

# Data Orchestration Buyers Guide

Software Provider and Product Assessment

EXECUTIVE  
SUMMARY

**\*ISG** Research



# Data Orchestration

While data-driven enterprises stand to gain a competitive advantage by responding faster to worker and customer demands for more innovative, data-rich applications and personalized experiences, this increasingly relies on a complex array of data pipelines to support agile, continuous data processing. Given the increasing complexity of evolving data sources and requirements, it is essential to automate and coordinate the creation, scheduling and monitoring of data pipelines as part of a DataOps approach to data management.

This is the realm of data orchestration, which ISG Research defines as providing the capabilities to automate and accelerate the flow of data to support operational and analytics initiatives and drive business value via capabilities for the monitoring and management of data pipelines and associated workflow.



**At the highest level of abstraction, data orchestration covers data collection, data transformation, and data activation.**

At the highest level of abstraction, data orchestration covers three key capabilities: data collection, including data ingestion, preparation and cleansing; data transformation, additionally including integration and enrichment; and data activation, making the results available to compute engines, analytics and data science tools, or operational applications.

This may sound very much like the tasks that data management practitioners have undertaken for decades. As such, it is fair to ask what separates data orchestration from traditional approaches to data management.

Viewing data management challenges through the lens of today's data-processing requirements is key to understanding why data orchestration is a necessary and differentiated approach that goes beyond traditional data management.

Being data-driven requires a combination of people, processes, information and technology improvements involving data culture, data literacy, data democracy and data curiosity. Encouraging workers to discover and experiment with data is a key aspect of being data-driven that requires new, agile approaches to data management.

Additionally, the increasing reliance on real-time data processing is driving requirements for more agile, continuous data processing. The rapid adoption of cloud computing has fragmented where data is generated and stored. Enterprise data is increasingly spread across multiple data centers and cloud providers.



Traditional approaches to data management are rooted in point-to-point batch data processing, whereby data is extracted from its source, transformed for a specific purpose and loaded into a target environment for analysis. These approaches are unsuitable for the demands of modern analytics environments, which instead require agile data pipelines that can traverse multiple data-processing locations and evolve in response to changing data sources and business requirements.

Given the increasing complexity of evolving data sources and requirements, there is a need to enable the flow of data across an enterprise through new approaches to the creation, scheduling, automation and monitoring of workflows. Traditionally, individual tasks related to these requirements have been addressed with a variety of specialist tools as well as manual effort, hand-coded scripts and expertise.



**Data orchestration tools are designed to automate and coordinate the sequential or parallel execution of a complete set of tasks via data pipelines.**

In comparison, data orchestration tools are designed to automate and coordinate the sequential or parallel execution of a complete set of tasks via data pipelines, typically based on directed acyclic graphs that represent the relationships and dependencies between the tasks. The capabilities delivered by data orchestration fall under three categories: pipeline monitoring, pipeline management and workflow management.

As is often the case with new approaches to data and analytics, the requirements for data orchestration were first experienced by digital-native brands at the forefront of data-driven business strategies. One of the most prominent data orchestration tools, Apache

Airflow, began as an internal development project within Airbnb, becoming an Apache Software Foundation project in 2016. Workflow automation platform Flyte was created and subsequently open-sourced by Lyft, and Metaflow was developed and open-sourced by Netflix.

Data orchestration is not just for digital natives, however, and a variety of software providers have sprung up with offerings based around these open-source projects, as well as other development initiatives, to bring the benefits of data orchestration to the masses.

In addition to stand-alone data orchestration software products and cloud services, data orchestration capabilities are also being built into larger data-engineering platforms addressing broader data management requirements, including data observability, often in the context of data fabric and data mesh.

Whether stand-alone or embedded in larger data-engineering platforms, data orchestration has the potential to drive improved efficiency and agility in data and analytics projects. ISG asserts that by 2027, more than one-half of enterprises will adopt data orchestration



technologies to automate and coordinate data workflows and increase efficiency and agility in data and analytics projects.

