The Business Analytics Directory

About this Directory

This directory is a work in progress and will be updated every month. It is free to download and distribute as desired.

Every supplier and product entry contains a hyperlink to the relevant web site.

Business Intelligence

KPI Dashboards

Key performance indicators (KPIs) provide measures of performance that are important to an organization, department, or even an individual. They are often represented on dashboards, where several KPIs can be seen together. Formats include a wide variety of charts including pie charts, histograms, dials, gauges, and even maps.

KPI dashboard software varies considerably. General purpose dashboard software from suppliers such as Qlik and Tableau are potentially more powerful, but are generally used as part of a larger business intelligence initiative. Specialized KPI platforms such as Bilbeo can provide specific KPI functionality, and may be more suitable for some organizations.

Bilbeo is perhaps the smartest KPI dashboard platform available.

- Cloud based service (on site options available) that predicts KPI performance and correlates KPIs to find leading indicators.
- API to access on-premises data sources, and alerts generated on KPI performance.
- KPIs rated from 'Very Bad' to 'Excellent' automatically, with drill down into individual KPIs.
- White labelling.
- Pricing based on size of data and number of users.

Bittle good value online service.

- Ideal for the SME with white labelling option.
- Collaboration features and mobile access.
- Connectors to many database sources, spreadsheets, online apps.

- Graphics library includes donut, pie, histogram, 3D, bar, gauge, funnel, maps, radar, treemap etc.
- Free single user subscription up to $600/mo for large data volumes.

ClicData easy to use but with sophisticated features

- Cloud based dashboard platform with easy to use interface.
- Sophisticated features such as conditional output, pivot tables and complex calculations.
- Wide range of charts including maps, bar, pie, tables, line, dials, gauges etc.
- Connectors for database, spreadsheets, social data, online apps, and support for SQL if needed.
- Free single user version through to the Enterprise at just $50 per month.

Cyfe is easy to use and extremely good value.

- Extensive support for internal and external data sources.
- Many graph and chart types including line, spline, area, area-spline, stacked area, column, … , pie, donut, list, funnel, tables – and so on.
- A top contender for people considering a social media monitoring dashboard.
- The whole environment is widget driven, and so it is very easy to connect with social sources such as Facebook, LinkedIn and Twitter (many others too).
- A free account is offered for people with modest needs, or the Premium service offers pretty well unlimited everything for $19 per month.

Dundas Dashboard - high end sophisticated dashboarding.

- Extensive collaboration features allowing comments to be attached to data points.

- Support for OLAP processing and complex metrics.

- Support for enterprise databases and applications.

- Above average visualization capabilities and part of broader BI suite.

- Pricing model unknown.

Geckoboard cloud service best for online data

- Online dashboards service with commonly used visuals and support for online and social data sources.

- Chart types include bar, bullet graph, funnel, dials, Highcharts charting library support, leaderboard, line, pie, maps and several others.

- Over 60 integrations with online apps and data sources, spreadsheet data, social sources.

- Custom widgets support push (data pushed to dashboard) and pull (data pulled by dashboard from datasource).

- Starts at $49/mo for single user with 2 dashboards, up to $399/mo for 15 users and 25 dashboards, plus more if needed.

iDashboards general cloud dashboard platform

- Cloud (or on-premises) dashboard platform with particularly attractive visuals.

- Data Integrator for blending data from multiple sources (databases, social, spreadsheet, web apps etc).

- Drag and drop interface and full drill down into data from dashboards.

- Good mobile support.

- Pricing unavailable.

Klipfolio - very capable and reasonably priced.

- Sophisticated online dashboard service with good data handling

- Hundreds of connectors for database, spreadsheet, online apps, social data.

- Ability to combine data sources and create complex formulas.

- Dozens of chart types - dials, line, pie, maps, gauges, tables etc.

- Pricing starts at $20/mo one user, to thousands of users.

Qlik Sense - sophisticated desktop and enterprise dashboards

- Very powerful desktop and enterprise dashboard and data visualization platform.

- Supports large number of data sources with an optional service for a large number of online data services.

- Associative engine makes combining data sources straight forward.

- Large number of chart types and can be extended if needed.

- Desktop version free with a free cloud dashboard sharing service. Enterprise edition requires server license.

SimpleKPI - specifically targeted at KPI applications.

- Complete suite of KPI apps for dashboards, reporting, analytics, tracking.

- Extensive sharing and collaboration.

- Easy to use with pre-built templates.

- Supports most data sources and comes with a rich set of charts.

- £11/mo for a single user, £467 for 100 users and upward.

Slemma is a cloud BI and dashboard software that allows devs and non-devs alike to do ad hoc visual analysis against data. Slemma plugs in directly to the most popular databases (including XMLA data sources), cloud storage and cloud services, and allows to bypass the need for a warehouse. It is easy to create a dashboard by using intuitive

drag-and-drop interface or writing and running SQLs to build charts.

With the dashboards created in Slemma, team and clients can access dashboards through the web. And they can slice and dice the data in real-time by dimensions and measures for their own needs. The platform also supports data transformation features like aggregation, calculations, filtering, sorting and ordering data.

A free version supports Excel and CSV files and integration with Google Drive, Dropbox and OneDrive. Paid subscriptions start at US $ 29 per month.

Tableau - rich graphics, cloud or on premises.

- Particularly rich graphics with desktop, server and cloud editions.

- Sharing, mobile and collaboration features.

- Fast in-memory processing for most data sets.

- Complete data visualization and discovery platform.

- Pricing based on number of users - typically around $2000 per seat.

BI with Advanced Analytics

Business Intelligence has traditionally delivered a look in the rear view mirror, providing a description of what has happened or is happening. Advanced analytics are more concerned with the future, allowing businesses to predict how customers, suppliers and other aspects of business will behave. Right now, advanced analytics is only supplied by a few BI vendors - eventually they will all have to deliver. The crucial property of advanced analytics in BI is ease of use - and this is what makes it difficult for suppliers to deliver. The platforms listed below offer varying degrees of support for advanced analytics, but new entrants such as BeyondCore set the pace to some degree.

Alteryx is an interesting product, filling a void that most other analytic platforms do not address. It allows skilled business users and analysts to analyse their data using a combination of data visualization and predictive analytics tools. It also supports spatial analytics where location is important.

BeyondCore - The four dominant modes of analysis are covered by BeyondCore – descriptive (traditional BI), discovery (looking for unknown facts), predictive (finding consistent patterns that can be used in future activities), and prescriptive (actions that can be taken to improve performance). What is particularly significant about BeyondCore is that all the complexity is hidden from the business user.

Birst is a cloud based BI platform that caters for self-service and production needs. The underlying data architecture is excellent, and places Birst ahead of the pack in this respect. It is one of only a few true enterprise BI platforms available as a cloud deployment (public or private). Advanced analytics are well supported via the Weka toolkit.

Microstrategy is a very deep and very broad business intelligence platform. It is most suitable for large businesses with extensive production reporting needs. The data visualization capabilities are above average, but not exceptional. Microstrategy does however support advanced analytics, and has done for some time.

SAS Visual Analytics is an easy-to-use, sophisticated analytics platform that stands apart from most of the competition. Several features guide and advise the user to make the experience informed and effective - something almost totally lacking in most other BI platforms.

Sisense majors on ease-of-use and performance. It is capable of handling very large data volumes at speed, and provides powerful data visualization capabilities. Sisense also supports good extensibility, and version 6 improves on this and adds support for advanced analytics via R.

Spotfire from TIBCO successfully combines advanced analytics with an easy-to-use interface. At the most basic level users can create charts and dashboards with considerable ease. As skills and experience increase so other forms of analytics are available. The implementations of additional analytic capabilities are of a very high quality, and Spotfire boasts one of the fastest R runtime engines available.

Yellowfin is an excellent all-round cloud enterprise BI solution with with good data connectivity and management, numerous mechanisms for effective governance and support for embedding and the creation of cloud based services. It also supports some forms of advanced analytics.

Visual Analytics Platforms

Visual analytics is the act of finding meaning in data using visual artifacts such as charts, graphs, maps and dashboards. In addition, the user interface is typically driven by drag and drop actions using wholly visual constructs. The up side to this is ease of use - the down side can be limited functionality. Many platforms include some form of scripting language for more complex requirements, and some of the newer tools also include help with interpretation. Visual analytics are very prone to misinterpretation, but guidance and significance are increasingly being incorporated into visual analytics software.

BeyondCore - The four dominant modes of analysis are covered by BeyondCore – descriptive (traditional BI), discovery (looking for unknown facts), predictive (finding consistent patterns that can be used in future activities), and prescriptive (actions that can be taken to improve performance). What is particularly significant about BeyondCore is that all the complexity is hidden from the business user.

BIME is a cloud based visual analytics platform that connects with a wide range of data sources – internal and external, it offers a very easy to use interface, supports extensive collaboration and does all the things that might be expected of a BI solution. All the usual charts are supported, in addition to highly functional maps, relational analysis, funnel charts – all in a very customizable environment.

ClearStory provides a cloud based platform that takes business users from data to visualization with minimal need for technical skills. In common with several other suppliers (Tamr, Paxata, Platfora and others) it uses machine learning techniques and Apache Spark in-memory processing, to take data from its raw state, to a state where business users can create the data visualizations they need.

DOMO is a cloud based BI solution that caters to management and executive information needs. It also comes with strong collaboration support, and is particularly easy to use. The downside of this is that DOMO provides relatively weak data visualization capability.

GoodData is a cloud based BI platform ideal for business users who need an easy-to-use interface. It supports the creation of charts and dashboards, and connectivity to a large number of data sources. This is not a particularly sophisticated platform, but does offer recommended visualizations and mechanisms to embed visualizations into other applications.

Inetsoft provides reporting and data visualization capability that is well appreciated by its OEM customers as well as enterprise users. It employs some novel techniques for high performance query execution, and addresses the needs of power business users in addition to production needs.

InfoCaptor is available through the cloud, or can be installed locally – in both cases it provides a browser interface. Technically this is a very competent offering, with a bewildering number of chart types, and the facility to easily create and distribute dashboards.

Logi Analytics provide Logi Info for operational reporting, dashboards and self-service needs, and Logi Vision for data visualization and exploration. It's a very broad range of functionality, and Logi offers excellent extensibility, developer tools, embedding and ease-of-use.

Looker bucks the trend by placing a powerful modeling language at the heart of its product architecture, and by directly processing data held in external databases. It is particularly well suite to businesses that need to analyze complex metrics.

Microsoft Power BI data visualization sets a standard for ease-of-use with its natural language query. It is also evolving at a rapid rate, and is starting to challenge established player in this space. Expect to see it become very widely adopted.

Microstrategy is a very deep and very broad business intelligence platform. It is most suitable for large businesses with extensive production reporting needs. The data visualization capabilities are above average, but not exceptional. Microstrategy does however support advanced analytics, and has done for some time.

Prognoz provides a broadly based BI capability with several types of analysis – specifically OLAP, time-series-analysis, and modeling and forecasting. It also provides end-user tools, and specifically the ability to create charts and dashboards.

Qlik Sense is an easy to use visual analytics platform with an excellent data discovery engine and can easily be extended for bespoke needs. It comes as a free stand-alone desktop version, an enterprise server edition and also as a cloud service.

SAP Lumira is the obvious data visualization tool for businesses with a large existing SAP investment. It doesn't distinguish itself in any particular way, and other users will find a large number of more suitable alternatives are available.

SAS Visual Analytics is an easy-to-use, sophisticated analytics platform that stands apart from most of the competition. Several features guide and advise the user to make the experience informed and effective - something almost totally lacking in most other BI platforms.

Sisense majors on ease-of-use and performance. It is capable of handling very large data volumes at speed, and provides powerful data visualization capabilities. Sisense also supports good extensibility, and

version 6 improves on this and adds support for advanced analytics via R.

Spotfire from TIBCO successfully combines advanced analytics with an easy-to-use interface. At the most basic level users can create charts and dashboards with considerable ease. As skills and experience increase so other forms of analytics are available. The implementations of additional analytic capabilities are of a very high quality, and Spotfire boasts one of the fastest R runtime engines available.

Tableau is the benchmark for easy to use data visualization tools, and has certainly caused other suppliers to improve their act. The scalability of the product is excellent, but it lacks the sophisticated developer support of some other products. It is a very good adjunct to BI platforms which do not have particularly good visualization capabilities.

ThoughtSpot BI platform is given the tag line of 'search driven analytics'. The user interface is driven by a search bar, much as a user of Google experiences, and natural language queries can be entered using the language of the business concerned.

Yellowfin is an excellent all-round cloud enterprise BI solution with with good data connectivity and management, numerous mechanisms for effective governance and support for embedding and the creation of cloud based services. It also supports some forms of advanced analytics.

