to saved reports and dashboards.

What is more rare but more useful is a search engine designed for numbers, one that can look directly at raw data and compute results on-the-fly with 100% accuracy.
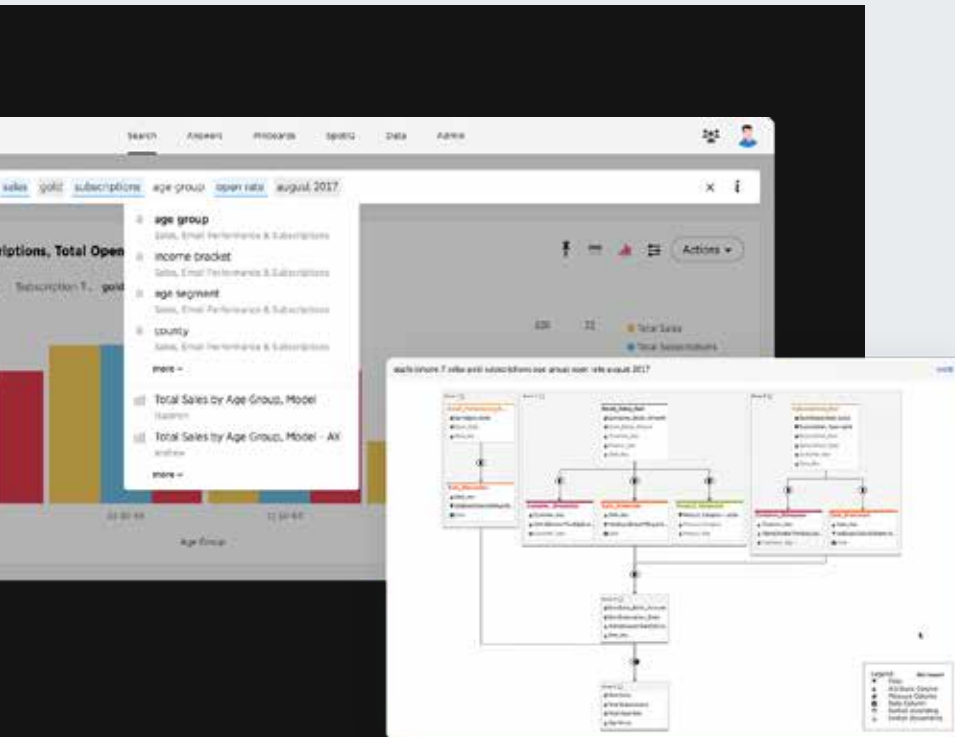
33%

