February 4, 2010

The Forrester Wave™: Predictive Analytics And Data Mining Solutions, Q1 2010

by James Kobielus
for Business Process & Applications Professionals

February 4, 2010

The Forrester Wave™: Predictive Analytics And Data Mining Solutions, Q1 2010

SAS, SPSS/IBM, KXEN, Oracle, And Portrait Software Lead, With Others Innovating

by James Kobielus
with Boris Evelson, Rob Karel, and Charles Coit

EXECUTIVE SUMMARY

In Forrester's 53-criteria evaluation of predictive analytics and data mining (PA/DM) vendors, we found that SAS Institute, SPSS (evaluated separately from new parent IBM's other PA/DM offerings), KXEN, Oracle, Portrait Software, and IBM (pre-SPSS acquisition PA/DM offerings) head the pack with mature, sophisticated, scalable, flexible, and robust solutions. SAS leads, providing the most feature-rich solution portfolio and, through its recent expansion of enterprise data warehouse (EDW) vendor partnerships, taking the industry lead in promoting in-database analytics as an emerging best practice for high-performance analytics deployments. SPSS is rapidly integrating its already strong PA/DM solution portfolio with new parent IBM's extensive data management solution family. KXEN stands out for its focus on content analytics, sentiment analysis, and social network analysis. Oracle distinguishes itself through the depth of its PA/DM tool's integration into its enterprise database and application portfolio. Portrait provides a comprehensive set of customer analytics offerings that integrate with its core PA/DM tool. TIBCO Software, FICO, and Angoss Software are Strong Performers in the PA/DM market.

TABLE OF CONTENTS

NOTES & RESOURCES

Forrester conducted product evaluations in Q3 2009 and interviewed nine vendor companies: Angoss Software, FICO, IBM, KXEN, Oracle, Portrait Software, SAS Institute, SPSS (a product group of IBM), and TIBCO Software. We also interviewed two customer references from each vendor.

Related Research Documents

"In-Database Analytics: The Heart Of The Predictive Enterprise"
November 12, 2009

"Business Intelligence (BI) Polishes Its Crystal Ball"
August 18, 2009

"Refresh Your Information Management Strategy To Deliver Business Results"
August 11, 2009

PREDICTIVE POWERS DRIVE BUSINESS OPTIMIZATION

Business is all about placing bets and knowing if the odds are in your favor. Success depends on your company being able to understand, analyze, and visualize likely futures and take appropriate actions as soon as possible. You must be able to predict future scenarios well enough to prepare plans and deploy resources so that you can seize opportunities, neutralize threats, and mitigate risks.

Predictive analytics can play a pivotal role in the planning and day-to-day operations of your business. It can help you focus strategy and continually tweak plans based on actual performance and likely future scenarios. These tools can sit at the core of your component-based service-oriented architecture strategy as business process and applications (BP&A) professionals embed predictive logic deeply into data warehouses, business process management platforms, complex event processing streams, and operational applications.[1] Forrester recommends the following best practices for selecting a platform and for deploying predictive analytics applications.

Focus Data Mining On Polishing Your Predictive Crystal Ball

Predictive analytics solutions are key components of enterprise decision support environments. Typically integrated with data mining tools to support forecasting against structured data sets, predictive analytics is a centerpiece of many organizations' advanced analytics environments. Forrester defines advanced analytics as:

> *Any solution that supports the identification of meaningful patterns and correlations among variables in complex, structured and unstructured, historical, and potential future data sets for the purposes of predicting future events and assessing the attractiveness of various courses of action. Advanced analytics typically incorporate such functionality as data mining, descriptive modeling, econometrics, forecasting, operations research, optimization, predictive modeling, simulation, statistics, and text analytics.*

Predictive analytics and data mining (PA/DM) solutions often leverage and extend companies' investments in business intelligence (BI); enterprise data warehousing (EDW); extract, transform, and load (ETL); and other key analytics platforms, tools, and applications.[2] In addition, many PA/DM tools often perform some functions associated with text analytics or supplement standalone tools that specialize in supporting predictive analysis against unstructured or semistructured information sets.

Forrester construes a PA/DM solution as any software, hardware, and other enabling component that allows BP&A professionals to, at the very least:

- **Implement a wide range of data preparation features.** The tool should offer the ability to discover, extract, profile, transform, cleanse, load, and otherwise prepare analytical data sets sourced from structured formats such as relational database tables.

- **Perform sophisticated statistical modeling.** The tool should enable development and validation of predictive models that contain multiple independent variables, using such approaches as clustering (assignment of a set of observations into groups based on similarities), classification (prediction of the group membership of a data instance), scoring (application of a predictive model to new data to make predictions about unseen behavior), and regression (statistical technique for building models applicable to continuous variables) against structured, historical data sets.

- **Enable flexible visualization of predictive models and their results.** The tool should allow modelers to interactively visualize, explore, and manipulate predictive models to support forecasting, prediction, simulation, what-if analysis, and other functions to assess alternative courses of action.

- **Leverage multiple database platform capabilities.** The tool should support storage and management of analytical data sets and predictive models in an EDW or other analytical data mart or database platform.

Model Multiple Predictive Scenarios With Diverse Information And Algorithms

Most important, BP&A professionals should assess whether PA/DM tools support the following capabilities for developing, validating, and deploying a wide range of predictive models:

- **Model multiple business scenarios.** You should be able to build complex models of multiple, linked business scenarios across different business, process, and subject-area domains using such key features as strategy maps, ensemble modeling, and champion-challenger modeling.[3]

- **Incorporate multiple information types into models.** You should be able to develop models against multiple information types, including unstructured content and real-time event streams. One key text analytics technique is sentiment analysis, which leverages natural language processing and computational linguistics to determine the mood, attitude, evaluation, or propensity of someone with respect to a topic. Another is social network analysis, which mines behavioral data to identify relationships among people, groups, organizations, computers, and other connected entities.

