January 16, 2018

Dresner Advisory Services, LLC

2018 Edition

Analytical Data Infrastructure Market Study (Excerpt)

Wisdom of Crowds® Series

Licensed to Snowflake

Disclaimer

This report should be used for informational purposes only. Vendor and product selections should be made based on multiple information sources, face-to-face meetings, customer reference checking, product demonstrations, and proof-of-concept applications.

The information contained in all Wisdom of Crowds® Market Study Reports reflects the opinions expressed in the online responses of individuals who chose to respond to our online questionnaire and does not represent a scientific sampling of any kind. Dresner Advisory Services, LLC shall not be liable for the content of reports, study results, or for any damages incurred or alleged to be incurred by any of the companies included in the reports as a result of its content.

Definitions

Business Intelligence Defined

Business intelligence (BI) is "knowledge gained through the access and analysis of business information."

Business Intelligence tools and technologies include query and reporting, OLAP (online analytical processing), data mining, and advanced analytics, end-user tools for ad hoc query and analysis, and dashboards for performance monitoring.

Source: Howard Dresner, *The Performance Management Revolution: Business Results Through Insight and Action* (John Wiley & Sons, 2007)

Analytical Data Infrastructure Defined

Analytical data infrastructure (ADI) defines a set of technology components for integrating, modeling, managing, storing, and accessing the data sets that serve as sources for analytic/BI consumers, e.g., analytic/business applications, tools, and users. In this study, we examine four common ADI use cases: 1) data science, 2) embedded analytic applications, 3) business user exploration, discovery, and 4) reporting and dashboards.

ADI data sources include, for example, SQL/Transnational sources, Excel/CSV files, events and log streams from operational applications, IoT data, metadata, multimedia etc. ADI development and support/administration includes tools for application development of data and analytic use case workflows with associated data life cycle management and governance, metadata, and security. Scale of applications can range from individual users (i.e., data science use case) to embedded applications supporting large numbers of users and involving several data sources.

ADI examples range from on-premises deployments such as a Microsoft SQL Server-based analytics platform system to a hybrid combination of on-premises and cloud service deployments of open source and commercial technology as in the Teradata Analytics Platform to pure cloud services (e.g., RedShift, Kinesis, etc.), deployed as a part of Amazon Web Services Data Analytics Products.

Executive Summary

- We see a trend potentially emerging in ADI use cases changing priorities. While respondents indicate "business user reporting and dashboards" and "business user discovery and exploration" are the top ADI use cases, both declined slightly in terms of driving respondents ADI priorities year over year.
- In 2018, performance and security are the top two selection priorities for ADI platforms, compared to performance and usability as the top two selection priorities in our 2017 report. Price is one of the lowest ADI selection criteria in 2018.
- On-premises deployments of ADI platforms lead in priority over cloud deployments, but the priority of deployment option varies by use case. Respondents' priority for hybrid deployment (a mix of on-premises and cloud) increases slightly year over year.
- Licensing preferences vary widely, with respondents indicating their top preference is concurrent use, and data volume-based licensing as their least preferred licensing option. Subscription-based licensing increases in preference year over year.
- Scale up / scale out capabilities (clustering, load balancing, high availability, etc.) and data life cycle management (lineage, impact analysis, metadata, governance) lead respondents' ADI development and deployment feature priorities.
- Transactional data is the top data type priority for ADI platforms in 2018, followed by Metadata, which increases in priority compared to 2017. Text, machine/log data, and images/video data type support also increase in year-over-year priority for ADI platforms.
- ETL and bulk load/transformation lead data preparation and loading priorities again this year. All other data preparation and loading capabilities increase in importance year over year.
- Use of R, machine learning, MapReduce, and Spark show the largest year-over-year increase in terms of ADI analytical feature priorities.
- Most functions place similar priorities on data preparation and loading functions; however, the BICC function places a slightly higher priority on "ELT / ETL – bulk load" capabilities.
- Year over year, all data model types other than SQL data increase in importance, reflecting the changing priorities towards more performance and non-SQL data model requirements of analytic applications.
- The ADI interface priority for MDX interfaces is low in general, indicating a potential low use of multi-dimensional data models.
- While the Apache interfaces are low in terms of relative priority in 2018, we see year-over-year growth in importance.
- Finance respondents place their highest priority on Excel/CSV interfaces and their lowest priority on Apache services. They also place a low priority on MDX interfaces.

