

## This story on the web

- [Flash page](#)
- [Non-Flash page](#)



### New York Police Dept. Real Time Crime Center

- [Let's go to the videotape](#)



### Threat and Fraud Intelligence

- [Everything you need to know](#)



### Smart Surveillance

- [In the news](#)
- [In the retail industry](#)
- [In banking](#)



### Fighting crime on the street

- [Santa Barbara Sheriff fights fraud](#)
- [Tyler, Texas police gets results](#)



### Stay on top of innovation that matters

- [More Ideas from IBM](#)

## Piecing together a crime in bits and bytes

While forensics coaxes the details of a crime out of the physical evidence, analytics can coax the story of a crime out of the data: names, ID numbers, photos, video, signatures, even information gathered decades ago. Take a look.



[Play video](#) (2:33min, 2MB)

## Global Name Recognition

### Video transcript

Name recognition is much more of an art than it is a technology and names are by far the most challenging attribute to recognize. Addresses and telephone numbers can be verified; however, there is no central verification source for names.

Names differ based on geography, culture, gender, tribe and family. Effective name recognition technology must take all of these variables into consideration.

Names can also contain a variety of optional information such as titles or nicknames, causing them to appear very different. Names aren't just a string of characters. They are data objects for which principles of similarity must be

determined. Names remain the single most important means for identifying a person. Malicious people around the world are learning how easily most name search systems can be circumvented.

IBM's Global Name Recognition Technology is optimized for the multicultural nature of most global name sets and is well equipped to deal with things such as name variations, nicknames, shortened names, multiple name prefixes, name order, titles, as well as genderizing and phonetic misinterpretations.

This technology is built upon a database of close to one billion names from around the globe and has been used extensively by United States security and intelligence agencies since 2001.

A combination of the names database, name-specific algorithms and user-interface tools offers industry-leading capabilities designed to give you the best understanding of names.

Let's look at an example from Southeast Asia. The Chinese characters shown here represent the person's name in the native Kanji. There are multiple transcription standards across the region as well as many dialects which result in extremely different versions of the Romanized name as we see on the right.

IBM has studied and encoded the various transcription standards into our software approach. The IBM Global Name Solution could detect that [Zang Q Su and Chang Tu Su] are likely to be the same person. Most US centric name-matching routines cannot do this because they do not recognize the variations stemming from cultural differences.

In Arabic, the issue can be even more complicated as is demonstrated here. Across the top, you see a man's name written in the Arabic script from right to left.

Each of the five regions pictured uses a different method of transliterating the same name from Arabic script into the Roman character set. Therefore, each of the variations of name Mohammed shown here represents an accurate rendering of the man's name in that particular region.

[End of segment]

---

## SEND US YOUR FEEDBACK:

We're interested in hearing what you think about this article.

Do you have a question, comment, suggestion or story idea to pass along?

Please e-mail [Ideas from IBM](#).

---

### IBM Corporation

330 Seventh Avenue, New York, New York 10001, United States