Navigating ESI Challenges in an FCPA Investigation

By John Tredennick and Robert Ambrogi
## Contents

- Contents .................................................................................................................. i
- Introduction: FCPA Actions Can Be E-Discovery Minefields ..................... 1
- The Law: Battling Bribery as a Business Tool .............................................. 1
- Prohibition on Bribery ...................................................................................... 2
- Enforcement Continues to Rise ...................................................................... 2
- ESI Challenges in FCPA Investigations ......................................................... 3
- Managing Multi-Language Data Collections ............................................... 4
- The Development of Unicode .......................................................................... 4
- Search and Review of Foreign Language Documents .................................... 5
- Dealing with Pictorial Languages ................................................................... 5
- Multiple Languages within Documents or Collections ............................... 6
- Handling Large Data Volumes .......................................................................... 6
- The Review Challenge ..................................................................................... 7
- Dealing with Multiple Custodians and ESI Formats ....................................... 8
- Meeting Tight Turnaround Times .................................................................... 9
- Using Search as a Navigational Aid ............................................................... 10
- Finding The Right Tools for an FCPA Investigation ..................................... 10
- About Catalyst ................................................................................................. 11
- About The Authors .......................................................................................... 12
- John Tredennick ............................................................................................. 12
- Robert Ambrogi ............................................................................................... 12
Introduction: FCPA Actions Can Be E-Discovery Minefields

For any enterprise that does business outside the United States, the Foreign Corrupt Practices Act is a minefield of potential dangers. One false step and a company can find itself facing an investigation by the U.S. Department of Justice or the U.S. Securities and Exchange Commission.

As if an FCPA investigation were not risky enough, it is made even more so by the nature of the electronically stored information at issue. By definition, an FCPA action is virtually certain to involve ESI in multiple languages, making it more difficult to process and search. Adding to the difficulty, the quantity of ESI is likely to be huge—multiple terabytes in some cases. And the data is apt to be housed under multiple custodians across any number of global locations and in any variety of formats.

Despite these challenges, counsel for a company in an FCPA investigation must meet tight time limits and ensure that data collection, processing and review are done in a manner that protects the integrity and defensibility of the company’s response. Ideally, all of this should be accomplished as economically as possible.

With a better understanding of the issues and the right tools to address them, the minefield can be safely navigated. This white paper reviews the challenges a company and its counsel face in managing ESI in an FCPA investigation and suggests how best to address those challenges.

The Law: Battling Bribery as a Business Tool

There is little debate that the purpose of the FCPA—to ban bribery of foreign officials in order to gain a business advantage—is commendable. In a global economy, however, compliance can be a challenge. Business customs vary widely from country to country. A practice shunned in one place might be common in another. Trip up and the consequences can be severe.

In December 2008, FCPA violations cost Siemens AG $1.6 billion in fines and disgorged profits. Two months later, in February 2009, Halliburton resolved...
FCPA charges with payment of a $402 million fine and disgorgement of $177 million in profits, for a total of $579 million.

These cases represent the extremes—so far, at least. But even lesser cases can result in serious penalties. For each violation by a corporation, the FCPA authorizes criminal penalties of up to $2 million and civil fines of up to $10,000. Individuals face criminal fines of up to $100,000 and imprisonment of up to five years for each violation. In some cases, the fines can be even higher—up to twice the benefit the company sought to obtain by making the bribe.

**Prohibition on Bribery**

The FCPA was a response to a series of SEC investigations during the 1970s. At the time, more than 400 companies admitted to making “questionable or illegal payments” to foreign entities, government officials, politicians and political parties. By enacting the FCPA in 1977, Congress sought to stop bribery of foreign officials and restore public confidence in the integrity of American business and commerce.

The act is a broad ban on bribery. It prohibits U.S. companies and citizens, foreign businesses that trade on any U.S. stock exchange, or any person acting while present in the United States, from offering to pay or corruptly paying money or anything of value to a foreign official to obtain or preserve business.

The FCPA also includes a reporting requirement. It applies to any company, including foreign entities, that trades securities on a U.S. stock exchange or is otherwise required to report periodically with the SEC. These companies must maintain accurate records of business transactions and enact strong internal controls.

