

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

by Boris Evelson
June 17, 2020

Why Read This Report

In our 26-criterion evaluation of document-focused AI-based text analytics platforms, we identified the nine most significant ones — AntWorks, EPAM, Expert System, Google, IBM, Micro Focus, OpenText, RavenPack, and SAS — and researched, analyzed, and scored them. This report shows how each provider measures up and helps AD&D professionals and their governance, risk, and compliance (GRC) colleagues select the right one for their needs.

Additionally, Forrester recommends that AD&D pros consider the text analytics platforms evaluated in “The Forrester Wave™: AI-Based Text Analytics Platforms (People Focused), Q2 2020,” which uses similar evaluation criteria for more people-oriented use cases.

Key Takeaways

IBM, SAS, And Micro Focus Lead The Pack

Forrester’s research uncovered a market in which IBM, SAS, and Micro Focus are Leaders; Google, OpenText, and Expert System are Strong Performers; and EPAM, RavenPack, and AntWorks are Contenders.

Knowledge-Based AI Is A Key Differentiator

As machine learning (ML) technology becomes more widespread, knowledge-based AI will dictate which providers lead the pack. Vendors that can provide a combination of knowledge- and ML-based text analytics capabilities will position themselves to successfully deliver document-focused text analytics applications to their customers.

Document-Focused Text Analytics Platforms Require Additional Capabilities

Document capture options and mining text from images and cursive writing in multiple languages are key differentiators for document-focused enterprise text analytics platforms, which focus on analyzing relatively long documents, such as contracts, insurance claims, invoices, and purchase orders.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up



by [Boris Evelson](#)
with [Srividya Sridharan](#) and Aldila Yunus
June 17, 2020

Table Of Contents

- 2 Variety Of Document Ingestion Options Is Key To Modern Text Analytics
- 3 Evaluation Summary
- 6 Vendor Offerings
- 6 Vendor Profiles
 - Leaders
 - Strong Performers
 - Contenders
- 11 Evaluation Overview
 - Vendor Inclusion Criteria
- 12 Supplemental Material

Related Research Documents

- [The Forrester Wave™: AI-Based Text Analytics Platforms \(People Focused\), Q2 2020](#)
- [The Forrester Wave™: Cognitive Search, Q2 2019](#)
- [Now Tech: AI-Based Text Analytics Platforms, Q1 2020](#)
- [Research Overview: Text Analytics Technology](#)



Share reports with colleagues.
Enhance your membership with
Research Share.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

Variety Of Document Ingestion Options Is Key To Modern Text Analytics

Many enterprises are well on their way to effectively leveraging their data, processes, and applications to derive insights from unstructured data — text-based data, specifically. But their efforts are still very siloed. Enterprises often use separate platforms for document capture, document categorization and classification, robotic process automation (RPA), and other document-focused applications, even though under the covers these are all largely based on similar text mining and text analytics technologies. The good news is that text analytics technology, specifically AI-based text analytics, has matured to a point where enterprises can use it for multiple use cases.

Due to market evolution and the specific requirements for document-focused text analytics apps, application development and delivery (AD&D) pros looking to select document-focused AI-based text analytics platforms should look for providers that:

- › **Capture and digitize documents at scale.** Earlier-generation document capture technology was largely template based, and it worked relatively well as long as all documents conformed to the same invoice or a purchase order template. But that simplistic approach did not scale, and new platforms now offer template- and zone-free capture of document objects, digitization of word pairs (a label and a value, such as a text string “date” and the actual date of the document), tables (by programmatically determining the number of rows and columns, identifying data types in each cell, intelligently handling merged cells, etc.), symbols (e.g., checkboxes), and myriad other complex document structures. And to ensure digitization quality, modern platforms use machine learning to stamp each digitized document with an OCR certainty level (the level of confidence that the text was captured correctly).
- › **Improve capture and digitization quality of messy documents.** In an ideal world, all documents would be perfectly scanned — the right side is up and the image is crisp. But the world is not perfect, and older technologies didn’t work when documents were scanned upside down, sideways, or were blurry. Newer technology addresses these unavoidable realities with image cleanup and enhancement. Leading text analytics platforms can enhance low-resolution images, convert color images to black and white, fix image tilting/skewness, recognize upside-down and sideways pages, and recognize and suppress watermarks. All of this innovation ensures higher-quality digitization and fewer breakages in the process.
- › **Digitize nontextual document content.** Unreachable to earlier-generation technology, text analytics platforms now use computer vision technology based on neural networks and deep learning to recognize logos and barcodes — ubiquitous in many modern documents. Some platforms can also tag image objects such as buildings and cars, including damage — this is especially useful in insurance claims processing. An increasing number of text analytics platforms also support digitizing cursive text, including but not limited to signatures.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

