IDOL™ Server

Intelligent Data Operating Layer (IDOL)

At the heart of Autonomy's infrastructure software is the Intelligent Data Operating Layer™ (IDOL). IDOL sits above an organization's data and serves as a scalable and secure platform for forming a conceptual and contextual understanding of all content in an enterprise. Offering far beyond simple retrieval, IDOL enables the enterprise to leverage all critical information for maximum benefit through its offering of over 500 functions, while promoting enterprise-wide compliance to minimize the risks inherent in the growth of content. Built on a Service Oriented Architecture (SOA), IDOL's flexible, modular design enables enterprises to match the benefits of IDOL functionalities to the organizations' needs.

IDOL provides a flexible framework incorporating the essential attributes of a pan-enterprise search platform:

- *Access to ALL data sources and file types: IDOL supports over 1,000 data types, including rich media, and connects to over 400 content repositories*
- *Language independence: IDOL's pattern-matching technology is fundamentally language independent*
- *Compatibility with all enterprise Operating Systems: IDOL is a cross-platform solution that supports all major systems*
- *FRCP compliance: IDOL provides FRCP-compliant search, as it searches all content in the enterprise with no "jump-out"*
- *Scalability: IDOL delivers linear scalability by use of its distribution model*
- *High Availability: IDOL offers a number of options to ensure redundancy and failover*
- *Security: IDOL's mapped security model is the only index security model empirically proven to scale in the enterprise*

Key IDOL Functions

Automatic Hyperlinking

Automatic Categorization

Automatic Query Guidance

Automatic Taxonomy Generation

Real-Time Sentiment Analysis

Automatic Clustering

Cluster Visualization

Directed Navigation

Eduction

Personalized Agents

Profiling

Social Search

Conceptual Retrieval

Full XML Support

Autonomy Business Console

IDOL Echo

IDOL Server has over 500 functions for processing information. On average, 5 to 15 functions are used per application. The key IDOL Server functions listed here are some of the most powerful and widely deployed.

Automatic Hyperlinking

IDOL allows manual and fully automatic linking between related pieces of information, regardless of their format. The concepts in a document can be linked automatically to those in another file. They can also be linked to related concepts within video or voice mail. Hyperlinks are generated in real-time at the moment a document is viewed, removing the need for any manual intervention and ensuring they are constantly up-to-date.

Automatic Categorization

IDOL's categorization features allow users to derive precise categories through concepts found within unstructured text, ensuring that all data is classified in the correct context with the utmost accuracy. Existing legacy taxonomies can be either maintained or enriched with contextual understanding. The information that IDOL has automatically aggregated and categorized is presented to users in the form of channels.

Automatic Query Guidance

IDOL's Automatic Query Guidance (AQG) automatically provides keyword-oriented users with the suggestions they need to find the most relevant information. AQG identifies the different meanings of a term by dynamically clustering the results into their most relevant groupings. Users are able to guide the engine simply by clicking on the context that most applies to what they are really interested in. AQG is unique because the guidance suggestions are created dynamically based on the query and the time of query. No manual process is required to teach, train or configure the software to provide suggestions.

## Automatic Taxonomy Generation

By understanding the information in your enterprise, IDOL automatically generates taxonomies and instantly organizes the data into a familiar child/parent taxonomical structure. IDOL automatically identifies, names and creates each node based on an understanding of the concepts within the data set as a whole. IDOL eliminates the need for any human intervention, preventing the errors and inconsistency that are frequently associated with manual categorization.

Real-Time Sentiment Analysis

By analyzing the structures and meaning of language, IDOL determines the positive and negative characteristics of each piece of information and creates relevant classification systems. IDOL can determine the degree to which a sentiment is positive, negative or neutral for the entire content or a segment of the content. In addition, administrators can apply multiple tagging functions and specific threshold cut-offs to determine the sensitivity of sentiment analysis.