percentage of users who click on the first Google link
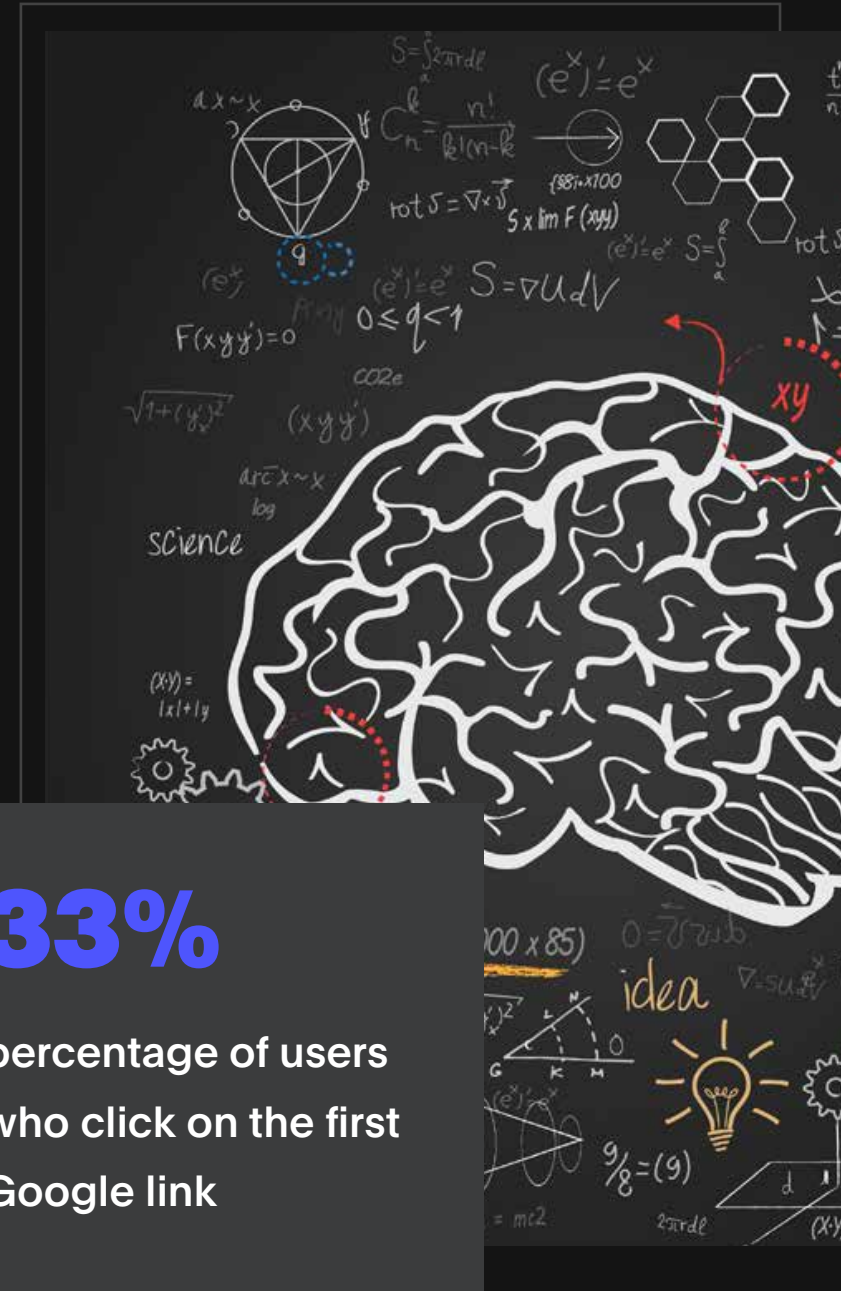
Accuracy Builds Trust. Trust Drives Adoption.

Business users need to be able to trust the numbers they get from a BI solution. A search-driven analytics engine should provide a single consistent and reliable answer - always.

Some methods such as NLP provide probabilistic results based on programmed algorithms that must be constantly refined. Even after months of tuning, they still have a 10-20% error rate.

Most users don't understand how all their data relates to each other, or which schema represents the underlying tables, or which joins are needed to find an answer. A smart search-driven analytics engine should hide all such complexity away from the user.

Users need a search experience that recognizes patterns, understands synonyms, has spell check, and offers suggestions as they type based on other users' activity - similar to Google's type-ahead feature.

It's also critical for a user to easily analyze results at different time granularity (daily, weekly, monthly, etc...) without waiting for the BI team to create new cubes or aggregate tables. Search-driven analytics solutions should do this automatically and compute results across billions of rows of data in under a second.

Finally, a good search-driven analytics experience should provide a way to verify how results were calculated, without requiring users to learn SQL or other programming languages.

Ask if search results are calculated on the fly or retrieved from pre-calculated aggregate tables. Are the results accurate or estimates?

4

Augmented Data Discovery

As consumers, AI is at work all around us. Playlist curation and content recommendations on sites like YouTube and Netflix are examples of AI and machine learning as the system automatically learns each user's preferences from his/her interaction with the content, without any explicit action from the user.

In the world of data and analytics, while data volume is growing exponentially, the volume of insights we're able to extract from it is fundamentally limited. That's because in today's analytics paradigm there's a huge gap between data supply and data demand.

