Intelligent Universal Search

Autonomy Intelligent Universal Search (IUS) empowers the enterprise with capabilities that reveal relevant business concepts, entities and experts contained in all forms of human friendly information. IUS can search across 1000+ file formats including rich media, 400+ repositories and 150+ languages. Powered by Autonomy IDOL (Intelligent Data Operating Layer), IUS understands and identifies the context and concepts within structured and unstructured information and delivers alerts based on news, social media or industry-specific data sources, without relying on manual tagging or keywords. Inside the enterprise, IUS can search intranets, wikis, email stores, SharePoint sites and content management systems for specific information or for experts who may be working on similar projects. Social features in IUS allow users to collaborate on content and provide ratings that add value and richness to enterprise content. Advanced capabilities such as automatic categorization, analytics and visualization tools allow users to hold conversations with their unstructured data, which uncovers unstated relationships.

Leveraging Human Information

From the time of the very first computers, the ability to process human information has posed a considerable challenge. As a result, a tremendous amount of manual effort has been put into sorting, distilling and tagging unstructured information. However, structuring information in this way is not a viable solution, not only because of the incredible amount of manpower required, but also because the richness and subtleties of the information are lost in the process. Today, approaches that leverage more intelligent and advanced technologies are transforming the way information is managed and used.

IUS understands the richness of human information by forming a conceptual and contextual understanding of any piece of data, including social media, blogs, IMs, media files, email, documents and web pages. IUS accepts search input in simple, human-friendly forms like sample data and returns results to conceptually related documents ranked by relevance to help people utilize their company’s wealth of knowledge. As a result, IUS increases the speed of information retrieval, while ranking quality and user satisfaction by applying what Autonomy calls Meaning Based Computing.

Autonomy IUS enables organizations to:
- Find the right information quickly and easily across all data repositories
- Locate and make connections with experts throughout the organization
- Identify experience and tacit relationships that can be leveraged to develop new business
- Alert users to new content that matches profiles or saved searches
- Automatically classify and categorize information and deliver a rapidly deployable, packaged and advanced retrieval application
User Experience Designed for Enterprise Use

Designed specifically for users, the intuitive, rich and uncluttered IUS interface offers innovative visualization techniques like concept clouds and clustering, and can automatically group conceptually related, unstructured information. As documents are previewed via the web interface, users are immediately presented with links to documents that are conceptually similar to the one being previewed. IUS combines structured and unstructured information to reveal rich, actionable information in a single interface. For example, internal and external corporate information can be viewed in a single “mash-up” that groups all related information from expertise banks, internal content management systems and documents. IUS also offers capabilities to hyperlink and further refine this information.

Rich Media Search of Video and Audio

Video is the only medium that is capable of capturing the complete spectrum of human information including gestures, tone and facial expressions. By its nature, video carries more meaning than can be narrowed down into a few lines of text. Because it requires a considerable number of hours to tag this type of content, many rich media assets within the enterprise, such as training videos or recorded WebEx sessions, remain unused.

Autonomy’s technology delves deep within the video file itself rather than relying on human-defined metatags which are subjective and limited in scope. Using advanced image and audio analysis engines that watch, listen to and read video signals in real time, IUS’s Rich Media Module can find information with pinpoint accuracy. Video can be automatically cross-referenced with other forms of information such as PowerPoint presentations, Word documents and web pages.

To enable analysis of audio files or video audio tracks, IUS’s Rich Media Module is able to understand the meaning of live or recorded communication. Sophisticated audio recognition and analysis technology process spoken interactions based on their conceptual content, not just by the way they sound. This approach provides the most accurate way of recognizing and finding speech because the technology understands what is being said.

Micro-Blogging and Social Tagging

IUS allows employees to collaborate and comment on content through the use of micro-blogging and social tagging. Users can rate the helpfulness of a document using a sliding scale. This information can be used to automatically adjust relevancy and allow users to see and learn from content that is most valued by their colleagues. Users can tag content with keywords that become part of the document’s metadata and fully searchable by the rest of the user community. Free-text comments can also be made regarding the concepts and context of content as well as on search results in a micro-blog environment.

