ANALYTICAL DATA
INFRASTRUCTURE
MARKET STUDY EXCERPT
WISDOM OF CROWDS® SERIES

Understand the use cases and workflows for
analytic data infrastructure features and capabilities.

Dresner
ADVISORY SERVICES

Learn more at www.incorta.com

September, 2020
Dresner Advisory Services, LLC

2020 Edition

Analytical Data Infrastructure Market Study Excerpt

Wisdom of Crowds® Series

Licensed to Incorta

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.

Reproduction and distribution of this publication in any form without prior written permission is forbidden.

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.

Executive Summary

Analytical Data Infrastructures provide the platform and technology for integrating, preparing, modeling, managing, storing, and providing analytic capabilities on analytic data. Previously, the market called analytic data infrastructures by names such as "data warehouses" or "analytic data bases." But today's diversity and scale of BI and analytic use cases require new technology, architectures, and data/analytic workflows at high scale and security and with new licensing models. Hence, the emergence of Analytic Data Infrastructures.

Various analytic and BI use cases consume the analytic content from the analytic data infrastructure. In our 2020 Analytical Data Infrastructure Market Study, we used the following use cases to analyze the market survey responses:
- Business user reporting and dashboards
- Business user discovery and exploration
- Data science (e.g., advanced and predictive workloads and workflows)
- Embedded analytics (e.g., analytic functions and data embedded within business applications for higher volume / low latency applications).

Each use case presents different combinations of data and analytical workflows to an ADI platform. The goal of this report is to better understand the priority of use cases and preferences for Analytical Data Infrastructure features/capabilities. Understanding these capabilities, uses, and adoption helps prioritization planning, developing, and executing a BI and analytics strategy for any size organization.

- When asked to prioritize the use cases for their ADI platforms, the use case most respondents identified as a top priority for ADI platforms is business user reporting and dashboards. The data science use case is a high priority for over 50 percent of respondents—up from 43 percent last year. Embedded analytics, i.e., a use case within business operational applications work and data flows and requiring low latency analytics and typically high data volumes, is a priority for evaluating ADI platforms for 45 percent of the respondents—up from last year.
- Cost and corporate standards are a low relative priority for ADI platforms compared to performance and security priorities. We think this will lead to further ADI platform fragmentation and associated data and analytics fragmentation across organizations.
- Most respondents utilize a single ADI platform that can support multiple use cases and workloads/workflows (e.g., it must provide capabilities required for serving combinations of use cases, i.e., business user reporting and dashboards as well as business discovery and exploration and/or data science analytic use cases).

- The majority of business function respondents prefer an ADI platform accessed/licensed via a cloud deployment ("as-a-cloud service" versus "on-premises software"). However, respondents from IT organizations do not have this preference.

- Across multiple use cases, the "as-a-cloud service" option is the top deployment priority for respondents. Few prefer the hybrid model, which is the reality for many organizations today; they have ADI use cases that prefer cloud services and others that prefer on-premises deployment. Multiple buying centers in most organizations ends up creating a "hybrid" deployment model for ADI platforms.

- Cross data center integration and management capabilities to support hybrid deployments (cloud and on premises) is a growing ADI platform priority. We expect cross data center integration and management tools for hybrid ADI platform deployments will continue to increase in priority and the range of capabilities and options will become extensive as the technology/market develops.

- The range of innovation and variety of ADI platform capabilities and diversity of use cases in the market today and the lack of priority on corporate standards and governance makes developing a business and technical strategy (for using data and analytics to drive business change) for larger scale, cross-functional, multi-use case BI and analytics projects more difficult than ever for business and technical leaders.

Analytical Data Infrastructure Use Cases

There is a diversity of ADI use cases. Different use cases often access different data sets for analysis and use different workflows and data processing pipelines for moving the data from the sources, storage, doing the analysis and delivery to the analytic consumer. Therefore, we see a diversity of buying requirements and priority. Buying requirements for ADI platforms often represent priorities for more than one use case.