Infusing AI into analytics workflows can transform your organization and bridge the supplier-consumer divide by giving everyone access to the tools they need to make data-driven decisions.
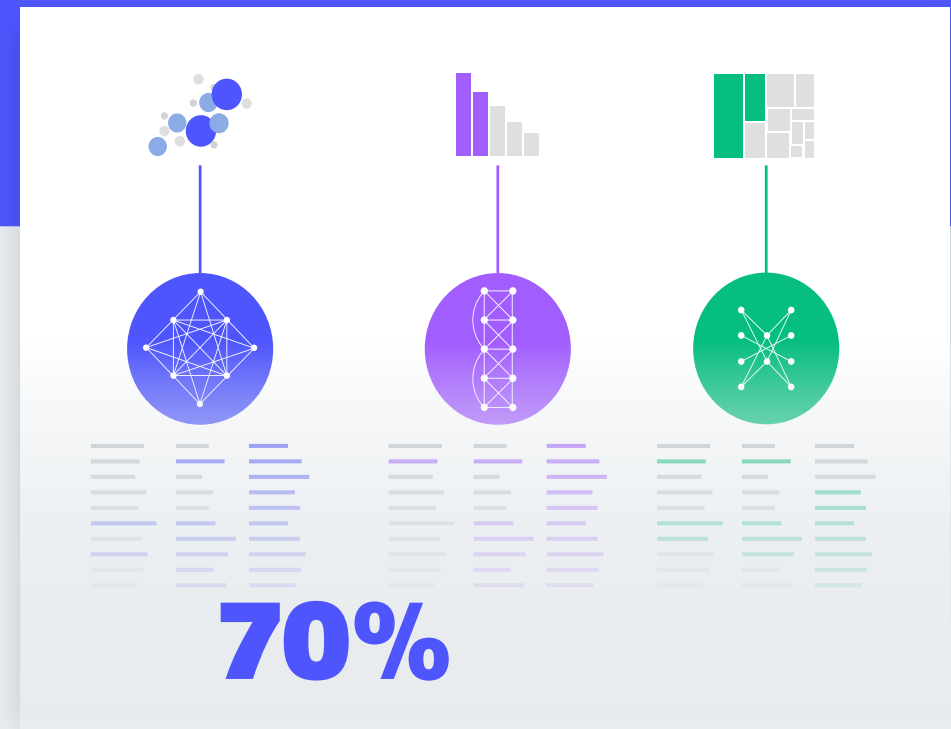
70%

percentage of content consumed on Netflix curated by automated recommendation engine

Personalized Automated Insights When It Matters Most

Finding the most relevant answer of your data questions is often a never-ending exercise of trying to find a needle buried deep in a haystack. It is not practical for a human to ask all possible questions on the data, let alone know all the questions to ask.

Now imagine if an intelligent and powerful machine could access numerous data sets, generate thousands of questions, analyze billions of data points, spot hidden trends and anomalies, and proactively push relevant and personalized insights to you, all in seconds - with a single click of a button. That is the power of augmented data discovery.

The number of possible questions to ask of data is often too much for any human. With automated data discovery technologies, business people can rely on machine-driven smarts to explore complex datasets with a few clicks and get insights explained to them in natural language, without the need for a trained analyst and the hours of time it would take them to explore the data manually and build a report. Instead, data experts can focus on data governance, building bulletproof data models, preparing new datasets for analysis.

Machine-generated insights also help to minimize errors in analysis and eliminate human bias, bringing to our attention