Adoption of data orchestration is still in the early stages and is closely linked to larger data transformation efforts that introduce greater agility and flexibility. If an enterprise's data processes and skills remain rooted in traditional products and manual intervention, then data



**Adoption of data orchestration is still in the early stages and is closely linked to larger data transformation efforts that introduce greater agility and flexibility.**

orchestration is not likely to be a quick fix. However, alongside the cultural and organizational changes involved in people, processes and information improvements, data orchestration has the potential to play a key role in the technological improvement involved in becoming more data-driven. All enterprises are recommended to explore how the orchestration of data pipelines can help increase the potential for improved data-driven decision-making as part of a broader evaluation of the people, processes, information and technology improvements required to deliver data-driven decision-making.

The orchestration of data pipelines is just one aspect of improving the use of data within an enterprise. In addition to the development, testing and deployment of data pipelines, DataOps also encompasses data observability, which has a complementary role to play in monitoring the health of data pipelines and associated workflows as well as the quality of the data itself.

The combination of healthy and well-orchestrated data pipelines and data observability is also complementary to developing and delivering data products, ensuring that data consumers can trust the provenance and quality of the data that is made available across the enterprise.

Data orchestration is also integral to the development and delivery of applications driven by artificial intelligence and generative AI, complementing MLOps, which serves the collection of artifacts and orchestration of processes necessary to deploy and maintain AI/ML models. Specifically, data orchestration can be used to automate and accelerate the flow of data from multiple sources, including existing applications and data platforms as well as the output of large language models and vector databases. Almost one-half (49%) of participants in ISG's 2023 Application Development and Maintenance Study expect to AI-enable applications by embedding AI and ML models into current applications and processes.

The ISG Buyers Guide™ for Data Orchestration evaluates software providers and products in key areas, including data pipeline management, workflow management and pipeline deployment. This research evaluates the following software providers that offer products to address key elements of data orchestration as we define it: Alteryx, Amazon Web Services, Astronomer, BMC, Cloudera, Dagster Labs, Databricks, DataKitchen, DataOps.live, dbt Labs,



Google, Hitachi, IBM, Informatica, Infoworks, K2View, Keboola, Matillion, Microsoft, Nexla, Prefect, Rivery, Saagie, SAP, Stonebranch, Y42 and Zoho.





## Buyers Guide Overview

For over two decades, ISG Research has conducted market research in a spectrum of areas across business applications, tools and technologies. We have designed the Buyers Guide to provide a balanced perspective of software providers and products that is rooted in an understanding of the business requirements in any enterprise. Utilization of our research



**ISG Research has designed the Buyers Guide to provide a balanced perspective of software providers and products that is rooted in an understanding of business requirements in any enterprise.**

methodology and decades of experience enables our Buyers Guide to be an effective method to assess and select software providers and products. The findings of this research undertaking contribute to our comprehensive approach to rating software providers in a manner that is based on the assessments completed by an enterprise.

The ISG Buyers Guide™ for Data Orchestration is the distillation of over a year of market and product research efforts. It is an assessment of how well software providers' offerings address enterprises' requirements for data orchestration software. The index is structured to support a request for information (RFI) that could be used in the request for proposal (RFP) process by incorporating all criteria needed to evaluate, select, utilize and maintain relationships with software providers. An effective product and customer experience with a provider can ensure the best long-term relationship and value achieved from a resource and financial investment.

In this Buyers Guide, ISG Research evaluates the software in seven key categories that are weighted to reflect buyers' needs based on our expertise and research. Five are product-experience related: Adaptability, Capability, Manageability, Reliability, and Usability. In addition, we consider two customer-experience categories: Validation, and Total Cost of Ownership/Return on Investment (TCO/ROI). To assess functionality, one of the components of Capability, we applied the ISG Research Value Index methodology and blueprint, which links the personas and processes for data orchestration to an enterprise's requirements.

