Cool Vendors in Analytics and Data Science

Published 10 May 2021 - ID G00746456 - 18 min read

By Julian Sun, David Pidsley, **and 2 more**

Organizations can achieve broader analytics capabilities in a self-service manner. To achieve more productive self-service analytics and data science, data and analytics leaders should assess these Cool Vendors to build synergies among automation, augmentation, contextualization and composability.

Overview

Key Findings

- Self-service has expanded its scope in the augmented era with more advanced capabilities, but not all organizations gain efficiencies from it.

- Emerging analytics capabilities come from cloud-first, leaving the organizations that deploy on-premises with concerns of being left behind.

- Business users struggle to get the contextualized insights, causing them to give up soon after trying self-service analytics.

Recommendations

- Improve the efficiency of analytics and data science activities by assessing the platform's coherence to ensure the seamless experience while using different self-service capabilities in one platform.

- Pilot innovative capabilities of cloud analytics by assessing the composability of a solution to integrate with emerging capabilities in the cloud, even if the majority of your deployments are still on-premises.

- Compose a tailored self-service analytics application to improve the business context by adding semantic models and knowledge graphs of the domain.

Strategic Planning Assumptions

By 2023, 60% of organizations will compose components from three or more analytics solutions to build business applications infused with analytics that connect insights to actions.

By 2023, overall analytics adoption will increase from 35% to 50%, driven by vertical- and domain-specific augmented analytics solutions.

By 2023, cloud architects will become key stakeholders when purchasing analytics and business intelligence (ABI) tools, as scalability and cohesive cloud ecosystems move into the top three key buying considerations.

Analysis

This research does not constitute an exhaustive list of vendors in any given technology area, but rather is designed to highlight interesting, new and innovative vendors, products and services. Gartner disclaims all warranties, express or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

What You Need to Know

Arguably, it is very hard to be cool in the analytics and data science space as a vendor. Cool and consumer-friendly features are expected as we move through the augmented era. For example, visual data discovery has been the hallmark of analytics and business intelligence capabilities for almost a decade, and came off the Hype Cycle for Analytics and Business Intelligence in 2020. Although visual data discovery is still extremely important to organizations day-to-day, these capabilities have been commoditized, and are no longer generating hype.

Based on Gartner 2021 Critical Capabilities for Analytics and Business Intelligence Platforms assessment, there is very little differentiation in product capabilities such as data visualization or data source connectivity. Instead, advanced analytics and natural language query are the battleground of analytics users.

Today's self-service means more about using basic capabilities. It increasingly includes more advanced data science and machine learning (DSML) capabilities. These capabilities redefine how organizations can achieve self-service analytics to solve more complex business questions in a productive way.
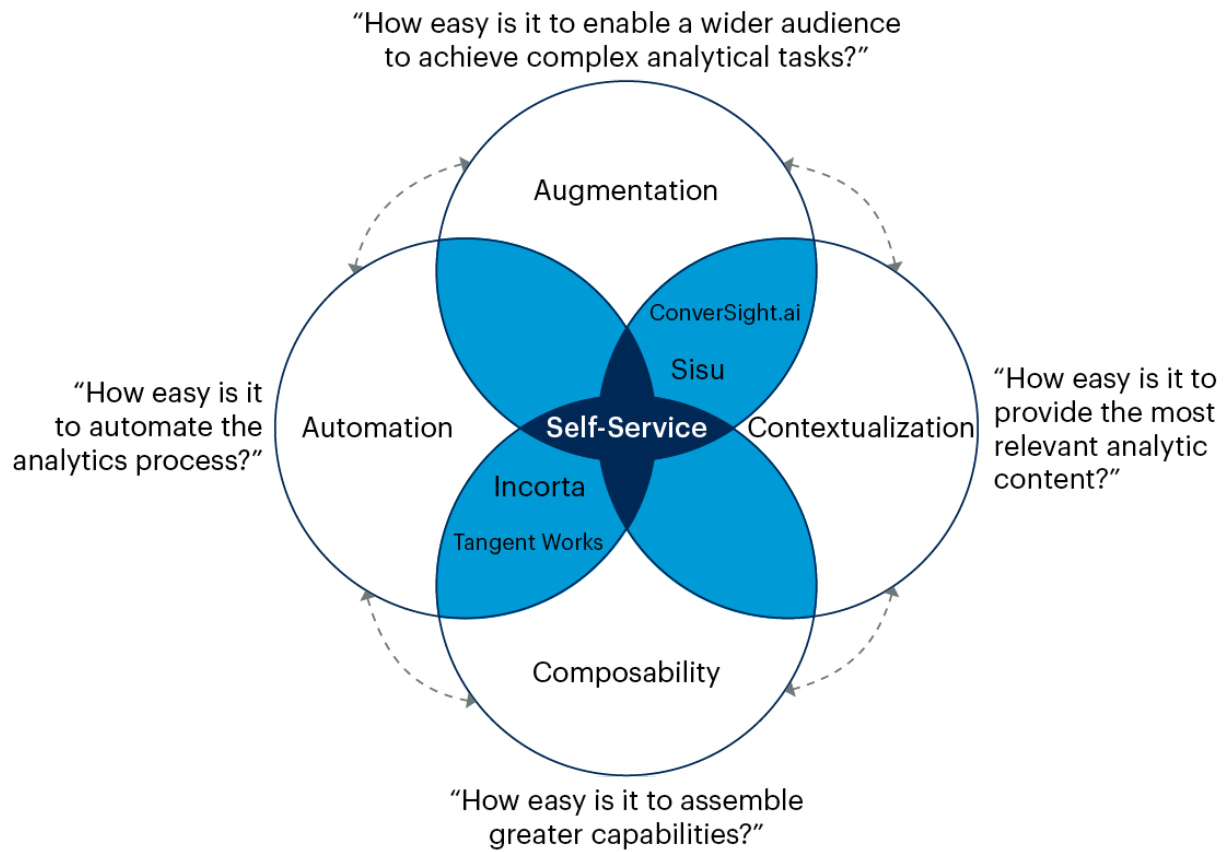
Self-service has been the essence of "how to enable more people to achieve capabilities more easily," and it is still important for organizations to stay ahead with data and analytics (D&A). However, the ability of technology to enable self-service doesn't guarantee the efficiency of analytics workflows. The cohesiveness of the analytics and data science capabilities in a platform is critical to the analytics experience, enabling users to adopt extensive capabilities and get the business value seamlessly.