new metrics and business drivers that weren't considered before. However, the key to adoption of AI-driven analytics is trust. When it comes to analytics, trust is created by delivering accurate, relevant, and transparent results. To do this, machines should not rely solely on their own built-in learning algorithms but must work together with humans, and learn from usage behavior to ensure every result meets these standards of trust.
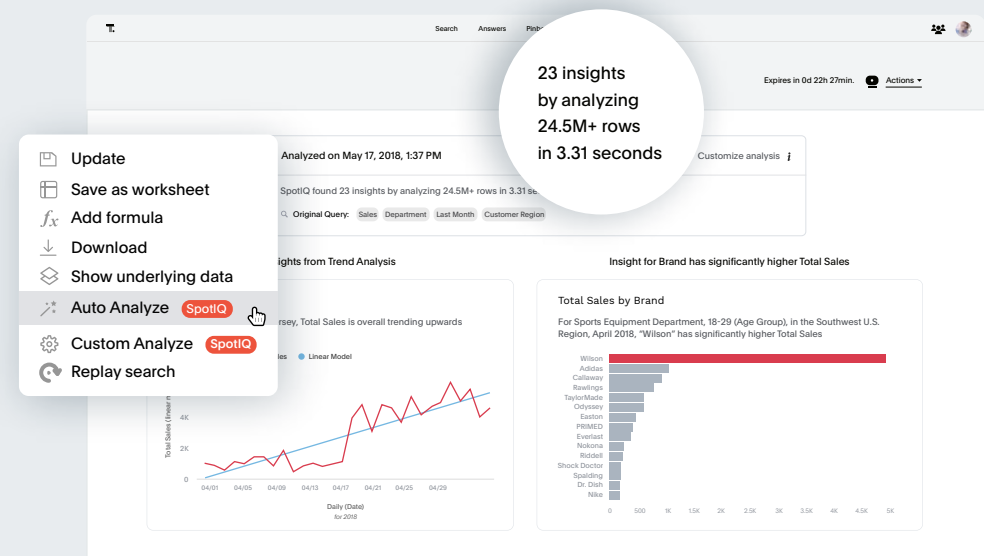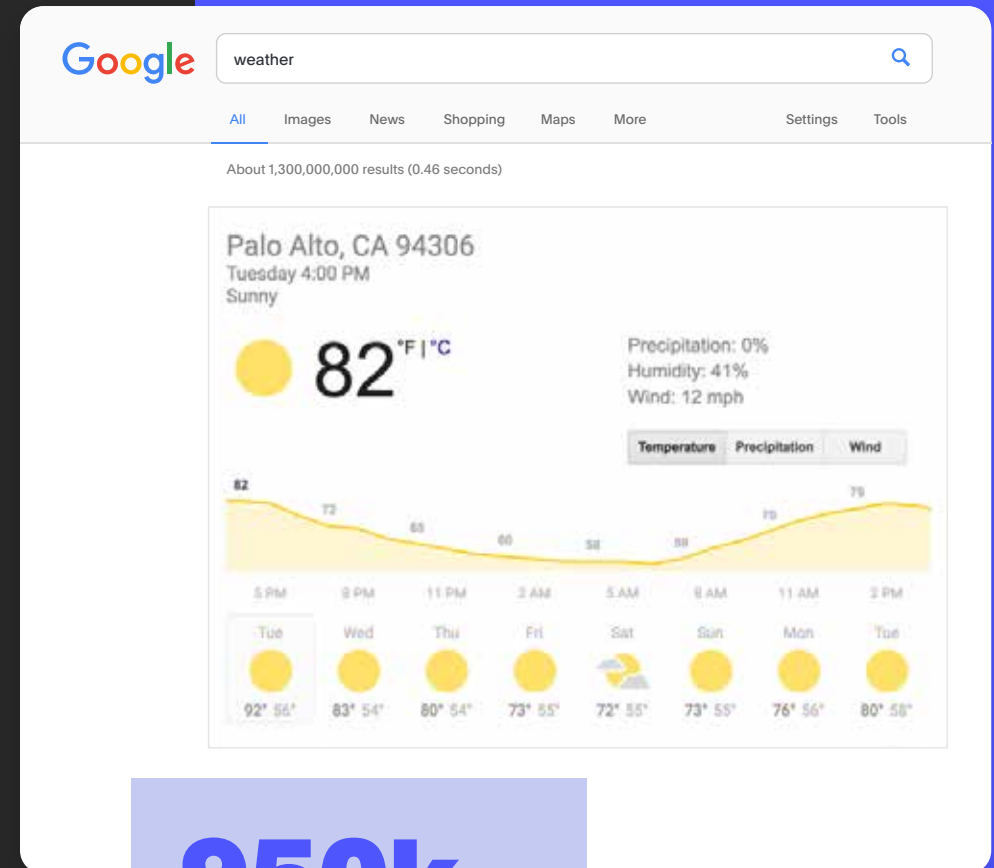
5

Chart Creation

Isn't it amazing when you type the term "weather" into Google you instantly get current and forecasted conditions for the city you're in along with a "card" visual showing you a picture of a sun or cloud? The app knows exactly what you're looking for and presents the information in the easiest way to consume it.

Contrast that with legacy BI products: after days of training, you still need to remember how to click eleven times in order to build a chart and then decide if it has the information you seek.

Google — weather

All | Images | News | Shopping | Maps | More — Settings | Tools

About 1,300,000,000 results (0.46 seconds)

Palo Alto, CA 94306
Tuesday 4:00 PM
Sunny

82 °F | °C

Precipitation: 0%
Humidity: 41%
Wind: 12 mph

Temperature | Precipitation | Wind

950k

hours per day saved by Google's Top Stories cards.

# The Best Visualizations Create Themselves

In today's world where search pervades our consumer experience, search and speed have become synonymous. If a search-driven analytics product is to be adopted widely, it needs to cut down any unnecessary wait time between the user's query and the visualized results.