The Department of Justice is the primary FCPA enforcement agency. It is responsible for all criminal and civil enforcement of the anti-bribery provisions with respect to domestic concerns, foreign companies and nationals. The SEC is responsible for civil enforcement of the anti-bribery and reporting provisions with respect to issuers.

**Enforcement Continues to Rise**

Although the FCPA has been on the books for more than 30 years, the federal government has never been so aggressive in its enforcement as in recent years. Since 2002—the year Sarbanes-Oxley came into law and renewed attention was given to corporate corruption—both the DOJ and the SEC have steadily increased FCPA investigations and enforcement actions.
The numbers of cases reflect the increase in enforcement. In the three years of 2004 to 2006, the DOJ and the SEC brought a total of 29 FCPA enforcement actions. The next year alone brought 38 cases. Last year, the government brought more than 100 cases. Increasingly, the government is targeting specific industries where it believes corruption is more common.

Recent years have seen a significant increase in FCPA prosecutions of individuals. In 2006, FCPA charges were brought against only six individuals. In 2009, the number of individuals charged rose to 44. In just the first month of 2010, the DOJ indicted 22 individual corporate executives—including one company's general counsel.

There is no sign that enforcement will let up. In February 2010, the head of the DOJ unit that prosecutes FCPA cases said he expects the unit's workload and staffing to grow substantially over next two years.

**ESI Challenges in FCPA Investigations**

Given heightened FCPA enforcement and the severity of the potential consequences if a violation is found, a company faced with an FCPA investigation must be thorough and precise in its response.

As the company makes its way across this minefield, the biggest danger it will face is its electronically stored information. Because FCPA cases by definition involve global business operations, the ESI presents challenge upon challenge:

- **Multi-language documents.** FCPA cases routinely involve documents in multiple languages. Companies must be prepared to handle processing, analysis and review of multi-language ESI quickly and efficiently to provide an accurate “first read” of the data.

- **Huge data collections.** ESI collections involving international business operations are inevitably composed of huge amounts of data. In the recent Siemens matter, for example, investigators secured five terabytes of data from Siemens offices.

- **Multiple custodians.** ESI collections in FCPA cases are likely to be pulled from multiple custodian files and corporate repositories, stored across any number of locations and in various media forms.

- **Time constraints.** When the DOJ requests documents, it wants them sooner, not later. To collect and process large volumes of ESI on a timely
basis, while protecting against non-responsive and privileged disclosures, a company faces a daunting task.

Each of these challenges is significant, but not insurmountable. With a full understanding of the issues and the right set of tools, a company and its counsel can safely navigate their way through the data.

**Managing Multi-Language Data Collections**

One factor, above all others, distinguishes an FCPA investigation from any other corporate investigation—the near certainty that it will involve documents in multiple languages. This means that, when initiating a collection and search strategy for an FCPA case, the company must be prepared to handle processing, analysis and review of multi-language ESI quickly, efficiently and accurately.

Here, technology can either be your best friend or your biggest obstacle. Many of the litigation-support systems and document-review platforms in use today are not able to support foreign languages. The weakness in these off-the-shelf systems, many of which were built in the 1990s, is that they understand ASCII code. ASCII is not up to the job of processing multi-language data.

The original ASCII, developed in the 1960s, allowed 128 characters. That covered the basic English alphabet, along with numbers and punctuation, but not much else. As business communications went global, ASCII proved insufficient to handle languages with special characters, such as German, Russian and Hebrew, let alone Asian character-based languages. Even when ASCII was extended to allow 256 characters, it still was not up to the job.

**The Development of Unicode**

A great leap forward came with the invention of Unicode. Compared to ASCII's 256-character limit, Unicode allowed more than 1.1 million characters. That enabled it to handle the difficult "CJK languages"—symbolic languages such as Chinese, Japanese, Korean and Thai. It quickly became the global standard for foreign-language documents.

Today, the most common form of Unicode is UTF-8. It can use up to four bytes to represent any possible language character. Because ASCII characters occupy the first byte, UTF-8 is backward compatible with ASCII. Its compatibility with more than 30 years of ASCII files makes UTF-8 the lingua franca of Unicode.
Unless your software understands Unicode, it will not properly handle foreign-language documents. Since most off-the-shelf litigation packages use ASCII, they fail to decode e-mail and documents with foreign language characters.