- Many respondents indicate they “don’t know” or rank as “unimportant” such ADI analytical features as sentiment analysis, path/link analysis, custom R packages, and Spark.
- Performance is the top priority for the embedded analytics and organizations having multiple use cases.
- Data preparation and loading priorities vary more broadly by respondents’ use cases compared to other dimensions we analyze (i.e., geography, functions, industry, organizational size).

Vendor Ratings

In this section, we offer ratings of analytical data infrastructure vendors. We rate vendors using 33 different criteria, on a five-point scale for each. Criteria covers sales /acquisition experience (eight criteria), value for price paid (1), quality and usefulness of product (12), quality of technical support (5), quality and value of consulting services (5), whether the vendor is recommended (1), and integrity (1).

As we explore vendor performance in more detail, it is important to understand the scale we use in scoring the industry and vendors:

- 5.0 = Excellent
- 4.0 = Very good
- 3.0 = Adequate
- 2.0 = Poor
- 1.0 = Very poor

Based on our scoring methodology, all vendors perform at a level that is considered more than “adequate” for all criteria categories.

Please note that “average score” is the mathematical mean of all items included in vendor ratings. Each column in the chart represents a scale consisting of varying numbers of items (for example, "sales" is a scale consisting of eight items, while "value for price paid” is one item). As such, each column is weighted differently (based upon the number of items represented and the number of respondents rating those items) in calculating the overall average rating. The average score cannot be calculated by simply averaging across the subscale scores.

Analytical Data Infrastructure Market Models

Starting in 2015, we began using two new models for examining and understanding the analytical data infrastructure market. Using quadrants, we plot aggregated user sentiment into x and y axes.

Customer Experience Model

The customer experience model considers the real-world experience of customers working with ADI products daily (fig. 1). For the x axis, we combine all vendor touch points—including the sales and acquisition process eight measures), technical support (five measures), and consulting services (five measures)—into a single "sales and service" dimension. On the y axis, we plot customer sentiment surrounding product, derived from the 12 product and technology measures used to rank vendors. On the resulting four quadrants we plot vendors based on these measures.

The upper-right quadrant contains the highest-scoring vendors and is named "overall experience leaders." Technology leaders (upper-left quadrant) identifies vendors with strong product offerings but relatively lower services scores. Contenders (lower-left quadrant) would benefit from varying degrees of improvement to product, services, or both.

User sentiment surrounding outliers (outside of the four quadrants) suggests that significant improvements are required to product and services.