Automatic Clustering

Organizations can analyze large sets of documents and even user profiles and automatically identify inherent themes or information clusters. IDOL even clusters the unstructured content exchanged in email, telephone conversations and instant messages. IDOL uses the most advanced heuristics, such as quantum clustering, to form these conceptual groupings. Some applications of this feature include: "What's Hot" clusters that automatically detect burning topics in an organization's information assets and "breaking news" clusters that alert users in real-time to new areas of information or individual interest.
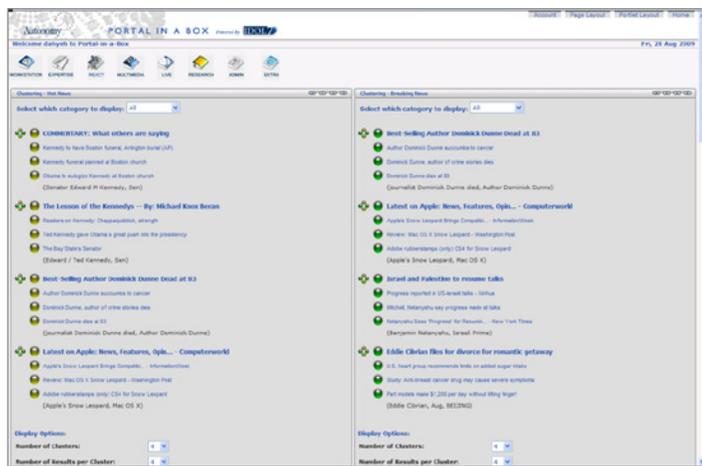

*Automatic clustering of breaking news and hot news*

Cluster Visualization

IDOL provides three intuitive Java-based user interfaces to make clusters visible.

Spectograph

This user interface displays the relationships between clusters in successive periods and sets of data. Clusters are presented as a JSP-based spectrograph, whereby the x-axis represents information over a given time period, while the y-axis represents the range of concepts defined within the knowledge base.
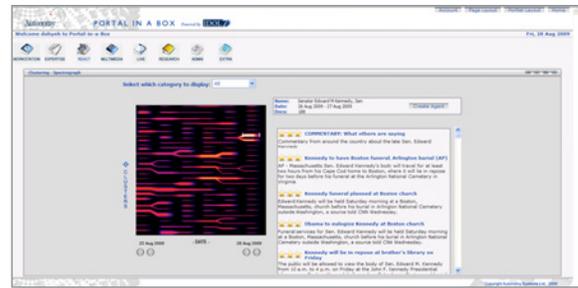
Automatic clustering of communications visualized through the Spectograph pane
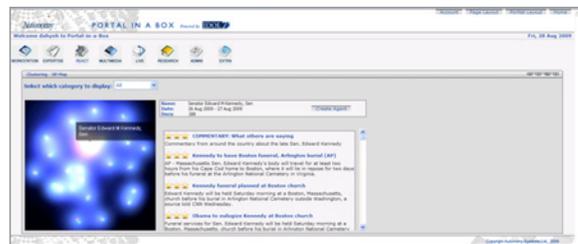
2D Cluster Map
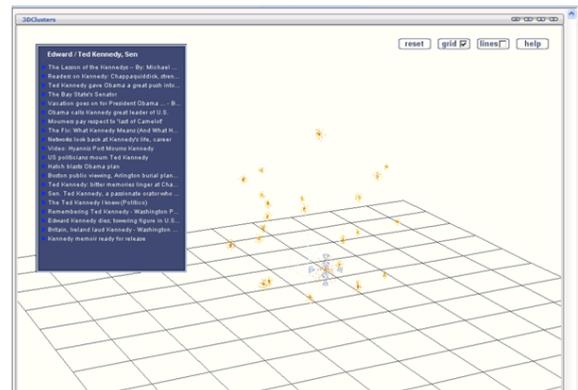
3D Cluster Map

2D and 3D Cluster Maps

The 2D and 3D Cluster Maps are used to identify conceptual similarities and differences between clusters.

Directed Navigation

IDOL dynamically presents navigable parameters based on document properties and all metadata attributes - explicitly defined, automatically generated and even fields that were not previously used. IDOL's Directed Navigation offers a way for organizations to integrate valuable information from structured databases with unstructured content. Combining information found within structured fields such as product name, size, price or manufacturer, with unstructured content such as product description, allows users to search and find products quickly and easily.

Eduction

IDOL goes beyond traditional entity extraction and enriches the extracted data based on the knowledge already held within the organization. Not only can IDOL extract specific data such as organizations, people, places and figures, it can conceptually relate this information to other data held within the enterprise and automatically form relationships. Eduction comes with out-of-the-box entities including: names of commercial organizations, people, places, postal/Internet addresses, phone numbers, dates, times, numbers, prices, Social Security numbers, job titles and holidays.

