

Rosette Entity Extractor

Locate and Extract Names, Places, Organizations, and Other Entities from Unstructured Text

Rosette® Entity Extractor (REX) enables search and text mining applications to automatically locate concepts, which cannot easily be found using only lists or rules, in unstructured text. REX uses advanced linguistics to find **entities** — names, places, dates, identifiers, and other unique words or phrases — that establish the real meaning in text.

Entity extraction enables the classification, management, and analysis of large volumes of unstructured text coming from sources such as e-mail, electronic documents, and the Internet. REX can be used both to locate concepts in unstructured data and as a step in preparing such data for structured analysis, a process known as “metadata annotation”.

REX locates generic terms such as “Vice President” and “Southeast Asia” as well as specific references such as “President Barack Obama” and “January 20, 2009.” These features enable the identification and extraction of important information from documents, for further analysis by other applications.

REX employs three types of entity detection technology — statistical modeling, regular expressions, and lists — to achieve industry-leading accuracy.

The REX Application Programming Interface (API) is an extension of the proven Rosette API, designed to quickly and efficiently integrate into a wide variety of applications. Unlike competing extractors, REX integrates seamlessly with the Rosette Name Indexer and Rosette Name Translator, enabling fully-automatic, cross-script resolution and translation of extracted foreign names.

APPLICATIONS

REX brings a deep level of linguistic analysis to software systems designed for information retrieval, CRM, computer forensics, intelligence analysis, competitive intelligence, social media monitoring, and countless other applications. Its functions are integral to taxonomy and categorization applications; text mining; trend analysis; and relationship extraction engines.

Kofi Atta Annan GCMG is a Ghanaian diplomat who served as the seventh Secretary General of the United Nations from January 1, 1997 to January 1, 2007, serving two five-year terms. Annan was the co-recipient of the Nobel Peace Prize in October 2001.

Kofi Annan was born in April 8, 1938 to Victoria and Henry Reginald Annan in Kumasi, Ghana. He is a twin, an occurrence that is regarded as special in Ghanaian culture. Efuwa Atta, his twin sister, shares the same middle name, which means 'twin'. As with most Akan names, his first name indicates the day of the week he was born: Kofi denotes a boy born on a Friday. The name Annan can indicate that a child was the fourth in the family, but in his family it was simply a name which Annan inherited from his parents.

In 1962, Annan started working as a Budget Officer for the World Health Organization, an agency of the United Nations. From 1974 to 1976, he worked as the Director of Tourism in Ghana. Annan then returned to work for the United Nations as an Assistant Secretary General in three consecutive positions.

REX is trained and tuned for each specific language, and is customizable to meet the particular entity needs of vertical applications such as health, medical, life sciences, financial, and manufacturing.

MIXED-LANGUAGE DOCUMENTS

REX is capable of analyzing documents which may be written in more than one language. By utilizing a technique known as “language boundary location”, REX is able to apply the correct language-specific linguistics to each distinct language region of a mixed-language document, thereby achieving greater accuracy on “real world” data.

SYSTEM SPECIFICATIONS

REX is portable and highly scalable, running on platforms ranging from laptop PCs to multi-CPU servers processing thousands of documents per second.

A fully-documented API is provided, and may be accessed from applications written in C, C++, Java, and other languages. A command-line interface is also available for testing purposes.

FOR MORE INFORMATION

For more information, please visit www.basistech.com. To request an evaluation copy, please write to info2010@basistech.com or call us at 617-386-2090 or 800-697-2062.

PREDEFINED ENTITY TYPES

REX natively supports the following entity types, as well as fundamental linguistic elements, such as sentence boundaries, parts of speech, and noun phrases. User-defined entity types (such as SKU numbers) are also supported.

- Person
- Organization
- Location
- Geopolitical Entity (GPE)
- Latitude/longitude coordinate
- Universal Transverse Mercator (UTM) coordinate
- Personal ID number (e.g. SSN)
- Nationality
- Email address
- Uniform Resource Locator (URL)
- Phone number
- Credit card number
- Percentage
- Date
- Time
- Facility

AVAILABLE LANGUAGES

REX natively supports the following languages. Additional languages are available via custom training.

Arabic	English	Japanese
Chinese (Simplified)	Farsi (Persian)	Korean
Chinese (Traditional)	French	Russian
Dutch	German	Spanish
Dari	Italian	Urdu

SYSTEM PLATFORMS SUPPORTED

Software development kits (SDKs) are available for popular architectures, operating systems, and development environments. Support for platforms not listed below is available by contacting your sales representative.