We asked respondents to rank the importance of four types of ADI use cases. The majority (82 percent) of respondents indicate "business user reporting and dashboards" as their highest use case driving ADI requirements and priorities, ranking it as "critical" or "very important" (fig. 1). The second-highest use case is "business user discovery and exploration," which 67 percent of respondents rank as "critical" or "very important." Fifty-one percent rank "data science (advanced and predictive analytics or data mining)" use cases as their highest priority, and 46 percent rank "embedded analytics with business applications" as their highest priority use cases.
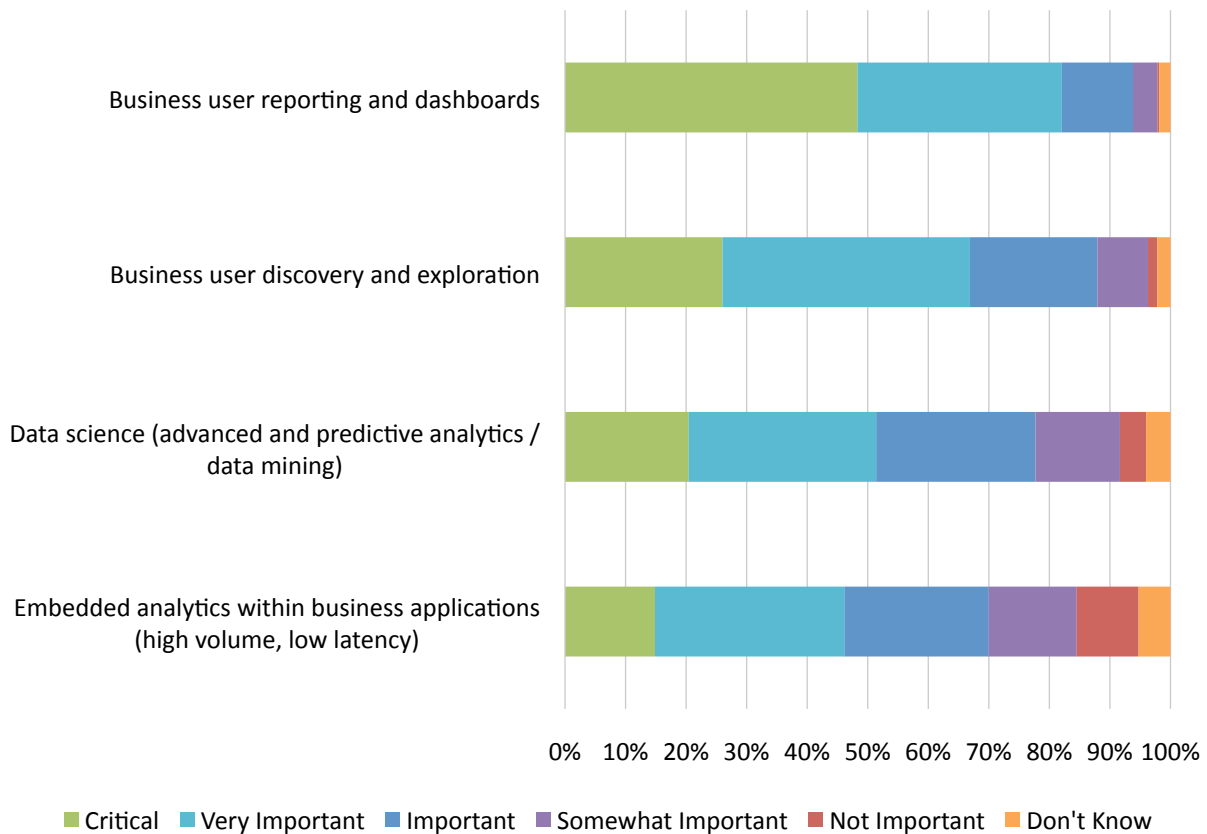
ADI Use Cases

Figure 1 – ADI use cases

Selection Priorities for Analytical Data Infrastructure

In this section, we discuss respondents' priorities covering price/performance product-related qualities such as scalability, usability, etc. Performance and security remain the highest selection priorities for ADI platforms in 2020. This is driven by the ever-growing volume of data, algorithms, and number of users arising from combinations of use cases. An ADI platform must support the volume and combinations of data, workloads, integrations, etc. required from the combinations of use cases and their workloads and workflows. Respondents place a relatively lower priority on ADI platforms' ability to meet compliance or regulatory requirements, ranking these features lower than most other ADI requirements (fig. 2). "Corporate standards" is a relatively low priority, demonstrating the diversity of ADI platforms/skills/workflows and that, for most organizations, no one ADI platform can be a "corporate standard."
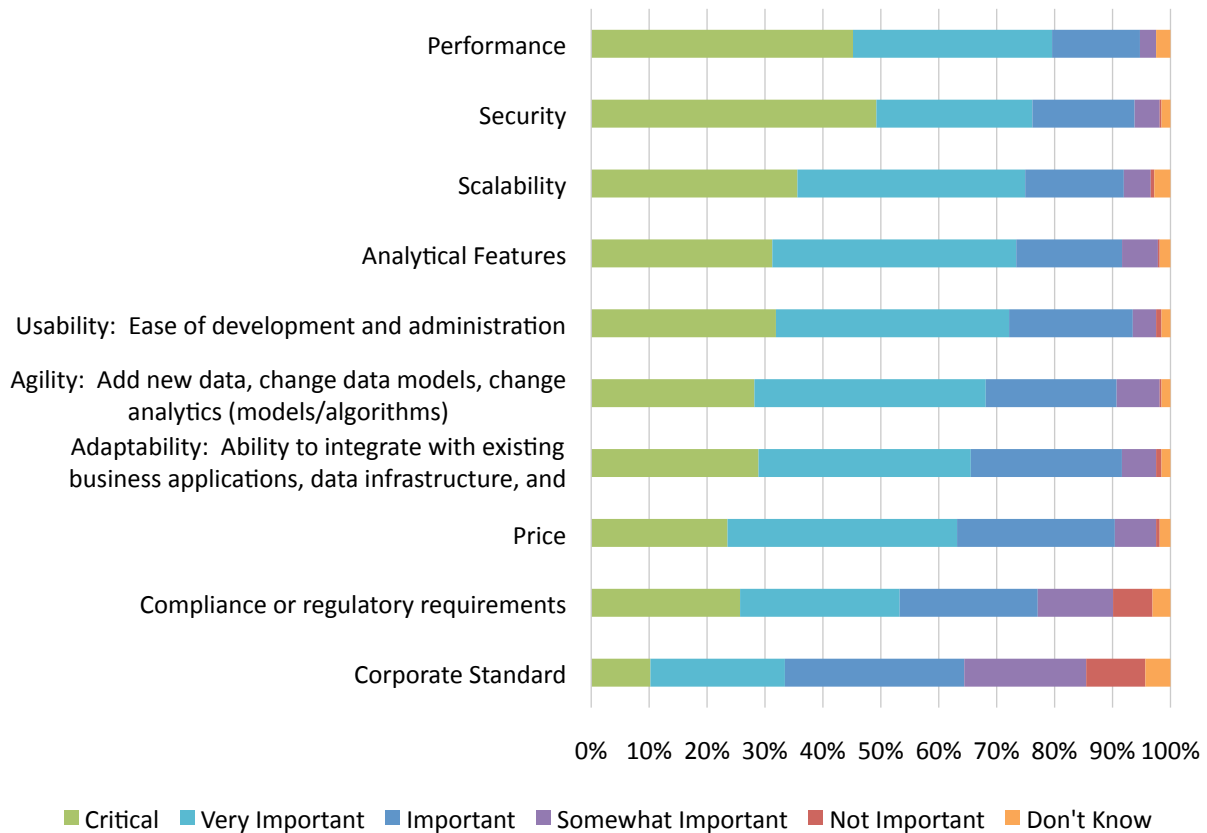
Overall Selection Priorities for ADI