Personalized Agents

IDOL allows users to set up Agents to monitor information 24x7 on specific topics from a range of data sources. An Agent represents a user's persistent interests and can be defined or trained either explicitly with a natural language description or Boolean expression. Most powerfully, an Agent can be trained or re-trained by example, simply by being shown a document, video, or verbal conversation that matches a user's interests. The Agent will then learn the concepts within the example and define itself accordingly. Once an Agent has been created it will monitor in real-time the changing information within the index, instantly alerting the user to anything new that matches the training.
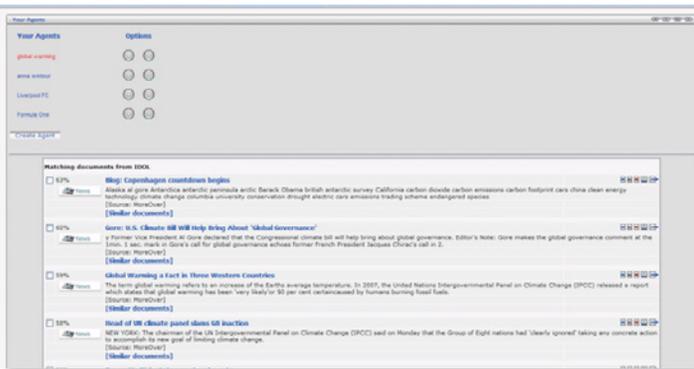
Personalized Agents
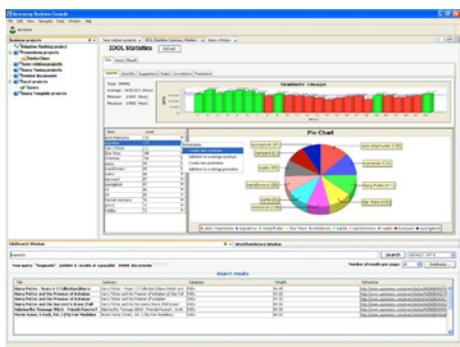
Autonomy Business Console

Profiling

IDOL accurately understands individuals' interests based on their browsing, content consumption and content contribution. Generating a multifaceted conceptual profile of each user based on both explicit (agents) and implicit profiles (click-through and submission), IDOL forms a very current understanding of users' interests and personalizes the relevancy model to deliver intent-based search results.

Social Search

IDOL provides a rich participatory platform for capturing unstructured, tacit knowledge created from Web 2.0, connecting people to related communities, projects and processes. Users can contribute and share tags, comments and votes, as well as share search folders and knowledge with their colleagues. All user activity, both explicit and implicit, can be incorporated to influence relevancy calculation.

Conceptual Retrieval

Built on an innovative pattern-recognition technology as demonstrated by over 170 patents, IDOL offers higher degrees of accuracy and sophistication using a scalable technology that recognizes concepts. Since conceptual search can find information based on words not located in the document, it does not fall prey to the limitations of legacy methods.

Full XML Support

IDOL allows organizations to eliminate the manual inefficiencies introduced by XML tags by understanding the content and purpose of either the tag itself, related information, or both. IDOL can automatically insert XML tags and links into documents based on the concepts contained in the information. IDOL's meaning-based technology also provides an infrastructure for complete and automatic interoperability between applications using different XML tagging rules.

Autonomy Business Console

The Autonomy Business Console (ABC) empowers business users to react quickly to changing usage patterns and provide highly targeted results to end-users without any programming. ABC provides a robust set of widgets, wizards and tools to simplify the management of IDOL functions, including promotions, synonym management and relevance boosting.

IDOL Echo

IDOL Echo allows fully auditable and accountable monitoring of information use. The Echo module enables the enterprise to forensically account, track and trace each piece of data that enters, leaves, is born or dies within the organization. Echo follows a file's path and history and can report on who or what the asset has influenced. It can not only follow a traffic pattern such as the path of an email attachment or voice mail (i.e. the what and when of who read, heard, forwarded and retained it), but also detect the influence of its content (i.e. who in the enterprise has taken, re-purposed, or even plagiarized).