- **Leverage multiple statistical algorithms and approaches in models.** You should be able to develop models using the widest, most sophisticated range of statistical and mathematical algorithms and approaches. Key among these approaches are regression, constraint-based optimization, neural networks, genetic algorithms, and support vector machines.[4]

- **Apply multiple metrics of model quality.** You should be able to score and validate model quality using multiple metrics and approaches to assess how well models fit the data. Key approaches include model quality scores, lift charts (graphical representation of the accuracy of the predictions of various models), and comparative model evaluation.[5]

- **Employ multiple variable discovery and assessment approaches.** You should be able to build and validate models using various approaches for variable discovery, profiling, and selection, including decision trees, feature selection, clustering, association rule learning, and outlier analysis.[6]

Deploy And Govern Your Predictive Models Effectively

While it's critical to build multiscenario predictive models against the complex information sets central to many advanced analytics projects, often that's not enough. Becoming a fully predictive enterprise demands much more. To instrument your organization for maximum predictive power, you should also tool your advanced analytics to support the following capabilities:

- **Support DW-integrated data preparation.** To speed up and standardize the most time-consuming predictive modeling project tasks, you should be able to leverage your existing data warehouse; ETL; data quality; and metadata tools to support a full range of data preparation features. These features include the ability to discover, acquire, capture, profile, sample, collect, collate, aggregate, deduplicate, transform, correct, augment, and load analytical data sets.[7]

- **Enable deep application and middleware integration.** To deliver models deeply into whatever heterogeneous SOA-enabled platform you happen to use, your predictive analytics tool should deploy on and/or integrate with a wide range of enterprise applications, middleware, operating platforms, and hardware substrate. You should be able to deploy models seamlessly into your data warehouse, business intelligence, online analytical processing, data integration, complex event processing, data quality, master data management, and business process management environments. And to play well in your SOA and other application environments, your predictive modeling tools should support application programming interfaces, languages (Java, C, C++), integrated development environments (IDEs, such as Eclipse and Visual Studio), Web services, open source statistical modeling environments such as R and Weka, as well as latest technologies such as SQL-MapReduce.

- **Provide consistent cross-domain model governance.** To avoid fostering an unmanageable glut of myriad models, your predictive analytics solution should support a wide range of tools, features, and interfaces to support a full life-cycle governance of models created in diverse tools. At the very least, your tools should enable model check-in/checkout, change tracking, version control, and collaborative development and validation of models. To realize this promise, it should support a full range of tools, standards, and interfaces for import and embedding of models from other tools, as well as export and sharing of models to other environments.[8]

- **Offer rich interactive visualization.** To deliver their precious payload — actionable intelligence — your advanced analytics tools should support interactive visualization of models, data, and results. Ideally, you should be able to visualize all of this in your preferred business intelligence tool, or in the predictive modeling vendor's integrated visualization layer. Of course, you have every right to

expect the full range of visualization techniques — histograms, box plots, heat maps, etc. — regardless of who provides the visualization layer.[9]

- **Deploy models flexibly.** To execute modeling functions — such as data preparation, regression, and scoring — on the widest range of data warehouses and other platforms, your tools should support in-database or embedded analytics. And to scale to the max, your predictive analytics tools should deploy models to massively parallel EDWs, software-as-a-service environments, and cloud computing fabrics. Your advanced analytics tools should also support development of application logic in open frameworks — such as MapReduce and Hadoop — to enable convergence of data mining and content analytics in the cloud.

PA/DM SOLUTIONS EVALUATION OVERVIEW

To assess the state of the PA/DM solutions market and see how the vendors stack up against each other, Forrester evaluated the strengths and weaknesses of top PA/DM solution vendors.

Evaluation Criteria Focus On Present And Future Solutions

After examining past research, user need assessments, and vendor and expert interviews, Forrester developed a comprehensive set of evaluation criteria. We evaluated vendors against these 53 criteria, which we grouped into three high-level buckets:

- **Current offering.** To assess the breadth and depth of each vendor's PA/DM product set, we evaluated each solution's overall architecture, as well as complete development and integration functionality.

- **Strategy.** We reviewed each vendor's strategy to assess how each vendor plans to evolve its PA/DM solution to meet emerging customer demands. We also evaluated each vendor's go-to-market approach, commitment, and direction strategies.

- **Market presence.** To establish each PA/DM product's market presence, we evaluated each solution provider's company financials, adoption, and partnerships.

Evaluated Vendors Must Meet Functional, Architectural, And Market Presence Criteria

Forrester included nine vendors in this assessment: Angoss Software, FICO, IBM (pre-SPSS acquisition), KXEN, Oracle, Portrait Software, SAS Institute, SPSS (an IBM product group whose PA/DM offering were evaluated separately from IBM), and TIBCO Software (see Figure 1). Each of these vendors meets the following inclusion criteria:

- **A comprehensive core PA/DM functionality.** We included vendors that offer one or more solutions that were generally available by August 25, 2009, and that provide at least the following core PA/DM functional components, tools, and features: 1) discover, extract, profile, transform,

cleanse, load, and otherwise prepare analytical data sets sourced from structured formats such as relational database tables; 2) develop and validate multivariate predictive models through clustering, classification, scoring, regression, and other statistical and mathematical functions against structured, historical data sets; 3) interactively visualize, explore, and manipulate predictive models to support forecasting, prediction, simulation, what-if analysis, and other functions to assess alternative courses of action; and 4) store and manage analytical data sets and predictive models in an EDW or other analytical data mart platform.