People Search

As organizations grow larger and expand geographically, it is increasingly important for managers to quickly identify colleagues with expertise in specific areas that are capable of helping with an existing offering or with new business. Traditionally, legacy search technology makes it difficult to locate these colleagues—leaving vast resources of tacit knowledge and experience untapped. To address this challenge, IUS builds communities of expertise to ensure that subject matter or industry experts are easily located throughout the enterprise to promote collaboration and fuel innovation. By forming a conceptual understanding of user interaction with information as it is consumed and created, IUS identifies tacit knowledge automatically and in context. IUS builds a conceptual understanding of the relationships between experts and the content with which they interact, automatically clustering similar people and resources into related groups. It can also automatically recommend an expert based on an understanding of the content that users consume and create—across all formats including email, IM, and voice and video data.
Social Media Search
IUS provides comprehensive access to all social media sources, enabling conceptual search for relevant business topics. IUS automatically filters out the mindless chatter of social media and only alerts people to relevant information. IUS’s extensive set of social media connectors include: RSS, Twitter, Facebook, Yelp, LinkedIn, TripAdvisor, Yahoo! Finance, C|Net Reviews, WebMD, IMDb, Kbb.com and Epicurious.

Connectivity as a Differentiator
The seamless connectivity between IUS and Content Management Systems (CMS) such as iManage WorkSite, SharePoint, Documentum, FileNet and other document management systems is a key differentiator that provides significant benefits for the end user and the organization. However, this connectivity is not easily accomplished. While most vendors claim a number of connectors or SDKs to develop more, these options do not equate to the ability to truly connect to all data types. Further, sophisticated technology is required to maintain existing enterprise security models while also supporting legacy file systems and future emerging standards.

Autonomy uniquely owns its own technology for accessing 400+ file repositories and 1,000+ individual file formats and is the clear and acknowledged OEM leader in this area, providing over 400 partners with technology that they integrate into their own solutions.

Intelligent Universal Search is Powered by IDOL
Autonomy’s IDOL platform is highly secure and used by the largest and most protected intelligence organizations in the world. With over 20,000 customers across the globe, IDOL is massively scalable with the largest installation exceeding 10 billion documents and delivering sub-second search times. The IDOL platform powers fast, comprehensive and secure search via the easy-to-use and intuitive IUS interface.

Security
The world’s largest and most secure intelligence organizations have deployed Autonomy’s Intellectual Asset Protection System (IAS) to safeguard their most sensitive information assets. Autonomy provides all aspects of security management, including front-end user authentication, back-end entitlement checking and secure encrypted communication between the IDOL Server and its client applications, with 128-bit Block Tiny Encryption Algorithm (BTEA). IDOL’s mapped security model is the only empirically proven index security model that scales in the enterprise.

Since IDOL’s architecture is inherently modular by design, it requires multiple sub-systems to communicate with each other, often across insecure networks. All communication between these processes may be encrypted (Secure Sockets Layer), to prevent packet sniffers from breaking past a firewall and reading the content of traffic between IDOL modules. All of the system’s modules are capable of operating in a secure communications mode, with minimal processing overhead, to provide 128-bit encryption protection. Additionally, IDOL can leverage SSL for both aggregation and querying of content, including access to SSL encrypted sites.

The following modules are available for extending the base IUS application:
• Social Media Package
• Classification Package
• Community Package
• Visualization / Advanced Analytics Package
• Rich Media Package

IUS Overcomes the Limitations of Other Approaches
IUS supports all legacy methods described below. However, Autonomy recognizes the limitations of these approaches in enterprise scenarios and uniquely offers conceptual retrieval to provide the most accurate and complete search results, requiring minimal manual intervention.