- › **Squeeze richer insights from documents compared to earlier-generation NLP.** AI-based text analytics platforms take natural language processing (NLP) to the next level — beyond extracting topics and entities (dates, people, places, and events). Natural language understanding (NLU) capabilities can tag individual phrases, paragraphs, and whole documents with scores for fraud (useful in insurance claims processing) or data privacy (key for GRC). Additionally, these platforms can resolve intradocument and interdocument conflicts when one part of a document negates something mentioned elsewhere in the document. For example, earlier-generation technology will treat a long list of appliances in a real estate rental contract as document topics and ignore a “not included in the contract” sentence that appears at the end of the list or on a different page. The icing on the cake? Modern AI-based text analytics platforms can do all of the above in multiple languages.¹

Evaluation Summary

The Forrester Wave evaluation highlights Leaders, Strong Performers, Contenders, and Challengers. It's an assessment of the top vendors in the market and does not represent the entire vendor landscape. You'll find more information about this market in our reports [“The Forrester Tech Tide™: Artificial Intelligence For Business Insights, Q3 2018,”](#) [“Now Tech: AI-Based Text Analytics Platforms, Q1 2020,”](#) and [“Research Overview: Text Analytics Technology.”](#)

We intend this evaluation to be a starting point only and encourage clients to view product evaluations and adapt criteria weightings using the Excel-based vendor comparison tool (see Figure 1 and see Figure 2). Click the link at the beginning of this report on Forrester.com to download the tool.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