Increasingly accessible via self-service, these approaches in analytics and data science have shifted its definition with broader capabilities in the areas of automation, augmentation, contextualization and composability (see Figure 1).

Build Synergies Among Analytics and Data Science Capabilities to Drive Self-Service

"How easy is it to enable a wider audience to achieve complex analytical tasks?"

Augmentation

ConverSight.ai

Sisu

"How easy is it to automate the analytics process?"

Automation — Self-Service — Contextualization

Incorta

Tangent Works

"How easy is it to provide the most relevant analytic content?"

Composability

"How easy is it to assemble greater capabilities?"

Source: Gartner
746456_C

Gartner.

These four areas should establish synergy with each other to achieve cohesiveness for more productive analytics and data science. For example, augmentation should be the path to achieve automation (see Augmented Analytics: Teaching Machines to Tell Data Stories to Humans).

A vendor's strength in only one of these areas is insufficient to be considered "cool." Combining these to provide efficient analytics and data science to users in a cohesive manner is what makes a vendor stand out.

Automation is the essential component when users require contextualized analytics in real time. Contextualization is a requirement for composed analytics applications for a tailored experience. Composability is the ease with which augmented analytics capabilities are able to be integrated with other applications. (see Composable Analytics Shapes the Future of Analytics Applications). While these areas can be found in on-premises solutions, they are enhanced while they are in the cloud. According to Gartner's 2020 Cloud Data and Analytics Survey, the share of analytics development and production-level advanced analytics run in the cloud would be doubled, and the major reasons are scalability and operations effort (see Adopt Cloud Analytics to Drive Innovation).

The Cool Vendors in this research demonstrate how they combine and build synergies among automation, augmentation, contextualization and composability:

- ConverSight.ai **augments** business users to get proactive insights by natural language. The insights will be more **contextualized** for different roles with adaptive learning of knowledge and vocabulary. Then the users can keep communicating the insights with other business roles in a conversational interface.

- Incorta's Direct Data Mapping platform has the **composability** with other ABI platforms. It **automatically** captures all of the relationships that exist within and across all different data sources and recognizes and implements the security provisions defined within the source database.

- Sisu accelerates the exploration of complex data and automates the diagnosis of key performance indicators (KPIs). The Sisu platform's personalized model for each user improves over time to provide **contextualized** insights while **augmenting** the user to achieve complex diagnostic analysis without requiring complex SQL or data modeling.

- Tangent Works combines feature engineering, model selection and tuning a one-step process with **automation** to provide time-series predictive analytics. The container-based architecture would further enable the **composability** to make the real-time predictive forecasting be portable and easily integrate with other solutions.