Keyword and Boolean Searches: Returns only those documents that contain the terms queried. This method is heavily reliant on user skill and adeptness with Boolean operators. It ignores the context in which the keywords were found. While weighting keywords only mitigates this issue, it does not remove the critical defect.

Page Rank: Determines the “importance” of a web page based on the number of pages linked to it. This ranking is then used with a keyword entered by a user to retrieve the most relevant results. The flaw with this approach is that search results rely on manually added hyperlinks which are rarely maintained in enterprise content and not always a good indicator of a file’s importance.

Parsing and Natural Language Analysis: Semantic analysis uses grammar and lexicons to explicitly understand textual information. Due to the inherent complexity of language, it cannot handle ambiguity. It can’t determine the relative importance of ideas. It is language-specific and grammar-based, which excludes slang or grammatically incorrect construction such as often cryptic social media messages, abbreviations and acronyms.

Meaning Based Computing: Unlike the above approaches, IUS has the ability to automate operations based on the context and concepts within content. Intelligent Universal Search’s ability to understand the meaning of content removes the dependence on manual efforts and metadata while performing relevant searches.
Language
In addition to accessing and processing almost any type of content, IDOL can connect to content in 150 languages, including double-byte languages. In the same way that IDOL builds a mathematical model of rich media content without relying on specific words, this same approach is applied to understanding languages; the individual words are unimportant. Autonomy can easily add additional language models simply by matching set linguistic data to indexed content.

Language agnosticism makes it possible for users to search in any language, enter a search term in one language and return results in another language and even provide a conceptual summary of content in other languages.

Scalability and Performance
The management of structured and unstructured content requires a platform that can meet the most rigorous performance requirements and easily scale to match business needs. 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, distributed architecture, it can handle massive amounts of data on commodity, dual-CPU servers. For instance, only a few hundred entry level enterprise machines are required to support ChoicePoint's 10 billion record footprint. By comparison, a competitor uses 150,000 machines to handle the same amount of data.

This enhanced scalability results in hardware cost savings as well as the ability to address larger volumes of content. In addition to scaling extremely well on commodity servers, IDOL's flexible architecture can further improve performance by taking full advantage of massive parallelism, SMP processing capabilities, 64-bit environments, software platforms (such as Solaris 10, Linux 64 and Win64), distributed server farms and all common forms of external disk arrays (such as NAS and SAN). This flexibility also allows organizations to leverage one or a combination of these different environments.

Flexible Deployment
IUS may be deployed on-site, or as an on-demand cloud computing offering. Many organizations have found that cloud deployment helps reduce the cost and traditional overhead of on-site software solutions. IUS is the most scalable enterprise search solution available on the market today and its highly flexible configuration options enable organizations to meet the most stringent business and technological needs.

About Autonomy
Autonomy Corporation, an HP Company, is a global leader in software that processes human information, or unstructured data, including social media, email, video, audio, text and web pages, etc. Autonomy’s technology manages and extracts meaning in real time from all forms of information, both unstructured and structured, enabling companies to leverage their data assets. Autonomy’s product portfolio helps power companies through enterprise search analytics, business process management and OEM operations. Autonomy also offers information governance solutions in areas such as eDiscovery, content management and compliance, as well as marketing solutions that help companies grow revenue, such as web content management, online marketing optimization and rich media management.

Autonomy’s solutions are used by more than 25,000 customers including 87 of the Fortune 100, 10 of the top 10 financial services firms, 75% of the global 100 law firms, 9 of the top 10 pharmaceutical companies and many government agencies. Over 400 of the world’s leading technology companies embed Autonomy’s technology in their products. Autonomy also owns the largest private cloud of diverse data, with 31 Petabytes of information.

Please visit www.autonomy.com to find out more.

“Autonomy outperforms all vendors on the technology front...scoring the category maximum in every assessment. In three categories; namely search and query capabilities, visualization and navigation capabilities, and interoperability and integration; the solution’s functionality is deemed to be the best-in-class.”