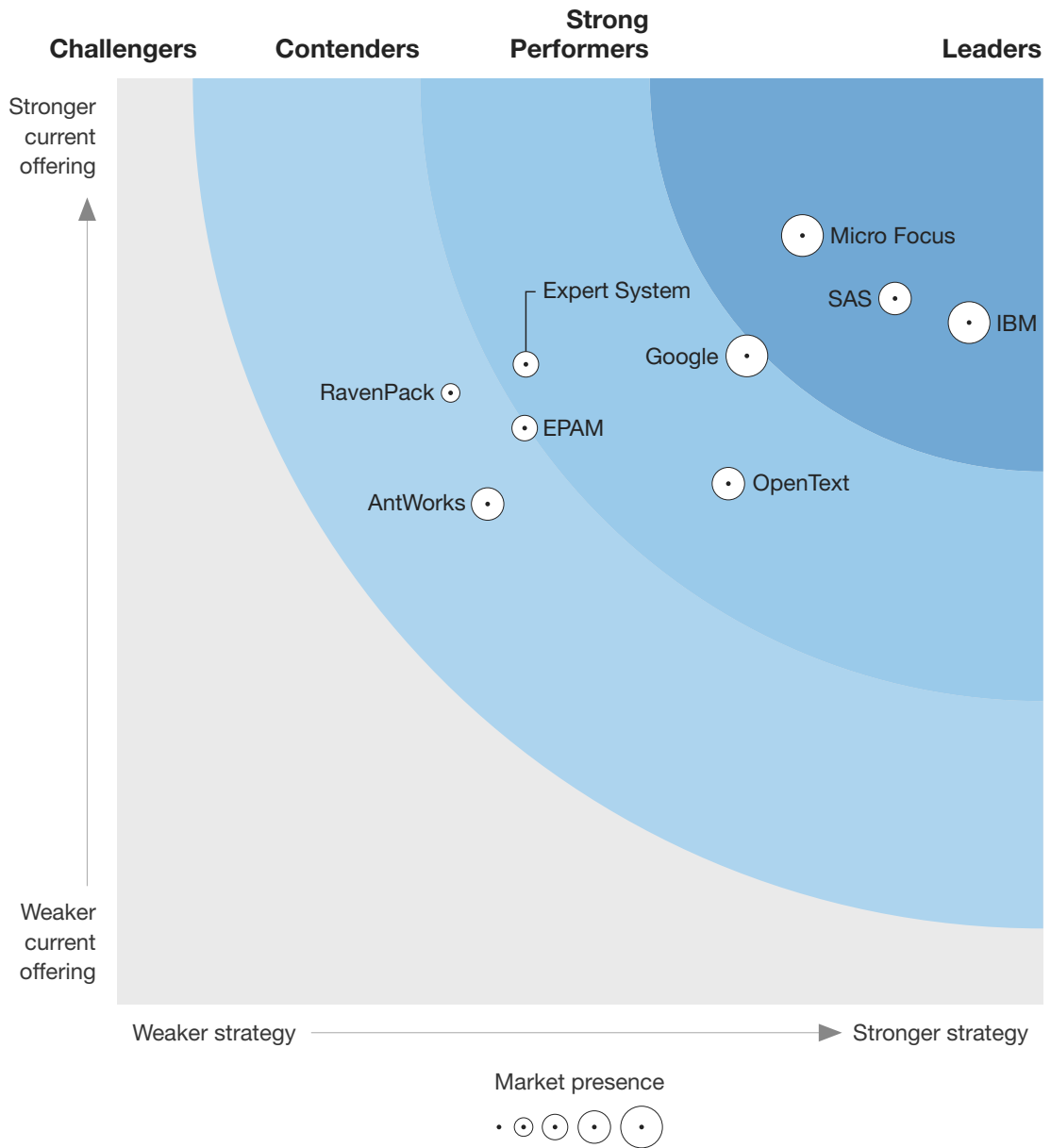
The Nine Providers That Matter Most And How They Stack Up

FIGURE 1 Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

THE FORRESTER WAVE™

AI-Based Text Analytics Platforms (Document Focused)

Q2 2020



The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

FIGURE 2 Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused) Scorecard, Q2 2020

	Forrester's weighting	AntWorks	EPAM	Expert System	Google	IBM	Micro Focus	OpenText	RavenPack	SAS
Current offering	50%	2.74	3.11	3.47	3.50	3.68	4.15	2.81	3.30	3.81
Data preparation	15%	3.00	3.40	1.00	5.00	3.40	5.00	2.40	2.00	2.40
Data enrichment	15%	1.20	3.00	4.80	3.00	3.00	5.00	3.20	4.60	3.00
Subdocument analysis: NLP	15%	1.00	3.00	5.00	3.00	3.00	5.00	3.00	3.00	5.00
Document/cross-document analysis	10%	4.00	3.00	5.00	3.00	3.00	5.00	5.00	1.00	3.00
Machine learning capabilities	10%	4.60	3.00	1.00	5.00	4.20	5.00	4.20	2.60	5.00
Applications	5%	2.00	4.00	3.00	1.00	3.00	4.00	2.00	2.00	1.00
GUI and collaboration	15%	3.00	3.00	5.00	3.00	5.00	1.00	1.00	5.00	5.00
Architecture	5%	3.00	1.00	3.00	3.00	5.00	3.00	3.00	3.00	5.00
Security	5%	5.00	3.00	1.00	5.00	5.00	5.00	3.00	5.00	5.00
Administration and usability	5%	3.00	5.00	3.00	3.00	3.00	3.00	1.00	5.00	3.00
Strategy	50%	2.00	2.20	2.20	3.40	4.60	3.70	3.30	1.80	4.20
Product vision	30%	1.00	3.00	3.00	3.00	5.00	5.00	3.00	1.00	5.00
Execution	5%	5.00	3.00	3.00	3.00	5.00	3.00	3.00	3.00	3.00
Performance	15%	5.00	1.00	1.00	3.00	5.00	3.00	1.00	3.00	1.00
Market approach	20%	1.00	3.00	3.00	3.00	5.00	1.00	3.00	3.00	5.00
Partnerships	20%	1.00	1.00	1.00	5.00	3.00	5.00	5.00	1.00	5.00
Supporting products and services	5%	1.00	3.00	1.00	5.00	5.00	5.00	5.00	1.00	5.00
Delivery model	5%	5.00	1.00	3.00	1.00	5.00	3.00	5.00	1.00	3.00
Market presence	0%	3.40	3.00	3.00	5.00	5.00	5.00	3.40	1.40	3.40
Revenue and customers	80%	3.00	3.00	3.00	5.00	5.00	5.00	3.00	1.00	3.00
Global presence	20%	5.00	3.00	3.00	5.00	5.00	5.00	5.00	3.00	5.00

