

Recognize customers, citizens, criminals,
across multiple cultural variations of name data



IBM **Information Management** software

IBM Global Name Recognition Technology



IBM provide state-of-the-art global name recognition technology for applications in mission critical situations including security, intelligence, law enforcement, fraud, financial compliance, and customer relationship management – when every name counts.

Highlights

- **Recognize customers, citizens, criminals, and other risks and threats across multiple cultural variations of name data**
- **See Beyond Name Order:**
Ovtsyuk, Nadezhda Ivanovna
- **See Beyond Multiple Titles:**
Dr., Rev, Haj, Sri., Col
- **See Beyond Cultural Variations:**
Jörg, Egor, Juergen
- **See Beyond Multiple Prefixes:**
Abdul, Fitz, O', De La,
- **See Beyond Transposition Errors:**
Dena, De"e"na, De"a"ne
- **See Beyond Nicknames:**
Hammed, Mogs, Skip

Advanced Name Recognition Technology

Names are often overlooked as miniature databases of knowledge, but that's precisely what they are. In our global society, the ability to accurately recognize and manage the building blocks of an individuals' name can provide the key to recognize identity across cultures, genders, and meanings. The ability to understand names across culture is a critical and difficult component of recognizing and managing threats and opportunities.

For example, a name is entered in one database with the surname "Rodgers," and in a different database as "Rogers." A person's name is recorded as "Dayton," but should actually be spelled "Deighton." The problem is greatly compounded with names originating outside North America. The same Chinese person may have one set of information recorded under the surname "Xue," and another under the surname "Hsueh."



IBM Global Name Recognition technology recognizes that these seemingly different names are really four different cultural representations of the same name across Southeast Asia.

The Cost of Failure

On February 18, 1993, The Washington Post reported “[Mir Aimal] Kansi, the Pakistani immigrant being sought in the shootings last month outside the CIA headquarters, passed through immigration checkpoints at John F. Kennedy International Airport in New York with a passport and business visa listing his name as Mir Aimal Kasi... Investigators say the discrepancy in Kansi’s name hampered their efforts to determine who he was and how he came to the United States.”

Unfortunately, in many organizations, very little has been done since 1993 to correct this problem, and it has been shown that similar border & law enforcements search failures occur every day!

IBM patented technology gives you the ability to see across these phonetical, spelling, and cultural differences between names to detect fraud, threat, and opportunities within your organization.

Leading the Industry in Advanced Name Recognition

IBM is the leader in providing high precision software for mission-critical name matching and searching. Since 1984, IBM (through Language Analysis Systems, Inc. which was acquired by IBM in March, 2006) has pioneered the use of computational linguistics expertise and technology to solve the complex problem of multi-cultural name matching and searching. It is no longer necessary to task programmers with “tweaking” older, key-based approaches to try to solve this persistent problem.

IBM offers off-the-shelf software and linguistics expertise for truly advanced solutions. Through a combination of computational linguistics and advanced software engineering, IBM’s name-recognition tools takes into account alternative spellings, cultural nuances and other linguistic issues to more effectively manage global data

sets or as part of an attempt to return the most relevant information for a search query, rather than a laundry list of close matches.

The technology can tell you the culture a name is from, the gender or possible marriage status of a person, the probable variations of the name in order of frequency, its literal meaning, and the countries in which it is most likely to be found.

For example, financial institutions face a variety of risks and responsibilities to combat international fraud and money laundering. The IBM global name management portfolio is essentially a comprehensive database that can trace the widest variety of names from the widest possible origins in order to cut down the chance of accidentally trading with a black-listed name.

In other business case scenarios, these tools can reduce the possibility of false positives that might result in unwarranted detentions, a credit card being blocked unnecessarily, or someone from not being allowed to fly.