Figure 2 – Overall selection priorities for ADI

Analytical Data Infrastructure Development and Deployment Features

We asked respondents to prioritize eight ADI development and deployment features. Not surprisingly, "scale up and scale out" capabilities lead as the highest priority for ADI development features, followed by "data life cycle management" capabilities. Like last year, "pre-built data models" is the lowest priority development capability/priority. Pre-built data models have a spotty history in analytic applications, given the difficulty of modifying/adapting the models to new/different requirements than what they were originally designed for. While multi-tenancy support is often an ADI capability used to promote cloud offerings, its priority is comparatively low.
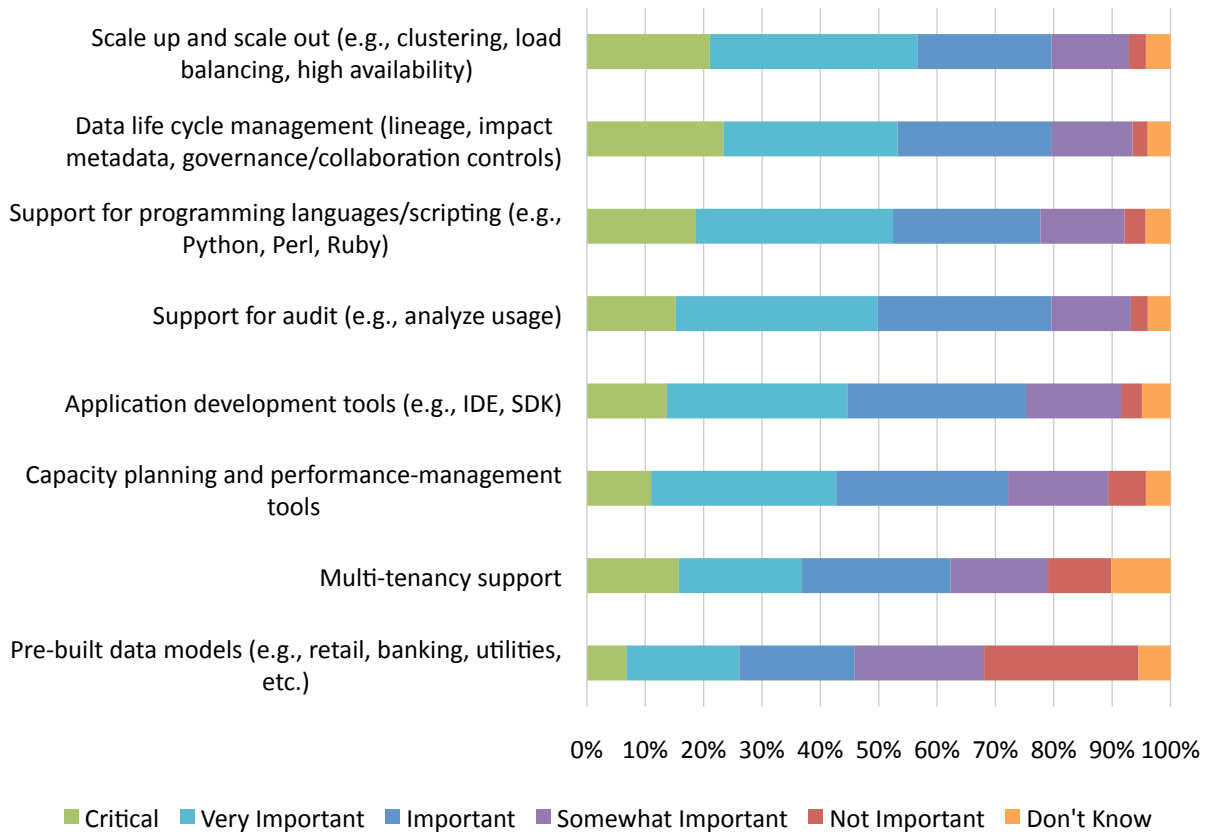
ADI Development and Deployment Features