An important part of this process is to decide intelligently the best chart type for the user's query and instantly return a visual along with an answer. But data is complicated. Picking axis and chart types is hard. This is a situation in which machines trump humans. Any assistance a user can get goes a long way toward adoption and insight. Then if the user wants to change the chosen chart type, they should always have the option.

> "
> **Only 23% of current BI users are comfortable creating charts & graphs.**
> *Source: TDWI*
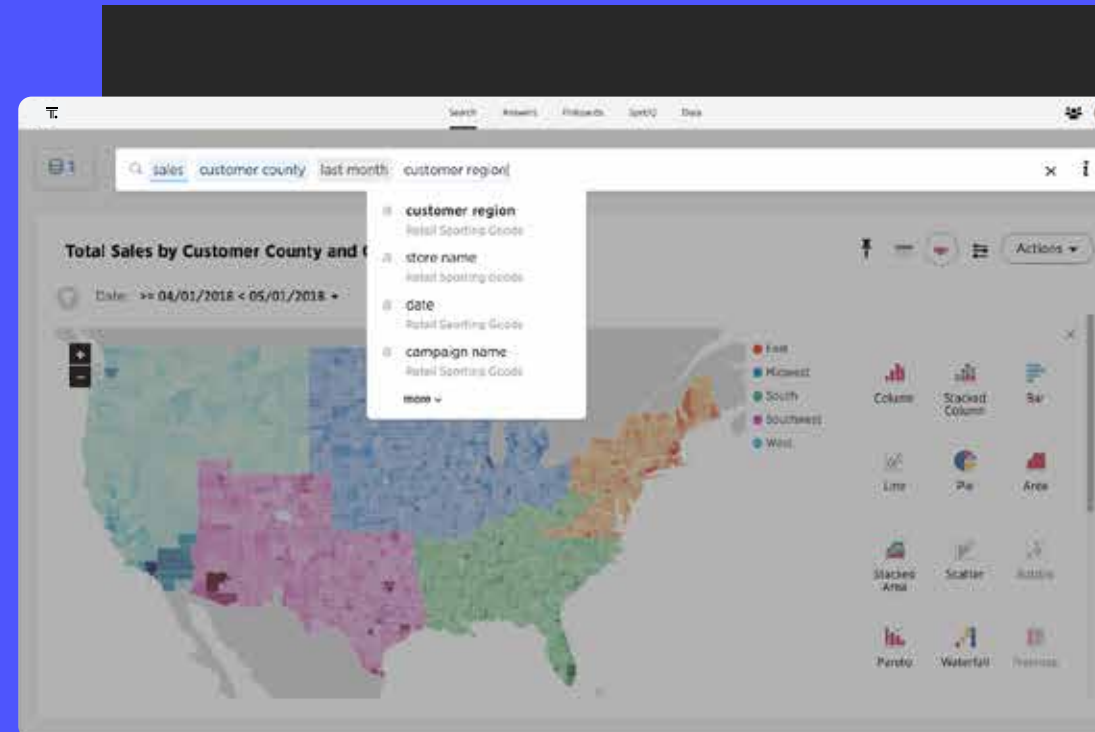
Count the number of clicks it takes to create a chart.

6
Speed
at Scale

The power of Google is that it delivers the one-two punch of a simple search experience done at massive scale. Using a search bar is simple and intuitive, but the most powerful part of Google is its ability to search everything across the web.

If Google was restricted to the files on your local machine it would be significantly less useful. Yet in the BI world, so many products offer restricted views into your data, that do not scale across the enterprise, across thousands of users, or across large volumes of data and data sources.

40%

percentage of people who abandon a website that takes more than 3 seconds to load

Speed at Scale is the Secret to Search and AI-Driven Insights

Mid-to-large size enterprises have hundreds of tables, billions of rows, and thousands of users. The key to providing insights is delivering a simple search experience at scale and still returning answers to the user in less than a second.

Studies have shown that if a user doesn't get a result from Google in less than three seconds, they abandon the page. Compare that statistic to waiting overnight for a big report to run in a legacy BI product and, again, it's not surprising that there's an adoption issue in the industry.

Meanwhile, some of today's faster more popular data visualization tools are desktop products that can't handle data sets larger than a few gigabytes. With hundreds of gigabytes created quarterly by the average enterprise, BI teams are faced with the challenge of determining which datasets are most important for different types of users. It's a continuous task that always leaves users wanting more.