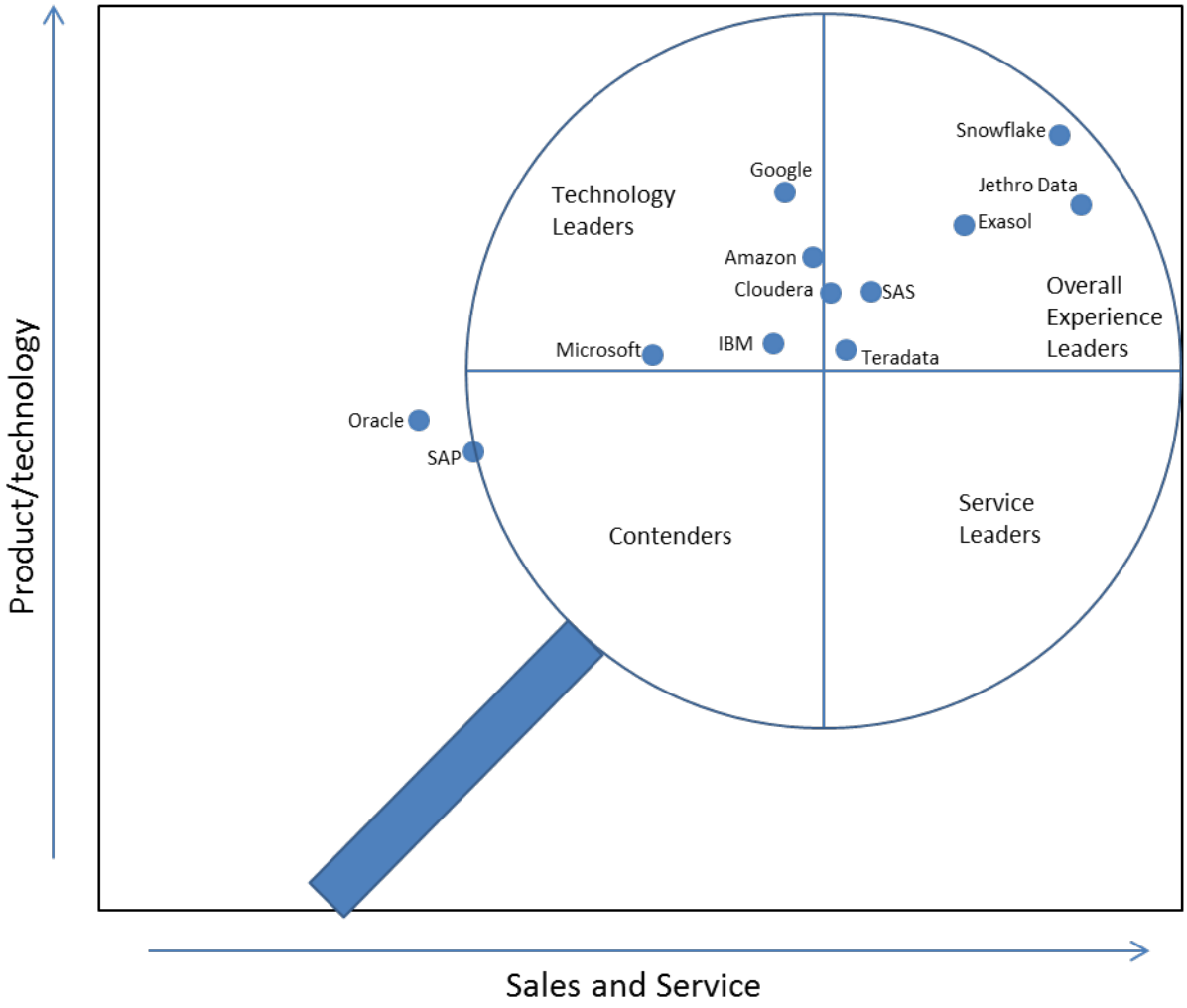

Figure 1 – Customer experience model

Vendor Credibility Model

The vendor credibility model considers how customers "feel" about their vendor (fig. 2). The x axis plots perceived value for the price paid. The y axis combines the integrity and recommend measures, creating a "confidence" dimension. The resulting four quadrants position vendors based on these dimensions.

The upper-right quadrant contains the highest-scoring vendors and is named "credibility leaders." Trust leaders (upper-left quadrant) identifies vendors with solid perceived confidence but relatively lower value scores. Contenders (lower-left quadrant) would benefit by working to improve customer value, confidence, or both.

User sentiment surrounding outliers (outside of the four quadrants) suggests that significant improvements are required to improve perceived value and confidence.