Figure 3 – ADI development and deployment features

Analytical Data Infrastructure Data Types

More than 80 percent of respondents have analytic workloads and workflows based on "transactional data" sources (i.e., from business applications). The next highest data type priority for ADI platforms and associated analytic use case workflows are Excel/CSV data (fig. 4) and metadata. Given the role of metadata in data lineage and impact analysis, the high priority for metadata echoes respondents' priority for "data life cycle management" capabilities in the ADI development and deployment sections (see fig. 32). In general use, images and video data type support was, historically, a lower priority for ADI applications / use cases, and this year's responses show a similar priority for ADI platforms. It is worth noting that "machine and events / log data" are an important priority for a large number of analytic use-case workflows/workloads.
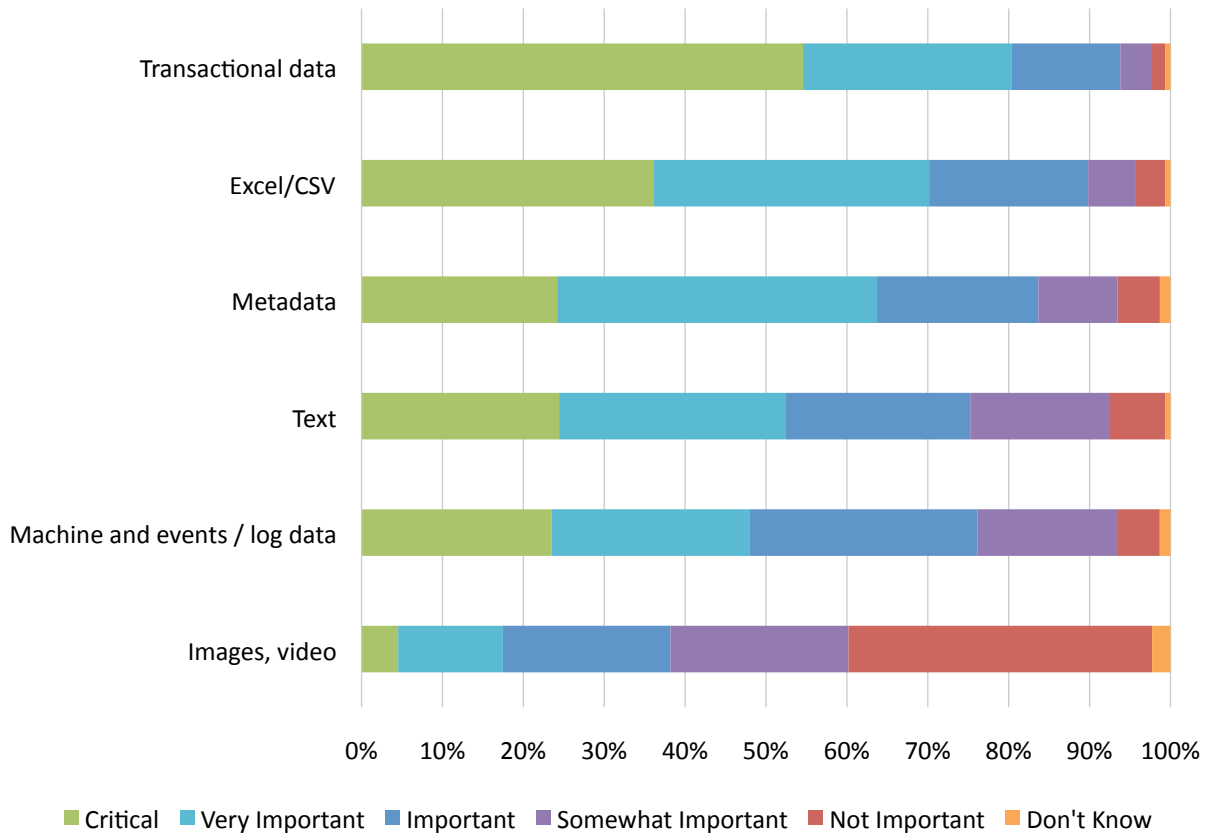
Data Type Priorities for ADI

Figure 4 – Data type priorities for ADI

Analytical Data Infrastructure Data Preparation and Loading

Earlier in this report, we note increasing growth in preferences for "cross data center integration and management" capabilities. This also reflects the relative priority of data integration as the second-highest strategic technology across BI use cases Traditional bulk load data preparation/transformation and loading capabilities such as ELT / ETL (Extract, Transform, Load) remain the highest priority associated with data preparation and loading for ADI platforms to support (fig. 46). Interestingly, metadata import and management is a lower priority in the context of data prep and loading compared to respondents' use of metadata in other ADI preferences and priorities. Slightly more than 46 percent indicate "real-time / streaming, trickle, increments / change capture" is a priority for data prep and loading of ADI. Less than 30 percent of respondents consider support for Apache big data services "critical" or "very important."