IDOL Architecture

The basic architecture requirements of a pan-enterprise search platform include:

FRCP-Compliant Search

The Federal Rules of Civil Procedure (FRCP) render all relevant Electronically Stored Information (ESI) discoverable, regardless of format or location. To be FRCP-compliant, pan-enterprise search platforms:

- *Need to search ALL repositories*
- *Cannot perform "jump-out" - a sleight of hand used by some vendors to feign performance, in which the search engine stops looking across an index as soon as it is believed a large enough group of results has been assembled*
- *Need to produce auditable results - hence ALL data needs to be searched fully*
- *Need to be able to pass results to a hold function - ensuring that relevant ESI is not altered in any way or deleted*

Scalability and Performance

IDOL scales to support the largest enterprise-wide and portal deployments in the world, with presence in virtually every vertical market. Since IDOL's scalability is based on its modular architecture, it can handle massive amounts of data on commodity dual-CPU servers. IDOL delivers linear scalability through a multi-threaded, multi-instance approach with load-balancing to intelligently distribute the indexing and query workload.

Mapped Security

The biggest single constraint on scalability within enterprise applications is the ability to manage entitlement checks in a scalable manner. IDOL stores security information in its native form directly in the kernel of the engine itself, with automatic updates to keep the security data current. This sharply contrasts with other security models that store security information in the original repositories, requiring communication between the search engine and the underlying repository for every potential result at the time of query.

Global Language Support

The need to manage content in varied languages has never been more acute for today's global enterprise. IDOL's language-independent technology develops a statistical understanding of the patterns of any language using sophisticated probabilistic modeling and pattern-matching techniques. It currently supports over 100 languages and provides cross-lingual search. Autonomy also provides optional language packs to further enhance localization, including stemming, stoplists, transliteration, multiple encoding support and term decomposition.
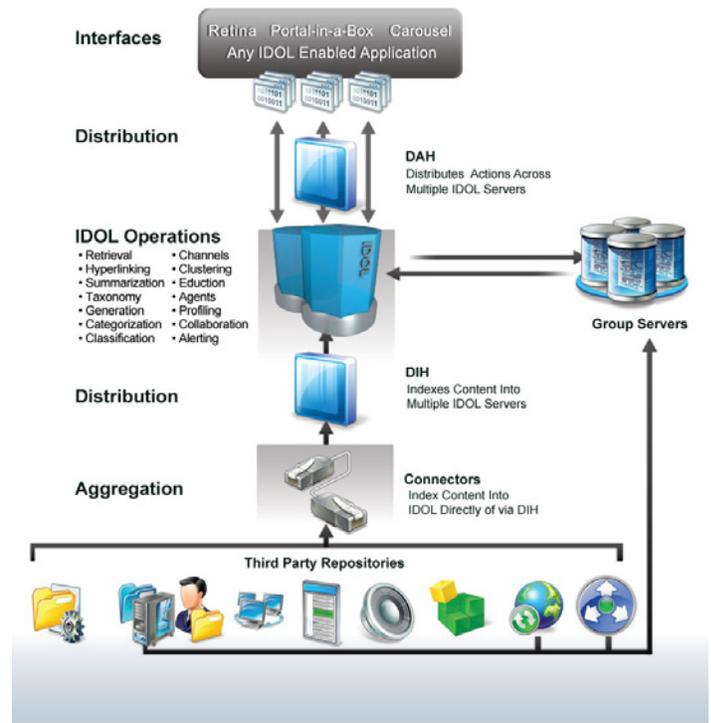
Requirements

Platforms Supported:

- Microsoft Windows 2000, XP, 2003 and Vista
- Linux Redhat AS 3.0, 4.0, 4.5, 5.0 and SUSE Enterprise 8.0, 9.0, 10.0
- Sun Solaris for SPARC version 8,9,10
- Sun Solaris for Intel version 9,10
- AIX version 5.2,5.3
- HP-UX for PA-RISC version 11i
- HP-UX for Itanium version 11i

Other POSIX compliant UNIX versions are available on request.

Minimum Server Specifications:

Dual Intel Xeon 3.0 Ghz
4 GB RAM
100 GB hard disk recommended

For specific sizing requirements, please consult the Autonomy Sizing Service.

www.autonomy.com