The structure of the research reflects our understanding that the effective evaluation of software providers and products involves far more than just examining product features, potential revenue or customers generated from a provider's marketing and sales efforts. We believe it is important to take a comprehensive, research-based approach, since making the wrong choice of data orchestration technology can raise the total cost of ownership, lower the return on investment and hamper an enterprise's ability to reach its full performance potential. In addition, this approach can reduce the project's development and deployment



time and eliminate the risk of relying on a short list of software providers that does not represent a best fit for your enterprise.

ISG Research believes that an objective review of software providers and products is a critical business strategy for the adoption and implementation of data orchestration software and applications. An enterprise's review should include a thorough analysis of both what is possible and what is relevant. We urge enterprises to do a thorough job of evaluating data orchestration systems and tools and offer this Buyers Guide as both the results of our in-depth analysis of these providers and as an evaluation methodology.



# How To Use This Buyers Guide

## Evaluating Software Providers: The Process

We recommend using the Buyers Guide to assess and evaluate new or existing software providers for your enterprise. The market research can be used as an evaluation framework to establish a formal request for information from providers on products and customer experience and will shorten the cycle time when creating an RFI. The steps listed below provide a process that can facilitate best possible outcomes.

1. Define the business case and goals.  
Define the mission and business case for investment and the expected outcomes from your organizational and technological efforts.
2. Specify the business needs.  
Defining the business requirements helps identify what specific capabilities are required with respect to people, processes, information and technology.
3. Assess the required roles and responsibilities.  
Identify the individuals required for success at every level of the enterprise from executives to frontline workers and determine the needs of each.
4. Outline the project's critical path.  
What needs to be done, in what order and who will do it? This outline should make clear the prior dependencies at each step of the project plan.
5. Ascertain the technology approach.  
Determine the business and technology approach that most closely aligns to your enterprise's requirements.
6. Establish software provider evaluation criteria.  
Utilize the product experience: Adaptability, Capability, Manageability, Reliability and Usability, and the customer experience in TCO/ROI and Validation.
7. Evaluate and select the technology properly.  
Weight the categories in the technology evaluation criteria to reflect your enterprise's priorities to determine the short list of software providers and products.
8. Establish the business initiative team to start the project.  
Identify who will lead the project and the members of the team needed to plan and execute it with timelines, priorities and resources.





# The Findings

All of the products we evaluated are feature-rich, but not all the capabilities offered by a software provider are equally valuable to types of workers or support everything needed to manage products on a continuous basis. Moreover, the existence of too many capabilities may be a negative factor for an enterprise if it introduces unnecessary complexity. Nonetheless, you may decide that a larger number of features in the product is a plus, especially if some of them match your enterprise’s established practices or support an initiative that is driving the purchase of new software.

Factors beyond features and functions or software provider assessments may become a deciding factor. For example, an enterprise may face budget constraints such that the TCO evaluation can tip the balance to one provider or another. This is where the Value Index methodology and the appropriate category weighting can be applied to determine the best fit of software providers and products to your specific needs.

## Overall Scoring of Software Providers Across Categories

The research finds Databricks atop the list, followed by Microsoft and Alteryx. Providers that place in the top three of a category earn the designation of Leader. Informatica has done so in five categories; Microsoft and SAP in three; AWS, Databricks and Google in two; and Alteryx, Astronomer, BMC, Keboola and Stonebranch in one category.

The overall representation of the research below places the rating of the Product Experience and Customer Experience on the x and y axes, respectively, to provide a visual representation and classification of the software providers. Those providers whose Product Experience have a higher weighted performance to the axis in aggregate of the five product categories place farther to the right, while the performance and weighting for the two Customer Experience categories determines placement on the vertical axis. In short, software providers that place closer to the upper-right on this chart performed better than those closer to the lower-left.