If the technology doesn't scale with speed, your BI project is destined for problems.
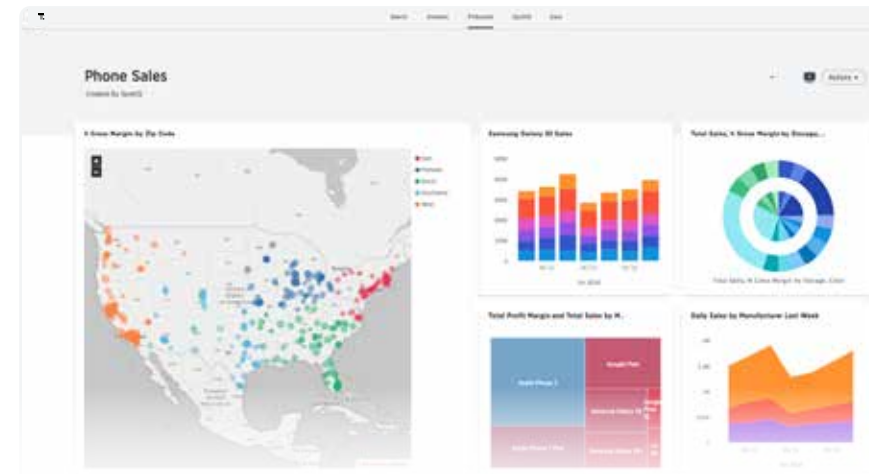
350 TB

average amount of data enterprises store

Ask how much data the product can handle. And how many users it can support simultaneously.

7

Data Modeling

IT teams spend too much time modeling data.

Data modeling headaches are the reason enterprises spend nearly 3 times more on BI software services than on software licences. It's why entirely new careers like "data wrangling" have emerged.

Creating cubes and aggregate tables for individual lines of business is not the best use of time for BI teams, especially when tactical dashboards may not have the answer an end business user needs.

Consumer search technologies have enabled untrained users to search through complex product catalogs, network graphs, and any type of document imaginable on the web. Why can't the enterprise user do the same with their data?

80%

percentage of time a data scientist spends modeling and preparing data for analysis.

Minimize Modeling to Reduce Professional Services Spend

A traditional BI environment takes months of modeling - building OLAP cubes or aggregate tables, and significant database tuning before any results can be exposed through a search interface. On an ongoing basis, these databases need maintenance and care, which sucks up even more time and resources.

Other systems based on NLP techniques require a significant professional services spend to build semantic search models for each implementation. Then, even after months of tuning from the world's top experts, they only yield 80-90% accuracy.

Meanwhile, some search-driven analytics products are schema-aware and able to remove a significant amount of modeling complexity. Schema-awareness means the search engine understands the relationships between different sources of data and it is able to relate them together automatically, even for complex models beyond traditional star or snowflake schemas.

A complicated product typically comes with an expensive professional services engagement in order to get it to work. Better products will free up BI teams to focus on higher value problems like data governance and data quality.
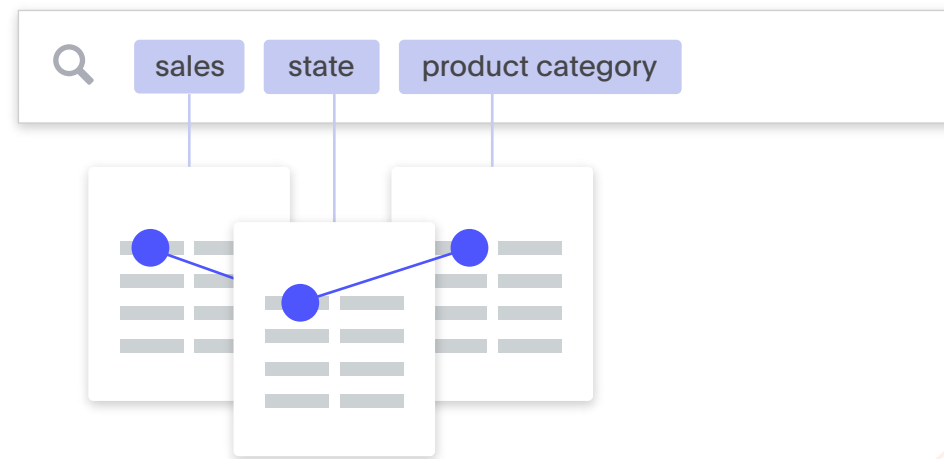
Find out how long a typical implementation takes before you can start using the product, and whether it can handle the complexity of your data model.