Free and Open Source Reporting Tools

Both free and open source reporting tools are listed here. The free versions are sometimes cut down versions of a commercial product, but some are fully functional.

BIRT (free BI suite) is an open source software project that provides the BIRT technology platform to create data visualizations and reports that can be embedded into rich client and web applications, especially those based on Java and Java EE. BIRT is a top-level software project within the Eclipse Foundation.

DataVision (free reporting tool) is an Open Source reporting tool similar to Crystal Reports. Reports can be designed using a drag-and-drop GUI. They may be run, viewed, and printed from the application or exported as HTML, XML, PDF, Excel, LaTeX2e, DocBook, or tab- or comma-delimited text files. The output files produced by LaTeX2e and DocBook can in turn be used to produce PDF, text, HTML, PostScript, and more.

DataVision is written in Java and runs almost anywhere. It can generate reports from databases or text data files. Any database with an available JDBC driver should work: Oracle, PostgreSQL, MySQL, Informix, hsqldb, Microsoft Access, Progress, and more. Columns read from text files can be separated by any character.

Report descriptions are stored as XML files. This means you can not only use the DataVision GUI but you may also edit reports using your favorite text editor.

dmyreports (free reporting tool) is a dynamic report generator for MySQL. End users will be able to select tables, fields and set conditions graphically using a web based interface to generate reports. These reports can then be saved or exported to other formats. (currently in alpha).

fyiReporting RDL Project is a powerful report and charting system based on Report Definition Language (RDL). Tabular, free form, matrix, charts are fully supported. Report output may be displayed as HTML, PDF, Excel, RTF, XML, .Net Control, Web Archive, and to a printer. Libraries exist for use in ASP.NET, Windows .Net, and command line applications. A WYSIWYG designer allows you to create reports without knowledge of RDL. Wizards are available for creating new reports and for inserting new Tables, Matrixes, and Charts into existing reports.

i-net Clear Reports is a full-blown reporting solution completely coded in Java and also available for .NET. It comes with a free report designer program to custom design report template files and is Platform Independent. It provides Crystal Reports Support, Ad Hoc Reporting, Report Repository and free mobile repository apps for Android OS and iPhone, iPod touch with iO.

icCube community edition is a free end-to-end BI solution covering reporting and analytics. The reporting is designed for business users and facilitates exploration of OLAP data via reports, graphs and dashboards. Analytics include an ETL layer and support for multidimensional analysis and MDX+. XMLA is also supported for connection to other reporting tools (eg Excel), and support for R and Java means other forms of analysis can be carried out.

Intellicus is a comprehensive suite of all BI & Reporting capabilities; Ad-Hoc & Standard reports, Dashboards, ETL, OLAP and more. Visualizations provide contemporary style of web controls allowing users to derive meaningful value from their data. Generate, format and analyze multiple Ad-hoc reports with ease. Use desktop and web studio to design complex & pixel perfect reports. Visualize your business data in dashboards. Free cut down version and full enterprise version.

Jaspersoft Community addresses the creation of reports and charts. The Community edition essentially provides a reporting and charting environment with supporting infrastructure. The report designer supports charts, images, crosstabs and sub-reports for sophisticated report layouts. Interactive report viewing is a browser based report viewer with sorting, filtering and formatting of report snapshot views. A centralized repository provides infrastructure for reporting and stores user profiles, reports, dashboards and analytic views.

JMagallanes is an open source end user application for Olap and Dynamic Reports written in Java/J2EE. It combines static reports (based on JasperReports), a Swing pivot table for OLAP analysis, and charts (based on JFreeChart). It reads from many data sources as SQL, Excel, XML, and others, and produces many outputs as PDF, XML, and application specific files for later off-line visualization of reports.

MyNeoReport library is a free and easy to use way to design and add reports to .NET applications. It supports a WYSIWYG report designer and the libraries have to be distributed with the application.

Pentaho Community (free BI suite) – main components are the reporting tools, the data integration platform, the ROLAP analytics platform and the data mining tools. With the Pentaho-Report-Designer you can create report-definitions in a graphical environment. Reports are usually published to the Pentaho-Platform, which allows you to manage, run and schedule the reports you created. If you are new to Pentaho-Reporting, you probably want to start with the Pentaho Report-Designer.

Internally, reports are executed by the Pentaho Reporting Classic Engine. Pentaho Reporting encompasses more than two dozen software projects that facilitate creating and publishing data-driven business reports.

Report Manager is a reporting application (Report Manager Designer) and a set of libraries and utilities to preview, export or print reports. Include native .Net and Delphi/C++Builder libraries, ActiveX component and also standard dynamic link library for use in any language like GNU C. The reporting engine includes a TCP enabled Report Server so thin clients can obtain reports processed in the server. Also a fully functional web report server application is available, generating Adobe PDF files on the fly.

This is a true net and web report server with no license fees and multiprocessor support. Report Manager is open source under the MPL license model, (including a GPL allowed use clause) so you can use it in your commercial software.

ReportServer (free reporting tool) – This provides an extremely flexible open source reporting and dashboard environment. It supports Eclipse Birt, JasperReports and SAP Crystal Reports reporting engines in addition to its own ad-hoc oriented reporting tools. The user interface is web based and it supports a wide range of admin tools.

Central to ReportServer is the Dynamic List. This is the preferred method supporting a wide range of functions such as column selection, filtering, sorting, grouping, sub-totals, calculation and so on. JasperReports and Eclipse Birt tend to be used for 'pixel perfect' reporting with output to a pdf file. Finally Script Reports are used for particularly complex reports, and require programming skills to use. Interactive dashboards are supported and are generally constructed for items called dadgets (Dashboard Gadgets) – these can be anything from a report to an interactive HTML5 app.

Currently supported data sources include Oracle, Microsoft SQL Server, IBM Informix, IBM DB2, mySQL, PostgreSQL, h2 and of course csv files.

Seal Report is an Open Source for the Microsoft .Net Framework entirely written in C#. It supports dynamic SQL sources, native pivot tables, HTML 5 charts and Microsoft charts, HTML rendering using the Razor engine, web report server and task scheduler.

SpagoBI (free BI suite) is essentially a very large collection of open source software brought together to create a broad business intelligence capability. In fact it goes beyond the traditional notion of BI to embrace domains such as data mining and BPM.

Support for 4 engines: JasperReport, BIRT, Accessible report, BO. SpagoBI allows to realize structured reports, using structured information views (e.g. lists, tables, crosstabs, reports) and to export them using several formats (HTML, PDF, XLS, XML, TXT, CSV, RTF).

Zoho Reports is an online reporting and business intelligence service that helps you easily analyze your business data, and create insightful reports & dashboards for informed decision-making. It allows you to create and share powerful reports. The free service limits users (2) and data, and has cut down functionality.

MySQL Reporting Software

These reporting tools tend to come as part of a more extensive BI suite or as stand-alone tools. Many are free and some are commercial products. In some cases the free products are pretty well unsupported, and suppliers offer commercial versions with training, and support.

BIRT (free BI suite) is an open source software project that provides the BIRT technology platform to create data visualizations and reports that can be embedded into rich client and web applications, especially those based on Java and Java EE. BIRT is a top-level software project within the Eclipse Foundation.

DataVision (free reporting tool) is an Open Source reporting tool similar to Crystal Reports. Reports can be designed using a drag-and-drop GUI. They may be run, viewed, and printed from the application or exported as HTML, XML, PDF, Excel, LaTeX2e, DocBook, or tab- or comma-delimited text files. The output files produced by LaTeX2e and DocBook can in turn be used to produce PDF, text, HTML, PostScript, and more.

DataVision is written in Java and runs almost anywhere. It can generate reports from databases or text data files. Any database with an available JDBC driver should work: Oracle, PostgreSQL, MySQL, Informix, hsqldb, Microsoft Access, Progress, and more. Columns read from text files can be separated by any character.

Report descriptions are stored as XML files. This means you can not only use the DataVision GUI but you may also edit reports using your favorite text editor.

dmyreports (free reporting tool) is a dynamic report generator for MySQL. End users will be able to select tables, fields and set conditions graphically using a web based interface to generate reports. These reports can then be saved or exported to other formats. (currently in alpha).

JasperReports (free BI suite) Server is a stand-alone and embeddable reporting server. It provides reporting and analytics that can be embedded into a web or mobile application as well as operate as a central information hub for the enterprise by delivering mission critical information on a real-time or scheduled basis to the browser, mobile device, printer, or email inbox in a variety of file formats. JasperReports Server is optimized to share, secure, and centrally manage your Jaspersoft reports and analytic views.

Jaspersoft ETL is a state-of-the-art data integration engine, powered by Talend. Extract data from various sources, transform the data based on defined business rules, and load into a centralized data warehouse or data mart for reporting and analysis.

A powerful eclipse-based report designer for JasperReports and JasperReports Server. Build reports from any data source, format the look and feel for print or on-screen reading, deploy to JasperReports Server or your own JasperReports implementation, and export to a wide range of formats.

JReport (commercial BI tool) from Jinfonet excels in embeddable reporting and data visualisation. It is used by some of the world's largest businesses, typically to provide online reporting facilities to customers, but is also used in many businesses to provide information directly to employees and partners. At the current time, around two thirds of Jinfonet's business comes from OEM, ISVs and other third parties who provide embedded intelligence in a variety of services and products. JReport supports dashboards, charts, reporting, and connects to a wide variety of data sources, including big data, relational databases and cloud based data.

myDBR (commercial reporting tool) is a web-based reporting system for relational databases that allows you build a feature rich, manageable and scalable reporting environment. myDBR manages the distribution and access rights automatically, so it also scales well to large scale installations. Users are able to access reports using just a web browser, no additional software installations is required.

With myDBR you can easily link reports together allowing you to building a report network. This gives your users a natural path for data mining and drill-down reports using business concepts close to them. Moving from individual reports to a report network makes sure you are getting the most out of your data.

MySQL Pivot Table Generator creates concise and informative web based pivot tables (cross tabulation) for any type of data summary without writing any code or conducting any database queries, and will simply and seamlessly export your pivot table's data into Excel. It automatically updates your web based pivot tables whenever the database has been updated with new information.

NextReports (free reporting tool) provides three, free open source utilities that support report creation using a variety of database platforms.

- NextReports Designer is an application to design in-grid reports, using connections to most popular databases including Oracle, MySql, MSSQL, PostgreSQL, Firebird and Derby. NextReports Designer has an intuitive interface that makes report creation a quick and easy task.

- NextReports Engine is a lightweight (330k) Java platform development library which can be used to run NextReports inside applications. This library is very easy to use, reports can be integrated with just a few lines of code.

- NextReports Server supports the scheduling of NextReports and Jasper reports and to automatically deliver reports via E-mail, FTP, SSH, Windows Share, WebDAV in most popular file formats (HTML, EXCEL, PDF, etc) or in real time dashboards.

Pentaho Community (free BI suite) – main components are the reporting tools, the data integration platform, the ROLAP analytics platform and the data mining tools. With the Pentaho-Report-Designer you can create report-definitions in a graphical environment. Reports are usually published to the Pentaho-Platform, which allows you to manage, run and schedule the reports you created. If you are new to Pentaho-Reporting, you probably want to start with the Pentaho Report-Designer.

Internally, reports are executed by the Pentaho Reporting Classic Engine. Pentaho Reporting encompasses more than two dozen software projects that facilitate creating and publishing data-driven business reports.

ReportServer (free reporting tool) – This provides an extremely flexible open source reporting and dashboard environment. It supports Eclipse Birt, JasperReports and SAP Crystal Reports reporting engines in addition to its own ad-hoc oriented reporting tools. The user interface is web based and it supports a wide range of admin tools.

Central to ReportServer is the Dynamic List. This is the preferred method supporting a wide range of functions such as column selection, filtering, sorting, grouping, sub-totals, calculation and so on. JasperReports and Eclipse Birt tend to be used for 'pixel perfect' reporting with output to a pdf file. Finally Script Reports are used for particularly complex reports, and require programming skills to use. Interactive dashboards are supported and are generally constructed for items called dadgets (Dashboard Gadgets) – these can be anything from a report to an interactive HTML5 app.

Currently supported data sources include Oracle, Microsoft SQL Server, IBM Informix, IBM DB2, mySQL, PostgreSQL, h2 and of course csv files.

Smart Report Maker (commercial reporting tool) – from MySQLReports provides three tools to generate reports, create charts and build forms:

- Smart Report Maker provides an environment to compile and manage an unlimited number of MySQL database reports based on tables and/or queries. It's a quick and easy PHP MySQL report generator that gives you the ability to prepare the reports that you need at the click of a button.

- Smart Chart Maker is very similar to Smart Report Maker in terms of usability and functionality. As previously noted, Smart Chart Maker supports unlimited dynamic charts that are fed directly from a MySQL database and situated around the tables and/or queries.