Make sure the software used to collect and process your data is Unicode compliant. Then it needs to be loaded into a database or other software that also supports Unicode.

**Search and Review of Foreign Language Documents**

Search software operates by creating an index of every word in every document within the data set. When you execute a search, the software searches the index, not the individual documents. That is how a powerful search engine such as Google can appear to search the entire Web in just seconds.

When computers index, they identify single words by the spaces and punctuation that surround them. This process is called tokenization. Computers do not see words as words, only as letter/number combinations, each a distinct token.

This is easy for most Western languages, because words are separated by spaces and punctuation. However, languages that do not use spaces or punctuation to delineate words pose a problem. In Chinese, Japanese and Thai, for example, characters run together without any clear breaks.

To define the individual words in a sentence and create a document index, search engines with special tokenizers must be used to recognize each language and determine where words begin and end.

**Dealing with Pictorial Languages**

A further difficulty in searching CJK languages is their use of pictorial characters called logograms. The search engine must figure out how to group these pictorial characters into words. Even more challenging, it must recognize how the characters are combined, because that affects their meaning.

For example, the traditional Chinese word for "Chinese" consists of three logograms that directly translate to "middle country people." To say "China" rather than "Chinese," you use two of these characters that represent "middle country."

A basic search engine would have no way of knowing whether these three characters should be read together or separately and therefore no way of discerning their meaning.
To overcome this challenge, linguists have designed software that can read characters and understand their context sufficiently to create an index. In order to search foreign language documents accurately, you need a search engine with a tokenizer that is capable of handling the peculiarities of the language.

**Multiple Languages within Documents or Collections**

To handle foreign language documents, an indexing engine must first determine the language being used. After identifying the language, the engine then applies the appropriate tokenizer to identify word boundaries. This task becomes more complex when a document set, or even a single document, contains more than one language, each having its own tokenization rules.

Many older review platforms deliver foreign language support through language packs. They apply tokenization rules for a single language to the entire collection. This means that searches run against the collection may miss applicable terms in documents that are in another language, jeopardizing search accuracy.

When using an older system, the common workaround for these problems calls for setting up a separate repository for each language in the collection (assuming you know what all of those languages are), running your search against each repository, and then manually combining the results.

Modern review platforms such as Catalyst CR take a different approach. They accept the concept that document collections, and even single documents, can contain multiple languages. They analyze documents in segments, identify the primary and secondary language in each segment, and apply the appropriate tokenizing schemes. With this sort of platform, a multi-language document set can be indexed, tokenized and searched—all within a single repository.

**Handling Large Data Volumes**

In an FCPA case, multiple languages is not the only challenge ESI presents. Further complicating e-discovery is the fact that the ESI collection might commonly involve a huge quantity of data—multiple terabytes in some cases. The need to handle such a large volume of data presents multiple challenges of its own.

One is processing. In an FCPA investigation, the task of processing and loading native files can present a speed-bump that delays analysis and review. The best route around this delay is to use a system that automates the processing of the data and that has the capacity to handle any volume.
An example is the Fast Track processing system in use at Catalyst. With Fast Track, users can submit multi-language e-mail and other electronic files directly into the Catalyst CR review platform. The automated Fast Track system processes, loads and indexes the files and makes them available for search, review and production. Because the system is automated, it dramatically reduces processing time.

By quickly processing large volumes of ESI into a multi-language hosting platform, analysis can begin concurrently with data indexing. That enables counsel to identify and target key custodians and specific date ranges, leading to an iterative search effort and the ability to focus the FCPA investigation.

The Review Challenge

Once data is collected and processed, a second challenge arises, that of reviewing all this data in a manner that is both thorough and efficient.

The recommended approach is to use a hosted enterprise system. A hosted system allows the investigation to get up and running quickly, with minimal IT involvement and minimal training. That allows counsel, early in the matter, to focus on what is most important. The system should be one that allows for realistic deliverables and planning with the legal team (counsel, consultants, project managers and others).