All scores are based on a scale of 0 (weak) to 5 (strong).

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

Vendor Offerings

Forrester included nine vendors in this assessment: AntWorks, EPAM, Expert System, Google, IBM, Micro Focus, OpenText, RavenPack, and SAS (see Figure 3).

FIGURE 3 Evaluated Vendors And Product Information

Vendor	Product evaluated	Product version evaluated
AntWorks	Cognitive Machine Reading (CMR)	ANTstein SQUARE
EPAM	InfoNgen	Version 8
Expert System	Cogito	14.5
Google	Google Cloud AI	
IBM	IBM Watson Discovery	
Micro Focus	Micro Focus IDOL	12.5
OpenText	OpenText Magellan	16.7
RavenPack	RavenPack News Analytics; RavenPack Text Analytics	1.0
SAS	SAS Visual Text Analytics (VTA)	3.5

Vendor Profiles

Our analysis uncovered the following strengths and weaknesses of individual vendors.

Leaders

- › **IBM Watson Discovery streamlines multiple products into a holistic platform.** Since our past evaluation, IBM has consolidated and rationalized its text analytics platform.² IBM Watson Discovery wraps IBM's NLU APIs (for extracting entities, categories, and concepts) and comes packaged with Watson Knowledge Studio for custom model development. The product's what-you-see-is-what-you-get UI to train models to categorize text-based docs and convert scanned and PDF docs to text is intuitive — you can visually track the step-by-step improvement of document “understanding” (sections, tables, etc.). Clients can use IBM Watson Visual Recognition (sold separately) to classify and analyze objects in images embedded in docs. IBM also positions

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

Watson Discovery as a [cognitive search](#) platform; however, IBM works with partners on other doc-focused solutions (such as eDiscovery and Contract Analytics). All text analytics apps require ample training to become highly accurate, and IBM clients can benefit from engaging with IBM Expert Labs or IBM GBS for professional services to train, customize, deploy, and scale applications based on Watson Discovery.

Although Watson Discovery offers handwritten signature detection, it can't recognize cursive text — an increasingly popular feature of document-focused apps. The product team has introduced new capabilities, including innovations from IBM Research. But IBM's sheer size — and the number of acquisitions, legacy products, and products it needs to integrate — diverts resources, sometimes at the expense of innovation. A few reference clients called out the lack of innovations in Watson Discovery. Clients often tell us it's challenging to navigate the IBM landscape. IBM needs to communicate better about its complex product catalog.