- PHP MySQL Form Maker allows web forms to be easily generated, and typically without programming.

SpagoBI (free BI suite) is essentially a very large collection of open source software brought together to create a broad business intelligence capability. In fact it goes beyond the traditional notion of BI to embrace domains such as data mining and BPM.

Support for 4 engines: JasperReport, BIRT, Accessible report, BO. SpagoBI allows to realize structured reports, using structured information views (e.g. lists, tables, crosstabs, reports) and to export them using several formats (HTML, PDF, XLS, XML, TXT, CSV, RTF).

Stimulsoft (commercial BI) Reports Server is a complete Business Intelligence, client-server solution that provides reporting and analytics. The front end provides facility to visualize data. The back end provides data processing and storing. Stimulsoft Reports Server can work as a standalone report server as well as an embeddable solution. The software provides all features appropriate for all reporting servers – creating and editing reports, storing them, scheduling and delivering them to the client.

Ubiq (commercial report tool) is an easy to use but powerful web based reporting tool for MySQL data. You can use it to analyze your business data, create web based reports & dashboards to track signups, sales, operations, and other business information without any programming. Data can be queried using drag and drop, and dashboards are updated automatically to reflect the latest information.

In drag and drop mode charts can be created by selecting fields from a list reflecting data available in connected databases. Functions can also be applied to data by selecting dropdown options, and filters and joins can be created with changes reflected in real-time. There is a choice of most charting types – table, line, bar (stacked, group, percent), scatter plots, stacked area, column (grouped, percent, stacked), pie, donut and funnel. Charts can also be created using sql syntax, and regardless of creation method can be exported to PDF, PNG or CSV.

Windward Solution (commercial reporting platform) is ideal for commercial software (OEM), system integrator and enterprise developers who need to add Reporting and DocGen functionality to desktop, online, SaaS or mobile products. The .NET or Java report engines connect to virtually any data source and can be quickly integrated in 14 lines of code, meaning faster time to market. Windward enables design without restrictions through AutoTag, the Microsoft Office-based template design tool. Any user can create and rapidly revise templates, freeing up dev resources.

Machine Learning

Machine Learning as a Service

Machine learning is a nebulous term used to address capabilities such as face recognition, speech recognition, natural language processing, predictive analytics and specific techniques such as deep learning (deep neural networks). All the services listed below provide APIs so the various services can be used, usually with the aim of embedding some form of intelligence into an application. Free accounts are offered by IBM, Amazon and Microsoft - usually throttled in some way, but giving an opportunity to evaluate. Some of the platforms also provide algorithm development and solution building tools (e.g. Microsoft, FICO and Amazon).

Algorithmia provides a could based platform for algorithm developers to share their work, and for application developers to incorporate algorithms into their applications. Hundreds of algorithms are already available addressing most conceivable tasks including text analytics, computer vision, graphs, machine learning and others. Costing is based on the frequency of algorithm usage and compute time.

algorithms.io provide a cloud hosted service to collect data, generate classification models and score new data. Code is added to web and portable device applications which stream data to the algorithms.io service, where it is captured and processed using random forest, support vector machine, K-Means, decision tree, logistic regression and neural network algorithms. The resulting model is then used to categorize new data. The results are passed back as a parsed data stream to power apps, or as reports and visualizations. A set of APIs are provided for developers to integrate machine learning into web and mobile applications. The algorithms are categorized as anomaly detection,

clustering, classification and collaborative filtering.

Amazon Machine Learning is a service that makes it easy for developers of all skill levels to use machine learning technology. Amazon Machine Learning provides visualization tools and wizards that guide you through the process of creating machine learning (ML) models without having to learn complex ML algorithms and technology. Once your models are ready, Amazon Machine Learning makes it easy to get predictions for your application using simple APIs, without having to implement custom prediction generation code, or manage any infrastructure.

BigML is a cloud based machine learning platform with an easy to use graphical interface. It also provides simple mechanisms to incorporate predictive models into production applications through its REST API. The platform combines supervised learning (to build predictive models), unsupervised learning (to understand behavior), anomaly detection (used in fraud detection), data visualization tools (scatter-plots and Sunburst diagrams) and many mechanisms for exploring data. The modest pricing will make it attractive to medium and large businesses who want the benefits associated with machine learning without large upfront costs and implementation delays. BigML is a pragmatic, low cost, easy to use platform for building powerful predictive models.

FICO arguably has the most experince of any supplier in the application of statistical and machine learning technologies to business problems. FICO Analytic Cloud embraces machine learning, statistics, optimization and business rules management, in the context of a well managed environment. It also serves as a marketplace for developers of analytic solutions and users who have a need for them.

Google Prediction API can integrate with App Engine, and the RESTful API is available through libraries for many popular languages,

such as Python, JavaScript and .NET. The Prediction API provides pattern-matching and machine learning capabilities.

HPE Haven OnDemand provides more than 60 APIs and services that deliver deep learning analytics on a wide range of data, including text, audio, image, social, web and video.

IBM's Watson Analytics offers predictive analytics and data visualization, and a conversational type interface. It automatically does the hard math to show the most relevant facts, patterns and relationships. A free version is offered with limitations on data volumes.

Microsoft Machine Learning Studio features a library of sample experiments, R and Python packages and best-in-class algorithms from Microsoft businesses like Xbox and Bing. Azure ML also supports R and Python custom code, which can be dropped directly into your workspace.

PurePredictive uses AI to automate the machine learning process. The platform automates the discovery of data transformations and higher order relationships between data features and automatically accommodates data drift. The cloud platform scales automatically for workloads so that data sets of virtually any size can be accommodated. The models are easily consumed through web services, and can be automatically maintained to deal with changes in business conditions.

Yottamine includes comprehensive capabilities for importing and applying models in a real-world setting. It is designed to allow users to take full advantage of scalable on-demand cloud computing, and eliminate the high costs of a dedicated infrastructure. The Yottamine Predictive Service allows for building models or making predictions in two simple steps. Via integration with scalable cloud computing it provides high speed and efficiency. It also conforms with the SSL

industry security standard and can exports to PMML those models that are supported by the standard. Data scientists can connect and control Yottamine Predictive Web Services using R programming language via YottamineR package.

Commercial Predictive Analytics Platforms

Alteryx is an interesting product, filling a void that most other analytic platforms do not address. It allows skilled business users and analysts to analyze their data using a combination of data visualization and predictive analytics tools. It also supports spatial analytics where location is important.

Angoss provides a broad suite of analytical tools and solutions which cover predictive analytics, text analytics, document exploration, scorecards and advanced modeling in an integrated environment. It has recently expanded the capability of its products considerably to meet general model building needs and provide model management capabilities. It is truly an enterprise solution for analytical needs, providing the infrastructure and management controls necessary to deploy predictive models into a production environment. Angoss provides an analytics platform of considerable breadth and capability, and joins an elite group of no more than five suppliers who truly offer enterprise capability.

Dataiku Data Science Studio (DSS) provides a productive data-to-production analytics workbench. Many of the time consuming steps that slow down analytic model production have been automated and streamlined, enabling skilled business analysts and data scientists to quickly prepare and understand data, build a model, and quickly integrate it into a production environment. Data scientists and business analysts alike will find that Dataiku DSS provides a productive, flexible, analytics workbench that is capable of addressing virtually all analytical needs. A free community version is available that is limited to 100,000 rows of data.

Datameer provides the means to bring large data sets, that display great diversity (text, relational, streaming data etc) into the Hadoop environment. Once there Datameer supplies the data wrangling tools necessary to profile and transform data into useable formats. Analysts and data scientists can then use the large set of algorithms provided by Datameer to create predictive models and perform other forms of quantitative analysis. Finally, business users can visualize data using a wide variety of charts and dashboards, and more advanced visualizations such as clustering and decision trees can be created via an easy-to-use interface.

FICO arguably has the most experience of any supplier in the application of statistical and machine learning technologies to business problems. FICO Analytic Cloud embraces machine learning, statistics, optimization and business rules management, in the context of a well managed environment. It also serves as a marketplace for developers of analytic solutions and users who have a need for them.

IBM SPSS provides statistical and data mining capability, and has associated predictive applications, particularly in marketing. Most likely to be of interest to existing IBM customers.

KNIME is arguably the premier open source platform for creating predictive models. It provides a drag and drop graphical interface for the creation of workflows of any complexity. The basic edition is free, and various extensions are available, along with support and training for business use.

RapidMiner is similar to KNIME, in that it delivers a graphical, drag and drop interface for the creation of predictive models. Like KNIME it provides hundreds of functions to prepare data, process data, find patterns, and visuals to display graphs and charts. The commercial edition comes with big data support and greater sophistication.

Salford Systems delivers a portfolio of products capable of traditional descriptive analytics and predictive analytics. What distinguishes this company is the lack of hype around the technology it offers and a willingness to discuss the pitfalls and traps associated with predictive analytics – which ironically is a prerequisite for successful analytics. The SPM Salford Predictive Modeler supports both traditional descriptive and predictive analytics. CART (Classification and Regression Tree) supports classification and the discovery of hidden relationships between attributes. It embodies a number of proprietary methods and patented extensions to the original work done in the eighties.

SAS predictive analytics is part of the very broad analytics capability offered by SAS. It not only offers the data preparation and data mining tools, but also a run-time environment. The main complaint is the cost.

Skytree will primarily appeal to large organizations with some experience in the use of machine learning technologies; and in fact Skytree positions itself as 'The Machine Learning Company'. Its Infinity platform provides the tools for analysts and data scientists to create predictive models, in a manner that is both productive and effective. Productivity benefits come from the automation of many tedious tasks that typically require weeks of fine tuning, and the effectiveness of the resulting predictive models is due mainly to the extensive data exploration tools, model performance monitoring, and the fact that most of the algorithms have been designed ground up, specifically for big data analytics.

Statistica, now part of Dell, is an integrated predictive analytics tool, which includes text analytics and statistical analysis. It has recently gone through a makeover in Version 13, with support for streaming data. There's a lot going on here and well worth a look.

Open Source Data Mining Platforms

Apache Mahout spark machine leaning supports mainly three use cases: Recommendation mining takes users' behavior and from that tries to find items users might like. Clustering takes e.g. text documents and groups them into groups of topically related documents. Classification learns from existing categorized documents what documents of a specific category look like and is able to assign unlabelled documents to the (hopefully) correct category.

Jubatus is the first open source platform for online distributed machine learning on the data streams of Big Data. Jubatus uses a loose model sharing architecture for efficient training and sharing of machine learning models, by defining three fundamental operations; Update, Mix, and Analyze, in a similar way with the Map and Reduce operations in Hadoop.

KEEL is an open source (GPLv3) Java software tool to assess evolutionary algorithms for Data Mining problems including regression, classification, clustering, pattern mining and so on. It contains a big collection of classical knowledge extraction algorithms, preprocessing techniques (training set selection, feature selection, discretization, imputation methods for missing values, etc.), Computational Intelligence based learning algorithms, including evolutionary rule learning algorithms based on different approaches (Pittsburgh, Michigan and IRL, …), and hybrid models such as genetic fuzzy systems, evolutionary neural networks, etc.

KNIME is arguably the premier open source platform for creating predictive models. It provides a drag and drop graphical interface for the creation of workflows of any complexity. The basic edition is free, and various extensions are available, along with support and training for business use.

MLlib is Spark's machine learning (ML) library. Its goal is to make practical machine learning scalable and easy. It consists of common learning algorithms and utilities, including classification, regression, clustering, collaborative filtering, dimensionality reduction, as well as lower-level optimization primitives and higher-level pipeline APIs.

mlpy is a Python module for Machine Learning built on top of NumPy/SciPy and the GNU Scientific Libraries. mlpy provides a wide range of state-of-the-art machine learning methods for supervised and unsupervised problems and it is aimed at finding a reasonable compromise among modularity, maintainability, reproducibility, usability and efficiency.

Orange is a very capable open source visualization and set of data mining tools with an easy to use interface. Most analysis can be achieved through its visual programming interface (drag and drop of widgets) and most visual tools are supported including scatterplots, bar charts, trees, dendograms and heatmaps. A large number (over 100) of widgets are supported.

R - is described as a project for statistical computing, but it might be more accurately described as the lingua franca of analytics. A large number of commercial analytics tools support R (Oracle, Microsoft, FICO, TIBCO, Angoss …), simply because it does pretty much everything. The out-of-the-box runtime environment is fairly slow, and so vendors such as TIBCO, Microsoft and Lavastorm provide speeded up runtime support. The great thing about R is that almost everything is possible, the downside is that it's a language and needs to be programmed. However various packages exist to make the whole thing more productive and easier.