By way of illustration, consider how a major AmLaw 100 law firm used the Catalyst CR platform in a recent FCPA investigation. The case involved data assembled from locations all over the world and in multiple formats. The data was in English, Russian, Chinese, Spanish, Italian and some 20 other languages. It needed to be loaded, analyzed and reviewed in under two months.

Catalyst CR enabled the review team to automatically determine primary and secondary languages within each of the documents and sort them based on their language characteristics. But because the firm had reviewers fluent in only four of the languages, documents in other languages had to be translated.

Catalyst offers translation on the fly for more than 32 languages. Here, with more than a million documents to translate, the client chose to use bulk translation. These translations were uploaded to Catalyst and linked to the native-language originals.

Documents identified through the searches were batched and assigned to reviewers using Catalyst’s review workflow module. The module allowed reviewers to tag documents in a variety of ways and to create and use multi-tag
macros for faster throughput. It also allowed them to tag documents for which the language was improperly identified or that contained multiple languages.

Second-level reviewers focused on documents marked Relevant and Hot. Their job was to review the Relevant documents and escalate appropriate ones to the Hot category. They also could add comments relating to the investigation and tag documents by issues and potential witnesses.

This was a two-step process. Documents marked as candidates for Hot were sent out for hand translation. These would be used during the investigation and given to witnesses for comments. The partners doing the questioning did not speak the languages and needed the highest-possible fidelity for the translations.

Ultimately, documents marked Hot were batched and placed in notebooks along with their translations. With Catalyst CR, both could easily be downloaded in batches for printing.

In this case, four million documents passed through the system. The fact that many were in languages other than English could have proven debilitating. The right tools—along with the right team—made the job manageable.

**Dealing with Multiple Custodians and ESI Formats**

In an FCPA investigation, not only is the quantity of data likely to be huge, but it is equally likely to be pulled from multiple custodian files and corporate repositories, stored across any number of global locations and in any variety of media forms.

With time of the essence, the best way to meet this challenge is to use a system with the multiple capabilities that are required. As noted earlier, at Catalyst, we use and recommend Fast Track processing. In an FCPA case, it offers key advantages:

- It handles native files of any type, including e-mail collections, Office files, PDFs, images and text files.
- It handles English and multi-language files.
- It allows simple uploading of data via FTP from anywhere in the world.
- Processing begins immediately upon upload and the data is quickly available for search and review.
- It runs against a grid of virtualized servers, thereby allowing for virtually unlimited capacity.

With data drawn from multiple custodians and multiple storage locations, the likelihood is greater that it will contain duplicate or near-duplicate documents. To reduce the population of documents for review, we recommend the use of Equivio's near-dupe and e-mail thread technology. It can help reduce a document population by as much as 50 percent.

**Meeting Tight Turnaround Times**

When a company faces an FCPA investigation into potential bribery or improper payments, its initial response is critical. A timely response indicates that it is acting diligently to assess the allegations and is committed to self-compliance.

With this shortened reaction time, the task of collecting, processing and assessing ESI from locations throughout the world can pose a serious obstacle. Counsel cannot wait indefinitely as data is identified and collected and the team prepares for the investigation and possible future litigation.

Due to the need to investigate quickly and assess data accurately, practitioners need to recognize the importance of employing an effective early case assessment (ECA) tool. The tool should enable counsel to spot key communications and explore strategic theories across all or a subset of the collection. With the added complexity of multiple languages, the ECA technology must also be able to address foreign language in a simple, efficient manner, allowing for multi-language search analysis.

ECA proves vitally important to accurately investigate the underlying allegations in a timely manner. By addressing quickly the concerns raised with alleged conduct, one can explore key custodian data, assess social networks detailed within ESI correspondence files, and iteratively assess and refine key search terms and concepts to build a factual chronology of events.

This information can be used to focus the investigation on specific communications, providing detail that can be applied during custodian interviews and leading to subsequent depositions, if required. This first-step analysis allows counsel to proceed more systematically through the remaining documents, stratifying the document populations based upon defined key words developed and confirmed through iterative search methods during early case assessment.
Using Search as a Navigational Aid

When boaters navigate treacherous waters, they rely on navigational aids to guide them to safety. In navigating the challenging terrain of ESI in an FCPA case, there is a parallel to the mariner's navigational aid—the professional search consultant.

