**CATALYST**
*Powerfully Simple E-Discovery*

Home        Products        Specialities        Resources        About        Contact

# E-Discovery Search Blog

## Technology, Techniques and Best Practices

Follow Us:

Home / Catalyst Language Services / Was Samsung Deal a Watershed for use of Machine Translation in FTC Second Requests?

### Was Samsung Deal a Watershed for use of Machine Translation in FTC Second Requests?
by Bob Ambrogi | May 15, 2012

Another e-discovery technology may recently have had its watershed moment, similar to the recent watershed marked by U.S. Magistrate Judge Andrew J. Peck's endorsement of predictive coding. This time, the watershed involved the use of machines in place of humans to translate foreign-language documents in a massive Federal Trade Commission Second Request.



The deal that was the subject of the Second Request was remarkable by any measure. It involved the nearly $1.4 billion sale by South Korea-based Samsung Electronics Co. Ltd., a world leader in digital consumer electronics and information technology, of its hard disk drive operations to Seagate Technology Plc, a major manufacturer of disk drives and storage products. After the deal closed in December, Samsung reported its largest-ever quarterly profit.

The deal was also momentous for the law firm that represented Samsung, Paul Hastings. Last month, the antitrust publication *Global Competition Review* named the Samsung/Seagate transaction as "Matter of the Year," recognizing both Paul Hastings for its representation of Samsung and Wilson Sonsini Goodrich & Rosati for its representation of Seagate.

But before the deal could be consummated, it had to await Second Request review by the FTC, a fast-track, discovery-like process that gives the FTC the documents and information it needs to evaluate whether the proposed transaction may violate antitrust laws.

## One Million Pages to Translate

For the attorneys at Paul Hastings, the Second Request process was further complicated by one simple fact—the vast majority of the documents they had to produce were in Korean, but the Second Request process requires all documents to be produced in English. Based on the firm's preliminary estimates, of roughly 350,000 documents they had collected, 200,000 were in Korean. At an estimated five pages per document, Paul Hastings estimated that they could be facing as many as a million pages of Korean-language text.

The attorneys faced two hurdles. Time was the first; as long as the parties have not complied with the Second Request, they are prohibited from closing the transaction, potentially resulting in significant business losses while the two companies continue to operate independently. Expense was the other; by their estimates, the cost to translate all 200,000 documents by hand could easily exceed $20 million.

Given these hurdles, the Paul Hastings attorneys decided early in the process to try machine translation (MT) as an alternative to hand translation. MT held the promise of both expediting the translation process and saving their client substantial expense.

In the field of e-discovery, MT is a technology that is gaining growing recognition and use. Sophisticated MT software uses linguistic and statistical techniques to translate text from one language to another. By seeding the software with key terms and phrases and updating the dictionary as the translation progresses, results can be impressive. While machine translation is far from perfect, in the e-discovery context, it can often eliminate or greatly reduce the need for human translation.

## Using MT for the Second Request

In a Second Request scenario, each party must certify that it has "substantially complied" with the government's requests, including the request that documents be provided in English.  With this substantial compliance standard in mind, the Paul Hastings team went forward with its plan, selecting an MT vendor, providing core terms and names to seed the dictionary, and working with the vendor to refine results. Simultaneously, certain documents identified as particularly relevant or important were singled out for hand translation to ensure maximum accuracy for this subset. When the translation process was complete, the lawyers delivered the translated documents to the FTC.

As the FTC was quick to note, the machine-translated documents did not have the same level of quality as the hand-translated materials.  However, Paul Hastings took the position that the translations were sufficient to allow the agency to identify relevant materials for further review and to seek hand translation of particular documents as necessary. In the end, the Second Request process ran its course and the FTC allowed the merger to close.

In this case, at least, machine translation was sufficient to allow the process to reach its end. Not only that, but by using MT, Paul Hastings saved its client a huge expense. Remember that estimated human translation cost of over $20 million? By contrast, the final cost of machine translation was less than 5% of that amount.