RapidMiner is similar to KNIME, in that it delivers a graphical, drag and drop interface for the creation of predictive models. Like KNIME it provides hundreds of functions to prepare data, process data, find patterns,

and visuals to display graphs and charts. This link is for the open source version. The commercial edition comes with big data support and greater sophistication.

Rattle (the R Analytical Tool To Learn Easily) presents statistical and visual summaries of data, transforms data into forms that can be readily modeled, builds both unsupervised and supervised models from the data, presents the performance of models graphically, and scores new datasets.

scikit learn provides many easy to use tools for data mining and analysis. It is built on python and specifically NumPy, SciPy and matplotlib.

TANAGRA is a free set of data mining tools for academic and research purposes. It proposes several data mining methods from exploratory data analysis, statistical learning, machine learning and databases area.

WEKA is set of data mining tools incorporated into many other products (Knime and Rapid Miner for example), but it also a stand-alone platform for many data mining tasks including preprocessing, clustering, regression, classification and visualization. The support for data sources is extended through Java Database Connectivity, but the default format for data is the flat file.

Free Data Mining Clustering Tools

Data mining clustering methods are usually used for exploratory purposes and to gain insights into data. The most commonly used methods include k-Means, agglomerative hierarchical clustering and increasingly self organizing maps. The tools listed below support a wide variety of methods, some of which are particular to a given toolset.

Databionic ESOM Tools is a suite of programs to perform data mining tasks like clustering, visualization, and classification with Emergent Self-Organizing Maps (ESOM).

KEEL is an open source (GPLv3) Java software tool to assess evolutionary algorithms for Data Mining problems including regression, classification, clustering, pattern mining and so on. It contains a big collection of classical knowledge extraction algorithms, preprocessing techniques (training set selection, feature selection, discretization, imputation methods for missing values, etc.), Computational Intelligence based learning algorithms, including evolutionary rule learning algorithms based on different approaches (Pittsburgh, Michigan and IRL, …), and hybrid models such as genetic fuzzy systems, evolutionary neural networks, etc.

KNIME is a general purpose data mining platform with over 1000 different operators. Its support for clustering includes k-Means, k-Mediods, Hierarchcial Clustering, Fuzzy c-Means and SOTA (self organizing tree algorithm).

mlpy is a Python module for Machine Learning built on top of NumPy/SciPy and the GNU Scientific Libraries. mlpy provides a wide range of state-of-the-art machine learning methods for supervised and unsupervised problems and it is aimed at finding a reasonable compromise among modularity, maintainability, reproducibility,

usability and efficiency. It includes k-Means and Hierarchical Clustering.

Orange is a (relatively) easy to use data mining platform with support for hundreds of operators. The clustering methods it supports include k-Means, SOM (self organizing maps), Hierarchical Clustering, and MDS (multi-dimensional scaling).

PermutMatrix provides data visualizations with clustering and seriation analysis. It supports hierarchical clustering.

RapidMiner Community Edition is perhaps the most widely used visual data mining platform and supports Hierarchical Clustering, Support Vector Clustering, Top Down Clustering, k-Means and k-Mediods.

scikit learn provides many easy to use tools for data mining and analysis. It is built on python and specifically NumPy, SciPy and matplotlib, and supports many clustering methods including k-Means, affinity propagation, spectral clustering, Ward hierarchical clustering, agglomerative clustering (hierarchical), Gaussian mixtures and Birch clustering.

Tanagra supports a variety of methods including k-Means, SOM, LVQ (Learning Vector Quantizers) and Hierarchical Clustering.

Weka is a platform whose algorithms are used in many other toolkits (RapidMiner for example). Clustering methods include DBSCAN, COBWEB (creates a classification tree), k-Means and the EM (expectation maximization) algorithm.

Text Analytics

Text Mining and Text Analytics Platforms

Virtually all the text analytics platforms listed below primarily address customer and social data. This is the signature application for text analytics in business, and so there are many options available. Many provide a cloud based service, and some support on-premises deployment. So instead of repeating on every entry in the list that customer based analytics is the focus, the information gives a summary of the text analytics capability offered.

AlchemyAPI (now part of IBM) provides cloud based text analytics services to support sentiment analysis, marketing, content discovery, business intelligence, and most tasks where natural language processing is needed. An on-site capability can also be provided if needed. The capabilities offered by AlchemyAPI go beyond those most large organizations could build in-house, and not least because the training set used to model language is 250 times larger than Wikipedia. Innovative techniques using deep learning technologies (multi-layered neural networks) also go well beyond most of the competition, and AlchemyAPI distinguishes itself by using the technology for image recognition in addition to text analytics.

KnowledgeREADER from Angoss is part of a broad suite of analytics tools and specifically addresses text analytics in the context of customer oriented and marketing applications. It majors on visual representation including dashboards for sentiment and text analysis, and also provides a somewhat unique map of the results of association mining to display words that tend to occur together. Many of the advanced features make use of the embedded Lexalytics text analytics engine – widely recognized as one of the best in class.

Entity, theme and topic extraction are supported along with decision and strategy trees for profiling, segmentation and predictive modeling. Sentiment analysis supports the visual graphing of sentiment trends and individual documents can be marked up for sentiment.

Attensity majors on social analytics, but also offers a general purpose text analytics engine. Four major components define the offering:

- Attensity Pipeline collects data from over one hundred million social sources as input for analysis.

- Attensity Respond provides a mechanism for responding to social comment.

- Attensity Analyze allows text in emails, call-center notes, surveys and other sources of text to be analyzed for sentiment and trend.

- Attensity Text Analytics provides an underlying engine that embraces several unique NLP technologies and a semantic annotation server for auto-classification, entity extraction and exhaustive extraction. It comes with good integration tools too so that the results of text analytics can be merged with structured data analytics.

Three horizontal solutions are offered for marketing, customer service and IT.

Basis Technology delivers a variety of products and services based on multilingual text analytics and digital forensics. The Rosette platform provides morphological analysis, entity extraction, name matching and name translation in fields such as information retrieval, government intelligence, e-discovery and financial compliance.

The Rosette search and text analytics technology comes in five distinct functional units:

- RLI – Rosette Language Identifier – automatic language and character encoding identification.

- RBL – Rosette Base Linguistics – many search engines have used RBL to provide essential linguistic services such as tokenization, lemmatization, decompounding, part-of-speech tagging, sentence boundary detection, and noun phrase detection. Currently supports 40 languages.

- REX – Rosette Entity Extractor – finds entities such as names, places, organizations and dates.

- RNI – Rosette Name Indexer – matches the names of people, places and organizations written in different languages against a single, universal index.

- RNT – Rosette Name Translator – provides multilingual name translation through a combination of dictionaries, linguistic algorithms and statistical inference.

A Rosette plug-in is available for Lucene and Solr search technologies and Basis Technology provides solutions for government, social media monitoring, financial compliance, e-discovery and enterprise search.

Buzzlogix provides cloud based natural language processing and machine learning API's to support sentiment analysis, data mining, content discovery, business intelligence, and most tasks where natural language processing is leveraged. Buzzlogix provides a free version as well as commercial versions.

The various functions are called via a REST API and address the following types of data applications:

- Sentiment Analysis – classifies text as positive, negative or neutral.

- Twitter Sentiment Analysis – classifies Twitter tweets as positive, negative or neutral.

- Subjectivity Analysis – categorizes text as subjective or objective based on the content and the writing style.

- Topic Classification – you can automatically tag text into topic categories based on the IAB QAG Taxonomy Standards.

- Gender Detection – This NLP API identifies whether content is written by or targets a man or woman based on the words, context and idoms found in the content.

- Keyword Extraction. – enables you to extract from an arbitrary document, webpage or data stream all the keywords and word-combinations along with their occurrences in the text.

- Entity Extraction – named entity recognition for identifying people, places, things, and other named items.

Clarabridge provides a text analytics solution with a customer experience focus. This embraces various sources of customer information including surveys, emails, social media and the call centre.

The technology addresses three essential steps in the analysis of textual information. It supports the aggregation of information from most sources imaginable, allows the information to be processed for linguistic content and the creation of categories, and finally provides a rich user interface so the results of analysis can be seen. There are three main areas of functionality:

- Clarabridge Analyze comes with the ability to tune classification models and the way sentiment is scored, and provides various reports and visualizations.

- Clarabridge Act provides a customer engagement environment for all customer facing employees by providing real-time dashboards and the mechanisms to address customer feedback.

- Clarabridge Intelligence Platform carries out analysis and is essentially a natural

language processing (NLP) engine. Connections to other applications in the organization can be facilitated by Clarabridge Connect, and includes out-of-the-box connectors for salesforce, Radian 6, Lithium and other applications.

Clustify, used mainly by legal firms, groups related documents into clusters, providing an overview of the document set and aiding with categorization. This is done without preconceptions about keywords or taxonomies — the software analyzes the text and identifies the structure that arises naturally. Clustify can cluster millions of documents on a desktop computer in less than an hour, bringing organization to large projects.

Clustify identifies important keywords used for clustering and reports frequency information so that clusters can be browsed which contain a set of specified keywords. It also identifies a representative document for each cluster, allowing decisions to be made on other documents in the same cluster.

Uses of Clustify include taxonomy development, search engine enhancement, litigation and ad targeting. The technology is built on proprietary mathematical models which measure the similarity of documents.

Connexor provides a suite of text analytics tools which embrace a wide variety of NLP methods. These include metadata discovery, name recognition, sentiment detection, language identification, automatic document summarization, document classification, text cleansing, language analysis (10 European languages) and machine translation.

Connexor's Machinese libraries transform text into linguistically analyzed structured data. This includes Machinese Phrase Tagger which splits text into word units, Machinese Syntax which shows the relationship between words and concepts and Machinese Metadata which will extract information in 10 languages.

Solutions are offered for organizations operating in defence and security, life sciences and media, and Connexor works with a wide variety of organizations (software houses, businesses, systems integrators etc.) to deliver NLP capability. A free NLP web service is available.

DatumBox provides a cloud based machine learning platform with 14 separate areas of functionality, much of which is relevant to text analytics. The various functions are called via a REST API and address the following types of application:

- Sentiment Analysis – classifies documents as positive, negative or neutral.

- Twitter Sentiment Analysis – specifically targeted at Twitter data.

- Subjectivity Analysis – classifies documents as subjective (personal opinions) or objective.

- Topic Classification – documents assigned to 12 thematic categories.

- Spam Detection – documents labeled as spam or nospam.

- Adult Content Detection.

- Readability Assessment – based on terms and idioms.

- Langauge Detection.

- Commercial Detection – commercial or non-commercial based on keywords and expressions.

- Educational Detection – based on context.

- Gender Detection – written by or targeting men/women based on words and idioms.

- Keyword Extraction.

- Text Extraction – extraction of important information from a web page.

- Document Similarity – to detect web page duplicates and plagiarism.

Eaagle provides text mining technology to marketing and research professionals. Data is loaded into Eaagle and a variety of reports and charts are returned showing relevant topics and words, word clouds, and other statistics. Both online and Windows based software is offered. The Windows offering is called Full Text Mapper with good visuals to explore topics and various word statistics.

ExpertSystem majors on semantic analysis, employing a semantic analysis engine and complete semantic network for a complete understanding of text, finding hidden relationships, trends and events, and transforming unstructured information into structured data. Its Cogito semantic technology offers a complete set of features including: semantic search and natural language search, text analytics, development and management of taxonomies and ontologies, automatic categorization, extraction of data and metadata, and natural language processing.

At the heart of Cogito is the Sensigrafo, a rich and comprehensive semantic network, which enables the disambiguation of terms, a major stumbling block in many text analytics technologies. Sensigrafo allows Cogito to understand the meaning of words and context (Jaguar: car or animal?; apple: the fruit or the company?) – a critical differentiator between semantic technology and traditional keyword and statistics based approaches.

Sensigrafo is available in different languages and contains more than 1 million concepts, more than 4 million relationships for the English language alone, and a rich set of attributes for each concept. The Cogito semantic network includes common words, which comprise 90% of all content, and rich vertical domain dictionaries including Corporate & Homeland Security, Finance, Media & Publishing, Oil & Gas, Life Sciences & Pharma, Government and Telecommunications, providing rich contextual understanding that improves

precision and recall in the process information retrieval and management.

The technology has found uses in CRM applications, product development, competitive intelligence, marketing and many activities where knowledge sharing is critical.

Fluent Editor 2014 from Cognitum is a comprehensive tool for editing and manipulating complex ontologies that uses Controlled Natural Language. Fluent editor provides a more suitable alternative to XML-based OWL editors. It's main feature is the usage of Controlled English as a knowledge modeling language. Supported via Predictive Editor, it prohibits one from entering any sentence that is grammatically or morphologically incorrect, and actively helps the user during sentence writing. Controlled English is a subset of English with restricted grammar and vocabulary in order to reduce the ambiguity and complexity of the language.