U.S. Magistrate Judge John M. Facciola, a noted authority on e-discovery, has written that search is an exacting science, one "clearly beyond the ken of a layman," including most lawyers and judges. "Whether search terms or 'keywords' will yield the information sought is a complicated question involving the interplay, at least, of the sciences of computer technology, statistics and linguistics." *U.S. v. O'Keefe*, Cr. No. 06-249 (D. D.C. 2/18/08).

In an FCPA case, use of consultants with specific expertise in search and review can serve both to enhance the integrity of the process and increase its efficiency. A key benefit of a search consultant is in helping to pare down a review set to a more manageable size, enabling the review to be completed more quickly and at less cost.

Cognizant of this, Catalyst employs a team of search consultants that is composed of professionals who are trained in both law and computer science. In one recent FCPA case, this Catalyst Consulting team was able to help a large multi-national reduce its review population by 76 percent.

After assessing the client's situation, the consulting team recommended the use of "Predictive Ranking." This technique combines statistical algorithms and data mining with review sampling. From an initial population of 3.2 million documents, Catalyst Consulting was able to cut the final count of documents requiring review to 750,000—a total reduction of more than 2.5 million documents.

With time of the essence in an FCPA investigation, the expertise of a search consultant can prove to be a potent weapon.

Finding The Right Tools for an FCPA Investigation

An FCPA investigation is risky enough without getting tripped up by electronically stored information. Even though the nature of the data in an FCPA case is likely to pose multiple challenges, the right tools can help navigate around them. To ensure the integrity, thoroughness and efficiency of the investigation, counsel should look for these characteristics:
• A Web-based system with enhanced search capabilities that enables international teams to access and analyze data any time of day, from any location in the world.

• Powerful search, culling and analytics to conduct early case assessment, enabling counsel to review custodian correspondence, explore developed theories through iterative search, engage multiple culling tools to “slice and dice” data, and generate a quick and accurate first read.

• Ease and accuracy of document culling and data analysis to allow legal teams to start analysis immediately—as soon as metadata is available for search.

• Ability to map relationships, tag subsets of data for specific attorney review, and chronicle timelines of a case with powerful date-restricted searches.

• Support for multiple languages with accurate language identification, translation services, and powerful search across multi-language documents.

• Ability to integrate bilingual managed review attorneys.

About Catalyst

Catalyst provides secure, scalable document repositories for electronic discovery and complex legal matters. For over a decade, the Catalyst team has helped corporations and counsel control litigation costs and structure legal teams to analyze and review more effectively and efficiently through a central, web-based platform to manage documents, coordinate deadlines and work together across organizations and locations. Catalyst’s enterprise repository systems and supporting services cover the entire litigation lifecycle, from processing and analytics, to review, production and trial.

The company delivers its software directly and through a global partner network that offers consulting, collection, project management and forensic services. The company manages a grid of hundreds of servers in two data centers and has hosted over 100 terabytes (over a billion pages) of confidential data for large corporations and insurance companies.

Catalyst products have been used by many of the largest organizations in the world, including five of the world’s 10 largest companies, four of the nation’s 10 largest insurers and 70 of the nation’s 100 largest law firms. Catalyst has also
hosted highly-sensitive terrorist information for the United States Justice Department and regularly assists with multi-party criminal matters.

**About The Authors**

**John Tredennick**

Prior to founding Catalyst, John Tredennick was a nationally-known trial lawyer and litigation partner with Holland & Hart. Active in the American Bar Association, Tredennick was chair of the Law Practice Management Section and served as editor-in-chief of Law Practice Management Magazine, which reached over 20,000 readers around the world. He also founded and for many years edited the ABA's successful Web Magazines: Law Practice Today and Law Technology Today.

He has written and edited five best-selling books and countless articles on litigation and technology issues. Recently, he was named one of the "Top 100 Global Technology Leaders" by London's CityTech magazine. He also serves as a member of the Short Course Faculty at the University of Virginia Law School where he teaches a course called "Electronic Discovery in a Global Environment.

**Robert 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.