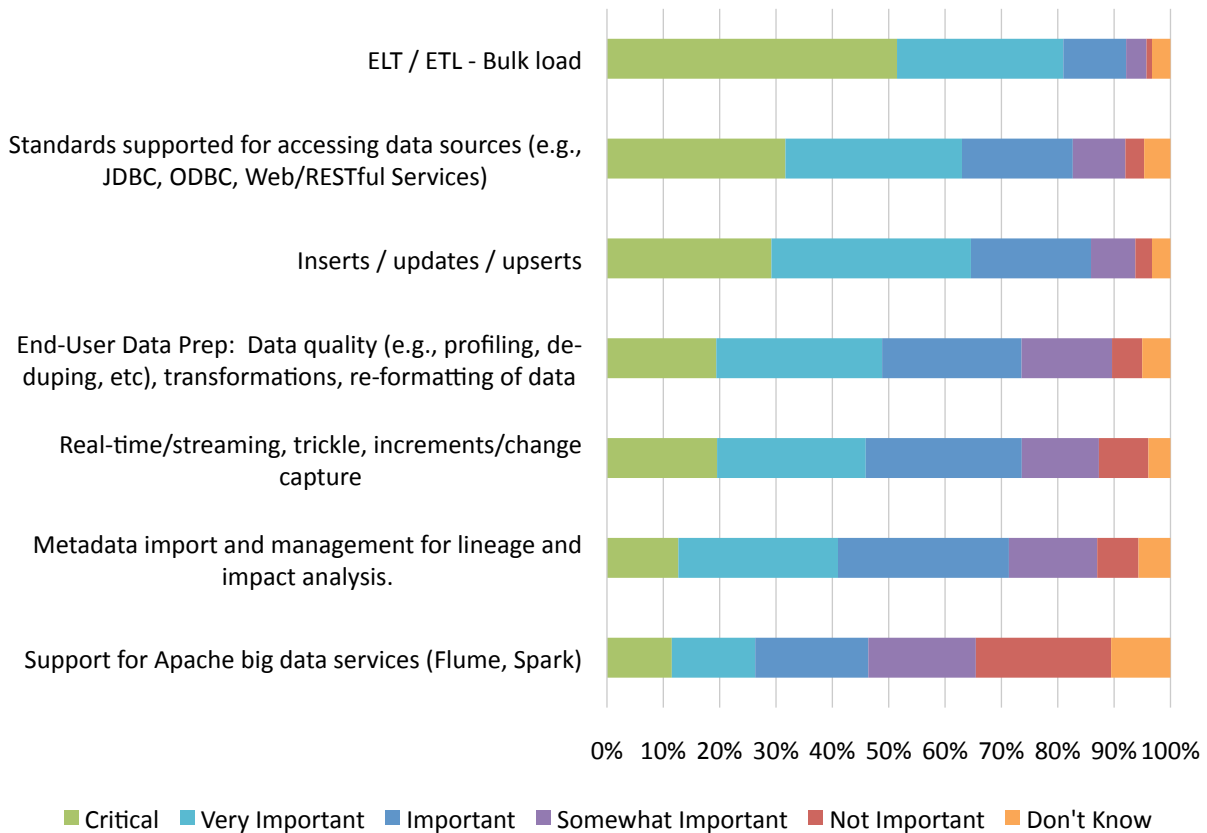## ADI Data Preparation and Loading

Figure 5 – ADI data preparation and loading

Analytical Data Infrastructure Data Model / Management of Data

SQL data capabilities are, by far, the top data model / management priority for an ADI platform (fig. 6). This matches the top priority respondents place on transactional data (see fig. 4).

Given the amount of market attention, in-memory data capabilities is a high priority for ADI platforms for more than 50 percent of respondents.