Intellexer provides a family of tools for natural language search, document management, document comparison and the summarization and analysis of documents and web content. Nine solutions are offered, all reasonably priced:

- Name recognition – extracts names (named entities) and defines relations between them.

- Summarizer – extracts main ideas in a document and creates a short summary.

- Categorizer – for automatic document categorization.

- Comparator – compares documents and determines the degree of proximity between them.

- Question-answering – looks for documents which answer a natural language query.

- Natural language interface – generates Boolean queries for any application.

- Related Facts – is an IE plugin for Google search and selects 5 main topics and supplements them with related facts.

- Summarizer plug-in for IE – summarizes web pages and extracts concepts.

- PDF Converter – to incorporate PDF documents into text processing.

KBSPortal provides an NLP capability which includes tagging and categorizing user submitted web site content, text summarization, document linking by entities, vulgarity detection, sentiment rating and association of sentiment with products and people. This functionality is available as a web service or through purchase of source code for in-house deployment.

Lexalytics is one of the forerunners in text analytics and its Salience text analytics engine is used in market research, social media monitoring, survey analysis/voice of customer, enterprise search and public policy applications. The functionality offered by Salinece includes sentiment analysis, named entity extraction, theme extraction, entity-level sentiment analysis, summarization and facet and attribute extraction. The Salience engine can be integrated into other business applications via a flexible set of APIs, and can be tuned for very specific tasks and high levels of performance.

Another essential component in the Lexalytics approach is data directories. This effectively provides a parameter driven environment with files to set up relationship patterns, sentiment analysis, and the creation of themes. Non-English support is provided through this mechanism. Each directory can be configured to support a particular task delivering considerable flexibility and power.

Leximancer uses 'concepts' as a primary analytic structure, and these are automatically identified by the software without need for existing structures such as taxonomies or ontologies. Analysis is presented through a variety of useful visualizations, with drilling down to individual documents. It is used in survey analysis, market research, social media monitoring, customer loyalty and forensic analysis.

Leximancer Enterprise runs on a multi-user server providing users with a browser interface, and also provides a REST web services interface for application integration. A desktop version is available as a stand-alone environment, or users can access the LexiPortal via a web browser for a web based service (charging based on usage). Moderately priced academic versions are also available.

Linguamatics provides a NLP capability with either in-house or cloud based implementation. A search engine approach to mining text comes with a good query interface and the ability to drill down to individual documents. A domain knowledge plug-in supports taxonomies, thesauri and ontologies.

The technology is widely used in life sciences and healthcare and the on-line service provides access to content in this domain. A web services API supports most programming languages.

Linguasys primarily satisfies the need to process text in multiple languages – and by multiple we mean English, Arabic, Chinese, German, French, Hebrew, Indonesian, Japanese, Korean, Malay, Spanish, Pashto, Persian, Portuguese, Russian, Thai, Vietnamese, Urdu and others under development. This may well be unique in the world of natural language processing, and is possible because all languages are transformed into a large collection of concepts, each with its own identifier. It is the concepts which link all the languages together. The concept 'mobile phone' for example has the same concept number in all languages and is given identifier 26300, along with all variants that mean the same thing – 'cellular phone' for example.

Luminoso is a cloud based text analytics service that calls upon a multi-lingual capability. Many of the current problems associated with text analytics (ambiguity for example) are at least partly solved by Luminoso. A variety of useful reports and visualizations provide users with a particularly good interface.

MeaningCloud comes with as an Excel add-in and cloud based service. It provides feature level sentiment analysis, supports multiple languages, is easily integrated into other applications and automatically codes and classifies documents of any kind. It comes with a generous free plan, otherwise fees are based on usage.

NetOwl provides both text and entity analytics in the cloud and in private deployments. Text analytics includes Extractor to perform entity extraction, DocMatcher which compares and categorizes documents according to user defined concepts, and TextMiner for mining large amounts of text. Entity analytics is used to accurately match and identify names – important in many areas, including CRM, anti-fraud and national security. This includes NameMatcher to identify name variants from large multicultural and multilingual name databases. EntityMatcher performs identity resolution on similar databases.

PolyAnalyst from Megaputer is a data and text mining platform which embraces the complete analytics lifecycle. Megaputer provides two separate software packages for text analysis. PolyAnalyst performs linguistic and semantic text analysis and coding, clustering and categorization of documents, entity extraction, visualization of patterns, automated or manual taxonomy creation, text OLAP, and generating interactive graphical reports on results. TextAnalyst provides a list of the most important keywords in a document, a set of related keywords for each word, and the ability to automatically summarize a document or perform natural language queries.

PolyVista provides easy-to-use software and services to improve customer experience, enable competitive analysis, and facilitate predictive analytics. PolyVista helps its customers extract actionable insights from social data. Without additional cost, PolyVista bundles its technology with professional services in a business model called Solution as a Service. PolyVista offers POC (proof of concept), as well as one-time, monthly, and multi-month contracts to meet its clients' needs and budgets. Several intuitive user-interfaces are offered, and the company has been in business since 2001.

Provalis provides a suite of text analytics tools, each of which facilitates a particular type of text analysis. QDA Miner (available in a free Lite version) supports qualitative analysis with coding, annotation, retrieval and analysis of document and image collections. WordStat on the other hand supports the extraction of themes and trends, taxonomy and ontology creation, clustering and proximity analysis, and machine learning tools for document classification. SimStat, as the name suggests provides statistical analysis tools for text analysis. These three components can be purchased separately or as ProSuite, and all components are integrated with each other. WordStat now integrates with Stata – the statistical analysis platform.

SAS Text Analytics is part of the very broad analytics capability offered by SAS. Several modules are provided including:

- SAS Contextual Analysis – for the creation of document classification models.

- SAS Enterprise Content Categorization – for automated content categorization, and various add-on modules add extra capability as needed.

- SAS Ontology Management – to define semantic relationships.

- SAS Sentiment Analysis

- SAS Text Miner – use of various supervised and unsupervised techniques.

SIFT is a powerful cloud-based text analytics and reporting platform that can scale to process thousands of text comments in a matter of minutes. SIFT uses natural language processing and machine learning technology to analyze unstructured customer feedback such as customer comments, emails, product reviews, call center transcripts and open-ended survey responses. Core features including the ability to import and export data, sentiment analysis, topic detection, integration with Salesforce, data visualization and report collaboration. There is no need for hard-coded logic or complex IT integrations, SIFT comes ready out-of-the box, and can be used by marketing, product and support teams regardless of their level of natural language processing expertise.

Statistica Text Miner is part of the extensive Statistica statistical analysis and data mining product set. Extensive pre-processing options are available with stemming and stub lists for most European languages. 'Bag of words' type analysis can be carried out with input to the data mining capabilities of Statistica.

Free NLP Machine Learning APIs

AlchemyAPI (developer) provides advanced cloud-based and on-premise text analysis infrastructure that eliminates the expense and difficulty of integrating natural language processing systems into your application, service, or data processing pipeline.

Free Natural Language Processing Service - 100% free service including sentiment analysis, content extraction, and language detection.

Idilia brings artificial intelligence to the cloud for the developer community. Idilia's core technology analyzes text (including social media streams and queries) and determines the meanings of words in context, including both common words and proper nouns. This core technology, a major breakthrough in AI, is the most accurate in the world. Third party application developers can now leverage three core components through our Web Services API. Free version throttled to 30 day quotas - e.g. 100K sense analysis.

Jeannie (Voice Actions) is a virtual assistant with over three million downloads, now also available via API. The objective of this service is to provide you and your robot with the smartest response to any natural language question, just like Siri. Jeannie API provides an interface to the standard functions that users demand of modern voice assistants, like chatting, controlling your devices, looking up information, creating messages and much much more.

MeaningCloud comes with as an Excel add-in and cloud based service. It provides feature level sentiment analysis, supports multiple languages, is easily integrated into other applications and automatically codes and classifies documents of any kind. It comes with a generous free plan, otherwise fees are based on usage.

ML Analyzer - Text Classification, Article Summarization, Sentiment Analysis, Stock symbol extraction, Person Names Extractor, Language Detection, Locations Extractor, Adult content Analyzer.

Sentiment works by examining individual words and short sequences of words (n-grams) and comparing them with a probability model. The probability model is built on a prelabeled test set of IMDb movie reviews. It can also detect negations in phrases, i.e, the phrase "not bad" will be classified as positive despite having two individual words with a negative sentiment.

Text-Processing - Sentiment analysis, stemming and lemmatization, part-of-speech tagging and chunking, phrase extraction and named entity recognition.

Tweet Sentiment API lets you send small chunks of text - for example Tweets or Facebook Posts - and returns information about whether the sentiment is probably positive, negative or neutral, along with a score indicating how strong that probability is. It does this using natural language processing, breaking the text down into component parts and comparing those parts to a massive database of human-rated historical data.

WebKnox REST API offers functionality such as Named Entity Recognition, Part-of-Speech Tagging, Question Answering, and Language Detection. WebKnox Keywords API gives you access to keyword data. You can find information about keywords such as how often they are searched, how many competing pages there are on Google. You may also get a list of related keywords to your given keywords. Several other APIs are available.

Wit.ai is a natural language processing platform recently acquired by Facebook. It allows developers to add voice interfaces to any application and it simply converts speech or text into structured data.

Decision Support

Free and Open Source Decision Support Software

Decision Explorer is a proven tool for managing "soft" issues – the qualitative information that surrounds complex or uncertain situations. It allows you to capture in detail thoughts and ideas, to explore them, and gain new understanding and insight. The result is a fresh perspective, and time saved through increased productivity, release of creativity and a better focus. A demo version can be freely downloaded that supports 30 concepts.

Lumina Analytica provides an intuitive interface to build influence diagrams and its Intelligent Arrays provide easy management of multidimensional arrays. Monte Carlo analysis provides estimation of risk and uncertainty, and a free version, Analytica 101 can be downloaded for perpetual use - with support for 101 variables and other objects.

OpenRules is a general purpose open source Business Rules and Decision Management System available as an Open Source product. It allows subject matter experts and software developers to create, test, execute, and maintain enterprise-class decision support applications.

Paramount Decisions is a cloud based decision tool that breaks down the decision-making process into 8 easy-to-follow steps. After users complete a decision, they can generate reports to communicate the rationale of the decision to other stakeholders. It comes with strong collaboration features, version control and strong security. A free subscription is available for up to 5 decisions, but does not support collaboration.

PROSUITE Decision Support System provides a coherent and consistent software platform for assessing economic, environmental and social aspects of technologies within a life cycle approach. The PROSUITE DSS is a practical, ready-to-use software tool that is available for free, as open source system application. It offers a consistent set of assessment tools, with a common user interface, that can be used either as stand-alone tools or integrated in the platform. It is designed in such a way that it can be updated on a regular basis within and beyond the project, in order to accommodate newly developed methods, and newly developed inventory data for current and future situations.

Social Network and Media Analysis

Free and Open Source Social Network Analysis Software

Social network analysis software generally uses network and graph theory to investigate social structures both analytically and visually. The main constructs are nodes (the entities we are interested in – typically people), and the ties or edges that connect them. Many of the products listed here are open source with a license that permits free use in commercial settings.

Cuttlefish is a network workbench application that visualizes the networks with some of the best known layout algorithms. It allows detailed visualizations of the network data, interactive manipulation of the layout, graph edition and process visualization as well as different input methods and outputs in tex using Tikz and PSTricks. It is developed by the Chair of Systems Design of ETH Zürich , a research group that applies a complex system approach to investigate economic and social networks.

Cytoscape is an open source software platform for visualizing complex networks and integrating these with any type of attribute data. A lot of Apps are available for various kinds of problem domains, including bioinformatics, social network analysis, and semantic web.

Gephi is an interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs. It runs on Windows, Linux and Mac OS X. Gephi is open-source and free. It supports all types of networks – directed, undirected and mixed, and is capable of handling very large network graphs of up to one million nodes. Various metrics are supported including betweenness, closeness, diameter, clustering coefficient, average shortest path, pagerank and HITS. Dynamic filtering allows edges and/or nodes to be selected based on network structure or data. Ideal for social network analysis, link analysis and biological network analysis. Perhaps the most advanced of the open source tools.

Graph-tool is an efficient Python module for manipulation and statistical analysis of graphs (a.k.a. networks). Contrary to most other python modules with similar functionality, the core data structures and algorithms are implemented in C++, making extensive use of template metaprogramming, based heavily on the Boost Graph Library. This confers it a level of performance that is comparable (both in memory usage and computation time) to that of a pure C/C++ library. Graph-tool has its own layout algorithms and versatile, interactive drawing routines based on cairo and GTK+, but it can also work as a very comfortable interface to the excellent graphviz package.

