

> Gain Valuable Insights with Text Analytics

SPSS Inc.'s PASW Modeler* data mining workbench enables organizations to achieve measurable results by basing decisions on patterns and associations found in their data. But did you know that up to 80 percent of your organization's data is contained in textual form?

The customer e-mails, call center notes, open-ended survey responses, Web forms, and other text sources that your organization captures—including content from RSS feeds, such as blogs and news feeds—contain up to four times as much valuable data as your organization's structured data stores. This means that data mining projects focusing only on structured data may use less than 20 percent of the information available.

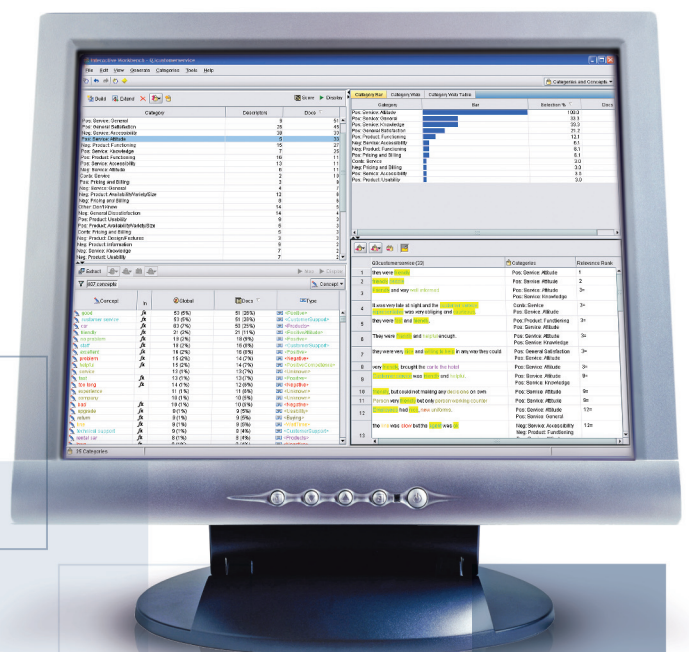
PASW Text Analytics* enables you to combine this valuable unstructured data with traditional structured data to significantly increase your understanding of customers, the public, and other groups. This product transforms PASW Modeler into a fully integrated data and text analytics workbench. You can perform both text analytics and data mining within the interactive, visualization-based PASW Modeler environment.

* PASW Modeler, PASW Text Analytics, and PASW Text Analytics for Surveys, formerly called Clementine®, Text Mining for Clementine, and SPSS Text Analysis for Surveys™, are part of SPSS Inc.'s Predictive Analytics Software portfolio.

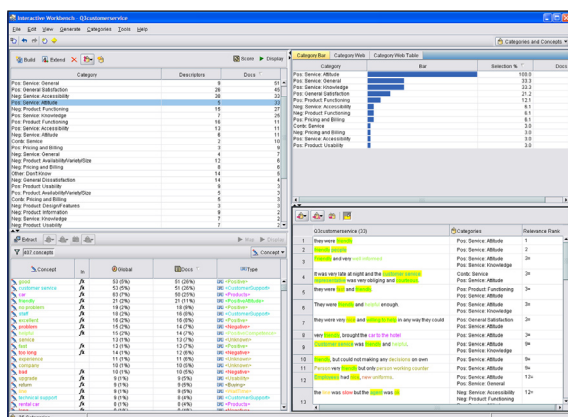
Powered by SPSS Inc.'s natural language processing (NLP) technology, PASW Text Analytics pulls key concepts from many types of unstructured data and groups them into categories. Extracted concepts, opinions, and categories are then combined with structured data and applied to predictive models to provide valuable insights into actions, behaviors, patterns, and associations.

Analyze all types of text from multiple languages

This version of PASW Text Analytics offers you the ability to create and share text analysis packages (TAPs), which are categories and linguistic resources that control the extraction process. Three pre-built TAPs are included for customer, employee, and product satisfaction analysis—or you can also build your own TAPs. All TAPs created in this version of PASW Text Analytics are available in English only and can also be used in the latest version of PASW Text Analytics for Surveys*, and vice versa.



Whether you're analyzing e-mails, RSS feeds, blogs, call center logs, or satisfaction surveys, you'll save time and effort while gaining a competitive advantage. You may also use customized libraries from an earlier text analytics project. Or, if you don't have any pre-defined categories or specific dictionaries to use, you can create your own and then use the software to refine them.



The new built-in customer satisfaction text analysis package, (TAP), has been used to automatically analyze and categorize customer feedback. Results can then be directly integrated to predictive models for better prediction of customer behavior.

PASW Text Analytics also supports Language Weaver™, a provider of software that automates human language translation. As a result, PASW Text Analytics' use of text link analysis (TLA) technology can help you quickly identify and extract sentiments, such as preferences and opinions, from text in more than 30 languages**. By capturing the attitudes and thoughts of individuals across more channels and markets, you can gain a more complete view of your customers and constituents and build more accurate predictive models faster.