ConverSight.ai

Carmel, Indiana, U.S. (conversight.ai)

Analysis by Julian Sun and James Richardson

Why Cool: ConverSight.ai's AI assistant, Athena, can understand user behavior, context and intent by combining machine learning, data science and an intuitive interface to deliver personalized insights and actions.

ConverSight.ai provides end-to-end supply chain analytics with text and chat capabilities. It can connect to multiple data sources (ERP, CRM) to uncover valuable, hidden and contextual insights and enable collaboration. With specialized domain and natural language understanding, it provides vertical- and industry-specific conversation and subject knowledge for supply chain. Athena senses and responds to changes of supply and inventory turnover on demand in real time. The tool is used across roles, including front-end operation worker, sales, logistics manager and more. Customers appreciate its ease of use and short learning curve. Organizations can use Athena daily to collaborate across the supply chain life cycle based on the insights. It offers personalized voice interaction that gives role-based custom interaction with dynamic dialogs. Users can get suggested additional questions to help get more insights. The adaptive learning of knowledge and vocabulary will improve the contextualization of insights as the use of the tool grows across the organization. The rich knowledge representation foundation provides a flexible experience to navigate rich information for the users.

Athena is equipped with an intelligent supply chain "control tower," a visualization tool to track performance in one place. It can deliver insights from a centralized data repository, proactively develop automated reporting and enable complete visibility for the supply chain. Users can receive proactive alerts on key metrics from visualization and communicate with various roles and teams by conversational interaction in a cohesive way. The ability to pin the report to the pinboard enables users to share and use it as a starting point for others. This type of traditional reporting capabilities enable the adoption of the tools.

Based on Gartner's social media analysis, the coolness perception for ConverSight.ai was driven by its Conversational AI Assistant that was positively mentioned to save operational costs for organizations.

Challenges: ConverSight.ai's major use cases are for supply chain analytics. Organizations in other domains might not get the same benefits as the ones in supply chain. An organization needs expertise to work with the ConverSight.ai team to build the correct semantics before use. Time is required for the system to learn context. Users need to be trained to ask the questions in a logical way.

Who Should Care: D&A leaders who want to improve the supply chain analytics capabilities should consider ConverSight.ai. It is especially well-suited for organizations that have already tried general ABI tools, but have failed to build domain-level analytics capabilities. Organizations that want to enable more front-end operational workers and other business roles in the supply chain life cycle should also assess its products.

Incorta

San Mateo, California, U.S. (www.incorta.com)

Analysis by James Richardson

Why Cool: Incorta offers a highly performant end-to-end ABI platform engineered to deal with complex data and analytics challenges. The core (and coolest part) of Incorta's offering is its Direct Data Mapping (DDM) engine, which automatically analyzes data and eliminates the need for data modeling, cubes, extraction, transformation and loading (ETL) batch jobs, or optimization. According to Incorta's customers, the platform delivers subsecond drill-down across billions of records, enabling them to handle complex data at detail.

Initially sold as a performance optimizer for analytics (see Market Guide for Analytics Query Accelerators), Incorta's platform now reaches end users directly with its own data visualization and reporting capabilities.

Direct Data Mapping is paired with an in-memory columnar analytics engine. Incorta can either query mapped data where it is, or ingest mapped data into Amazon S3, Google Cloud Storage and Microsoft Azure Data Lake Storage Gen2. When loaded, data drill-down is supported to the lowest level of detail. Incorta's customer references highlight the ability to analyze complex, full-fidelity business data in real time without requiring data to be reshaped and aggregated to fit an analytical or dimensional model. As Incorta loads data, it generates a dataset that is smaller than the source data, as opposed to most dimensional modeling schemes that create a dataset that is larger than the source data. This means efficient cluster data transfers, and fast data loads into memory. Incorta maintains a columnar and compressed data organization, both on disk and in memory. This makes it possible to analyze 10 to 20 times more data than would normally fit into available memory. To avoid being memory-bound, data is dynamically loaded from disc as needed, unless a data table has been specifically configured to always remain in memory. Incorta's end-user self-service capabilities are optional, loosely coupled with its data management technology, but used by the lion's share of customers, generally alongside third-party ABI tools. Its Analyzer offers a range of data visualizations. These are extensible, using a React-based framework for rapid updates and to support custom visualizations. In addition, it offers visual data prep and an ML pipeline tool for business users and data analysts. Incorta can be run on-premises, hosted by a cloud provider, or delivered as a fully managed cloud service. The platform can run as a complete stand-alone data and analytics pipeline, or as a component within a larger ABI technology portfolio.