GraphChi can run very large graph computations on just a single machine, by using a novel algorithm for processing the graph from disk (SSD or hard drive). Programs for GraphChi are written in the vertex-centric model, proposed by GraphLab and Google's Pregel. GraphChi runs vertex-centric programs asynchronously (i.e changes written to edges are immediately visible to subsequent computation), and in parallel. GraphChi also supports streaming graph updates and removal of edges from the graph. The promise of GraphChi is to bring web-scale graph computation, such as analysis of social networks, available to anyone with a modern laptop.

JUNG – the Java Universal Network/Graph Framework–is a software library that provides a common and extendible language for the modeling, analysis, and visualization of data that can be represented as a graph or network. It is written in Java, which allows JUNG-based applications to make use of the

extensive built-in capabilities of the Java API, as well as those of other existing third-party Java libraries.

The JUNG architecture is designed to support a variety of representations of entities and their relations, such as directed and undirected graphs, multi-modal graphs, graphs with parallel edges, and hypergraphs. It provides a mechanism for annotating graphs, entities, and relations with metadata. This facilitates the creation of analytic tools for complex data sets that can examine the relations between entities as well as the metadata attached to each entity and relation.

libSNA is an open-source library for Social Network Analysis, licensed under the LGPL. This library is under active development by Abe Usher in hopes that it will serve as a catalyst for improving the field of Social Network Analysis.

- Easy to use Python API

- Flexible data import options

- Scalable

- Built-in reports

- Built-in data export capabilities

- Open source – easily extended

- Fast processing time (efficient use of graph algorithms)

MeerKat is suitable for many types of network analysis, including social networks. If provides filtering mechanisms, interactive editing, support for dynamic networks, various metrics and automatically detects communities.

- network visualization with multiple layouts

- interactive network editing

- support for dynamic networks (multiple timeframes)

- network filtering

- computes different measures of centrality (network metrics and statistics)

- automatically detects communities (community mining)

- shows community dynamics in time (community event analysis and visualization)

Netlytic is a cloud-based text analyzer and social networks visualizer. Netlytic can automatically summarize large volumes of text and discover and visualize social networks from conversations on social media sites such as Twitter, Youtube, blog comments, online forums and chats. It is designed to help researchers and others to understand an online group's operation, identify key and influential constituents, and discover how information and other resources flow in a network.

NetworKit is a growing open-source toolkit for high-performance network analysis. NetworKit is a Python module. It implements efficient graph algorithms, many of them parallel to utilise multicore architectures. These are meant to compute standard measures of network analysis, such as degree sequences, clustering coefficients and centrality.High-performance algorithms are written in C++ and exposed to Python via the Cython toolchain.

NetworkX is a Python language software package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. Features include:

- Python language data structures for graphs, digraphs, and multigraphs.

- Many standard graph algorithms

- Network structure and analysis measures

- Generators for classic graphs, random graphs, and synthetic networks

- Nodes can be "anything" (e.g. text, images, XML records)

- Edges can hold arbitrary data (e.g. weights, time-series)

NodeXL is a free, open-source template for Microsoft® Excel® 2007, 2010 and 2013 that makes it easy to explore network graphs. With NodeXL, you can enter a network edge list in a worksheet, click a button and see your graph, all in the familiar environment of the Excel window.

Flexible Import and Export Import and export graphs in GraphML, Pajek, UCINet, and matrix formats. Direct Connections to Social Networks Import social networks directly from Twitter, YouTube, Flickr and email, or use one of several available plug-ins to get networks from Facebook, Exchange, Wikis and WWW hyperlinks.

Zoom and Scale Zoom into areas of interest, and scale the graph's vertices to reduce clutter. Flexible Layout Use one of several "force-directed" algorithms to lay out the graph, or drag vertices around with the mouse. Have NodeXL move all of the graph's smaller connected components to the bottom of the graph to focus on what's important.

Easily Adjusted Appearance Set the colour, shape, size, label, and opacity of individual vertices by filling in worksheet cells, or let NodeXL do it for you based on vertex attributes such as degree, betweenness centrality or PageRank.

Dynamic Filtering Instantly hide vertices and edges using a set of sliders—hide all vertices with degree less than five, for example.

Powerful Vertex Grouping Group the graph's vertices by common attributes, or have NodeXL analyse their connectedness and automatically group them into clusters. Make groups distinguishable using shapes and colour, collapse them with a few clicks, or put each group in its own box within the graph. "Bundle" intergroup edges to make them more manageable.

Graph Metric Calculations Easily calculate degree, betweenness centrality, closeness centrality, eigenvector centrality, PageRank, clustering coefficient, graph density and more. Task Automation Perform a set of repeated tasks with a single click.

Pajek – suite of programs for analysis and visualization of very large networks.

R is a general purpose analytics software, but several libraries are available for social network analysis. These include degreenet, RSeina, PAFit, igraph, sna network, tnet, ergm, Bergm, hergm, latentnet and networksis. Each provides specialized functionality and for people familiar with R represent a rich set of resources.

Social Networks Visualizer (SocNetV) is a cross-platform, user-friendly tool for the analysis and visualization of Social Networks. It lets you construct networks (mathematical graphs) with a few clicks on a virtual canvas, or load networks of various formats (GraphML, GraphViz, Adjacency,Pajek, UCINET, etc). Also, SocNetV enables you to modify the social networks, analyse their social and mathematical properties and apply visualization layouts for relevant presentation.

Furthermore, random networks (Erdos-Renyi, Watts-Strogatz, ring lattice, etc) and known social network datasets (i.e. Padgett's Florentine families) can be easily recreated. SocNetV also offers a built-in web crawler, allowing you to automatically create networks from links found in a given initial URL.

The application computes basic graph properties, such as density, diameter, geodesics and distances (geodesic lengths), connectedness, eccentricity, etc. It also calculates advanced structural measures for social network analysis such as centrality and prestige indices (i.e. closeness centrality, betweeness centrality, information centrality, power centrality, proximity and rank prestige), triad census, cliques, clustering coefficient, etc.

SocNetV offers various layout algorithms based either on prominence indices or dynamic models (i.e. Spring-embedder) for meaningful visualizations of social networks. There is also comprehensivedocumentation, both online and while running the application, which explains each feature and algorithm of SocNetV in detail.

Socioviz is a web-based Twitter analytics platform powered by Social Network Analysis metrics. Allows user to query Twitter conversations and find most influential people based on who replies to whom and who mentioned whom. Social Network graphs (user mention and hashtag copresence) are visualized and can be exported in Gephi format (gexf) for further analysis.

statnet is a suite of software packages for network analysis that implement recent advances in the statistical modeling of networks. The analytic framework is based on Exponential family Random Graph Models (ergm). statnet provides a comprehensive framework for ergm-based network modeling, including tools for model estimation, model evaluation, model-based network simulation, and network visualization. This broad functionality is powered by a central Markov chain Monte Carlo (MCMC) algorithm.

SUBDUE is a graph-based knowledge discovery system that finds structural, relational patterns in data representing entities and relationships. SUBDUE represents data using a labeled, directed graph in which entities are represented by labeled vertices or subgraphs, and relationships are represented by labeled edges between the entities. SUBDUE uses the minimum description length (MDL) principle to identify patterns that minimize the number of bits needed to describe the input graph after being compressed by the pattern. SUBDUE can perform several learning tasks, including unsupervised learning, supervised learning, clustering and graph grammar

learning. SUBDUE has been successfully applied in a number of areas, including bioinformatics, web structure mining, counter-terrorism, social network analysis, aviation and geology.

Tulip is an information visualization framework dedicated to the analysis and visualization of relational data. Tulip aims to provide the developer with a complete library, supporting the design of interactive information visualization applications for relational data that can be tailored to the problems he or she is addressing. Written in C++ the framework enables the development of algorithms, visual encodings, interaction techniques, data models, and domain-specific visualizations. One of the goal of Tulip is to facilitate the reuse of components and allows the developers to focus on programming their application. This development pipeline makes the framework efficient for research prototyping as well as the development of end-user applications.

Visone is a software tool intended for research and teaching in social network analysis. It is specifically designed to allow experts and novices alike to apply innovative and advanced visual methods with ease and accuracy. The main features include:

- interactive graphical user interface, tailored to social networks

- innovative network visualizations

- support of unconfirmed relations

- available in Java for Windows, Linux, and MacOS

- import and export of standard formats for social network data

- publication-quality export in JPEG, PDF, SVG, Metafile, and other formats

Commercial Social Network Analysis Tools

Social network analysis tools generally use network and graph theory to investigate social structures both analytically and visually. The main constructs are nodes (the entities we are interested in – typically people), and the ties or edges that connect them. Commercial products tend to be more targeted at the end-user, and are generally easier to use than open source products (but not always).

Centrifuge Analytics is a big data discovery technology that provides the power and flexibility to connect, visualize and collaborate without complex data integration, costly services or a data science degree. It combines sophisticated link-analysis, interactive visualizations and discovery features to dramatically simplify data pattern and connection recognition.

Commetrix is a Software Framework for Dynamic Network Visualization and Analysis that supports Community Moderators, Members, and Network Researchers. It provides easy exploratory yet comprehensive access to network data and allows for:

- Extracting virtual communities in electronic communication networks

- Analyzing dynamic network change, properties, lifecycles, and structures

- Creating rich expert network maps or recommendation systems from communication logs or other network data sources (including surveys)

- Searching, filtering, navigating social corpora, like e-mail, discussions

- Understanding and utilizing your social networks

- Trace dissemination of topics or properties through the network

- Extendable to all sources of network data (e.g. collaborative work on electronic documents or contents, electronic project collaboration, VoIP telephony/Contact Centers, Instant Messaging, E-Mail, Discussion, …)

InFlow performs network analysis and network visualization in one integrated product – no passing files back and forth between different programs like other tools. What is mapped in one window is measured in the other window – what you see, is what you measure. InFlow excels at what-if analysis - change the network, get new metrics - just 2 clicks of the mouse. You do not need to be an expert in statistics to use InFlow. InFlow runs under Windows 7, 8, and 10, and also on the Mac.

Keynetiq is an innovative platform for Organizational Network Analysis. It's designed to map, visualize and analyze networks of people and relations between them, revealing how organizations really operate in day-to-day business. By mapping and analyzing the patterns of information flow, knowledge exchange, collaboration, decision flow and other key interactions, Keynetiq gives customized, actionable insights. At an individual level, Keynetiq identifies employees with key network roles - informal opinion leaders, change agents, rising stars, natural mentors, connection brokers, top influencers and knowledge hubs. Understanding their strengths and network position allows optimal resource allocation and process improvement.

Netlytic is a cloud-based text analyzer and social networks visualizer. Netlytic can automatically summarize large volumes of text and discover and visualize social networks from conversations on social media sites such as Twitter, Youtube, blog comments, online forums and chats. It is designed to help researchers and others to understand an online group's operation, identify key and influential constituents, and

discover how information and other resources flow in a network.

NetMiner is a premium software tool for Exploratory Analysis and Visualization of Network Data. NetMiner allows users to explore a network data visually and interactively, and helps detect underlying patterns and structures of the network. NetMiner has the comprehensive data model expressing various types of nodes, links, node attributes and link attributes. Through its data model, NetMiner is able to represent most social, natural and physical phenomena as network data.

Polinode has the ability to map, visualize and analyze network data. Targeted mainly at organizations and consultants, applications range from identifying change agents and finding critical links through to promoting diversity and improving workplace layouts. Features include:

- Upload any existing network data you have easily in either Excel, JSON or GEXF formats

- Completely customize the color and appearance of the network visualization, including node size, edge thickness, node colors, background color, label appearance and a lot more

- Animated and computationally efficient Force Atlas layout algorithm

Sentinel Visualizer provides users with insight into patterns and trends hidden in data. Its database driven data visualization platform lets users quickly see multi-level links among entities and model different relationship types. Advanced drawing and redrawing features generate optimized views to highlight the most important entities.

Social Network Analysis (SNA) metrics reveal the most interesting suspects in complex webs. With advanced filtering, squelching, weighted relationship types, shortest path analysis, timelines, and integrated geospatial

features, Sentinel Visualizer helps users maximize the value of their data.

- Integrated knowledgebase, link analysis, social network analysis, geospatial, and timelines

- Industry standard database format

- Supports network multi-user environments

- Laptop deployable

- Multiple monitor support

- Originally created for the intelligence community

SVAT is for data visualization, fraud investigation, and more. It provides user-friendly, cost-effective visualization of links and flows between subjects. A chronological overview of the visualized dataset is crucial in many cases. SVAT supports two different timeline views with a lot of options to choose from. It can mine data from structured or unstructured sources and crunches them to reveal hidden patterns.

TouchGraph Navigator is 100% java technology. A range of different relationship types are supported and edges can be directed, undirected, and can show flow in both directions. Text and numerical attributes can be associated with nodes and edges. Tables display the attributes and allow sorting, and mages can be associated with nodes. Advanced cluster computation reveals inherent groupings and co-citations and co-occurrence analysis clarifies dense networks. Desktop and server editions available.