Add value throughout your organization

Text analytics can be used in nearly any business or research situation that involves unstructured data. Here are just some examples of PASW Text Analytics applications:

- **Product development and refinement**—Identify trends in complaints or requests by analyzing call center logs, customer e-mails, open-ended survey responses, and RSS feeds, including blogs and news feeds. Use this information to improve existing products and services and develop successful new offerings.

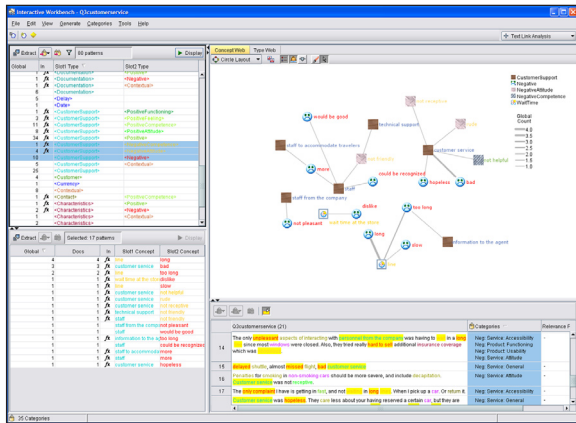
- **Marketing campaigns**—Improve campaign effectiveness and revenue. For example, analyze inbound customer calls in real time to provide better product and service recommendations. Complement Net Promoter® Score analysis by really understanding why your customers are promoters or detractors.
- **Cross-selling**—Improve sales by using information about customer preferences to better target products to existing customers
- **Churn prevention**—Discover why some customers leave—and take steps to prevent defection—by analyzing customer communication records for recurring problems or complaints that precede churn
- **Drug discovery**—Find relationships in chemical and biomedical databases
- **Market intelligence**—Survey competitor and industry Web sites, RSS feeds, including blogs and news feeds, and other publicly available textual information to maintain a current view of your competitors
- **Security**—Discover potentially suspect behavior by analyzing Web site content, chat rooms, e-mails, blogs, and other sources of information, and identify patterns and associations in the data
- **Fraud**—Identify potentially fraudulent patterns in banking transactions or insurance claims and use investigative resources more effectively

Boost productivity, improve predictive quality and accuracy

Text analytics is an iterative process. Regardless of what type of content you analyze, you'll likely re-extract concepts and re-categorize responses using different category definitions or coding schemes, different term or synonym definitions, or different groupings of responses. You may repeat this process several times before you're satisfied with the results.

PASW Text Analytics simplifies and speeds this process. The new Category Rule Editor enables you to create rules for categorizing complex concepts and filtering out errors. The software automatically and quickly classifies concepts, extracts results, and groups them. You can even train

**** Translate your non-English data into English by purchasing a Language Weaver license.**



Users can navigate through sentiment analysis maps and directly get back to a customer's own words.

the extractor to easily handle blogs and their related URLs. The end result? PASW Text Analytics provides more accurate categorization and extraction of a wider variety of text—and the potential for more sophisticated analysis—than manual methods.

Unlock the power of text—no linguists required

Unlike other text analytics products, you do not need a linguistic background to use PASW Text Analytics. And because PASW Text Analytics uses an interactive interface within PASW Modeler, text analytics is straightforward and efficient. For example, interactive graphs enable you to explore and display text data and patterns for instant analysis.

You can easily customize concept dictionaries for a particular domain area by using the Resource Editor, an integrated resource for managing the text extraction process. This enables you to find relevant concepts and associations faster.

Create customized templates and libraries for specific business applications directly from the main PASW Modeler toolbar. You can reuse these valuable resources with other products and applications, including PASW Text Analytics for Surveys. The linguistic resources in PASW Text Analytics support a range of industries and applications, including sentiment analysis, CRM, security and intelligence, market intelligence, life sciences (genomics and MESH), and IT.

For example, a company that wants to analyze call center notes can use the Resource Editor to adjust the software's dictionaries to reflect acronyms, abbreviations, and slang typically found in call transcripts. Or a pharmaceutical company can use the Resource Editor to set the included genomics dictionary in PASW Text Analytics as its default.

Deploy into operational systems

In order to make the best use of your textual data, you need to be able to use it throughout your organization. Text insights deployed through the software's predictive models to operational databases provide value to areas throughout your organization. And you can deploy any part of the text or data mining process by using PASW Modeler Solution Publisher***.

To help manage your analytical assets and automate analytical processes, use PASW Text Analytics with PASW Collaboration and Deployment Services***.

You can also use insights derived from text data to achieve more accurate results with SPSS Inc.'s predictive applications. For example, improve real-time and batch scoring, provide real-time recommendations to inbound callers, or speed insurance claim processing.

If your organization collects customer insight through an enterprise feedback management (EFM) solution, PASW Text Analytics can help you understand the opinions, attitudes, and preferences of your customers, employees, citizens, business partners, and others.

PASW Text Analytics' deployment capabilities make textual insights available to business users in critical areas, so your entire organization benefits from a comprehensive, 360-degree view of customers or, for government agencies, the citizens they serve.

For examples of text analytics, visit www.spss.com/success and choose "Text Mining" under "By Solution."