- **Standalone solutions incorporating original, cross-domain PA/DM functionality.** We included vendors that offered generic PA/DM offerings that are not technologically or functionally tied or limited to particular functional or horizontal applications — such as ERP or CRM — or to a particular business intelligence (BI); business performance solution (BPS); data warehousing (DW); extract, transform, load (ETL); or middleware stack. The vendor must be marketing, selling, and implementing the solutions as self-sufficient, general-purpose PA/DM workbenches, tool sets, or development platforms that can stand alone, meaning that it does not need to be embedded in other applications. The solutions must not be a repackaging, rebranding, and/or extension of PA/DM libraries sourced from an OEM partner. The solutions must not be functionally constrained to mining and modeling in only one analytical data subject-domain area, such as only customer or financial data, though this may be their primary application.

- **Significant market presence with substantial and referenceable customer base.** We included vendors that report at least US$5 million in PA/DM-specific revenues in the latest fiscal or calendar year, where at least 80% of revenues are from PA/DM solutions (software licenses, maintenance contracts, and/or subscription offerings), not consulting or other professional services. The vendors must each have at least 50 in-production PA/DM solution customers that span more than one major geographical region (Americas, Europe/Middle East/Africa, and Asia-Pacific); and represent five or more industry verticals. The vendor must have more than 10% of installations with more than 10 users (statistical analysts, data modelers, and/or data mining specialists) each.

- **Client inquiries that put the vendors on Forrester's radar.** The vendors included in this Forrester Wave™ are often discussed in client inquiries, appearing on client vendor selection shortlists.

Note that any PA/DM vendor solutions, versions, revisions, or feature enhancements that were released after the Forrester Wave cutoff date of August 25, 2009 were scored indirectly in "strategy/direction" criteria in the Forrester Wave model rather than directly in the "current offering" segment.

Figure 1 Evaluated Vendors: Product Information And Selection Criteria

Vendor	Product evaluated	Product version evaluated	Date evaluated
Angoss Software	Angoss KnowledgeSEEKER	7.0	Q3 2009
	Angoss KnowledgeSTUDIO	7.0	Q3 2009
	Angoss StrategyBUILDER	7.0	Q3 2009
FICO	FICO Model Builder	3.6	Q3 2009
	FICO Decision Optimizer	5.7	Q3 2009
	FICO Xpress Optimization Suite	7.0	Q3 2009
IBM	InfoSphere Warehouse	9.7	Q3 2009
	InfoSphere Balanced Warehouse		Q3 2009
	IBM Smart Analytics System		Q3 2009
KXEN	KXEN Analytic Framework	5.0	Q3 2009
	KMF Department	1.0	Q3 2009
Oracle	Oracle Data Mining with Oracle Database	11gR2	Q3 2009
Portrait Software	Portrait Customer Analytics	5.3	Q3 2009
	Portrait Uplift Optimizer	5.4	Q3 2009
	Portrait Self-service Analytics	2.1	Q3 2009

Source: Forrester Research, Inc.

Figure 1 Evaluated Vendors: Product Information And Selection Criteria (Cont.)

Vendor	Product evaluated	Product version evaluated	Date evaluated
SAS Institute	SAS Enterprise Miner	6.1	Q3 2009
	SAS Forecast Server	3.1	Q3 2009
	SAS/ETS	9.2	Q3 2009
	SAS/STAT	9.2	Q3 2009
	SAS Model Manager	2.2	Q3 2009
	SAS Simulation Studio	1.4	Q3 2009
	SAS Operations Research	9.2	Q3 2009
	SAS Scoring Accelerator	1.5	Q3 2009
	SAS Credit Scoring for Enterprise Miner	6.1	Q3 2009
	SAS Enterprise Guide	4.2	Q3 2009
	SAS Add-in for Microsoft Excel	4.2	Q3 2009
	JMP	8	Q3 2009
	SAS Text Miner	4.1	Q3 2009
	SAS Data Integration Studio	4.2	Q3 2009
	SAS/IML Studio	3.2	Q3 2009
	SAS Enterprise BI Server	9.2	Q3 2009
SPSS	IBM SPSS Modeler	13	Q3 2009
	IBM SPSS Text Analytics	13	Q3 2009
	IBM SPSS Text Analytics for Surveys	3	Q3 2009
	IBM SPSS Statistics	18	Q3 2009
	IBM SPSS Collaboration and Deployment Services	4	Q3 2009
	SPSS Event Builder		Q3 2009
	SPSS Interaction Builder		Q3 2009
	SPSS Risk Control Builder		Q3 2009

Source: Forrester Research, Inc.

Figure 1 Evaluated Vendors: Product Information And Selection Criteria (Cont.)

Vendor	Product evaluated	Product version evaluated	Date evaluated
TIBCO Software	TIBCO Spotfire Miner	8.1	Q3 2009
	TIBCO Spotfire S+	8.1	Q3 2009
	TIBCO Spotfire S+ Server	8.1	Q3 2009
	TIBCO Spotfire Professional — Windows client	3.0	Q3 2009
	TIBCO Spotfire Web Player — browser-based client	3.0	Q3 2009

Vendor selection criteria

Report at least US$5 million in PA/DM-specific revenues in the latest fiscal or calendar year, where at least 80% of revenues are from PA/DM solutions (software licenses, maintenance contracts, and/or subscription offerings), not consulting or other professional services.

