Earnings Guidance after Regulation FD

By

Ronen Feldman
Bar-Ilan University
(Details)

Ron Lazer and Joshua Livnat

Department of Accounting
Stern School of Business Administration
New York University
311 Tisch Hall
40 W. 4th St.
NY NY 10012
(212) 998 – 0022

rlazer@stern.nyu.edu
jlivnat@stern.nyu.edu

The authors acknowledge the access to data provided by Factset Information Services, Inc., and the assistance of Scott Starkin. The authors also gratefully acknowledge the contribution of Clearforest for providing

Earnings Guidance after Regulation FD

Abstract

This study investigates market reactions to voluntary earnings guidance provided by managers after the enactment of Regulation FD, which requires companies to disseminate material news to all investors simultaneously. More managers now issue their guidance to the public instead of a selective group of analysts, in conformity with Regulation FD. We examine a very large set of earnings guidance based on identification of these announcements using text mining techniques.

Our results indicate that guidance provided with the disclosure of earnings is not associated with significant market reactions, but guidance provided between earnings releases is associated with significant negative reactions. We further show that market reactions are consistent with the trend implied by management even when it is in the form of qualitative disclosure. Finally, we show that market reactions are stronger for NASDAQ firms than NYSE or AMEX firms, larger firms, and when the disclosure involves revenues and not earnings.

Earnings Guidance after Regulation FD

Regulation FD was supposed to change the way publicly-listed companies release information to market participants. If in the past managers could disclose their assessments and forecasts of future results to selected groups of analysts, under the new rules all material information must be disclosed to the public simultaneously. As a result, and concurrently with the development of webcasts, many companies began opening their conference calls to the public, so some of the information about future events and management expectations is reported to all potential investors who choose to listen in. Furthermore, many managers have begun issuing public guidance to the market in press releases, consistent with the requirements of Regulation FD, although companies vary considerably in the manner and timing of earnings guidance.

The purpose of this study is to examine earnings guidance practices of companies after Regulation FD and their associated effects on stock returns. In particular, this study investigates whether market reactions differ among earnings guidance that occurs at the time of the preliminary earnings announcements and those that occur outside the earnings announcements. The study also compares the effects of earnings guidance on stock returns for companies with different market capitalizations and exchange listing; presumably, smaller companies are less closely followed by analysts and management's guidance may be more effective in changing investors' assessments and consequently prices. Finally, we examine whether investors react differently to guidance about earnings and revenues.

To examine a large sample of earnings guidance disclosures, we use a text-mining approach to identify such announcements from press releases issued by companies. Text mining is similar to data mining, except that it relates to analysis of text documents using linguistic rules, instead of the typical analysis of numerical data typically used in data mining. Using text mining, we are able to identify over 3,400 earnings guidance disclosures during the period October 2000 through July 2002. Most studies to date have used substantially smaller samples, as described below.

The results of this study indicate that market participants react negatively to earnings guidance that occurs between earnings releases, consistent with the intuition that these types of voluntary disclosure are typically negative in nature, warning investors that future prospects are worse than previously expected. In contrast, earnings guidance provided with the earnings release or up to three days afterwards is unassociated with abnormal returns. Our results show that even when the guidance is set qualitatively, i.e., by indicating that earnings will be higher (lower), market participants react by increasing (decreasing) prices significantly. We further show that reactions are stronger to announcements made by NASDAQ firms (as compared to NYSE or AMEX firms), but are weaker for smaller companies than for larger companies. Finally, market reactions seem more pronounced to revenue disclosures than to earnings disclosures, possibly because managers who wish to forewarn investors about deteriorating conditions do so more often in terms of revenues than earnings.

II. Institutional Setting and Prior Research

Some companies find it valuable to provide market participants guidance about either revenues, earnings, earnings per share, or some other measures of profits (such as operating profits). Note that this is voluntary disclosure; management has no responsibility to pre-announce their expectations about future operations to investors. They are likely to do so if the benefits of this disclosure exceed the costs of the disclosure. In particular, managers who possess negative information about future financial results may provide guidance to lower levels of expectations in an attempt to reduce expected payouts in law suits brought against them, claiming they had material negative information which had not been timely released to the market. Some companies provide guidance consistently, whereas others do so occasionally or rarely. The restrictions imposed on private dissemination of guidance by Regulation FD are likely to cause an increase in the frequency of earnings guidance provided by companies.[1] Let us examine several examples of earnings guidance announcements.

On Oct. 17, 2000 Intel Corp. ("Intel") announced record results for the third quarter of 2000 and revised its outlook publication procedures in connection with the adoption of Regulation FD. Intel said it would keep its Outlook forward-looking statements and risk factors statements publicly available on its web site. Towards the end of each fiscal quarter, it would have a "Quiet Period" when it would not update Outlook, but prior to the start of the Quiet Period, the public can continue to rely on the Outlook as reflecting Intel's most current expectations The Quiet Period extends to the day when Intel's next quarterly earnings release is published. Six months later, On April 17, 2001

[1] One of the authors, Ron Lazer, is currently investigating this issue in another study.

Intel announced that “Beginning this quarter (second quarter, 2001) Intel will have a mid-quarter Business Update to the Outlook provided…”. Currently, during Intel’s earnings quarterly press releases, Intel also announces the date for its mid-quarter update. On July 17, 2001 Intel announced the results for the second quarter of 2001 and its business outlook for the third quarter of 2001: “Revenue in the third quarter of 2001 is expected to be between $6.2 billion and $6.8 billion.” In addition, Intel announced that it “plans to provide a mid-quarter Business Update to the Outlook provided below on Sept. 6.” On Sept. 6, 2001 Intel announced the mid quarter update, stating it “expects revenue for the third quarter to be within the previous expectation and slightly below the midpoint of the range provided on July 17.” The release included updates regarding other expenses and capital expenditures. Thus, Intel provides systematic forecasts with the earnings release, and an update between earnings releases with a known date for the update.

