

Use Technology to Streamline Processing, Review and Analysis

In the age of globalization, documents and emails created or consumed in different countries and different languages can be highly relevant to legal investigations. Litigation often spans national boundaries, and legal teams are flooded with hundreds of thousands of documents which may contain languages other than English. All of these documents need to be identified, processed, evaluated, and analyzed.

The growing importance of multi-language discovery creates new challenges for attorneys and their technology partners. E-Discovery is already complex, and that complexity grows by orders of magnitude when it involves documents in different languages, writing systems, and character sets. Yet the need for meticulous investigation is as crucial as ever for lawyers to provide the best outcomes for their clients.

What is Your Multilingual E-Discovery Plan?

Good News for Lawyers and Their Technology Partners

Basis Technology helps the legal community meet its multilingual discovery challenges head-on. We provide the core technologies necessary to power comprehensive electronic discovery solutions that uncover evidence buried in terabytes of unstructured multilingual text—accurately, quickly, and cost-effectively. We do this by using the most advanced linguistics software in the industry—software found at the core of virtually all leading multilingual search engines and information retrieval applications.

Our multilingual E-Discovery solutions are based on Rosette®, a linguistics platform proven in hundreds of commercial and government environments. Interoperable Rosette software components are configured as building blocks for multilingual E-Discovery, working seamlessly within discovery workflows and information retrieval applications while identifying and handling many different languages, character sets, and data sources.

Make Your Discovery Application Multilingual

Our industry-leading linguistics software is complemented by ease of integration within data mining, reviewing, search, and other discovery applications used by legal teams. By plugging the Rosette API into an application, users get instant access to unique E-Discovery tools covering major European, Asian, and Middle Eastern languages. For legal professionals, it means the ability to examine multilingual text with unparalleled accuracy and efficiency.

Benefits

- Finds evidence buried in terabytes of unstructured multilingual text
- Identify document language and enable search
- Supports over 55 languages, including the major Asian, European, and Middle Eastern language Multilingual E-Discovery in One Complete Package

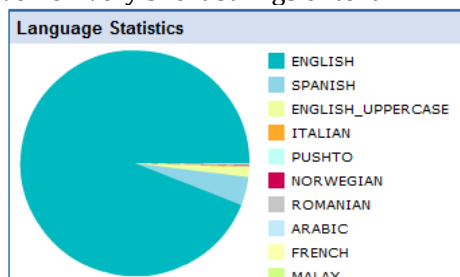
Multilingual E-Discovery in One Complete Package

Step 1: Identify the language(s) and encoding in a document, and convert to Unicode

Rosette Component: Rosette Language Identifier (RLI)

RLI identifies the language(s) of a document and its encoding so content can be filtered and processed. The text is converted to Unicode so that discovery and information retrieval applications can display and process the text. Using a feature called the Rosette Language Boundary Locator (RLBL), documents containing more than one language are segmented into language regions that can be routed to separate

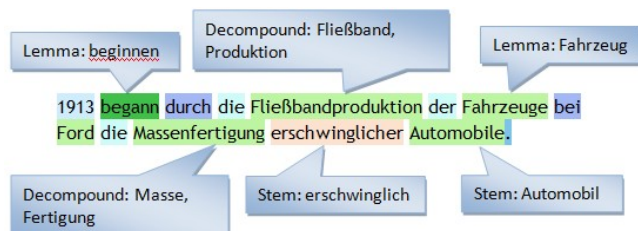
processes. RLI identifies 55 languages with extreme accuracy, even on very short strings of text.



Step 2: Apply linguistic intelligence to identify word forms, parts of speech, and sentence structure

Rosette Component: Rosette Base Linguistics (RBL)

RBL examines documents and performs a complete morphological analysis so text can be accurately filtered, analyzed, and searched. RBL identifies word breaks, parts of speech, sentence boundaries, compound words, and other linguistic components within a document, in European, Asian, and Middle Eastern languages. The technology and linguistic data in RBL results from over 10 years of development and use in web and enterprise search engines.



Step 3: Extract the items of interest (including those you didn't know about)

Rosette Component: Rosette Entity Extractor (REX)

REX sifts through unstructured text and identifies people, places, dates, and other items that establish the true meaning of a document for further analysis. REX locates generic terms as well as custom entities such as specific names, phone numbers, and email addresses. Statistical modeling helps determine if an entity resides within a document, rather than simply referring to a list of possibilities and risk overlooking a variation. The result is entity extraction technology that lets you find what you know—and also what you didn't know.

Named Entity (# instances)	Text
IDENTIFIER:MONEY 1	The fictitious Indian philosopher Abhay Urjavaha Revati was born on January 3, 1624 at 4:00 PM , in Aranasi , a town in the Nagpur province of India . His father, Mr Krupali Setar , a theologian of some repute, together with his mother, Mrs. Sushanti Singh , worked at the Central Bank of India . Abhay was the vice-president of the United Humanist Philosophers Association , which is located in Thiruvananthapuram .
IDENTIFIER:NUMBER 1	
IDENTIFIER:URL 1	
LOCATION 8	
NATIONALITY 1	
ORGANIZATION 2	
PERSON 9	He married his wife Madam Savita Narvas on Dec. 5th, 1657 in Mumbai . They had several children together, all of whom were born in Southeast Asia , with the exception of their youngest daughter Kishana who was born in Boston, Massachusetts .
RELIGION 1	
TEMPORAL:DATE 2	
TEMPORAL:TIME 2	
TITLE 1	Abhay's Hindu background, the influence of his parents, and his favorite teacher, Professor Sungul Kahani all contributed his great work in religious philosophy.

His daughter **Kishana** took **two years** off from work to dedicate a memorial website to her father. The website (www.revati.com) raised over **\$5 million** for children's hospitals across the country. About **one million** went to **Deodhar Children's Hospital**.

Make multilingual e-Discovery work for you. Rosette finds and solves the language problems in your documents.

© 2010 Basis Technology Corporation. All rights reserved. "Basis Technology", "Rosette", and "We Put the World in the World Wide Web" are registered trademarks. All other company and product names mentioned are trademarks or registered trademarks of their respective owners. (2010-03-10)



Browse
www.basistech.com

Write
info@basistech.com

Call
617-386-2090 or 800-697-2062

Boston
One Alewife Center, Cambridge, MA 02140

San Francisco
171 Second Street, San Francisco, CA 94105

Washington, D.C.
13800 Coppermine Road, Herndon, VA 20171