



LJN'S

LEGAL TECH

Newsletter®



Volume 25, Number 9 • December 2007

PRACTICE TIP

A Primer on Foreign Language e-Discovery

By Ari Kaplan

While e-discovery may be Greek to many, it is those documents written in Chinese, Japanese, Korean and Russian that cause much of the trouble. These “multi-byte” languages have exponentially more characters than the 26 letters and few other punctuation marks that Latin languages like English, Spanish, French and German need. In fact, the number of Chinese characters included in the Kangxi dictionary is over 47,000 (though only 3,000-4,000 are reportedly necessary for full literacy). The impact on e-discovery is significant considering the increased sophistication necessary for case evaluation.

At the most basic level, computers think in ones and zeros, with a one or zero being a bit. Eight bits is a byte. There are 256 different combinations of numbers you can create using a byte (2 (bits) to the eighth power). For languages that are not based solely on letters, *i.e.*, those where symbols represent a concept or a syllable, you need to add bytes (256 x 256, which equals 66,536). That is the essence of multi-byte vs. single-byte languages — single-byte languages have 256 possible combinations, while multi-byte languages have 66,536.

continued on page 8

Maximizing the Power of Virtual Data Rooms

By Joel Ginsberg

A virtual data room (“VDR”) is a secure Web-based version of a traditional physical data room used to facilitate document exchange and review in connection with corporate transactions. The technology emerged within the early 1990s to facilitate mergers and acquisitions and loan syndication transactions, and today supports a range of legal and financial transactions, such as IPOs and other securities transactions, real estate financings, private equity, venture capital, corporate restructuring and litigation matters.

VDRs offer legal and financial professionals an array of advantages. Overall, VDRs make it possible for lawyers and other deal advisers to focus more on the substantive work to be done and less on procedural aspects of deal management. By making the data room accessible over the Internet, VDRs help to accelerate the pace of the due diligence process. VDRs provide interested parties with convenient and instant access to deal-related documentation while reducing administrative expenses. Furthermore, VDRs provide professionals with the ability to monitor data room activity, including who accesses data room information and who views specific documents.

Professionals typically outsource VDR technology and support services to third-party VDR providers. For most firms, it would be considered cost prohibitive and logistically challenging to develop their own VDR technology platform. With a growing number of VDR providers and recent advances in technology, professionals have more options than ever when selecting a VDR solution for their transactional needs. When deciding to utilize VDR technology, professionals must understand VDR terminology, and the features and options available to better evaluate and maximize their technology investment.

As one evaluates the functional benefits of VDRs, there are intangible factors to consider. With the greater adoption of technology such as Blackberries, e-mail and voice over IP, professionals tend to gravitate towards standardized Web behaviors. The intrinsic behaviors learned with these types of applications serve as underlying

continued on page 2

In This Issue

Maximizing the Power of Virtual Data Rooms 1

A Primer on Foreign Language e-Discovery 1

Redlining and PDF Utilities Extraordinaire 3

Turbo-Charge Your Table of Authorities . . 5

PRSR STD
 AUTO
 US POSTAGE PAID
 PHILADELPHIA, PA
 PERMIT #1437

Foreign Language

continued from page 1

Confused? Then let's address codings. An encoding is a grammatical translation of what you input to what you get on the screen. The problem is when you have multiple encodings. For example, when analyzing an Outlook 2000 e-mail file (PST format) under a Japanese operating system that you then convert to an English-language machine for review, there will be problems because the native data in Japanese is corrupted due to linguistic differences.

Ari Kaplan, the principal of Ari Kaplan Advisors, is a lawyer and a writer based in the New York area. He teaches a CLE course at bar associations and law firms nationwide on the mechanics of getting published for business development. He also coaches individual lawyers and provides strategic writing assistance to companies that service the legal community. He can be reached at www.arikaplanadvisors.com.

Unicode was created to solve some of these problems and offer a universal solution; however, it is only available for files created on newer systems, making legacy data a continuing area of concern. "Each language family has its own unique set of problems and solutions," says Thomas Barnett, Special Counsel for Sullivan & Cromwell, LLP.

In fact, "in some parts of the world, you are not allowed to take the data out of the country due to local data protection laws," adds Brian Kim of PriceWaterhouseCoopers LLP. He adds that certain countries also have native applications that are more popular than those commonly used in the United States, requiring additional evaluation of your program inventory.

Whether your data is in Unicode or not, proper preservation is the key. While Microsoft Windows NT, 2000, XP and subsequent versions support Unicode, many archiving or compression tools do not support it. This could result in missing files that may or may not be reported in the error

logs. For that reason, you must test carefully, notes Kim. Also, to ensure correct extraction, properly align the regional settings.

Some languages overlap in terms of characters, *e.g.*, Chinese and Japanese, and others do not use spacing, which makes search more complicated. And many corporate documents will combine English with another language as well.

To avoid mistakes and enhance defensibility, consider organizing data for review beyond keyword searching, given the difficulty in establishing such terms for foreign languages. Also bear in mind that the expense incurred for translation is substantial. While expert translators, ordinary native speakers and native machine translators are options the issue is often one of timing and the reliability of the end product.

Remember, e-discovery is only Greek to you if you don't know the code.