Non-SQL and hierarchical files (e.g., Hadoop HDFS) are high priorities for less than 25 percent of respondents.
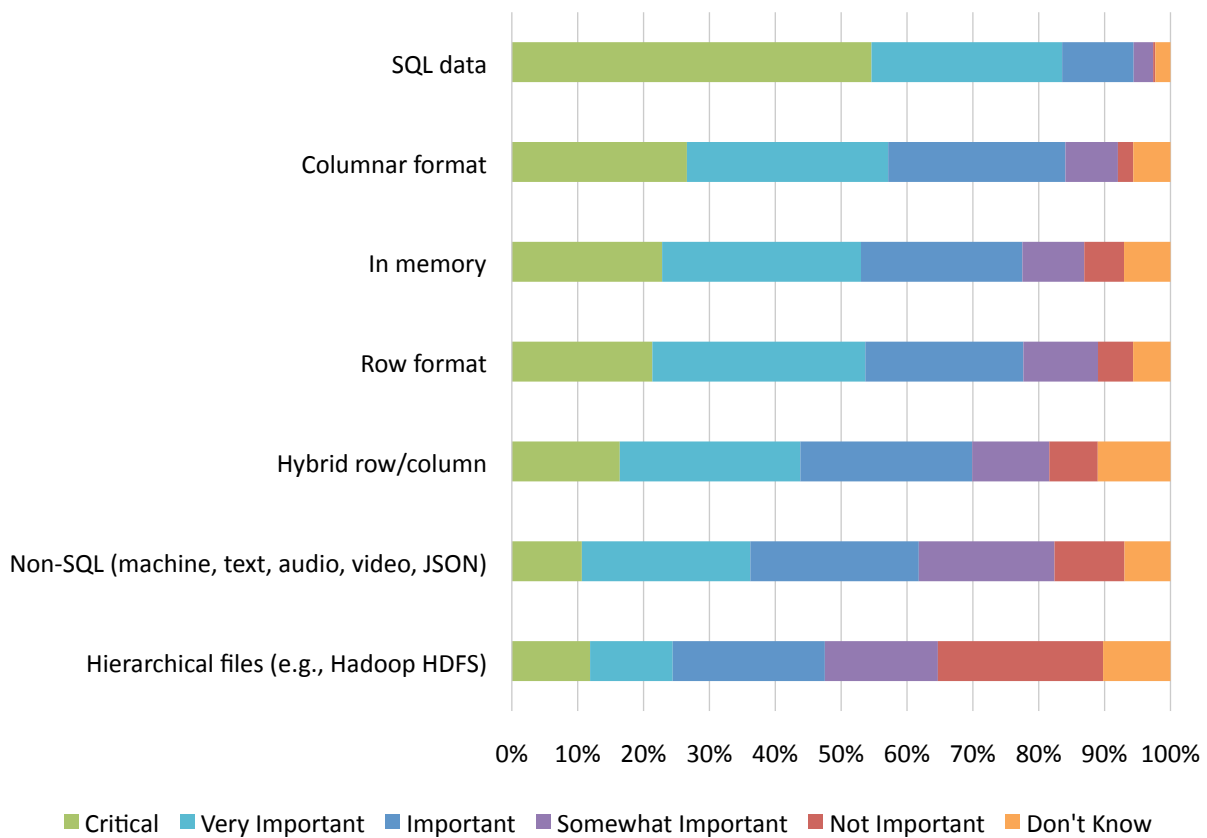
ADI Data Model / Management of Data

Figure 6 – ADI data model / management of data

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. 7). 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.
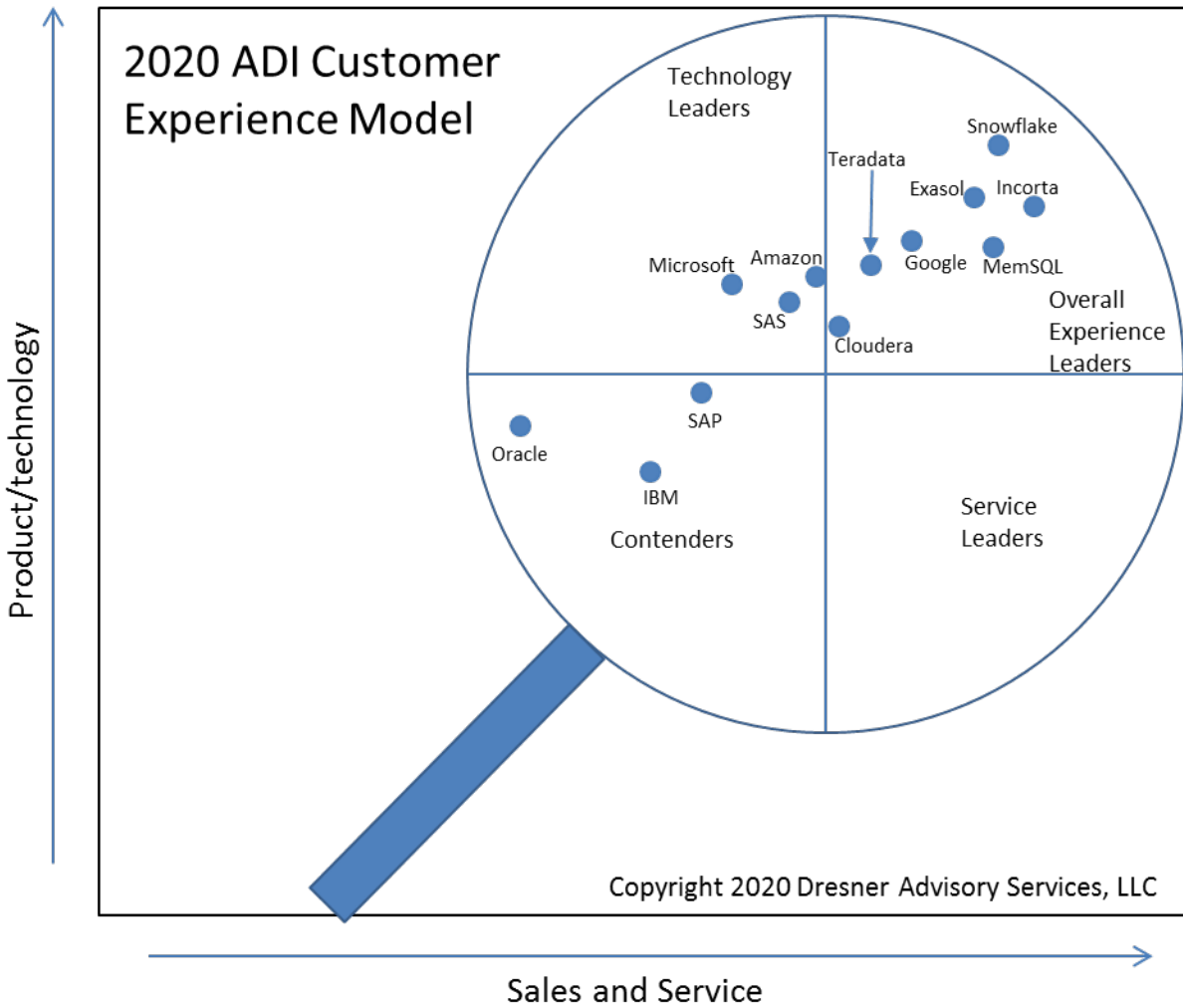
Figure 7 – Customer experience model