- › **SAS Visual Text Analytics bolsters SAS's family of formidable analytics products.** SAS Visual Text Analytics is one of several applications built on the SAS Viya platform, where all applications share data and model management, a business intelligence (BI) and analytics GUI, and other microservices, resulting in consistent UX. Additionally, SAS Visual Text Analytics customers will benefit from augmented BI capabilities — including an especially useful feature based on natural language generation (NLG) that explains results and what-ifs. This feature is available in SAS Visual Analytics (SAS's [enterprise BI platform](#)), which is the visualization/analytics component of SAS Visual Text Analytics. Core features for OCR and image/cursive writing analytics are available natively in Visual Text Analytics; however, for customized models, clients will have to purchase separate products: SAS Event Stream Processing for high-speed document capture and SAS Visual Data Mining and Machine Learning (VDMML) for data scientists looking to build and customize native SAS code.³

Most competitors package domain-specific ontologies with their products. SAS chooses to sell “industry packs” separately, but the vendor includes predefined concepts and sentiment taxonomies out of the box (OOTB), at no additional charge. While SAS Visual Text Analytics' UX is seamless, the UX becomes less harmonious for users diving deep into building and customizing models in SAS VDMML and SAS Studio.

- › **Micro Focus IDOL remains a spaceship of unstructured data analytics.** At most organizations, unstructured analytics is still federated — teams use different tools for people- and document-focused applications, as well as different tools for text, speech, and computer vision analytics. Organizations ready for one platform to address all of the unstructured data analytics should put Micro Focus IDOL on their shortlist. IDOL comes OOTB with more data connectors (based on Apache NiFi and comprehensive data ingestion, transformation, and routing capabilities), more language support, and more domain-specific ontologies than most of its competitors. Companies can instantly leverage IDOL as an enterprise [cognitive search](#) platform; however, most of the other document-focused solutions are sold as separate products or made available through indirect channels (such as via partners or OEMs).

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

Users conscious of data privacy will like all of IDOL's capabilities that treat personally identifiable information (PII) data beyond just masking text — IDOL has APIs to blur faces and other sensitive information embedded in document images. The product's breadth and depth are also its weakness — it's not an easy tool to use for a non-IT professional. Users must maneuver an overwhelming number of screens as they develop data ingestion processes, build models, and analyze the results, and the UX is less than seamless. Client references confirmed the platform's complexity and gave IDOL lower marks for ease of use and time-to-value compared to competitors.

Strong Performers

- › **Google Cloud AI turns an assortment of NLP services into a platform-like offering.** Google's text analytics platform is a collection of multiple Google Cloud AI services and other services. It includes NLP APIs for entity and sentiment extraction, Document AI and Vision AI for document and image classification, OCR, and forms processing. Other relevant services include Data Fusion for data ingestion and preparation. It also offers AutoML, which enables developers and citizen data scientists with limited machine learning expertise to build custom models for document classification, entity extraction, and sentiment analysis. These services natively integrate with Google data lake (Google Cloud Storage) and BigQuery and Looker for analysis. A huge benefit of using Google's text analytics services is its crowdsourced knowledge graph, which competitors will struggle to keep up with. Last but not least, Google offers text analytics services on a pay-per-use basis, which is still a differentiator.

Even though Google has done a decent job of pulling together all of the relevant services into a platform-like look and feel, navigating its UX may be a daunting task for non-IT, nondata professionals. Also, Google does not offer any OOTB document-focused solutions — its approach to solutions is partner dependent. Lastly, all Google services are mostly available as cloud software-as-a-service (SaaS), which could be a concern for clients that require an on-premises deployment (a limitation confirmed by customer references).⁴

- › **OpenText Magellan aspires to be your one-stop shop for analytics.** Text analytics is one of the three core capabilities of OpenText Magellan, the other two being an [enterprise BI platform](#) and a [predictive analytics and machine learning platform](#) — all based on a big data Spark architecture. While some competitors prefer to work with BI vendor partners for comprehensive data visualization, analysis and slice-and-dice functionality, and ML-based predictive analytics, OpenText Magellan Text Analytics can simply function as a data ingestion and transformation step for its BI component. A user can switch between building text mining processes, analyzing results, and building predictive models without switching platforms. OpenText customers can use Captiva (a separately sold product) for OCR but will have to use partner products for image and cursive writing recognition.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

A bonus to OpenText customers — they will find Magellan embedded and integrated into multiple OpenText products for content management and business networks, such as enterprise content management, content capture, intelligent filing, information archiving, eDiscovery, and smart contracts. And because of OpenText's long tenure in the information management space, clients can reuse all of the vendor's relevant policies and rules (like record retention) in Magellan. OpenText mostly grows by acquisitions, and product integration is its specialty. However, based on the sheer number of products it acquires and frequency of acquisitions, integration does take time. As a result, a different UX sometimes bursts through the seams during the end-to-end text mining and analysis process.