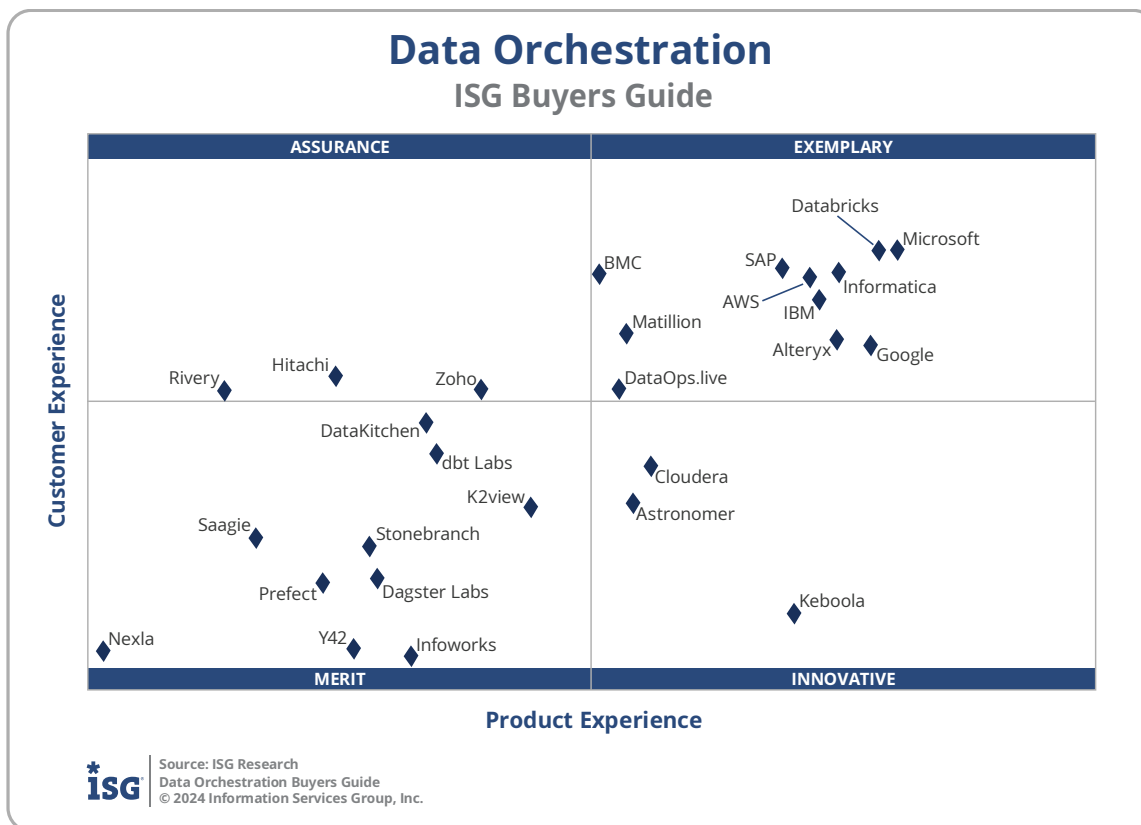
**Data Orchestration**  
Overall

Providers	Grade	Performance
Databricks	A-	<b>Leader</b> 82.8%
Microsoft	A-	<b>Leader</b> 82.7%
Alteryx	B++	<b>Leader</b> 81.1%
Google	B++	80.3%
IBM	B++	79.6%
Informatica	B++	79.2%
AWS	B++	79.2%
SAP	B++	78.1%
BMC	B++	75.7%
Keboola	B+	74.7%
DataOps.live	B+	73.9%
Matillion	B+	73.1%
Astronomer	B+	72.5%
Cloudera	B+	72.3%
DataKitchen	B+	69.8%
K2view	B	68.8%
Zoho	B	67.6%
dbt Labs	B	66.2%
Hitachi	B	66.0%
Dagster Labs	B	64.6%
Y42	B-	62.3%
Rivery	B-	62.2%
Stonebranch	B-	62.2%
Prefect	B-	62.0%
Infoworks	B-	61.7%
Saagie	B-	60.8%
Nexla	C++	55.1%

**ISG** Source: ISG Research  
Data Orchestration Buyers Guide  
© 2024 Information Services Group, Inc.



The research places software providers into one of four overall categories: Assurance, Exemplary, Merit or Innovative. This representation classifies providers' overall weighted performance.