Offer at least the following functional components, tools, or features in at least one PA/DM solution that is generally available by the time of Forrester's hands-on evaluation (August 25, 2009):

• Preparation, transformation, profiling, cleansing, and loading of analytical data sets extracted from structured sources such as relational database tables

• Development, segmentation, classification, clustering, simulation, scoring, and multivariate analysis of statistical models that incorporate structured data sets

• Interactive visualization, exploration, drilldown, and manipulation of predictive models that identify meaningful patterns among variables in historical data sets and support simulation, forecasting, and prediction of future events and trends to support assessment of alternative courses of action

• Storage and management of data sets and models in an enterprise data warehouse (EDW) or analytical data mart platform provided by the same vendor or a third party

Show that these commercially components, tools, or features:

• Constitute generic PA/DM offerings that are not technologically or functionally tied or limited to particular functional or horizontal applications, such as ERP or CRM, or to a particular business intelligence (BI), business performance solution (BPS), data warehousing (DW), extract transform load (ETL), or middleware stack

• Are being marketed, sold, and implemented as integral features of a self-sufficient, general-purpose PA/DM workbench, tool set, or development platform that can stand alone, meaning that it does not need to be embedded in other applications

• Are not functionally constrained to mining and modeling in only one analytical data subject area, such as only customer or financial data, though this may be their primary application

• Are not largely a repackaging, rebranding, and/or extension of PA/DM libraries sourced from an OEM partner

Substantiate at least 50 in-production customers that:

• Span more than one major geographical region (Americas, Europe/Middle East/Africa, and Asia-Pacific)

• Represent five or more industry verticals

• Have more than 10% of installations with more than 10 users (statistical analysts, data modelers, and/or data mining specialists) each

Sufficient interest from Forrester business process and applications (BP&A) clients, with at least 10% of PA/DM-related customer inquiries and/or advisory/consulting projects mentioning, addressing, or concerning the vendor's solutions

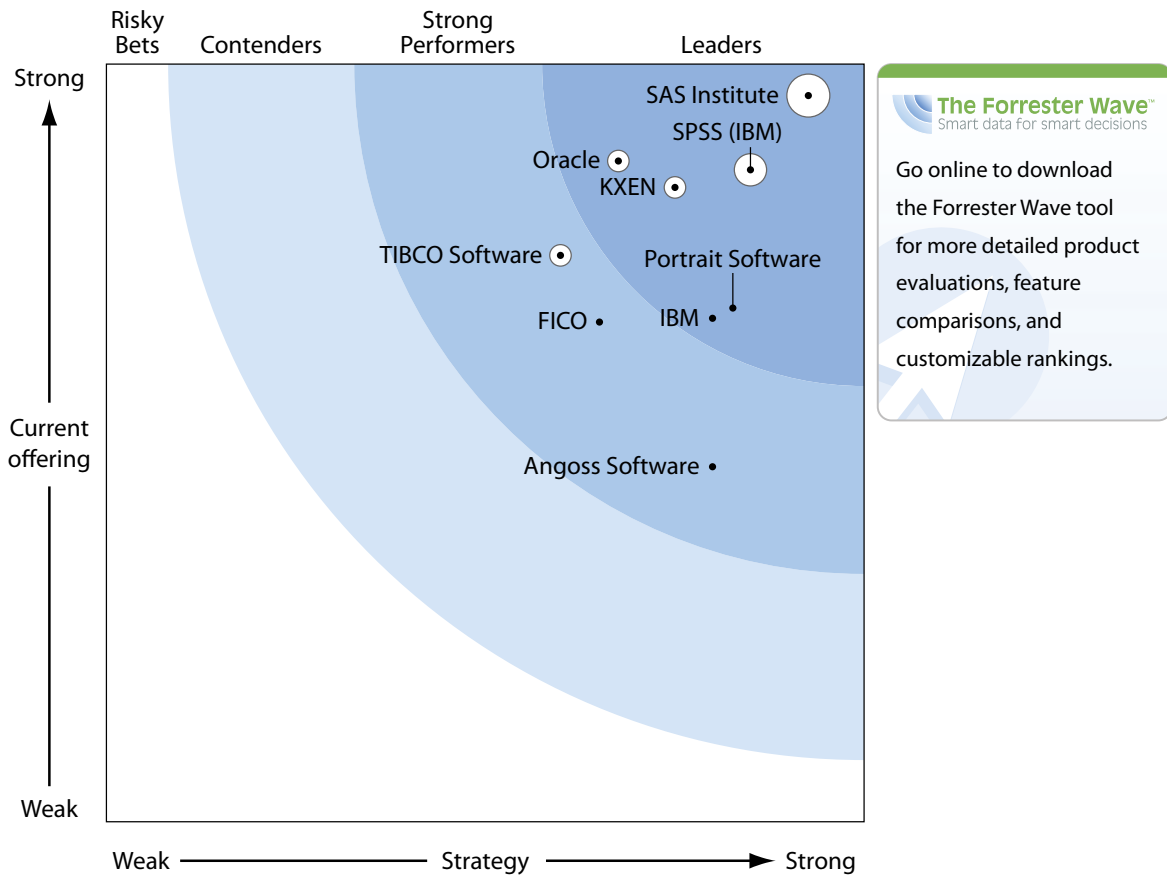Source: Forrester Research, Inc.

LEADING PA/DM SOLUTIONS ARE EMPHASIZING IN-DATABASE ANALYTICS

The evaluation uncovered a market in which (see Figure 2):

- **SAS, SPSS, KXEN, Oracle, Portrait Software, and IBM lead the pack.** All of the leading PA/DM vendors have staked their future solution evolution on in-database analytics. SAS has taken the industry lead in promoting in-database analytics as an emerging best practice for high-performance analytics deployments. SPSS, now an IBM product group, is rapidly integrating its already strong PA/DM solution portfolio with its new parent's extensive family of EDW, ETL, BI, and other data management offerings, while maintaining and deepening SPSS's longstanding support for heterogeneous in-database analytics. IBM, which will converge its current PA/DM offerings with those of SPSS, continues to push the PA/DM industry envelope by providing solutions that deploy predictive logic into complex event processing (CEP) streams. KXEN is at the forefront of the convergence of PA/DM, content analytics, and social network analysis for Web 2.0 environments. Portrait Software provides one of the most comprehensive PA/DM solution portfolios for life-cycle customer analytics, supporting real-time scoring, interaction optimization, uplift optimization, and campaign management. Oracle has recently rolled out a new generation of its Exadata technology, supporting execution of PA/DM, text analytics, and transactional application logic in an optimized storage layer.