- › **Expert System Cogito bets on out-of-the-box accuracy with its knowledge graph.** While AI will undoubtedly outpace humans in building a base of knowledge someday, that time has not yet come. For now, Expert System's knowledge graph, built over 20 years, beats purely ML-based text analytics apps during the initial months of client deployments, often achieving faster time-to-value, more predictable results, and lower overall cost of ownership (references gave Cogito high scores for its multilingual capabilities, ontologies, and time-to-value). Especially impressive is Cogito knowledge graph's availability in 12 languages, supporting full linguistic NLP functionality, including term disambiguation. In addition to its main knowledge-graph-based components (Discover for text mining and Intelligence for analytics and visualization), Cogito packages Studio for building custom ML-based text analytics models. Expert System also sells domain-based apps (e.g., for claims, underwriting, clinical research, and life sciences) and packages Cogito to support eDiscovery and contract analytics. While it lacks a native OCR and has only a basic forms recognition capability, orgs should consider Cogito for RPA apps, specifically involving long, complex documents. Cogito earns this recommendation based on its ability to identify near-duplicates and duplicates across documents, resolve other intradocument and interdocument conflicts, identify intradocument and interdocument entity links, and provide document summaries.

Machine learning, for now, is the product's Achilles' heel. While Expert System is increasing its investments in supervised and unsupervised ML models, in the near term, it will lag behind competitors that invested in ML technologies earlier.

Contenders

- › **EPAM InfoNgen remains a solid choice if you look for software/services combination.** No matter what any vendors claim, no text analytics application is as accurate on day one as it will be after a few weeks or months of diligent system training and customization. And this is precisely where EPAM's sweet spot lies — it is a large, global [systems integrator](#). InfoNgen is one of several products EPAM productized over a decade ago, and it is a respectable competitor to software pure-plays. The product comes with all of the capabilities an enterprise text analytics platform may require: numerous data connectors, ontologies, support for multiple languages via linguistic rules, and ML algorithms. While InfoNgen can be used for virtually any text analytics application, it's an especially good fit for RPA (via its advanced forms recognition and workflow automation).

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

EPAM also goes to market with domain-focused applications like Regulatory Compliance & Risk Avoidance and Counterfeit & Fraud Detection, as well as applications like Enterprise Search. We found InfoNgen's "newsletter" application, used to distribute insights derived from text mining, to be useful and unique. Additionally, end users will find a feature allowing them to validate platform tags (or not) and submit the results for review quite handy.

Reference clients give EPAM's text analytics professional services top scores; this is definitely a key strength. The InfoNgen platform is only available as cloud SaaS, which is often an issue for clients that are not ready to move sensitive data to public clouds.

- › **RavenPack can take over your text-heavy research as a BPO service.** RavenPack News Analytics is a business process outsourcing (BPO) service for market and competitive intelligence, risk management, and compliance.⁵ Its sweet spot is in financial services, specifically capital markets. Impressively, data scientists use News Analytics results as one of the inputs into securities trading models. News Analytics is based on an ontology and knowledge graph of corporate actions, events, and news — M&A, equity or debt issuing, changes in management positions, dividends, etc. Most of RavenPack's clients are in financial services, but recently the vendor started going after other industries with its general-purpose Text Analytics product. Since transparency is critical for highly regulated industries like financial services, the platform is based on a mix of rules and supervised machine learning so that it is always clear to users how the application produces the results.