XANALYS specialise in providing powerful software capabilities. From threat assessment, Investigative major case management and advance crime and fraud analytics. It helps to manage multi-jurisdiction major crime investigations, evaluate and analyse suspicious financial transactions, capture and act upon intelligence reports, and disclose evidence in a court-ready format to ensure successful outcomes.

Developer and Operations

Free Python Data Visualization Tools

Several python data visualization tools - some aimed at scientific work, and others with a more commercial touch.

Bokeh is a Python interactive visualization library that targets modern web browsers for presentation. Its goal is to provide elegant, concise construction of novel graphics in the style of D3.js, but also deliver this capability with high-performance interactivity over very large or streaming datasets. Bokeh can help anyone who would like to quickly and easily create interactive plots, dashboards, and data applications.

ggplot is a plotting system for Python based on R's ggplot2 and the Grammar of Graphics. It is built for making professional looking, plots quickly with minimal code. Making plots is a very repetitive: draw this line, add these colored points, then add these, etc. Instead of re-using the same code over and over, ggplot implements them using a high-level but very expressive API. The result is less time spent creating your charts, and more time interpreting what they mean.

matplotlib is a python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. matplotlib can be used in python scripts, the python and ipython shell (ala MATLAB®* or Mathematica®†), web application servers, and six graphical user interface toolkits.

Mayavi2 is a general purpose, cross-platform tool for 3-D scientific data visualization. Its features include:

- Visualization of scalar, vector and tensor data in 2 and 3 dimensions.

- Easy scriptability using Python.

- Easy extendibility via custom sources, modules, and data filters.

- Reading several file formats: VTK (legacy and XML), PLOT3D, etc.

- Saving of visualizations.

- Saving rendered visualization in a variety of image formats.

- Convenient functionality for rapid scientific plotting via mlab

Pygal is a python SVG chart generator with support for bar charts, line charts, XY charts, pie charts, radar charts, box plots, dot charts, pyramid charts, funnel charts, gauge charts, worldmap charts and country charts. There are many options for chart customization.

Seaborn is a library for making attractive and informative statistical graphics in Python. It is built on top of matplotlib and tightly integrated with the PyData stack, including support for numpy and pandas data structures and statistical routines from scipy and statsmodels. Some of the features that seaborn offers are:

- Several built-in themes that improve on the default matplotlib aesthetics

- Tools for choosing color palettes to make beautiful plots that reveal patterns in your data

- Functions for visualizing univariate and bivariate distributions or for comparing them between subsets of data

- Tools that fit and visualize linear regression models for different kinds of independent and dependent variables

- Functions that visualize matrices of data and use clustering algorithms to discover structure in those matrices

- A function to plot statistical timeseries data with flexible estimation and representation of uncertainty around the estimate

- High-level abstractions for structuring grids of plots that let you easily build complex visualizations

Vispy is a high-performance interactive 2D/3D data visualization library. Vispy leverages the computational power of modern Graphics Processing Units (GPUs) through the OpenGL library to display very large datasets.

Applications of Vispy include:

- High-quality interactive scientific plots with millions of points.

- Direct visualization of real-time data.

- Fast interactive visualization of 3D models (meshes, volume rendering).

- OpenGL visualization demos.

- Scientific GUIs with fast, scalable visualization widgets (Qt or IPython notebook with WebGL).

Free ETL Tools for MySQL

Apatar provides connectivity to many popular applications and data sources (Oracle, MS SQL, MySQL, Sybase, DB2, MS Access, PostgreSQL, XML, InstantDB, Paradox, BorlandJDataStore, Csv, MS Excel, Qed, HSQL, Compiere ERP, SalesForce.Com, SugarCRM, Goldmine, any JDBC data sources and more). Supports bi-directional integration, is platform independent and can be used without coding via the Visual Job Designer. An on-demand version supports Salesforce and QuickBooks.

CloverETL supports a wide range of data sources including CSV, Excel, databases via JDBC drivers, LDAP, Lotus Notes, Quickbase, Infobright, web services, XML and JSON. Functionality includes filters, joins, lookup, aggregate, sort, dedup, rollup, normalize, pivot and much more. Interface is primarily visual without coding. Free community edition and various commercial packages.

Jaspersoft ETL is easy to deploy and out-performs many proprietary ETL software systems. It is used to extract data from your transactional system to create a consolidated data warehouse or data mart for reporting and analysis.

KETL is a premier, open source ETL tool. The data integration platform is built with portable, java-based architecture and open, XML-based configuration and job language. KETL features successfully compete with major commercial products available today. Highlights include:

- Support for integration of security and data management tools

- Proven scalability across multiple servers and CPU's and any volume of data

- No additional need for third party schedule, dependency, and notification tools

- KETL is open sourced under a combination of the both the GNU Lesser Public License (LGPL) and the GNU Public License (GPL).

Pentaho's Data Integration, also known as Kettle, delivers powerful extraction, transformation, and loading (ETL) capabilities. You can use this stand-alone application to visually design transforms and jobs that extract your existing data and make it available for easy reporting and analysis.

Talend Open Source Data Integrator provides multiple solutions for data integration, both open source and commercial editions. Talend offers an Eclipse-based interface, drag-and-drop design flow, and broad connectivity with more than 400 pre-configured application connectors to bridge between databases, mainframes, file systems, web services, packaged enterprise applications, data warehouses, OLAP applications, Software-as-a-Service, Cloud-based applications, and more.

Free Data Ingestion Tools

Chukwa is an open source data collection system for monitoring large distributed systems. Chukwa is built on top of the Hadoop Distributed File System (HDFS) and Map/Reduce framework and inherits Hadoop's scalability and robustness. Chukwa also includes a flexible and powerful toolkit for displaying, monitoring and analysing results to make the best use of the collected data.

Cloudera Morphlines is an open source framework that reduces the time and skills necessary to build or change Search indexing applications. A morphline is a rich configuration file that simplifies defining an ETL transformation chain. These transformation chains support consuming any kind of data from any kind of data source, processing the data, and loading the results into Cloudera Search. Executing in a small embeddable Java runtime system, morphlines can be used for Near Real Time applications, as well as batch processing applications.

Databus is a source-agnostic distributed change data capture system, which is an integral part of LinkedIn's data processing pipeline. The Databus transport layer provides latencies in the low milliseconds and handles throughput of thousands of events per second per server while supporting infinite look back capabilities and rich subscription functionality.

Fluentd is an open source data collector, which lets users unify the data collection and consumption for a better use and understanding of data. Fluentd tries to structure data as JSON as much as possible: this allows Fluentd to unify all facets of processing log data: collecting, filtering, buffering, and outputting logs across multiple sources and destinations. 300+ community-contributed plugins connect dozens of data sources and data outputs.

Flume is a distributed, reliable, and available service for efficiently collecting, aggregating, and moving large amounts of log data. It has a simple and flexible architecture based on streaming data flows. It is robust and fault tolerant with tunable reliability mechanisms and many failover and recovery mechanisms. It uses a simple extensible data model that allows for online analytic application.

Gobblin is a universal data ingestion framework for extracting, transforming, and loading large volume of data from a variety of data sources, e.g., databases, rest APIs, FTP/SFTP servers, filers, etc., onto Hadoop. Gobblin handles the common routine tasks required for all data ingestion ETLs, including job/task scheduling, task partitioning, error handling, state management, data quality checking, data publishing, etc. Gobblin ingests data from different data sources in the same execution framework, and manages metadata of different sources all in one place. This, combined with other features such as auto scalability, fault tolerance, data quality assurance, extensibility, and the ability of handling data model evolution, makes Gobblin an easy-to-use, self-serving, and efficient data ingestion framework.

Heka is a tool for collecting and collating data from a number of different sources, performing "in-flight" processing of collected data, and delivering the results to any number of destinations for further analysis. Heka is written in Go, but Heka plugins can be written in either Go or Lua. The easiest way to compile Heka is by sourcing (see below) the build script in the root directory of the project, which will set up a Go environment, verify the prerequisites, and install all required dependencies. The build process also provides a mechanism for easily integrating external plug-in packages into the generated

A single Kafka broker can handle hundreds of megabytes of reads and writes per second from thousands of clients. Kafka is designed to allow a single cluster to serve as the

central data backbone for a large organization. It can be elastically and transparently expanded without downtime. Data streams are partitioned and spread over a cluster of machines to allow data streams larger than the capability of any single machine and to allow clusters of coordinated consumers.

Apache Sqoop is a tool designed for efficiently transferring bulk data between Apache Hadoop and structured datastores such as relational databases. You can use Sqoop to import data from external structured datastores into Hadoop Distributed File System or related systems like Hive and HBase. Conversely, Sqoop can be used to extract data from Hadoop and export it to external structured datastores such as relational databases and enterprise data warehouses.

IoT Platforms

These IoT platforms address a wide range of needs, from connection and management of consumer devices through to plant monitoring and management. Most address the development of apps, data analytics and connectivity - but with each product the emphasis is different.

Arrayent is a SaaS IoT platform primarily targeted at allowing manufacturers to create connected products. Customers can control products via mobile devices, and collected data is used for warranty and understanding customer behavior. The platform is comprised of an embedded software agent to connect to the Arrayent cloud, cloud software for analytics and toolkits for mobile and web application development.

Axeda provides an advanced Cloud-Based Service and Software for managing connected products and machines and implementing innovative Machine-to-Machine (M2M) and Internet of Things (IoT) applications. It supports the conversion of machine and sensor data into business information, machine and sensor data integration, and M2M and IoT application development.

B-Scada provides a software system used to automate and/or monitor industrial processes in various vertical markets: manufacturing, transportation, energy management, building automation, and any other field where real time operational data is used to make decisions. HMI/SCADA data visualization solutions work by giving operators access to real-time operational data in graphic form through a Human Machine Interface (HMI), allowing them to remotely monitor and control the processes from a centralized location.

Carriots is a platform as a service (PaaS) for M2M and IoT projects. It supports the collection of data from connected things, and the building of applications. It also provides a development environment, APIs and hosting

for IoT projects. The functionality includes device management, listeners, rules, triggers, SDK application engine, custom alarms and custom control panels.

[EVRYTHNG] is an award-winning IoT cloud platform that connects any consumer product to the Web and manages real-time data to drive applications. Smart products don't just deliver connected experiences and support services, they share data with enterprise systems and other device clouds for smarter ROI.

EVRYTHNG manages digital identity data in an intelligent IoT 'smart products' cloud to connect consumer products to the Web and drive real-time applications.

Exosite platform supports the creation of connected devices, processing in the cloud and data management tools. It uses well-established, open Internet protocols that are thoroughly tested, widely suetheriospported, and continuously improved upon. It also provides example code for a variety of devices to help developers communicate with the platform. Exosite's APIs allow users to easily get data to and from any existing platform or business system, so they can use data in the way they need to.

IoT Eclipse is an implementation of IoT standards such as MQTT, CoAP, LWM2M and OneM2M. The Eclipse IoT Working Group is an industry collaboration of companies who invest and promote an open source community for IoT. Eclipse IoT provides building blocks that sit on top of open standards and protocols and provide additional services and frameworks for device management, wired/wireless communication and vertical solutions like home automation.

IoT-Ticket has a history going back to 1999 and has considerable experience in connecting devices and machines to the IoT. It can be used to create condition, maintenance or status reports to keep track of assets, and provides a web-based

Desktop to monitor and control them in real time.

Jasper provides a cloud-based software platform for the Internet of Things (IoT) and, more specifically, to enable product businesses to become IoT service businesses. The platform is designed to help organizations launch, manage, and monetize the deployment of the Internet of Things worldwide.

MachineShop is an API-centric IoT platform company that simplifies the way organizations access, understand and act upon real world events in real time. MachineShop is a collection of hundreds of unique services with RESTful APIs - like building blocks - to create and integrate applications and solutions.

Muzzley is a platform that enables users to access connected devices via mobile devices. It's a single entry point for consumers to interact, in one single app, with all their connected devices, access to their activities and receive important notifications from them. IoT developers and/or business can get started with Muzzley by creating their own connected apps and widgets (or use the ones Muzzley provide off-the-shelf) and use them to engage with users and/or customers.

Raco Wireless supports the creation, deployment and management of M2M solutions. The Omega Management Suite is a powerful cloud-based dashboard that provides partners of RacoWireless with custom control of their connected solutions. The Position Logic platform focuses on making it easy for business to address enterprise asset management, in-house asset tracking solutions, GPS asset tracking in a fully-customizable solution. With Omega DevCloud, web developers with common programming knowledge can transfer their skills to quickly build IoT/M2M applications in a matter of minutes.