- **TIBCO, FICO, and Angoss offer innovative enterprise solutions.** All of the other PA/DM solutions vendors in the Forrester Wave also offer solutions with important differentiators. TIBCO provides feature-rich statistical analysis and interactive visualization tools that can be used both by "black belt" data miners and business analysts. FICO, which has sunk deep roots in financial services, provides a flexible PA/DM tool set that can be used for development of predictive models in many industries. When deployed in conjunction with the vendor's Blaze Advisor rules engine, FICO's predictive modeling tool can support sophisticated real-time scoring applications. Angoss supports integrated modeling of multiple business scenarios through such features as strategy maps, champion-challenger modeling, and automatic best-model selection, thereby standing in the vanguard of PA/DM vendors that help customers become more thoroughly predictive enterprises.

This evaluation of the PA/DM solutions market is intended to be a starting point only. We encourage readers to view our detailed product evaluations and adapt the criteria weightings to fit their individual needs through the Forrester Wave Excel-based vendor comparison tool.

Figure 2 The Forrester Wave: Predictive Analytics And Data Mining Solutions, Q1 '10

Figure 2 Forrester Wave™: Predictive Analytics And Data Mining Solutions, Q1 '10 (Cont.)

	Forrester's Weighting	Angoss Software	FICO	IBM	KXEN	Oracle	Portrait Software	SAS Institute	SPSS (IBM)	TIBCO Software
CURRENT OFFERING	50%	2.35	3.30	3.33	4.19	4.35	3.40	4.80	4.30	3.74
Architecture	25%	3.00	3.00	3.50	4.00	4.00	4.00	4.50	4.00	2.00
Functionality	50%	2.34	3.08	3.21	4.44	5.00	3.10	5.00	4.50	4.45
Development	20%	1.40	3.80	3.00	3.60	3.00	3.00	4.60	4.00	3.80
Professional services	5%	3.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
STRATEGY	50%	4.00	3.25	4.00	3.75	3.38	4.13	4.63	4.25	3.00
Licensing and cost	25%	5.00	3.00	3.00	5.00	3.00	5.00	5.00	3.00	3.00
Commitment	25%	3.00	4.00	3.00	4.00	4.50	3.50	3.50	4.00	3.00
Direction	50%	4.00	3.00	5.00	3.00	3.00	4.00	5.00	5.00	3.00
MARKET PRESENCE	0%	1.68	1.65	1.15	2.88	2.23	1.68	4.38	3.48	2.48
Company financials	25%	1.00	1.00	1.00	3.00	1.00	2.00	5.00	3.00	1.00
Adoption	50%	1.75	2.00	0.20	1.95	1.45	1.45	3.95	3.45	2.35
Partnerships	25%	2.20	1.60	3.20	4.60	5.00	1.80	4.60	4.00	4.20

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

Source: Forrester Research, Inc.

VENDOR PROFILES

Leaders: In-Database Analytics With Strong Predictive Modeling And Content Analytics

The PA/DM Leaders have supported large enterprise customers' advanced analytics needs for many years. All offer mature, high-performance, scalable, flexible, and robust PA/DM solutions that combine a wide range of statistical algorithms with integrated support for in-database analytics and a broad range of information types. Clearly, BP&A professionals have strong PA/DM solution options to meet their most demanding requirements.

- **SAS Institute, a longtime PA/DM market leader, goes deep on in-database analytics.** With an increased focus on pushing down more data mining functions for scalable execution on partner EDW platforms, SAS positions its mature, widely adopted offerings as the premier analytical modeling offerings for the enterprise market. SAS provides a diversified portfolio that includes a comprehensive feature-rich set of statistical, data mining, forecasting, and optimization tools complemented with a model and data governance platform for analyzing complex structured and unstructured information. The SAS offerings integrate out of the box with wide range of third-party DW platforms and with the vendor's own analytics platform and wide range of prepackaged analytical applications. SAS continues to extend its deep domain expertise and technology solution set to drive customer innovation and performance.

SAS's key current offering differentiators are an integration of its PA/DM tools with a wide range of third-party data preparation tools; semi-structured and unstructured information integration; strategy maps; ensemble modeling; champion-challenger modeling; text analytics; sentiment analysis; social network analysis; constraint-based optimization; and support for wide range of statistical algorithms, model quality metrics, and variable selection techniques. In addition, SAS supports a wide range of model deployment options; supports Eclipse, multilanguage, SOA-based, and R-based development; and supports interactive visualization in BI tools, rich model governance, and scalability through in-database analytics integration in third-party massively parallel processing database management systems (MPP DBMSes). However, SAS lacks integration of support vector machines and integration of any business rules engine into its middleware portfolio to support real-time business process integration.

- **SPSS is integrating with the information management portfolio of its new parent, IBM.** With its longtime commitment to in-database analytics in heterogeneous DBMSes, SPSS positions its broad portfolio of PA/DM, text analysis, and statistical analysis tools as a comprehensive enterprise tool set. SPSS (evaluated in this Forrester Wave independently of IBM's legacy PA/DM solution portfolio) provides a PA/DM solution portfolio with comprehensive feature-rich modeling/mining and statistical analysis tools, as well as model governance platform, for mining complex structured and unstructured information. SPSS's products integrate with a wide range of third-party DW/DBMS platforms and with the vendor's own wide range of prepackaged predictive applications.

  With its recent acquisition by IBM, SPSS will focus on continuing its active product development and partnering initiatives while also leveraging IBM's considerable research and development resources. SPSS will also focus on integration with IBM's Cognos, InfoSphere, and other information management product groups. SPSS's key (pre-IBM acquisition) current offering differentiators are strategy maps; ensemble modeling; champion-challenger modeling; model comparison; text analytics; sentiment analysis; constraint-based optimization; and support for a wide range of statistical algorithms and variable selection techniques. In addition, SPSS supports a wide range of model deployment and software deployment options, multilanguage and SOA-based development, interactive visualization in BI tools, rich model governance, and scalability through in-database analytics integration in third-party MPP DBMSes. However, SPSS (pre-IBM acquisition) lacks social network analysis, genetic algorithms, and R-based development.

