2004 Socha-Gelbmann Electronic Discovery Survey

This article first appeared in the August, 2004 issue of Law Technology News

This is the second year we have conducted a survey to gauge EDD development. First, a bit of history: Noticeable expansion of the U.S. electronic data discovery (EDD) market began in 1999, with about $40 million in revenues. The market grew to approximately $70 million in 2000, $150 million in 2001, and at least $270 million in 2002.

Based on our most recent poll, we estimate that 2003 U.S. EDD revenues were in the range of $430 million, about a 60 percent increase from the previous year. According to the people we talked with — providers and consumers — we predict that the EDD market will grow at rates exceeding 60 percent in 2004 and 2005, with growth slowing down to just under 60 percent in 2006.

As the overall market size has grown, so has the number of providers that offer EDD services or software. We counted 67 companies that entered the electronic discovery market in 2003, compared to 29 in 2002, 32 in 2001 and 14 in 2000.

While the top providers' revenues will continue to grow in absolute terms, we expect as a group they will lose market share to the rapidly-increasing number of other providers in the market. In part, that's because the newcomers will gain expertise and add capacity, and in part due to the sheer number of new players.

We also expect to see steady growth among the do-it-yourselfers – law firms and corporate legal departments that attempt to carry out at least some of their EDD work internally.

Market Leaders

In 2003, as in 2002, no single vendor dominated the market. According to our survey results, the top five market leaders were (in alphabetical order) Cricket Technologies, Daticon, Electronic Evidence Discovery Inc., Kroll Ontrack Inc. and Lexis-Nexis Applied Discovery.

With market shares ranging from about 13 percent to approximately 5 percent, these five vendors accounted for 36 percent of the market. The next 14 providers controlled an additional 34 percent of the market, leaving 30 percent of the market for a couple hundred other providers as well as the do-it-yourselfers.

An important note: Consumers and providers alike should view this top-five list with more than a little skepticism. By its nature, it is weighted to the best-known providers – the ones who do the most work – the ones with the highest revenues. This does not mean, however, that you should only consider using providers who make it to the top of this list (or to the top of anyone else's). Some of the providers we find most impressive did not come close to making this list. Some fell short because they serve a limited geographical area, others because they concentrate on a single part of the overall electronic discovery process, and yet others because they have not been in the market long enough to gain broader recognition.

Software Usage

We asked consumers to identify the software programs they used to work with EDD. For the individual titles they named, we asked them to tell us how much they used the software in 2003, how much they though they would use it in 2004, and how satisfied they were with it.

Fifty-three programs were mentioned. Based on 2003 usage responses, the top 10 programs, starting with the most popular, were:

1. Microsoft Access
2. Microsoft Excel
3. Concordance
4. Summation
5. Microsoft Word
6. Trial Director
7. Sanction
8. iConect
9. dtSearch
10. DolphinSearch

Access seemed to be the tool which pleased users the most. Satisfaction levels for Excel, Concordance and Summation were somewhat lower and less uniform; they dropped off slowly after that.

Issues Influencing the Market

One of the key sets of questions we asked in our interviews with both consumers and providers was their opinions on the state of the electronic discovery industry. These questions focused on services, strengths, weaknesses and concerns.

The top six sets of issues influencing the market, according to survey results, were:

1. The need for standards
2. The dearth of good project management
3. Controlling and containing costs
4. The value of high quality work
5. The ever-increasing demands for and on capacity
6. The need for better education

Of special note was the gulf we saw between what consumers said mattered most to them, and the areas providers thought were important.

When it came to the concerns most frequently heard about the EDD industry, for example, providers cited out-of-control costs as far and away the most important issue. Consumers agreed that cost issues matter greatly, but they placed equal emphasis on an area hardly mentioned by providers at all — how to differentiate between "real providers and charlatans."

Scope

This survey addresses only those activities revolving around the handling of materials that had been "born electronic." By that we mean materials that started out in electronic form and stayed that way until they entered the electronic discovery pipeline. This survey also is limited to activities in the United States and to commercial activities.

We decided to include all aspects of EDD activities — services, software, and consulting. As a practical matter, the services costs accounted for almost all of the revenues (for the providers) and expenditures (for the consumers) reported to us.