Incorta has also grown into one of the few new ABI platform offerings that is competing effectively against the dominant players in the space, albeit with a narrower proposition than some of its competitors.

Challenges: When it launched in 2013, Incorta was primarily positioned as a provider of software complementary to the ABI tools covered in the Gartner Magic Quadrant for Analytics and Business

Intelligence Platforms — using its DDM functionality as a performance optimization layer for tools like Power BI and Tableau. However, Incorta's offering is increasingly used as a full ABI platform in its own right, becoming a viable alternative. This broadened value proposition brings Incorta into direct competition with these commonly used ABI platforms, in what is a brutally price-competitive market with a dominant vendor in Microsoft. Currently a strong partner of Microsoft's Synapse team (Incorta was named U.S. Startup Partner of the Year by Microsoft in 2020), Incorta may find itself competing more with Microsoft than cooperating with it.

Incorta's value proposition resonates with IT buyers, but its brand is little known in the business, which may slow evaluations and adoption. In addition, its front-end data visualization, dashboarding and collaboration capabilities, while developing, still have some basic gaps in functionality when compared to its erstwhile partners. It has no offering at all when it comes to the differentiated capabilities, such as data storytelling and natural language processing (NLP), that form the core of the augmented capabilities and are directional in the ABI market for evaluators.

Who Should Care: D&A leaders that are looking to deliver fast drill-down to transactional data should consider Incorta. Organizations planning to remove traditionalolder-wave 1 BI technologies and deliver modern reporting that's faster to deploy should also be evaluating the product, in many cases as a complement to their existing modern ABI platform portfolio.

Sisu

San Francisco, California, U.S. (sisudata.com)

Analysis by David Pidsley

Why Cool: Sisu is an augmented analytics tool for diagnostic analysis that helps D&A leaders and technical professionals accelerate the exploration of complex, cloud-scale data. Its analytics engine tests hundreds of millions of hypotheses in seconds and automatically surfaces key drivers in the data that are having a significant impact on KPIs. This enables its clients to query large transaction-level datasets in cloud data warehouses in near real time.

Sisu is designed for monitoring and fast drill-down into the highest-impact dimensions of data to get a more complete and contextualized view of why KPIs that are most relevant to each individual user are changing. Just like services like Netflix and Facebook learn relevance based on user feedback and engagement, Sisu learns a personalized model for each user that improves over time based on their interactions with trends in the data.

Sisu applies machine learning and statistical analysis (including efficient feature engineering, feature selection, and relevance modeling for structured and semistructured data) based on research from Stanford University. Sisu received $66 million since its inception to accelerate its expansion and it has several marquee customers. Clients report time to production of six weeks or less.

Challenges: Sisu competes in the growing augmented analytics market with Anodot, Orbiter and Outlier, as well as incumbent vendors that are adding automated insight and key driver analysis features as part of broader ABI platforms. Compared to Microsoft, Tableau or Qlik, Sisu provides lightweight data preparation capabilities to support the creation of wide, flat tables in cloud data warehouses.

Sisu is a cloud-based tool with data source connectivity to cloud data warehouses (Amazon, Google, Microsoft), Delta Lake, PostgreSQL and Snowflake (and its Data Marketplace for third-party data). It may not be suitable for primarily on-premises data and analytics deployments, although secure on-premises connections and virtual private cloud deployments are possible. Clients report spending time working on their data definitions, and Sisu is now offering services to accelerate this.

Despite being strong in diagnostic analytics, Sisu is not focused on re-creating the descriptive, predictive or prescriptive analytics capabilities of an ABI platform. Its roadmap includes forecasting and scenario analysis features as it develops toward its vision of a decision-making hub for collaboration and action. Sisu currently has no API for third-party developers or cloud ecosystem partners to build upon.

Sisu is not a data science and machine learning platform. Although suitable for data analysts needing an intuitive user interface, it still has a steep learning curve for line-of-business decision makers. Poor explainability of AI features makes interpreting results more difficult.

Who Should Care: D&A leaders should look at Sisu for their analyst teams if their organization makes extensive use of cloud data warehousing for large, highly dimensional datasets, and want a diagnostic analytics tool that generates automated insights about key drivers.

Sisu clients include direct-to-consumer, subscription and high-volume SaaS businesses such as retail, e-commerce and financial services, which have lots of transactional and high-dimensional customer journey data.

Users of ABI platforms such as Google's Looker should look at Sisu if seeking greater automated insight capabilities or suffering a shortage of analysts able to identify key drivers in cloud data warehouses.