- **KXEN offers strong usability within a feature-rich PA/DM portfolio.** With sophisticated PA/DM features and thorough integration of wizard-driven automation, KXEN positions its PA/DM tools as key enablers for maximum productivity by data-mining specialists and business analysts alike. KXEN provides a PA/DM solution portfolio with automated modeling/mining and sophisticated life-cycle model governance for high-productivity model building, validation, and deployment by professional data miners doing real-time scoring in in-production inline models.

KXEN focuses on expanding the use of PA/DM within analyst communities, making them more productive and more responsive to business needs, as well as among business consultants with no particular knowledge of statistics. KXEN's key current offering differentiators are text analytics; sentiment analysis; social network analysis; ensemble modeling; model comparison; constraint-based optimization; time-series analysis; support for respectable range of statistical algorithms; and support for a wide range of model quality metrics and variable selection techniques. In addition, KXEN provides support for a wide range of model deployment and import/export options, Eclipse-based development tools, support for SOA-based development, interactive visualization in BI tools, feature-rich model governance, and scalability through in-database analytics in third-party MPP DBMSes. However, KXEN lacks strategy maps, champion-challenger modeling, genetic algorithms, neural networks, support vector machines, support for R-based development, and an integrated same-vendor application/middleware stack.

- **Oracle offers deep database and application integration of its PA/DM tools. With** extensive in-database pushdown of its full range of sophisticated advanced analytics functions to its widely adopted enterprise RDBMS, Oracle positions its PA/DM portfolio as a key enabler for both traditional data mining as well as the new world of real-time decision automation, content analytics, and social network analysis. Oracle provides a PA/DM solution portfolio that is built into its own widely adopted DBMS, DW, data integration, and BI platforms, with a wide range of prepackaged predictive applications, and it provides a powerful assortment of algorithms for mining complex structured and unstructured information types.

  Oracle focuses on in-database mining in the Oracle Database, on integration of Oracle Data Mining into the kernel of that database, and on leveraging that technology in Oracle's branded applications. Oracle's key current offering differentiators are DBMS-integrated data-preparation tools; semistructured and unstructured information integration; text analytics; strategy maps; ensemble modeling; champion-challenger modeling; sentiment analysis; social network analysis; and support for a competitive range of statistical algorithms and variable selection techniques. Other key features include support for multilanguage and SOA development, interactive visualization in BI tools, and integrated same-vendor application/middleware stack. However, Oracle lacks genetic algorithms; time-series forecasting; support for R-based development. Also, though supporting an otherwise wide range of software deployment options, Oracle's PA/DM tool is built into the Oracle Database and lacks the ability to deploy over any other database.

- **Portrait offers a feature-rich PA/DM portfolio for comprehensive customer analytics.** With strong core PA/DM features, a solid group of PA/DM-enabled customer analytics applications, and a highly differentiated "uplift modeling" capability, Portrait positions its PA/DM tools and applications as mature, feature-rich offerings for enterprises across all verticals. Portrait provides a user-friendly, feature-rich PA/DM solution portfolio in support of real-time scoring, interaction optimization, uplift optimization, and campaign management for customer analytics.

Portrait focuses on the customer interaction market, providing intelligence and insight to support such applications. It has invested to penetrate customer analytics opportunities in the financial services, insurance, telecommunications, travel & leisure, catalog retail, and expanding partner channels. Portrait's key current offering differentiators are semistructured and unstructured information integration; model combination; comparative evaluation; constraint-based optimization; neural networks; genetic algorithms; support for wide range of statistical algorithms and variable selection techniques; model quality scores; lift charts; a wide range of model-deployment options; and interactive visualization in third-party BI tools. However, Portrait lacks DBMS-integrated data preparation tools, integration with third-party data preparation tools, strategy maps, ensemble modeling, champion-challenger modeling; text analytics; sentiment analysis; social network analysis; support vector machines; time-series forecasting; association rules and market basket analysis; support for R-based development; and feature-rich model governance.

- **IBM prepares to transition to an SPSS-centric PA/DM portfolio.** With its widely adopted analytics, DW, DBMS, data integration, and other information management tools, IBM positions its current PA/DM tools as an optional add-on software component that is tightly integrated with its InfoSphere platform. IBM (evaluated in this Forrester Wave independently of its newly acquired SPSS PA/DM solution portfolio) offers a PA/DM tool that is integrated and bundled with its leading BI, DBMS, DW, DI, and DQ platform. IBM greatly increased its presence in the PA/DM market through its acquisition of SPSS. PA/DM is central to IBM's Smarter Planet and Information-Led Transformation Initiatives, which span all IBM business units and affect how IBM's software, hardware, and services division delivers offerings to market. IBM will continue to integrate and extend its PA/DM offerings among its rich portfolio of recent acquisitions such as Cognos, Exeros, ILOG, SPSS, and others.

  IBM's PA/DM solution (pre-SPSS acquisition) provides strong functionality in several PA/DM current offering criteria. IBM's key current offering differentiators are DBMS-integrated data preparation tools; text analytics; and support for a respectable range of statistical algorithms, model quality metrics, and variable selection techniques. Other key differentiators include support for a wide range of model deployment, model import/export, and Eclipse-based development tools; support for multilanguage and SOA-based development; interactive visualization in BI tools; and integrated same-vendor application/middleware stack. However, IBM (pre-SPSS acquisition) lacks strategy maps, ensemble modeling, champion-challenger modeling, sentiment analysis, social network analysis, constraint-based optimization, genetic algorithms, support vector machines, and R-based development. IBM also has a software deployment limitation: only deploying on a single DBMS (DB2).