2003 TOTAL ELECTRONIC DISCOVERY MARKET	
Tier 1	$ 300,000,000
Tiers 2 & 3	$ 94,000,000
Do-It-Yourself	$ 35,000,000
Total	**$ 429,000,000**

We ultimately divided providers into three groups. Tier 1 providers are providers who are established in the market, have revenues of at least $1M, are national in scope and they are widely recognized in the market. They offer a broad range of services; in other words, they are not one-trick ponies.

Tiers 2 and 3 providers consists of approximately 170 additional electronic discovery providers — local or regional, specialized, companies doing related work (such as scanning and imaging, data recovery, investigative) who also are doing electronic discovery work.

It also includes an additional group of at least 100 additional small providers, generally with only a few employees, who do related work but also do a limited amount of electronic discovery.

Do-it-yourselfers are law firms and corporations who are doing electronic discovery internally, work that otherwise might have been sent to a vendor.

Methodology

We gathered our primary data through telephone interviews and an online survey, collecting information from more than 60 organizations. We interviewed top personnel at more than 20 EDD providers and key decision makers at nearly as many law firms, as well as a handful of corporations and one governmental entity.

The interviews, conducted according to 10-page-long sets of talking points, lasted from 30 minutes to two hours.

We supplemented the interviews with an online survey whose content and organization tracked that of the telephone surveys with allowances made for the differences inherent in the two approaches.

We also gathered information from a variety of secondary sources. We collected data from Web sites for more than 180 organizations purporting to offer some form of electronic discovery services or software, as well as from a variety of business and information resources.

We entered much of the data we gathered into a database we developed to track and evaluate the information as well as spreadsheet models we prepared to allow us to rate and weight the information and develop aggregated results. For purposes of analysis, we divided electronic discovery activities into seven groupings:

• *Data identification:* This is the process of determining what electronic data should be obtained during discovery.

• *Data collection:* This process revolves around the actual preservation and collection of electronic information identified during the previous sub-segment's activities.

• *Data restoration:* This involves restoring electronic information from whatever medium it is on when it reaches the service provider, and making it accessible for use by the service provider and possibly the service provider's customer.

• *Data processing:* Data processing constitutes converting information from less-accessible forms to forms that can be more readily used, as well as reducing the volume of the dataset to a more useable size and limiting it to data with more meaningful content.

• *Data searching:* This area consists of additional filtering and searching of the electronic information, sometimes while it is still with the vendor and sometimes after it has been made available through a hosted or similar system. This area also includes work done to make sense of the data that has been gathered, looking for significant patterns and locating relevant personnel.

• *Data delivery:* At the end of the day, vendors generally need to deliver data to various recipients for use in other systems and on various media.

• *Consulting:* The final segment we included was consulting not tied to other revenue-generating projects.

We set out six sets of factors around which we built our analysis of the data collected during the survey. The sets of factors and their weightings were:

• *Experience/reputation* (20 percent): We considered the length and depth of the vendors' experience in the electronic discovery industry as well as their reputation among competitors and consumers.

• *Capacity* (10 percent): We looked at the level of resources that the vendors dedicated to the electronic discovery market.

• *Computer forensics/investigative* (30 percent): We examined the breadth and depth of the vendors' services dedicated to computer forensics and investigations. These services include data identification, collection, restoration, conversion and pre-processing. This area also includes forensic software.

• *EDD/litigation support* (20 percent): We examined the breadth and depth of the vendors' services dedicated to conversion of data to more useable formats and to searching, retrieving and analyzing electronic information. These services include data hosting, filtering, searching and analysis. This area also includes search software.

• *Consulting* (5 percent): We included advice given by vendors to customers about approaches to take with respect to retention and discovery of electronic information, but only if it was done apart from processing or similar engagements.

• *Revenue* (15 percent): We took into account estimated annual revenues for the vendors, calculated to exclude revenues for activities other than electronic discovery.

George Socha is an attorney and consultant, based in Minneapolis.
E-mail: george@sochaconsulting.com.

Tom Gelbmann is an information technology management consultant focused on the legal sector
E-mail: tom@gelbmann.biz.

For more information about the 2004 Socha-Gelbmann Electronic Discovery Survey visit www.sochaconsulting.com/2004survey.htm.