Vendor Credibility Model

The Vendor Credibility Model considers how customers "feel" about their vendor (fig. 8). 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.
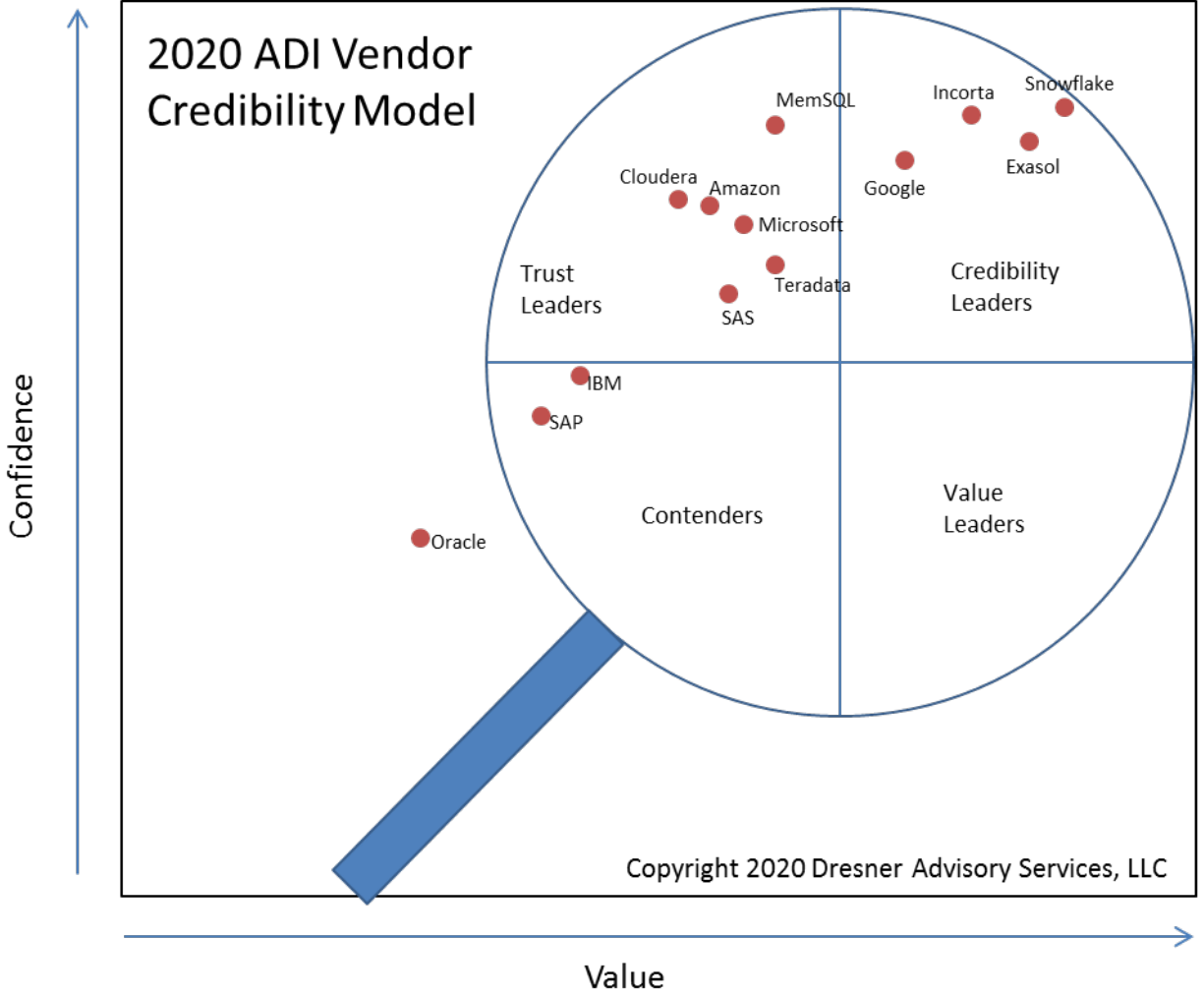
Figure 8 – Vendor credibility model