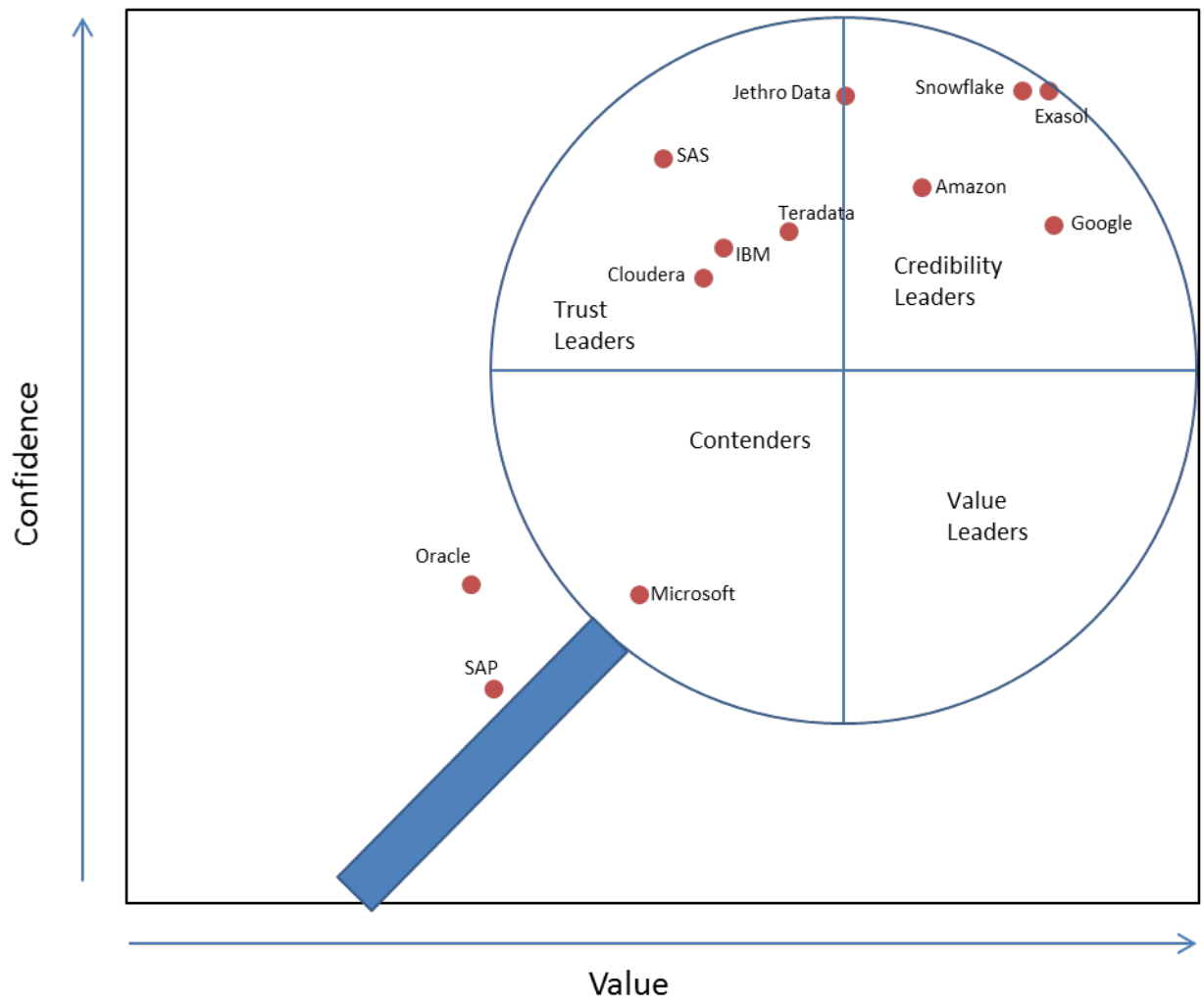

Figure 2 – Vendor credibility model

Detailed Vendor Ratings

In this section, we offer detailed vendor scores. Using our 33-criteria evaluation model (table 1), we compare each vendor's performance to their previous year's performance and to the average for all vendors (all records in the study population).