**Exemplary:** The categorization and placement of software providers in Exemplary (upper right) represent those that performed the best in meeting the overall Product and Customer Experience requirements. The providers rated Exemplary are: Alteryx, AWS, BMC, Databricks, DataOps.live, Google, IBM, Informatica, Matillion, Microsoft and SAP.

**Innovative:** The categorization and placement of software providers in Innovative (lower right) represent those that performed the best in meeting the overall Product Experience requirements but did not achieve the highest levels of requirements in Customer Experience. The providers rated Innovative are: Astronomer, Cloudera, and Keboola.

**Assurance:** The categorization and placement of software providers in Assurance (upper left) represent those that achieved the highest levels in the overall Customer Experience requirements but did not achieve the highest levels of Product Experience. The providers rated Assurance are: Hitachi, Rivery and Zoho.



**Merit:** The categorization of software providers in Merit (lower left) represents those that did not exceed the median of performance in Customer or Product Experience or surpass the threshold for the other three categories. The providers rated Merit are: Dagster Labs, DataKitchen, dbt Labs, Infoworks, K2view, Nexla, Prefect, Saagie, Stonebranch and Y42.

We warn that close provider placement proximity should not be taken to imply that the packages evaluated are functionally identical or equally well suited for use by every enterprise or for a specific process. Although there is a high degree of commonality in how enterprises handle data orchestration, there are many idiosyncrasies and differences in how they do these functions that can make one software provider's offering a better fit than another's for a particular enterprise's needs.

We advise enterprises to assess and evaluate software providers based on organizational requirements and use this research as a supplement to internal evaluation of a provider and products.



## Product Experience

The process of researching products to address an enterprise’s needs should be comprehensive. Our Value Index methodology examines Product Experience and how it aligns with an enterprise’s life cycle of onboarding, configuration, operations, usage and maintenance. Too often, software providers are not evaluated for the entirety of the product; instead, they are evaluated on market execution and vision of the future, which are flawed since they do not represent an enterprise’s requirements but how the provider operates. As more software providers orient to a complete product experience, evaluations will be more robust.

The research results in Product Experience are ranked at 80%, or four-fifths, of the overall rating using the specific underlying weighted category performance. Importance was placed on the categories as follows: Usability (10%), Capability (25%), Reliability (15%), Adaptability (15%) and Manageability (15%). This weighting impacted the resulting overall ratings in this research. Microsoft, Databricks and Google were designated Product Experience Leaders.

### Data Orchestration Product Experience

Providers	Grade	Performance
Microsoft	A-	<b>Leader</b> 66.4%
Databricks	A-	<b>Leader</b> 66.0%
Google	A-	<b>Leader</b> 65.7%
Informatica	B++	64.8%
Alteryx	B++	64.7%
IBM	B++	64.1%
AWS	B++	63.9%
Keboola	B++	63.4%
SAP	B++	63.1%
Cloudera	B+	59.2%
Astronomer	B+	58.6%
Matillion	B+	58.4%
DataOps.live	B+	58.2%
BMC	B+	57.4%
K2view	B+	55.6%
Zoho	B	54.1%
dbt Labs	B	52.8%
DataKitchen	B	52.5%
Infoworks	B	52.0%
Dagster Labs	B	51.0%
Stonebranch	B	50.8%
Y42	B	50.3%
Hitachi	B-	49.8%
Prefect	B-	49.4%
Saagie	B-	47.4%
Rivery	B-	46.5%
Nexla	C++	42.9%



Source: ISG Research  
Data Orchestration Buyers Guide  
© 2024 Information Services Group, Inc.



## Customer Experience

The importance of a customer relationship with a software provider is essential to the actual success of the products and technology. The advancement of the Customer Experience and the entire life cycle an enterprise has with its software provider is critical for ensuring satisfaction in working with that provider. Technology providers that have chief customer officers are more likely to have greater investments in the customer relationship and focus more on their success. These leaders also need to take responsibility for ensuring this commitment is made abundantly clear on the website and in the buying process and customer journey.