IBM’s patented name recognition technology is exactly suited for each type of problem. Some of the high-

lights include the ability to:

- Identify name by culture and relative frequency
- Search for multi-cultural names in a database
- Parse a name into Surname and Given Name
- Generate frequency statistics for name tokens
- Generate all variants of a name
- Generate additional attributes such as gender
- Quickly train field personnel in advanced multicultural name searching techniques
- Utilize rich name data gained from the comprehensive study of over 1 billion names from around the world

“... The chief investigators at LAS [now IBM Global Name Recognition] are by far the most knowledgeable people working in the name search problem domain.”

6 month United States Government study of name search technology

Global name recognition's industry applications

Global name recognition technology has historically been and continues to be widely used within the federal government to support the identity recognition needs of solutions ranging from border control to international law enforcement. Today IBM's global name recognition technology has seen increased utilization in the private sector, where it is being used to increase the accuracy of watch list checks, ferret out fraud and enable global customer facing applications through better recognition and verification of customer names to ensure that operations like call centers and direct marketing run smoothly. Industry applications of the technology include:

- Intelligence and Security
- Financial Services
- OFAC Compliance
- Law Enforcement
- Homeland Security
- Marketing Services

IBM Global Name Analytics

IBM Global Name Analytics is designed to address the specific needs and demands of managing multicultural data sets. Unlike traditional data cleansing capabilities

that have been designed primarily to manage data assets in westernized, romanized cultures, IBM Global Name Management is designed to meet the unique demands of organizations and governments that rely upon data sets from cultures around the globe.

IBM Global Name Analytics identifies and classifies what cultural background a given name comes from and recognizes whether a name is predominantly male or female. It automatically parses culturally diverse personal name information into surname and given name components to ensure name data is consistent and accurate across your systems. These capabilities enable organizations to improve data quality, retain customer information, and treat multiple cultures sensitively by accurately parsing and storing customers' names within their automated systems.

IBM Global Name Scoring

IBM Global Name Scoring allows a user to identify identical and fuzzy clear text and phonetic name matches more effectively, overcoming the vagueness and inexactness of transliteration, pronunciation, and the wild profusion of naming and syntactical schemes.

IBM Global Name Scoring enables users to search for multi-cultural names in a database and provides the most likely variations more effectively, improving the accuracy of name searching and the quality of identity verification initiatives. This capability overcomes the vagueness and inexactness of transliteration, as well as the wild profusion of naming and syntactical schemes that make it difficult to distinguish the Saddam Hussein's from the Prince Hussein's. Users can search and recognize foreign names, screen potential threats, and perform background checks across multiple geographies and cultures.

IBM Global Name Scoring provides a phonology-oriented search capability that provides ranked search results based on similarity of pronunciation. Phonetic matching applies language-specific letter-to-sound rules in order to identify potential pronunciations for names, so that two superficially dissimilar names can be matched by a shared spoken form.

IBM Global Name Reference

Encyclopedia

The IBM Global Name Reference Encyclopedia is a comprehensive, interactive database of names, name use, and variations for use by analysts, investigators, and researchers within global public and private organizations. It contains culture-specific information about names, their use, meanings, and patterns of spelling variations. Each name is automatically analyzed to show: cultural/ethnic classification, most prominent spelling variants, gender associations/probabilities, titles, affixes, qualifiers, and countries where name occurs most frequently.

The IBM Global Name Reference Encyclopedia includes information and analyses for names from around the world: Anglo/European, Arabic, Chinese, Hispanic, French, German, Indian, Korean, Pakistani, Russian/Slavic, Thai, Japanese, Western African cultures, and more.

For more information

For additional information about IBM's global name recognition technologies, please visit:

www.ibm.com/data/globalname



© Copyright IBM Corporation 2006


IBM (United States of America)
Entity Analytic Solutions
6600 Bermuda Rd, Suite A
Las Vegas, Nevada
United States of America, 89119

Printed in the United States of America
7/06
All Rights Reserved.

DB2, IBM, the IBM logo, and the On Demand logo are trademarks of International Business Machines Corporation in the United States, other countries or both.

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.

 Printed in the United States of America on recycled paper containing 10% recovered post-consumer fiber.