8

Data Environment

When it comes to data access within the enterprise, the last mile is always the hardest, even more so when the data is split across several sources requiring different data integration tools. The entire process of getting useful data into the hands of business users can take months, which no company can afford to waste.

Businesses need to gather insights from external data sources just as easily as they would from their internal systems. Google compiles search results from a variety of sources. Why should enterprise BI tools be any different?

Search and AI-driven analytics should accomplish this with the same ease of use we expect from consumer technology.

6%

organizations that have all their data in one place.

Search Should Analyze Any Source

The ability to search data at scale from a variety of sources is essential to a productive business user. In the same way Google combines search results from across the entire web, search-driven analytics solutions should be capable of analyzing search results across tables from different databases, applications, spreadsheets, or Hadoop clusters.

For this to happen, the search-driven analytics solution has to be compatible with your existing data environment - different types of data sources, as well as different data integration or ETL technologies.

Instead of learning to use different BI products for different types of data sources, one search-driven experience for all data sources lowers the bar for business users and makes significant adoption more likely.
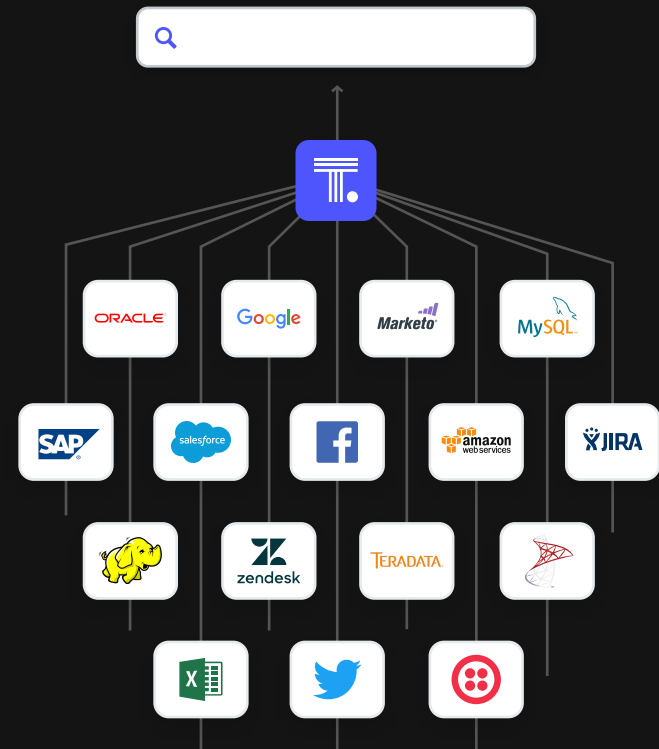
Ensure the product can search through data from any source you might need to analyze.

> "
> *Speed of insight and breadth of data sources are the critical factors to help stand out in the marketplace.*
>
> *Source: TDWI*

18

9

Data Security & Governance

Securing data within the enterprise is a solved problem. The best BI vendors already offer that. But packing all of those security requirements into a sophisticated search bar? Now that's a different story.

How do you ensure that even the search suggestions obey security restrictions? In other words, how do you secure the search intelligence at a user level?

This is a unique challenge in the enterprise that even the likes of Google haven't had to tackle for consumer search.

90%

percentage of IT professionals that say data security is a top concern.

Security Should Be Built Into the Results & Search Box

A good search interface needs to be able to access all data across the enterprise, while limiting access to only what each user is supposed to see. It should be able to integrate easily into the existing directory services through LDAP or similar protocols.

The underlying data needs to be secured at a row, column, and table level. An employee table might have a compensation column that is visible only to select users. A sales table might have rows of sales information by region that can be seen only by reps in that region. And table level protection should ensure that departments can see only their own tables.

An enterprise-class search-driven analytics experience needs to honor access privileges, while accessing billions of rows of data, and returning results in under a second.

Verify that both the search box and search results obey your access rules and users see only what they are allowed to see.