News Analytics works best in domains with a significant amount of data — its rules and models may not deliver accurate analysis if there's little data available about a company or event. Prospects in financial services should shortlist RavenPack News Analytics if the BPO model aligns with their preference to outsource and tolerance for limited control of the application — clients need to submit a request to the vendor to add new data sources and make changes to the software, rules, and models. Prospects from other industries should proceed cautiously, as the general-purpose RavenPack Text Analytics is relatively new.

- › **AntWorks can turbocharge your document-based RPA.** AntWorks ANTstein SQUARE's Cognitive Machine Reading (CMR) is a document classification and text mining platform based on fractal science that learns from repeating image patterns. It has impressive capabilities to capture scanned documents, correct document images (remove image noise, autorotate inverted images and text, unskew images, sharpen blurred text, and fix slanting and misalignments), and convert data pairs (label and value) and forms to text (not based on templates and not tied to a location in the document). These features make the product attractive for RPA use cases, especially when ingesting and processing different types of documents in a single batch. The GUI for training is intuitive, appealing, and very different from competitive products.

CMR is 100% based on machine learning, so it does not come packaged with any domain-specific ontologies. As such, it requires some training — supervised learning — to be accurate.⁶ Currently, it does not come with an OOTB capability to fix spelling and grammar errors. The functionality to analyze the results of text image and text mining processes — slicing and dicing the data — is mostly limited to basic reporting and needs an upgrade.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

Evaluation Overview

We evaluated vendors against 26 criteria, which we grouped into three high-level categories:

- › **Current offering.** Each vendor's position on the vertical axis of the Forrester Wave graphic indicates the strength of its current offering. Key criteria for these solutions include variety of data ingestion options (not limited to text files but also including OCR, image, and cursive writing capture); data enrichment using domain-specific ontologies; subdocument, document, and cross-document NLP capabilities; machine learning functionality; ease of use for end users and administrators; and modern architecture (microservices, containers, serverless, and security).
- › **Strategy.** Placement on the horizontal axis indicates the strength of the vendors' strategies. We evaluated vendor and product vision, execution, performance, market approach, partnership ecosystem, supporting products and services, and delivery model.
- › **Market presence.** Represented by the size of the markers on the graphic, our market presence scores reflect each vendor's product revenue and geographical presence.

Vendor Inclusion Criteria

Forrester included nine vendors in the assessment: AntWorks, EPAM, Expert System, Google, IBM, Micro Focus, OpenText, RavenPack, and SAS. Each of these vendors has:

- › **A standalone, general-purpose text analytics platform.** We excluded vendors that embed text analytics functionality into solution-specific platforms like eDiscovery, smart contracts, RPA, and others.
- › **Document-focused text analytics capabilities.** Although all vendors in this evaluation can support the majority of text analytics use cases, this evaluation focuses on vendors' document-focused capabilities. These include capturing and digitizing documents via scanning or OCR; recognizing text in forms, tables, and images; and others.
- › **Significant market presence.** Included vendors demonstrated significant market presence via revenues, customers, global presence, and interest from Forrester clients.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

Engage With An Analyst

Gain greater confidence in your decisions by working with Forrester thought leaders to apply our research to your specific business and technology initiatives.

Analyst Inquiry

To help you put research into practice, connect with an analyst to discuss your questions in a 30-minute phone session — or opt for a response via email.

[Learn more.](#)

Analyst Advisory

Translate research into action by working with an analyst on a specific engagement in the form of custom strategy sessions, workshops, or speeches.

[Learn more.](#)

Webinar

Join our online sessions on the latest research affecting your business. Each call includes analyst Q&A and slides and is available on-demand.

[Learn more.](#)



Forrester's research apps for iOS and Android.

Stay ahead of your competition no matter where you are.

Supplemental Material

Online Resource

We publish all our Forrester Wave scores and weightings in an Excel file that provides detailed product evaluations and customizable rankings; download this tool by clicking the link at the beginning of this report on Forrester.com. We intend these scores and default weightings to serve only as a starting point and encourage readers to adapt the weightings to fit their individual needs.