Strong Performers: Innovative Predictive Analytics Tools From Established Vendors

Today's PA/DM Strong Performers are aggressively challenging the Leaders with mature, fast-evolving solution portfolios that have gained significant adoption among large enterprises in many verticals. Just as important, these Strong Performers have established themselves as innovators with functionality in such key PA/DM areas as wizard-driven development automation, multibusiness scenario modeling, interactive visualization, content analytics, sentiment analysis, social network analysis, in-stream analytics, and open-source modeling languages. These vendors' substantial, loyal, and longtime customer bases suggest plenty of opportunity for well-differentiated PA/DM solutions for various enterprise and midmarket opportunities. BP&A professionals can rest assured that these and other substantial PA/DM solution vendors have the staying power, resources, and vision to weather the ups and downs in today's turbulent IT market.

- **FICO integrates its business rules engine more deeply into its PA/DM solution.** With its focus on decision automation for operational applications, FICO positions its PA/DM tools as an integrated solution portfolio that is time-proven in the financial services industry but has the flexibility of horizontal applications in many industries. FICO provides a PA/DM solution portfolio with sophisticated ensemble modeling and is integrated with a business rules engine (BRE) for decision automation in financial analytics.

  FICO is growing its presence in and beyond financial services by continuing to introduce business-oriented analytic solutions through offerings that shorten the time required to discover, implement, test, and deploy predictive models by extending the precision of analytics across all decision areas and by bringing more advanced decision modeling and mathematical optimization within reach of decision management professionals.

  FICO's key current offering differentiators are support for champion-challenger modeling; semistructured information types; ensemble models; support for a wide range of statistical algorithms and approaches (including constraint-based optimization); an application/middleware stack that includes a tightly integrated business rules engine; feature-rich model governance; and scalability . However, FICO lacks DBMS-integrated data-preparation tools; strategy maps; unstructured information integration; text analytics; sentiment analysis; social network analysis; Eclipse-based development tools; support for multilanguage, SOA, or R-based development; and interactive visualization in BI tools.

- **Angoss offers a PA/DM portfolio for enterprises and for partner solutions.** With strong core PA/DM features, a solid group of PA/DM-enabled business applications, and an expanding range of platform and application partnerships, Angoss positions its PA/DM tools as user-friendly, high-productivity enablers for building and managing models to support real-time scoring in customer analytics. Angoss provides a PA/DM solution portfolio with strong strategy maps, automation, and real-time scoring for customer analytics.

Angoss focuses on widening the accessibility of PA/DM in part by producing software with an emphasis on flexibility and ease of use, and in part by focusing on the operational deployment of software and solutions. Angoss' key current offering differentiators are strategy maps; champion-challenger modeling; model combination; automatic best-model selection; support for semistructured information types; constraint-based optimization; and support for a wide range of statistical algorithms, model quality metrics, and variable selection techniques. However, Angoss lacks several important features, without which it cannot be considered a comprehensive PA/DM solution provider: It lacks DBMS-integrated data preparation tools and the ability to integrate with third-party data preparation tools. Other deficiencies include lack of ensemble modeling, unstructured information integration, text analytics, sentiment analysis, and social network analysis.

- **TIBCO Software offers interactive visualization and comprehensive inline analytics.** With tight integration of its interactive BI visualization into its S+ technology (acquired from Insightful in 2008), TIBCO positions its now Spotfire-branded PA/DM tools as suited both to traditional data mining specialists and to the legions of non-technical business analysts.[10] TIBCO provides a PA/DM solution portfolio that supports in-memory and out-of-memory, open source language, and visualization for analytic applications, in-CEP, in-BPM, in-SOA, and other inline deployment scenarios.

  TIBCO Spotfire is focusing its PA/DM solutions on helping customers discover new and actionable insights in information at rest or in motion, while proactively detecting and responding to events. It is integrating its PA/DM offerings with the parent's broad middleware portfolio. TIBCO's key current offering differentiators are strategy maps; ensemble modeling; champion-challenger modeling; comparative model evaluation; model combination; content analytics; social network analysis; constraint-based optimization; and support for a wide range of other statistical algorithms. Other key features include support for Eclipse, multilanguage, SOA-based, and R-based development; interactive visualization in BI tools; a range of model import/export options; and an integrated same-vendor application/middleware stack. However, TIBCO lacks DBMS-integrated data preparation tools, sentiment analysis, and feature-rich model governance.

SUPPLEMENTAL MATERIAL

Online Resource

The online version of Figure 2 is an Excel-based vendor comparison tool that provides detailed product evaluations and customizable rankings.

Data Sources Used In This Forrester Wave

Forrester used a combination of the following data sources to assess the strengths and weaknesses of each solution:

- **Vendor surveys.** Forrester surveyed vendors on their capabilities as they relate to the evaluation criteria. Once we analyzed the completed vendor surveys, we conducted vendor calls where necessary to gather details of vendor qualifications.

- **Product demos.** We asked vendors to conduct demonstrations of their product's functionality. We used findings from these product demos to validate details of each vendor's product capabilities.

- **Customer reference calls.** To validate product and vendor qualifications, Forrester also conducted reference calls with two of each vendor's current customers.

The Forrester Wave Methodology

We conduct primary research to develop a list of vendors that meet our criteria to be evaluated in this market. From that initial pool of vendors, we then narrow our final list. We choose these vendors based on: 1) product fit; 2) customer success; and 3) Forrester client demand. We eliminate vendors that have limited customer references and products that don't fit the scope of our evaluation.

After examining past research, user need assessments, and vendor and expert interviews, we develop the initial evaluation criteria. To evaluate the vendors and their products against our set of criteria, we gather details of product qualifications through a combination of lab evaluations, questionnaires, demos, and/or discussions with client references. We send evaluations to the vendors for their review, and we adjust the evaluations to provide the most accurate view of vendor offerings and strategies.