The research results in Customer Experience are ranked at 20%, or one-fifth, using the specific underlying weighted category performance as it relates to the framework of commitment and value to the software provider-customer relationship. The two evaluation categories are Validation (10%) and TCO/ROI (10%), which are weighted to represent their importance to the overall research.

The software providers that evaluated the highest overall in the aggregated and weighted Customer Experience categories are Databricks, Microsoft and SAP. These category leaders best communicate commitment and dedication to customer needs. While not Leaders, Informatica and BMC were also found to meet a broad range of enterprise customer experience requirements.

Software providers that did not perform well in this category were unable to provide sufficient customer case studies to demonstrate success or articulate their commitment to customer experience and an enterprise's journey. The selection of a software provider means a continuous investment by the enterprise, so a holistic evaluation must include examination of how they support their customer experience.

### Data Orchestration Customer Experience

Providers	Grade	Performance
Databricks	A-	<b>Leader</b> 17.4%
Microsoft	A-	<b>Leader</b> 17.4%
SAP	A-	<b>Leader</b> 17.2%
Informatica	A-	17.1%
BMC	A-	17.1%
AWS	A-	16.9%
IBM	A-	16.7%
Matillion	B++	16.2%
Alteryx	B++	16.1%
Google	B++	16.0%
Hitachi	B++	15.5%
Rivery	B++	15.2%
Zoho	B++	15.2%
DataOps.live	B++	15.2%
DataKitchen	B+	14.9%
dbt Labs	B+	14.4%
Cloudera	B+	14.3%
Astronomer	B	13.7%
K2view	B	13.7%
Saagie	B	13.2%
Stonebranch	B	13.1%
Dagster Labs	B	12.6%
Prefect	B	12.5%
Keboola	B-	12.2%
Y42	B-	11.6%
Nexla	B-	11.5%
Infoworks	B-	11.4%

**ISG** | Source: ISG Research  
Data Orchestration Buyers Guide  
© 2024 Information Services Group, Inc.



## Appendix: Software Provider Inclusion

For inclusion in the ISG Buyers Guide™ for Data Orchestration in 2024, a software provider must be in good standing financially and ethically, have at least \$10 million in annual or projected revenue verified using independent sources, sell products and provide support on at least two continents and have at least 50 workers. The principal source of the relevant business unit's revenue must be software-related, and there must have been at least one major software release in the past 18 months.

The software provider must provide a product or products that support agile and collaborative data operations and marketed as addressing at least one of the following functional areas, which are mapped into the Buyers Guide capability criteria: pipeline management, workflow management and pipeline monitoring.

Data orchestration enables the flow of data across the organization via capabilities for pipeline monitoring, pipeline management and workflow management. Given the increasing complexity of evolving data sources and requirements, it's essential to automate and coordinate the creation, scheduling and monitoring of data pipelines.

To be included in this Buyers Guide requires functionality that addresses the following sections of the capabilities document:

- Pipeline management
- Workflow management
- Pipeline monitoring

The research is designed to be independent of the specifics of software provider packaging and pricing. To represent the real-world environment in which businesses operate, we include providers that offer suites or packages of products that may include relevant individual modules or applications. If a software provider is actively marketing, selling and developing a product for the general market and it is reflected on the provider's website that the product is within the scope of the research, that provider is automatically evaluated for inclusion.

All software providers that offer relevant data orchestration products and meet the inclusion requirements were invited to participate in the evaluation process at no cost to them.

Software providers that meet our inclusion criteria but did not completely participate in our Buyers Guide were assessed solely on publicly available information. As this could have a significant impact on classification and ratings, we recommend additional scrutiny when evaluating those providers.