East Manager

Q sales this quarter in |

Q sales this quarter in New York

Q sales this quarter in New Hampshire

Q sales this quarter in New Jersey

East Manager

Q sales this quarter in |

Q sales this quarter in California

Q sales this quarter in Oregon

Q sales this quarter in Nevada

10
Cost

Business users today often wait months to get access to new BI products thanks to lengthy deployment cycles. Cobbling together different pieces of infrastructure to get your BI environment up and running is a nightmare for most IT organizations. There's a huge cost to implementing and an arguably even greater opportunity cost to waiting for insights.

Best-of-breed BI solutions should work right out of the box, with minimal implementation headaches - just like your personal computer or favorite consumer app.

80%

percentage of BI dollars spent on services to make the software work.

# Understand the True Cost of Democratization

Time to value is the first thing to evaluate. Will the product take months to deploy? Weeks? By eliminating data modeling, cube building, semantic modeling, and hardware tuning, search and AI-driven analytics products can be up-and-running in a matter of hours.

Beyond implementation and licensing, the true cost of many BI solutions include hardware, tuning and storage costs, training costs, IT maintenance and support, and user training costs. These occur after the initial implementation and can have a major impact on ROI. Modern search-driven products drastically reduce these costs.

Then there's the financial impact of user adoption. For many BI products today, more than half of the usage is attributed to simple report and dashboard viewing. This means the user logins are simply replacing emailed PDF reports - thereby making the cost of those user licenses hard to justify.

A modern, well-designed search experience should go far beyond scheduled reports and give business users the ability to answer ad hoc questions on the fly. It should be addictive and spread quickly within an enterprise.

As adoption builds, it's important to evaluate the per user costs and not artificially penalize new users. When software works well, adoption should be both contagious and economically beneficial.
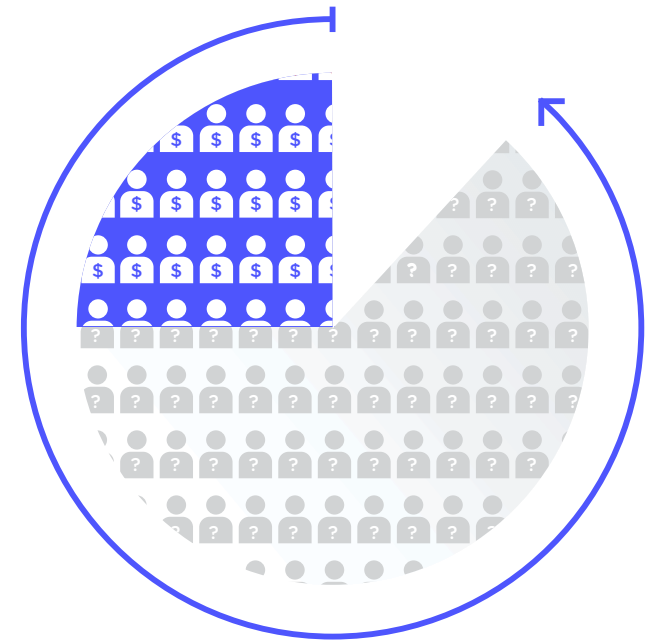
Understand hidden implementation and maintenance costs. Ensure that wide adoption is not gated by high per-user license costs.
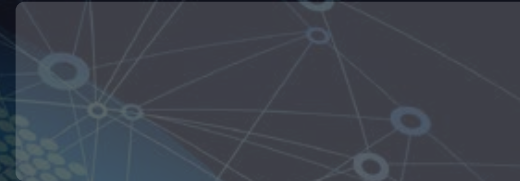
Conclusion

Search and AI has infiltrated every aspect of our consumer tech lives and is now making bold new strides into enterprise software. Products that offer search and AI-driven analytics are poised for rapid growth because they bring both speed (instant results) and scale (billions of rows) to business intelligence. With so many approaches, it is critical to understand the differences between vendors before making a significant investment. We hope this framework proves useful as you begin delivering instant answers to every business user in your company.

ThoughtSpot

ThoughtSpot, the leader in search & AI-driven analytics for enterprises, is helping the largest companies in the world succeed in the digital era by putting the power of a thousand analysts in every business person's hands. With ThoughtSpot's next-generation analytics platform, business people can use Google-like search to easily analyze complex, large-scale enterprise data and get trusted insights to questions they didn't know to ask, automatically - all with a single click. ThoughtSpot connects with any on-premise, cloud, big data, or desktop data source, deploying 85 percent faster than legacy technologies.

For more information please visit www.thoughtspot.com.