- AIX 5.3/6.1, PPC
- HP-UX 11i, PA-RISC/IA64
- Linux CentOS 4.x/5.x, IA32/AMD64 (GCC 4.1/4.2)
- Linux Debian 5.0, IA32/AMD64
- Linux Red Hat 3.0, IA32/AMD64 (GCC 3.2/3.4/4.0)
- Linux Red Hat 4.0, IA32/AMD64 (GCC 3.4/4.0)
- Linux Red Hat 5.0, IA32/AMD64 (GCC 4.1/4.2)
- Linux Ubuntu 8.04/9.x/10.x, IA32/AMD64
- MacOS (GCC 4.0)
- Solaris 9, SPARC32/64
- Solaris 10, SPARC32/64
- Solaris 10, IA32/AMD64
- Windows XP/Vista/7, IA32 (MSVC 7.1)
- Windows XP/Vista/7, IA32/AMD64 (MSVC 8.0)

(서울-연합뉴스) 이승관 김경희 기자 = 청와대가 최근 **한나라당** 시, 도당 여성 당직자를 초청, 간담회를 갖는 과정에서 복당 의원들을 일괄 배제해 "복당파는 뒷전"이라는 불만이 제기되고 있다.

한나라당 한 관계자는 **27일 연합뉴스와의** 통화에서 "지난달부터 **총청**, 호남권 등을 시작으로 **대통령** 부인 **김윤옥** 여사가 각 시, 도당별 여성 당직자들을 **청와대에** 초청해 다과회를 갖고 있다"면서 "당원협의회별로 여성 당직자들을 선별하다보니, **복당 의원** 지역의 경우는 총선에서 **한나라당** 공천을 받은 당원위원장이 추천권을 갖게 됐다"고 설명했다.

문제는 지난달 **9일 총청권을** 시작으로 두 달 가까이 계속돼 온 행사 자체가 복당 의원들에게 양해는 물론 통보조차 되지 않았다는 것.

게다가 당원위원장이 대부분 남성이어서 그 부인들이 여성 당직자들을 인솔해 **청와대를** 방문해야 하고, 그러다 보니 같은 **한나라당** 국회의원 부인 가운데에도 복당 의원들의 부인만 행사에서 소외되는 상황까지 발생하고 있다.

이제까지는 **총청과 호남, 강원, 인천, 제주** 등 복당 인사들이 그다지 많지 않은 지역을 중심으로 행사가 진행됐기 때문에 특별한 잡음이 발생하지 않았지만, 복당 인사들이 상당수를 차지하는 **부산과 대구 지역 간담회**가 오는 **29일로** 다가오면서 이와 관련한 불만과 문제제기가 여기 저기서 터지고 있다.

صربيا تسحب سفيرها من واشنطن بعد اعترافها باستقلال كوسوفو
GMT+04:00 - 18/02/08, 2358

(CNN)-- في رد فعل فوري على إعلان وزارة الخارجية الأمريكية اعتراف الولايات المتحدة بكوسوفو كدولة مستقلة عن صربيا، وعموما إقامة علاقات دبلوماسية معها، قررت حكومة بلغراد سحب سفيرها من واشنطن.

جاءت هذه الخطوة، التي أعلنها وزير الخارجية الصربي، فوك جيرميتش الاثنين، بعد قليل من إعلان وزارة الخارجية الأمريكية، كوندوليزا رايس، في بيان مكتوب، أن الولايات المتحدة تعترف بدولة كوسوفو، كأحد دولة مستقلة في أوروبا.

وجاء في البيان الذي صدر بواشنطن: "نتقدم بالتهنئة إلى شعب كوسوفو بهذه المناسبة التاريخية"، مضيفة قولها إن "الرئيس بوش قرر الاستجابة على الفور، لطلب من حكومة كوسوفو، بشأن إقامة

朝鲜外务省发言人11月1日在平壤宣布，朝鲜将重返六方会谈，但前提条件是朝鲜与美国在六方会谈框架内讨论解除美国对朝鲜金融制裁问题。针对朝鲜方面的动向，各方均表示欢迎。美联社11月1日报道说：“长期以来一直拒绝与平壤进行直接对话的美国总统布什认为，各方达成一致、同意恢复六方会谈应归功于中国的斡旋。

朝鲜中央通讯社11月1日报道，朝鲜外务省发言人当天在接受该社记者采访时，对10月31日北京举行的双边及多边会晤问题作出回答。朝鲜最近对美国的金融制裁采取了自卫性措施。10月31日在北京举行了双边及多边会晤。会晤讨论了寻求恢复六方会谈途径的事宜。朝鲜决定重返六方会谈，前提是朝鲜与美国将在六方会谈框架内就解除金融制裁问题展开讨论并作出安排。在朝鲜宣布重返六方会谈前，中国外交部10月31日发布消息说，根据中方建议，六方会谈团长在北京举行非正式会晤。三方就继续推进六方会谈进程坦率地交换了意见。三方一致同意在六方方便的近期举行六方会谈。美国总统布什10月31日对新闻界说，他对中国、美国和朝鲜三方同意于近期恢复六方会谈表示欢迎，对朝鲜核问题取得进展感到非常高兴。此外，参加六方会谈另外三方韩国、俄罗斯、日本以及联合国秘书长科菲·安南都对此作出积极回应。



One Alewife Center
Cambridge, MA 02140

171 Second Street
San Francisco, CA 94105

13800 Coppermine Road
Herndon, VA 20171

9-6 Nibancho, Chiyoda-ku
Tokyo 102-0084