The detailed criteria are below. We added "clock" position information to assist in locating specific scores.

Table 1 - Detailed vendor rating criteria

- **Sales/acquisition experience** *(12 - 2 o'clock)*
  - Professionalism
  - Product knowledge
  - Understanding our business/needs
  - Responsiveness
  - Flexibility/accommodation
  - Business practices
  - Contractual terms and conditions
  - Follow-up after the sale
- **Value for price** *(3 o'clock)*
- **Quality and usefulness of product** *(3 - 7 o'clock)*
  - Robustness/sophistication of technology
  - Completeness of functionality
  - Reliability of technology
  - Scalability
  - Integration of components within product
  - Integration with third-party technologies
  - Overall usability
  - Ease of installation
  - Ease of administration
- **Quality and usefulness of product (continued)**
  - Customization and extensibility
  - Ease of upgrade/migration to new versions
  - Online forums and documentation
-
- **Quality of technical support** *(8 - 9 o'clock)*
  - Professionalism
  - Product knowledge
  - Responsiveness
  - Continuity of personnel
  - Time to resolve problems
- **Quality and value of consulting services** *(9 - 10 o'clock)*
  - Professionalism
  - Product knowledge
  - Experience
  - Continuity
  - Value
- **Integrity** *(11 o'clock)*
- **Whether vendor is recommended** *(12 o'clock)*