The Forrester Wave Methodology

A Forrester Wave is a guide for buyers considering their purchasing options in a technology marketplace. To offer an equitable process for all participants, Forrester follows [The Forrester Wave™ Methodology Guide](#) to evaluate participating vendors.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

In our review, we conduct primary research to develop a list of vendors to consider for the evaluation. From that initial pool of vendors, we narrow our final list based on the inclusion criteria. We then gather details of product and strategy through a detailed questionnaire, demos/briefings, and customer reference surveys/interviews. We use those inputs, along with the analyst's experience and expertise in the marketplace, to score vendors, using a relative rating system that compares each vendor against the others in the evaluation.

We include the Forrester Wave publishing date (quarter and year) clearly in the title of each Forrester Wave report. We evaluated the vendors participating in this Forrester Wave using materials they provided to us by April 8, 2020, and did not allow additional information after that point. We encourage readers to evaluate how the market and vendor offerings change over time.

In accordance with [The Forrester Wave™ Vendor Review Policy](#), Forrester asks vendors to review our findings prior to publishing to check for accuracy. Vendors marked as nonparticipating vendors in the Forrester Wave graphic met our defined inclusion criteria but declined to participate in or contributed only partially to the evaluation. We score these vendors in accordance with [The Forrester Wave™ And The Forrester New Wave™ Nonparticipating And Incomplete Participation Vendor Policy](#) and publish their positioning along with those of the participating vendors.

Integrity Policy

We conduct all our research, including Forrester Wave evaluations, in accordance with the [Integrity Policy](#) posted on our website.

Endnotes

- ¹ Forrester advises clients to dig deeper than “how many languages does the platform support,” especially for the languages that matter for your geographic presence. For instance, ask questions about the number of languages the platform supports with full NLP, including word disambiguation, domain-specific ontologies, or spelling and grammar error correction. Usually, these fall into three tiers. Tier 1 languages support all of the above. Tier 2 language support is mostly based on text tokenization (finding sentences and words based on periods, spaces, and other delimiters) and can only analyze words based on sequences and proximity — a less accurate technique than tier 1 language support offers. Tier 3 language support is based on machine translation and is the least accurate option.
- ² At the time of 2018 Forrester Wave evaluation of multimodal predictive analytics and machine learning solutions, IBM had five different text analytics products and API libraries. It has consolidated most under Watson Discovery. IBM SPSS, evaluated separately, is still sold as a separate product and has text mining capabilities. See the Forrester report [“The Forrester Wave™: Multimodal Predictive Analytics And Machine Learning Solutions, Q3 2018.”](#)
- ³ SAS Studio — a interface to build and customize models — comes with all SAS Viya applications.
- ⁴ Clients can deploy their own version of all required Google services on-premises in containers — Google Anthos — but the approach carries multiple implications, such as the need to keep all relevant services in sync with Google Cloud updates.

The Forrester Wave™: AI-Based Text Analytics Platforms (Document Focused), Q2 2020

The Nine Providers That Matter Most And How They Stack Up

⁵ In this context, BPO differs from SaaS since the users have little to no control of the software — they hand the data over to the provider and get the results back.

⁶ The vendor claims that text analytics based on fractal science requires less training than other ML techniques, but Forrester cannot verify such a claim.

We work with business and technology leaders to drive customer-obsessed vision, strategy, and execution that accelerate growth.

PRODUCTS AND SERVICES

- › Research and tools
- › Analyst engagement
- › Data and analytics
- › Peer collaboration
- › Consulting
- › Events
- › Certification programs

Forrester's research and insights are tailored to your role and critical business initiatives.

ROLES WE SERVE

Marketing & Strategy Professionals

CMO

B2B Marketing

B2C Marketing

Customer Experience

Customer Insights

eBusiness & Channel Strategy

Technology Management Professionals

CIO

› Application Development & Delivery

Enterprise Architecture

Infrastructure & Operations

Security & Risk

Sourcing & Vendor Management

Technology Industry Professionals

Analyst Relations

CLIENT SUPPORT

For information on hard-copy or electronic reprints, please contact Client Support at +1 866-367-7378, +1 617-613-5730, or clientsupport@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.