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)*
    - o Professionalism
    - o Product knowledge
    - o Understanding our business/needs
    - o Responsiveness
    - o Flexibility/accommodation
    - o Business practices
    - o Contractual terms and conditions
    - o Follow-up after the sale

- **Value for price** *(3 o'clock)*

- **Quality and usefulness of product** *(3 - 7 o'clock)*
    - o Robustness/sophistication of technology
    - o Completeness of functionality
    - o Reliability of technology
    - o Scalability
    - o Integration of components within product
    - o Integration with third-party technologies
    - o Overall usability
    - o Ease of installation
    - o Ease of administration

- **Quality and usefulness of product (continued)**
    - o Customization and extensibility
    - o Ease of upgrade/migration to new versions
    - o Online forums and documentation

-

- **Quality of technical support** *(8 - 9 o'clock)*
    - o Professionalism
    - o Product knowledge
    - o Responsiveness
    - o Continuity of personnel
    - o Time to resolve problems

- **Quality and value of consulting services** *(9 - 10 o'clock)*
    - o Professionalism
    - o Product knowledge
    - o Experience
    - o Continuity
    - o Value

- **Integrity** *(11 o'clock)*

- **Whether vendor is recommended** *(12 o'clock)*
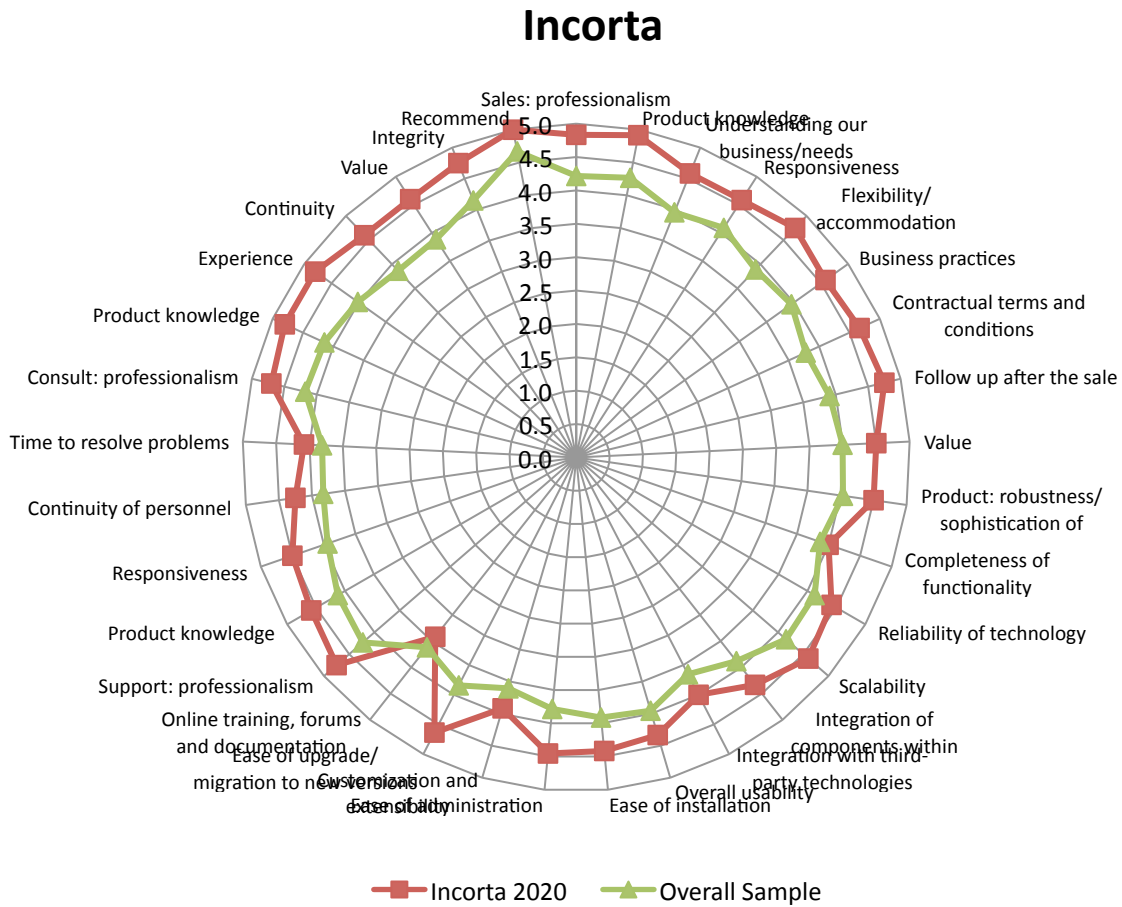
Incorta Detailed Score

Figure 9 – Incorta detailed score

This is the first year we evaluate Incorta as an ADI vendor. In general, it is above the overall sample for all measures and is considered an Overall Leader in both the Customer Experience and Vendor Credibility models. It is best in class for consulting product knowledge, experience, and continuity. It 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 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 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 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 BI Market Study
- Cloud Computing and Business Intelligence
- Data Preparation
- Data Science and Machine Learning
- IoT Intelligence®
- Location Intelligence
- Self-Service BI

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.