*** PASW Modeler Solution Publisher and PASW Collaboration and Deployment Services, formerly called Clementine Solution Publisher Runtime and SPSS Predictive Enterprise Services™, are part of SPSS Inc.'s Predictive Analytics Software portfolio.

PASW Text Analytics features*

Methodology

PASW Modeler and PASW Text Analytics support the CRoss-Industry Standard Process for Data Mining (CRISP-DM), which enables analysts to focus on solving business problems, rather than on programming. PASW Text Analytics enables you to merge unstructured data with structured data during the CRISP-DM process. In addition, PASW Text Analytics uses a proven combination of natural language processing (NLP) techniques and predictive analytics to efficiently extract meaningful information from unstructured data.

Linguistic extraction

- Extract text data from files or a database
- Work with multilingual text. Select from seven native language extractor options: Dutch, English, French, German, Italian, Portuguese, and Spanish
- Translate more than 30 languages into English using Language Weaver add-on options
- Manage errors in punctuation and spelling
- Extract domain-specific concepts such as uniterms, expressions, abbreviations, acronyms, and more
- Calculate synonyms using sophisticated linguistic algorithms and embedded or user-specified linguistic resources
- Name concepts by person, organization, term, product, location, and other user-defined types
- Extract non-linguistic entities such as address, currency, time, phone number, and social security number (SSN)
 - Templates for non-linguistic entities are available for all seven languages
- Included opinion, market intelligence, security intelligence, and genomics dictionaries enable relationship extraction; the opinion dictionary also enables sentiment extraction (such as likes and dislikes)
- Cache extraction results during an interactive workbench session for reuse later

Web feed node

- Easily retrieve and analyze text from RSS feeds, such as blogs and news feeds
- Enhanced retrieval of blog times, date handling, and a new filtering algorithm to strip out advertising and other unwanted text

Text analytics modeling node

- Use pre-built Text Analytics Packages (TAPs), categories and linguistic resources, for customer/product/employee satisfaction analysis**
- Build your own TAPs
- Create rules to categorize complex information and filter out errors
- Create clusters based on term co-occurrence using concept clustering algorithms, which provide an at-a-glance view of main topics and the way in which they are related
- Intelligently group text documents and records based on content, using text classification algorithms
 - Aggregates concepts from a wide variety of unstructured text data and groups them into a small number of categories
 - Includes new grouping techniques to create categories from scratch
 - Reuses categories, enabling the scoring of any new text documents and records based on the text they contain
 - Accelerates and improves data management
- Drag and drop text records to better teach categories
- Identify best matching documents at a glance with the new relevance ranking
- Enable advanced concept selection and deselection for use in subsequent PASW Modeler predictive modeling applications
- List extracted concepts by type, frequency, document coverage, and other user-defined classifications
- Discover how concepts are interrelated, with the new concept map
- Highlight synonyms used for each selected concept
- Convert selected concepts to structured form for use in PASW Modeler predictive modeling algorithms
- Access all text analytics capabilities through the text analytics node palette in PASW Modeler

Text link analysis

- Identify and extract sentiments (for example, likes and dislikes) from text in Dutch, English, French, German, and Spanish
- Identify links and associations between, for example, people and events, or diseases and genes
- Identify and extract content from URLs within blogs

- Include opinions, semantic relationships, and linked events in deployable predictive models
- Reveal complex relationships through interactive graphs that show multiple semantic links between two concepts

Resource Editor

- Support specific industries and applications using verticalized templates and libraries for sentiment analysis, CRM, security and intelligence, market intelligence, life sciences, and IT
- Create and edit custom libraries directly in the PASW Text Analytics interface
- Define and edit:
 - Domain-specific terms
 - Non-linguistic entities
 - Synonyms
 - Concept libraries
- Share library resources with other products and applications, including PASW Text Analytics for Surveys
- Edit the CRM, opinion, market intelligence, security intelligence, and genomics dictionaries

Deployment

- Deploy the entire data mining process automatically with Real Time Scoring Service
- Send text analytics extractions directly to PASW Modeler for export into PASW Decision Management Services, and then into specific business-related applications
- To help manage your analytical assets and analytical processes, use PASW Text Analytics with PASW Collaboration and Deployment Services

Open integration

- Easily integrates external linguistic resources
- Accesses any text file or relational database via PASW Modeler's high-performance ODBC drivers
- Supports the industry-standard Predictive Modeling Markup Language (PMML)
- Supports standard text document formats, including plain text, PDF, HTML, Microsoft® Office, and XML
- Works with multilingual text natively (Dutch, English, French, German, Italian, Portuguese, and Spanish), integrates with Japanese in other products, and uses translations via Language Weaver add-on options

* Features subject to change based on final product release.

** English-language only

□ Symbol indicates a new feature.



To learn more, please visit www.spss.com. For SPSS office locations and telephone numbers, go to www.spss.com/worldwide.

SPSS is a registered trademark and the other SPSS Inc. products named are trademarks of SPSS Inc. All other names are trademarks of their respective owners.
© 2009 SPSS Inc. All rights reserved. MTA13SPC-0509