We set default weightings to reflect our analysis of the needs of large user companies — and/or other scenarios as outlined in the Forrester Wave document — and then score the vendors based on a clearly defined scale. These default weightings are intended only as a starting point, and we encourage readers to adapt the weightings to fit their individual needs through the Excel-based tool. The final scores generate the graphical depiction of the market based on current offering, strategy, and market presence. Forrester intends to update vendor evaluations regularly as product capabilities and vendor strategies evolve.

ENDNOTES

[1] Predictive analytics goes beyond traditional BI by empowering organizations to create and test detailed, data-driven models of future scenarios. By contrast, traditional BI applications provide deep insights into past and present events, by helping business people to identify historical patterns in sales, marketing, supply chain, financial, and other key performance indicators. But traditional BI tools do not offer much assistance with proactive concerns such as, "What will happen?," "What may happen next?," and "What could possibly happen if I take action X versus action Y?" See the August 18, 2009, "Business Intelligence (BI) Polishes Its Crystal Ball" report.

[2] Predictive analytics is essential to a mature enterprise implementation of BI technologies, especially if predictions are integrated with historical data as presented in reports, dashboards, and queries. Other key

elements of mature analytics environments include support for strategic and operational decision-making, online analytical processing (for physical aggregation and for business functionality such as time series analysis, allocations, etc), and free-form analysis (not limited by data models). See the September 11, 2009, "Forrester's BI Maturity Assessment Tool" report.

3   Strategy maps are visual modeling representations of the strategy of an organization. Ensemble modeling refers to machine learning methods that leverage the power of multiple models to achieve better prediction accuracy than any of the individual models could on their own. Champion-challenger modeling refers to a multiscenario approach in which one develops a principal, or "champion," model and one or more alternatives, or "challengers," that each differ from the champion in some measurable and defined way.

4   Constraint-based optimization supports identification of actions that will produce the best results while operating within resource limitations and other relevant restrictions. Neural networks support dynamic learning relationships among multiple variables when the exact nature of those relationships is not known in advance. Genetic algorithm uses an approach derived from Darwin's biological model of natural selection to solve optimization problems using parallelized search techniques. Support vector machines help to automatically detect complex nonlinear patterns by clustering, classifying, and ranking data.

5   Common model quality or fitness scores in the industry include KS (Kolmogorov-Smirnov), ROC (Receiver Operating Characteristic), GINI (which refers to an accuracy ratio summary index for cumulative accuracy profiles), and Area Under the Curve (AUC).

6   Decision trees are statistical methods that use tree-like representations to classify predictive variables for regression analysis. Feature selection refers to approaches for selecting the most appropriate predictive variables from a large list of candidates without assuming that linear relationships exist between the predictors and the dependent or outcome variables. Association rule learning is an approach for discovering interesting relations among variables in very large, complex data sets. Outlier analysis supports identification and assessment of records in a data set that are significantly different from expectations.

7   EDW platforms should be the centerpiece of a well-developed enterprise PA/DM operation. Some commercial EDW platforms bundle and integrate tightly with ETL, DQ, and other key data preparation tools provided by the same vendor. Most integrate to varying degrees with leading third-party offerings in those and other data integration segments. See the February 6, 2009, "The Forrester Wave™: Enterprise Data Warehousing Platforms, Q1 2009" report and see the May 2, 2007, "The Forrester Wave™: Enterprise ETL, Q2 2007" report.

8   Realizing maximum business value from PA/DM investments requires that enterprises implement governance and other enterprisewide analytics best practices that are defined and enforced through a "business intelligence solution center" (BISC). A BISC is a permanent, cross-functional organizational structure responsible for governance and processes necessary to deliver or facilitate delivery of successful BI solutions, as well as being an institutional steward, protector, and forum for BI best practices. Enterprises with more successful BI implementations often implement some form of BISC practices. There are a wide range of BISC implementation options however, and not all of them are appropriate for every scenario, such as PA/DM and other advanced analytics initiatives. See the November 4, 2008, "Implementing Your Business Intelligence Solutions Center" report.

[9] PA/DM demands interactive visualization to support analyst discovery, exploration, assessment, modeling, and manipulation of complex historical data sets. Traditional OLAP visualization approaches are insufficient for the new generation of advanced analytics, which demand the ability to visualize large, complex information sets and models with hundreds or thousands of dimensions. This new post-OLAP visualization paradigm — essentially both to traditional "black belt" PA/DM tools and to enriched BI environments — requires automated information discovery, complex unstructured information sets, flexible dimensionality, dynamically contextualized presentation, multiple linked visualizations, and guided analytics. See the November 7, 2008, "OLAP: In Fashion Or Old-Fashioned?" report. Additional source: Boris Evelson, "How To Differentiate Advanced Data Visualisation Solutions", *The Forrester Blog For Business Process & Applications Professionals* (http://blogs.forrester.com/business_process/2009/11/how-to-differentiate-advanced-data-visualisation-solutions.html).

[10] More information on TIBCO's acquisition of Insightful is available on TIBCO's Web site. Source: "TIBCO Announces Agreement to Acquire Insightful Corporation", TIBCO press release, June 19, 2008. (http://www.tibco.com/company/news/releases/2008/press913.jsp).

FORRESTER®

Making Leaders Successful Every Day

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.

Forrester Research, Inc. (Nasdaq: FORR) is an independent research company that provides pragmatic and forward-thinking advice to global leaders in business and technology. Forrester works with professionals in 20 key roles at major companies providing proprietary research, customer insight, consulting, events, and peer-to-peer executive programs. For more than 26 years, Forrester has been making IT, marketing, and technology industry leaders successful every day. For more information, visit www.forrester.com.

FORRESTER®