## What Does this Mean?

Recently, I interviewed Michael S. Wise, a Paul Hastings litigation associate who was heavily involved in the Samsung deal and the translation project.

There is no definitive way to determine whether MT has ever before been used to such an extent in an FTC second request, Wise said. And the circumstances of the case prevent any solid inference that the FTC now endorses MT.

Still, the implications are significant for large corporations headquartered outside the United States, he believes. "We hear multiple war stories of companies paying millions for hand translations of materials in the Second Request process. That seems ludicrous from a cost perspective."

Despite his firm's success with the process in the Samsung case, Wise cautions that MT remains far from perfect and therefore is not always a substitute for human translations, particularly not in an adversarial context where precision may be demanded by the opposing party.

"But in the agency context, where there's no wrongful conduct being alleged, where it only involves the economic effects of the transaction, you have a strong argument that the government has an obligation to employ reasonableness in the costs it imposes on a company," Wise said. "With MT, you can control the cost and still be able to run search terms and find what you're looking for."

Wise says he would use MT again when the circumstances warrant it. He also says he would refine the approach based on the lessons he learned here. For example, he found that having separate vendors for translation and document review/hosting resulted in technical compatibility issues; he would prefer a vendor that offers a single platform to handle both MT and review in the future. Also, he would budget more time for the MT process and start with a larger seed set than the vendor recommended in this case.  He believes that to maximize accuracy in his next case, he would look to have double the number of seed documents and an extra month built into the timeline which could be used to work with the vendor on fine tuning its translation engine.

So was this a watershed moment for MT? Wise believes it was.

"Five years ago, maybe less, you had to hand-translate everything. We are getting to the point now where, done properly, there is the opportunity to employ these technologies to substantially reduce costs."

The story of how lawyers can use technology to streamline e-discovery is not just about predictive coding. With machine translation, the story now has another chapter.

*(Note: Although Catalyst offers a full range of* language services, *including MT and human translation, it was not involved in this case.)*

Filed Under: Catalyst Language Services Tagged With: Machine Translation, Translation

**About Bob Ambrogi**

A lawyer and veteran legal journalist, Bob advises Catalyst on strategic communications and marketing matters. He is also a practicing lawyer in Massachusetts and is the former editor-in-chief of The National Law Journal, Lawyers USA and Massachusetts Lawyers Weekly. A fellow of the College of Law Practice Management, he also writes the blog LawSites.

**Share Your Thoughts**

Name *

Email *

Website

Post Comment

**Welcome**

We are passionate about the role of search in electronic discovery. With digital content growing at an increasing pace, search becomes ever more important and interesting. Our mission is to chronicle the rise of search in the discovery process — from law to technology, techniques and best practices. And to write about other things that interest us as we go along.

Search Is The Word

**Search**

Search

**Browse By:**
**Category**

Select Category

**Tags**

Select Tag

**Authors**

Read the posts from your favorite expert:



**John Tredennick**
Founder & CEO

View Bio & Posts



**Robert Ambrogi**
Director of
Communications

View Bio & Posts



**Ron Tienzo**
Senior Consultant

View Bio & Posts



**Jim Eidelman**
Senior Consultant

View Bio & Posts

**Additional Contributors**

Bruce Kiefer
Vice President of Hosting Applications Division and Director of Research and Development

Jeremy Pickens
Senior Applied Research Scientist

Christopher Toomey
Director of Alliances and Channel Development

Patty Daly
Director of Marketing & Training

Nirupama Bhatt
Search and Analytics Consultant

**Products**

Catalyst Insight
Insight Predict
Catalyst CR

**Specialties**

Corporate Enterprise
Global Language Services
Search Consulting
Catalyst Asia

**Resources**

News
Events & Webinars
Media Library
Training & Support
Catalyst Gear

**About**

Why Catalyst
Who We Are
Partners
Leading Edge Research
Careers

**Contact**

Contact a Rep
Live Demo

Patty Daly
Director of Marketing & Training

Nirupama Bhatt
Search and Analytics Consultant