Snowflake Detailed Score

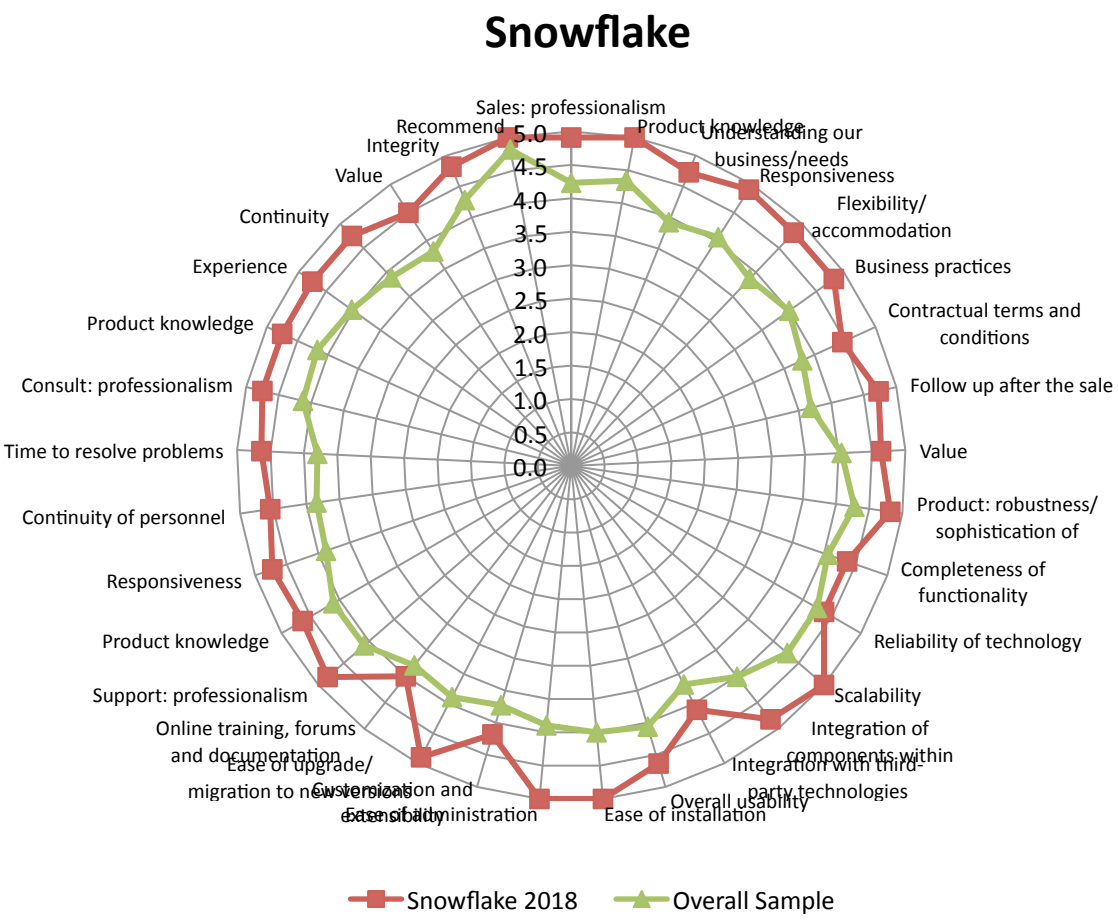

Figure 3 – Snowflake detailed score

This is the first year Snowflake has been included in these ratings. With scores substantially above the overall sample, Snowflake is an Overall Leader in both the Customer Experience and Vendor Credibility models. It is best in class for most sales, product/technology, and consulting measures. It was also best in class for integrity and has a perfect recommend score.

About Bill Hostmann

Bill Hostmann is a Research Fellow with Dresner Advisory. His area of focus includes trends in Analytic Data Infrastructures (ADI)—integrating and managing the information and information models used by BI, Advanced Analytics, and CPM/PM applications.

Bill has more than 20 years of product management experience at the intersection of business intelligence / analytics and data analytics infrastructure, including positions in product and general management at Gemstone Systems, Informix, and Informatica.

He spent 14 years as a research analyst at Gartner, including several years as a VP and Distinguished Analyst for BI/Analytics.

Bill served as conference chair of the Gartner BI/Analytics Conference for many years, growing the number of conference attendees from hundreds to several thousand attendees.

About Kathleen Goolsby

Kathleen Goolsby is Executive Editor with Dresner Advisory Services.

Kathleen has been a writer and editor for more than 20 years. In addition to her bylined content featured in various publications, she is an accomplished ghostwriter and has ghostwritten several business books as well as many feature articles, blogs, reports, and white papers.

She has interviewed more than 1,000 top executives at companies, associations and government entities around the world and written articles about their business issues and expertise.

About Howard Dresner and Dresner Advisory Services

DAS market studies are conceived, designed and executed by Dresner Advisory Services, LLC—an independent advisory firm—and Howard Dresner, its President, Founder and Chief Research Officer.

Howard Dresner is one of the foremost thought leaders in business intelligence and performance management, having coined the term "Business Intelligence" in 1989. He has published two books on the subject, *The Performance Management Revolution – Business Results through Insight and Action* (John Wiley & Sons, Nov. 2007) and *Profiles in Performance – Business Intelligence Journeys and the Roadmap for Change* (John Wiley & Sons, Nov. 2009). He lectures at forums around the world and is often cited by the business and trade press.

Prior to Dresner Advisory Services, Howard served as chief strategy officer at Hyperion Solutions and was a research fellow at Gartner, where he led its business intelligence research practice for 13 years.

Howard has conducted and directed numerous in-depth primary research studies over the past two decades and is an expert in analyzing these markets.

Through the Wisdom of Crowds® Business Intelligence market research reports, we engage with a global community to redefine how research is created and shared. Other research reports include:

- Wisdom of Crowds "Flagship" Business Intelligence Market study
- Advanced and Predictive Analytics
- Business Intelligence Competency Center
- Cloud Computing and Business Intelligence
- Collective Insights®
- Embedded Business Intelligence
- End User Data Preparation
- IoT Intelligence®
- Location Intelligence

Howard (www.twitter.com/howarddresner) conducts a weekly Twitter "tweetchat" on Fridays at 1:00 p.m. ET. The hashtag is #BIWisdom. During these live events the #BIWisdom community discusses a wide range of business intelligence topics.

You can find more information about Dresner Advisory Services at www.dresneradvisory.com.