Tangent Works

Gooik, Belgium (tangent.works)

Analysis by Shubhangi Vashisth

Why Cool: Tangent Works' Tangent Information Modeler (TIM) is an automatic model-building engine for time-series forecasting and anomaly detection. It engineers features and applies a highly efficient algorithm in a single step. As compared to AutoML, where users need to do feature engineering, model selection and tuning, the TIM engine covers all these steps in a one-step process. Users still

need to prepare their datasets prior to feeding it to TIM, which then provides automated selection of the likely most relevant input variables (with an explanation of the impact of each predictor) and automated model generation. Time-series forecasting presents unique challenges, such as identification of significant features, how to address changing dynamics in time-series data, dynamic data availability, multipoint and multisituational forecasts, and so on.

The TIM solution architecture includes the TIM Engine, the TIM API and the TIM Studio (a web-based user interface). The TIM engine is fully containerized and supports parallel processing of workloads. In other words, users can execute multiple models simultaneously with different features (such as time horizons), and view and overlay the outcomes to find the most optimal parameters for their time-series models. This is particularly important when dealing with a large number of parameters at scale. Under the hood, the TIM Engine is built on a field of mathematics called information geometry, an interdisciplinary field that uses differential geometry techniques to study probability theory and statistics. TIM engine includes two modules — InstantML and RTInstantML (real-time InstantML). Both the modules are explainable and provide full insight into the model and the feature combinations.

The TIM Engine can be used as a cloud service, in a user's own cloud, on-premises, or on an Internet of Things (IoT) device or on an edge device. TIM is available as an add-on module with platforms such as Alteryx, Qlik and others.

Challenges: Time-series forecasting is a capability being supported by many existing DS and ML platforms; however, not all of these offer augmented/automated capabilities. In a crowded market, it is important for Tangent Works to partner with right companies who can broaden the customer reach of TIM modeling framework and make it available to the right audiences. Its GTM strategy includes working with technology, implementation, reselling and OEM partners.

Users still need to do data preparation and preprocessing before they input data to the TIM engine. For complex visualizations, users would want to view the outcome of the TIM engine in another analytics platform.

Complex time-series predictions that may involve non-numeric features or domain-specific feature engineering would be difficult to run on the TIM engine alone, and would require involvement from expert data scientists.

Who Should Care: Time-series data is everywhere. The current time-series efforts are often stuck for complexity, a lack of skill or funding. D&A leaders looking for a composable business model should consider the TIM engine for its simplicity, ease of use, accuracy and explainability in use cases for forecasting and anomaly detection.

TIM is an important tool for both business and technical users across various industries, including retail and consumer packaged goods, banking, manufacturing, utilities, telecom, healthcare, and

many more. TIM can be used for complementing what-if and simulation scenarios for budget exercises, doing risk assessments in your business transformation process and more.

Where Are They Now?

Aible

Palo Alto, California, U.S. (www.aible.com)

Analysis by Julian Sun

Profiled in Cool Vendors in Analytics, 2019

Why Cool Then: Aible is cool because it enables business people to build machine learning (ML) and artificial intelligence (AI) models and to take into account the fact that all costs and benefits are not equal when assessing which model is best.

Where They Are Now: In addition to promoting its AutoML capabilities in the data and analytics market, Aible is also now providing low-code/no-code augmented ML and ModelOps capabilities to serve application developers in the cloud AI developer market. It bridges data and analytics and application development teams with collaboration features to achieve more comprehensive composable analytics.

Who Should Care: Data and analytics leaders and software engineering leaders should consider Aible if they want to give business analysts or application developers without data science skills the ability to build and test the business impact of predictive models.

Evidence

Social Media Analytics Data

Approved Methodology: Gartner conducts social listening analysis leveraging third-party data tools to complement or supplement the other fact bases presented in this document. Due to its qualitative and organic nature, the results should not be used separately from the rest of this research. No conclusions should be drawn from this data alone as it may not be entirely market representative. Social media data in reference is from 17 February 2019 through 21 February 2021 in all geographies (except China) and recognized languages.

Sources Covered: By default, social media sources considered for analysis include Twitter, Facebook (publicly available information only), aggregator websites, blogs, news, mainstream media, forums and videos (comments only); unless and until specified.

The Team: The Social Media Analytics Team contributed to this research.

Gartner's Cloud Data & Analytics survey was conducted online from 17 June through 29 June 2020 with 86 members of Gartner's Research Circle — a Gartner-managed panel. Fifty-six percent of

qualified respondents are involved in setting the cloud data and analytics management strategy of their organization.

The survey was developed collaboratively by a team of Gartner analysts and was reviewed, tested and administered by Gartner's Research Data and Analytics team.