## Products Evaluated

Provider	Product Names	Version	Release Month/Year
Alteryx	Analytics Cloud	N/A	October 2024
Astronomer	Astro	N/A	October 2024
AWS	Amazon Managed Workflows for Apache Airflow	N/A	September 2024
BMC	Control-M	9.0.21.300	October 2024
Cloudera	Data Platform - Data Engineering	1.22.0-h1	August 2024
Dagster Labs	Dagster+	1.8.12	October 2024
Databricks	Data Intelligence Platform	N/A	October 2024
DataKitchen	DataOps Automation	1.2.9	February 2024
DataOps.live	DataOps.live	October 2024	October 2024
dbt Labs	dbt	October 2024	October 2024
Google	Cloud Data Fusion Cloud Dataflow	N/A N/A	October 2024 September 2024
Hitachi	Pentaho Data Integration	10.2	September 2024
IBM	Cloud Pak for Data	5.0	September 2024
Informatica	Intelligent Data Management Cloud	October 2024	October 2024
Infoworks	Infoworks	6.1.0	September 2024
K2view	Data Product Platform	8.1.1	October 2024
Keboola	Keboola	N/A	November 2024
Matillion	Data Productivity Cloud	N/A	October 2024
Microsoft	Fabric	October 2024	October 2024



Nexla	Nexla	N/A	October 2024
Prefect	Prefect Cloud	3.0	September 2024
Rivery	Rivery	October 2024	October 2024
Saagie	Saagie	September 2024	September 2024
SAP	Data Intelligence Cloud Datasphere	N/A 2024.20	April 2024 September 2024
Stonebranch	Universal Automation Center	7.7.0.0	October 2024
Y42	Y42	N/A	July 2024
Zoho	DataPrep	2.0	October 2024



## Providers of Promise

We did not include software providers that, as a result of our research and analysis, did not satisfy the criteria for inclusion in this Buyers Guide. These are listed below as “Providers of Promise.”

<b>Provider</b>	<b>Product</b>	<b>Annual Revenue &gt;\$10M</b>	<b>Operates on 2 Continents</b>	<b>At Least 50 Employees</b>	<b>GA Product/ Documentation</b>
Ascend	Data Automation Cloud	No	Yes	No	Yes
Datacoves	Datacoves	No	Yes	No	Yes
Datafold	Datafold	No	Yes	No	Yes
Kestra	Kestra Enterprise	No	Yes	No	Yes
Orchestra Technologies	Orchestra	No	Yes	No	Yes
Promethium	Promethium	No	Yes	No	Yes
PurpleCube AI	PurpleCube AI	No	Yes	No	Yes
Saturam	Qualdo, Piperr	No	Yes	Yes	No
Switchboard Software	Data Automation	No	Yes	No	Yes
Torana	iceDQ	No	Yes	Yes	No



## About ISG Software Research

ISG Software Research provides expert market insights on vertical industries, business, AI and IT through comprehensive consulting, advisory and research services with world-class industry analysts and client experience. Our ISG Buyers Guides offer comprehensive ratings and insights into technology providers and products. Explore our research at [www.isg-research.net](http://www.isg-research.net).

## About ISG Research

ISG Research provides subscription research, advisory consulting and executive event services focused on market trends and disruptive technologies driving change in business computing. ISG Research delivers guidance that helps businesses accelerate growth and create more value. For more information about ISG Research subscriptions, please email [contact@isg-one.com](mailto:contact@isg-one.com).

## About ISG

ISG (Information Services Group) (Nasdaq: [III](#)) is a leading global technology research and advisory firm. A trusted business partner to more than 900 clients, including more than 75 of the world's top 100 enterprises, ISG is committed to helping corporations, public sector organizations, and service and technology providers achieve operational excellence and faster growth. The firm specializes in digital transformation services, including AI and automation, cloud and data analytics; sourcing advisory; managed governance and risk services; network carrier services; strategy and operations design; change management; market intelligence and technology research and analysis. Founded in 2006 and based in Stamford, Conn., ISG employs 1,600 digital-ready professionals operating in more than 20 countries—a global team known for its innovative thinking, market influence, deep industry and technology expertise, and world-class research and analytical capabilities based on the industry's most comprehensive marketplace data. For more information, visit [isg-one.com](http://isg-one.com).