Thingsquare is a software platform that product makers use to connect their products with smartphones. The wireless firmware – which runs on a wireless chip inside a product – automatically creates a self-healing wireless network that automatically detects and heals wireless problems. The Thingsquare platform has out-of-the-box support for Android and iOS apps that deal with device discovery and setup, user signup and login sessions. Apps can be either mobile web HTML5 apps, to hit the market quickly, or as native apps, for a more in-depth investment.

Xively is an enterprise IoT platform and application solution for connected businesses. Xively simplifies the way companies securely and robustly CONNECT their products & users, MANAGE IoT data at scale, and ENGAGE more closely with their customers, users & partners.

Data Management

Data Quality Tools

AB Initio provides significant data quality tools as part of a broader suite of products for building, running and integrating enterprise applications. The end-to-end approach to data quality is based on design patterns using Ab Initios coupled technologies. They are all architected together including the Co>Operating System, the Enterprise Meta>Environment (EME), the Business Rules Environment (BRE), and the Data Profiler. Using Ab Initio, a company can implement a complete data quality program including detection, remediation, reporting, and alerting.

Blazent's 5-step data evolution process begins with data atomization, which breaks down IT data, regardless of its source, to its most granular level. It then enriches the data with identity management, relationship analysis, purification, and historicity. To create the master source of truth, Blazent integrates with more than 230 discrete data sources, from ITSM systems like ServiceNow to procurement, billing, operational tool stacks, or even shadow IT sources like spreadsheets. Powered by high-performance technologies including Active MQ, Cassandra, Hadoop and Spark, Blazent's big data engine is optimized for scalability and near real-time data processing.

Data Ladder's data quality tools offer very high levels of matching speed and accuracy for the business user at an affordable price. The company recently beat IBM and SAS in matching accuracy and speed for enterprise level data cleansing in an independent study. The company's flagship software suite DataMatch includes the following features:

- Clean, deduplicate, and match data with advanced technology previously available only in high-end customized software solutions

- Quickly combine customer, vendor, and sales lead information

- Big data capability on data sets up to 100 million records

- Advanced record linking technology provides ability to create data warehouses

- Quick data profiling tool

- Scalable configurations for deduplication and record linking, suppression, enhancement, extraction, and standardization of business and customer data

Link and consolidate customer data quickly and easily. Data Ladder offers a free trial for users.

Data Manager is a program which allows you to process and manipulate your data in a easy and logical manner using a graphical interface. It reads and writes delimited files such as comma separated files (CSV) and also can read data from ODBC Data Sources. It also allows you to construct a conceptual design on how you are going to process your data and transform it into another form. You form your design by adding functional nodes and linking them such that the links form the data flow through nodes on a graphical work area.

Each node performs a single function on your data, once it completes it passes your data to the node it is linked to and the process continues until the data encounters a output node. You can form a simple design or a complicated design with hundreds of nodes and multiple input and output nodes.

Datamartist is a fast, easy to use, visual data profiling and transformation tool. It includes a data profiling tool for analyzing format, types, completeness and value counts. Understand data quality issues clearly and quickly. Data can be transformed in a graphical ETL environment with a library of different data

blocks. Export out to files or directly to databases.

DataPreparator is a free software tool designed to assist with common tasks of data preparation(or data preprocessing) in data analysis and data mining. DataPreparator provides:

- A variety of techniques for data cleaning, transformation, andexploration

- Chaining of preprocessing operators into a flow graph (operator tree)

- Handling of large volumes of data (since data sets are not stored in the computer memory)

- Stand alone tool independent of any other tools

- User friendly graphical user interface

DQ Cloud Services from Uniserv is unusual in that it is a cloud based service. Connectors exist for a wide range of business applications, including Microsoft Dynamics CRM, Oracle Siebel, Salesforce and SAP Business Suite. It is able to enhance data in applications (telephone numbers, location etc) by accessing public databases. These functions cover bank data checks, email validation, entity titles (businesses or individuals), and address correction.

DQGlobal provides a suite of data quality software including deduplication, data migration, an API with a set of data quality improvement functions, and specific utilities (e.g. formatting addresses).

iManageData is an essential tool for commercial data pre-processing. iManageData helps you create cleaner, more useful information from your data. With its comprehensive selection of data sources, filtering, data conversions and mathematical transformations, iManageData provides quality data for any analytical application.

OpenRefine (formerly Google Refine) is a powerful tool for working with messy data:

cleaning it; transforming it from one format into another; extending it with web services; and linking it to databases like Freebase.

Paxata is an enterprise platform providing the tools to significantly speed up data cleansing, and offers a contemporary solution that employs a big data infrastructure and automated techniques which exploit machine learning methods. The net result is a self-service data preparation platform that can be used by business analysts and skilled business users to considerably speed up the data preparation task.

The core capability of Paxata leverages Hadoop and specifically Spark, so that large scale in-memory processing is available for the machine learning algorithms that give Paxata much of its power. Paxata can be deployed on premises or accessed as a cloud service. The on-premises deployment requires a Hadoop environment (either dedicated or shared).

Syncsort offers a variety of data processing tools including fast data sort, ETL, ETL Optimization, SQL migration - and several others.

Talend's open source data quality tools are embedded in Talend Open Studio for Data Quality, a popular open source data quality application. Main features include:

- Free to download and use under an Apache license.

- Very easy to learn, with an Eclipse-based graphical workspace geared toward drag 'n drop functionality.

- Versatile enough to work in any IT architecture, with more than 400 built-in connectors that enable easy access to major databases, file formats, and package enterprise applications.

- Comprehensive data quality improvement functionality, including support for data standardization, de-duplication, and enrichment.

- Support for the design and deployment of reusable enterprise data quality services, including real-time data cleansing services to keep up the quality of incoming data.

- A web-based data quality monitoring and reporting portal to help spread data quality awareness and a data quality culture across your enterprise.

The data quality tools in Talend Open Studio for Data Quality allow users, without having to write any code, to perform data quality analysis tasks ranging from simple statistical profiling, to analysis of text fields and numeric fields, to validation against standard patterns (email address syntax, credit card number formats) or custom patterns of a user's own creation.

Tamr provides a sophisticated, but easy to use, data mapping, cleansing, integration and unified data access platform. Underneath the hood there is some very smart technology. Support for big data and semi-structured data types will be welcomed by many users employing those data sources and types.

TS Quality is the data cleansing and standardization component of the Trillium Software System, a robust, scalable, highly available and easily deployable solution for mission-critical enterprise data quality. Trilliums data quality services deploy in batch or real-time through an on-site or hosted solution, using the same rule sets and standards across an unlimited number of applications and systems.

Uniserve Data Quality Service Hub augments and corrects data in many business applications. Correction of address information, email address checking, telephone number checking, and bank data verification are all included. The platform can operate in batch and/or real-time. Data Analyzer establishes the current state of data quality, Data Cleansing corrects data, Data Protection ensures updates are adequate and

from authorised people, and Data Governance detects unusual activity.

WinPure's Clean & Match data quality software provides advanced data matching and sophisticated data cleansing at an affordable price for any business size. Clean & Match has been specially designed to be used by anyone, not just IT professionals. Combining a simple-to-use interface and powerful features its ideal for cleaning, correcting and deduplicating mailing lists, databases, spreadsheets and CRM's. Main features include:

- Fast fuzzy matching technology to identify duplicate records, previously only available in enterprise software.

- Real-time statistics with 3D charts and scoring system to help populate missing data.

- Complete suite of data quality tools including standardizing, profiling, filtering and de-duplication.

- Works with business and consumer data, local and international.

- Scalable editions for any business size.

- World-class customer support.

WinPure offer a free 21 day trial so you can try the software using your own data.

Solutions

Sales Recommendation Engines

The availability of large volumes of data, and the development of analytics technologies mean there is every reason to add intelligence to sales and marketing processes. Here is a list of over 30 platforms that include recommendation engines.

4-Tell offers several products including Recs4 Web for ecommerce up-sell and cross-sell, Recs4 Mobile, Recs4 Email, Recs4 Ads and Target4 Ads and Lists. The 4-Tell Boost Dashboard is a powerful merchandising tool to view, manage and control the recommendations.

Apptus eSales' personalisation algorithms work across all your screen real-estate, and recommendations will select the right products to promote to customers.

Attraqt is an online visual merchandising tool that helps retailers deliver a step change in their conversion rate, and drive online sale success. It supports product sequencing, search technology that learns from customer behaviour and product recommendation.

Barilliance creates a personalised shopping experience. Every step of the purchase journey is supported with real time website personalisation, and will generate cart abandonment emails and personalised product recommendations.

Baynote supports collection of customer data, A/B testing, and reporting for online merchandising.

bitREC boosts eCommerce conversion rates by providing Shoppers personalized and context-aware shopping experience.

blueknow is a SaaS solution that transforms a complex product into a simple and easy to use service. The recommendation engine is simple to put in place, and will yield result in the first month. Users also benefit from client support, which is free, immediate, multicanal and multilingual.

BrainSINS provides the right personalisation tools to increase eCommerce conversion and sales

Celebrus collects, contextualizes and delivers data about how individuals are behaving across a brand's digital channels including websites, mobile apps, social and streaming media. The tagging-free software streams this online behavioral data in real-time, or near real-time, into the client's chosen target database or technology. Global blue-chip clients use the data captured by Celebrus to deepen their customer knowledge through digital intelligence, uncover hidden insight through big data analytics and power individualization through contextual marketing.

Certona delivers customer experiences tailored to each individual using continuous behavioural profiling and predictive technology, resulting in increased engagement and conversions.

Clerk increase sales and conversions with an intelligent search engine, and increases loyalty and customer lifetime value with personalized retention emails and offers that make customers return.

click2customer QuickStart Recommendations provides targeted recommendations to your customers. QuickStart Rescues allows users to save customers who are just about to abandon their shopping carts.

Commerce Sciences supports demand generation, e-commerce optimisation, A/B testing, and customer onboarding.

DirectedEdge web services API allows users to add E-Commerce recommendations to an existing shop. Language support for PHP,

Ruby, Python and Java (along with our generic REST API) to allow users to offer related product and personalized recommendations to customers. Also offer a social API and recommendation engine.

Ezako offer a recommendation engine in a totally integrated version that is directly embedded in the web pages blended to the site design.

FACT-Finder combines all sales channels with one solution - online shop, in-store, point of sales, terminals, mobile and social.

Hugefly incorporates best-in-class machine learning technology that helps make ecommerce engaging. Use our superior real-time search, recommendations and personalization to offer relevant, trustworthy buying experience.

Magnetic brings together big data, patented algorithms, predictive intelligence and proprietary analytics. The platform is powered by a unique combination of in-market and in-store data, and the technology responds to shifts in people's behaviors in real-time: when they're on the web, on site, in their inbox, and on any device.

Monetate makes it easy to create and deploy super-relevant, multi-channel experiences for unlimited segments.

Peerius personalization engine helps the world's leading brands increase revenue by creating a personalized experience for their customers, through a suite of powerful next-generation technology.

Prediggo provides intelligent search, semantic merchandising and 1 to 1 marketing. Ontology Filtering - a patented algorithm, integrates semantic product understanding with transactional analysis to support personalization strategies.

Prudsys allows organizations to personalize product and content recommendations as well as search result lists in real time. Optimize the customer experience by

generating relevant content at every touchpoint of the customer journey and increase sales and customer loyalty through personalized recommendations. It also supports marketing automation and dynamic pricing.

Qubit provides AB & MV Testing, Personalisation, Social Proof, Recommendations, Voice of Customer, Landing Page Optimisation, Merchandising, Mobile Apps, Digital Analytics, Mobile Analytics, Marketing Attribution, Customer Feedback

RichRelevance enables personalized search and browse with the most relevant products. Also create highly relevant content experiences for each customer.

Sajari adds search and recommendations to any app or web site.

SmartFocus supports personalization, market intelligence, social insights, email and mobile notifications.

Softcube provides personalization without any programming effort.

Sparkow is a SaaS solution that integrates with all the software on the market. It connects to all product and customer data sources: e-commerce platforms, CRM, ERP, PIM, DOM, WMS, social networks, web analytics… Users can create business rules that personalise the buying experience. Access to the application is via a simple web interface.

Steelhouse supports creation of unlimited audience segments based on purchase patterns, device, search terms, location, and more. First-party data is gathered in real time, allowing users to target any and all audience segments with relevant offers immediately.

Strands Retail makes intelligent product recommendations to customers – based on what other users bought, unique shopper behavior, and upsell, cross-sell.

Swogo provides automated bundles to sell more high margin products.

TargetingMantra Snowflake, segments visitors at the push of a button, based on Browsing behavior, Purchase history, Location and much more. Create curated campaigns on the fly.

UNBXD supports a personalized shopping experience with site search and recommendations.