On Oct. 29, 2001 FedEx Corp. (“FedEx”) updated its earnings outlook for the second quarter, ending Nov. 30, 2001. In its press release, FedEx announced it “expects to earn 40 cents to 45 cents a share excluding its slice of the aid package and 61 cents to 66 cents a share including the $101 million in government assistance.” On Dec. 19, 2001 FedEx reported the results for the second quarter (ending Nov. 30, 2001). Earnings per share (not included the compensation from the Air Transportation Safety and System Stabilization Act) were 57 cents per share. In its press release, FedEx also included earnings forecasts for the third and the fourth quarters: “we now expect earnings for the third quarter to be $0.25 to $0.35 per diluted share, and earnings for the fourth quarter to be $0.70 to $0.80 per diluted share.“ The first announcement on October 29, 2001 occurred between earnings announcement and was intended to provide investors with a profit warning. The second guidance coincided with the earnings announcement. Unlike for Intel, investors could not have anticipated when the first announcement would be made, if at all.

Prior Research:

Prior research has generally focused on the motivation for companies' pre-announcements, and on the market reaction to these pre-announcements. Soffer, Thiagarajan and Walther (2000) examine the factors influencing the decision of a firm to voluntarily accelerate the release of earnings via a "pre-announcement". They find that firms are more likely to preannounce earnings if the consensus of analyst' forecasts is very different from actual earnings, if the dispersion of these forecasts is high, and if the firm has negative news. Skinner (1994) argues that managers pre-announce to prevent lawsuits caused by large stock price declines at the bad earnings announcement, and to protect their reputation with analysts and institutional investors by not delaying bad news. Using a random sample of 93 NASDAQ firms during 1981-90, he finds that earnings-related voluntary disclosure occur infrequently (on average, one disclosure for every ten quarterly earnings announcements), and that good news disclosures tend to be point or range estimates of annual earnings-per-share, whereas bad news disclosures tend to be qualitative statements about the current quarter's earnings. He also indicates that the stock price responses to bad news disclosures are stronger than the responses to good news disclosures. In his sample, quarterly earnings announcements that convey large negative earnings surprises are preempted about 25% of the time by voluntary corporate disclosures while other earnings announcements are preempted less than 10% of the time. Skinner suggests that managers may attempt to avoid a negative surprise on the earnings announcement date because of litigation-related costs, as well as costs of diminished reputation with money managers and institutional investors.

Kasznik and Lev (1995) look at management's discretionary disclosures prior to firms that experience large earnings surprises. They find that less than ten percent of their large-surprise firms published quantitative earnings or sales forecasts, and 50% of the firms remained, providing no earnings guidance. Heflin, Subramanyam and Zhang (2001) examine voluntary disclosure after the implementation of Regulation FD and find (a) lower return volatility around earnings announcements; (b) some improvement in the speed with which the price before the earnings announcement converges to its post announcement level; (c) no reliable evidence of changes in various aspects of analysts forecast bias, accuracy, and dispersion; and (d) an increase in the quantity of firms' voluntary forward looking disclosures.

Miller (2001) examines a sample of 80 companies experiencing an extended period of seasonally adjusted earnings increases. Using 416 observations of forecasts and earnings pre-announcements over a period of 3 years, He finds an increase in disclosure during the period of increased earnings. The increase tends to be bundled with earnings announcements, and the market responds positively to this disclosure. Coller and Yohn (1997) use a sample of 278 quarterly earnings forecasts, including point estimates, range estimates, and upper and lower bound estimates to examine whether the decision to issue a management earnings forecast is related to information asymmetry in the market for the firm's stock, and whether the forecasts reduce the information asymmetry. They find support to the view that managers release forecasts to reduce information asymmetry. Johnson, Kasznik and Nelson (2001) use 1135 forecasts in 1994 and 1996 to evaluate corporate voluntary disclosure of forward looking information under the Safe Harbor provision of the Private Securities Litigation reform Act of 1995. They find that managers engage in more disclosure after the rule enactment, and that the increase in disclosure is an increasing function of the firms' ex-ante risk of litigation.

Need to rearrange according to motives, and then discuss market reactions better.

III. Text Mining and the Ability to Identify Earnings Guidance

Text mining is a new area of research and application that addresses the information overload problem. It uses techniques from the general field of data mining to process text documents, extract relevant information from these documents, and provide the user with this information obtained using report or graphical representations. The application described in this study processed all the text documents of Comtex (***Ronen, we need to explain and provide a source) related to publicly-listed business entities. It identified those documents that included earnings guidance, and was used to form our sample.

A typical Text Mining system begins with collections of raw documents, without any labels or tags. Documents are first automatically tagged by Categories, Terms or Relationships extracted directly from the documents. Next, extracted Categories, Entities and Relationships are used to support a range of data mining operations on the documents. Text categorization is concerned with partitioning a large collection of documents into subsets that are interrelated by some pre-defined criteria. For instance, the Yahoo web-browser categorizes the whole web into areas such as “News and Media”, “Science”, “Arts” etc. Each document in this large collection is tagged by words characteristic of categories, which enables the association of the document (or web-site) with its relevant categories. Limiting the set of documents for mining to certain relevant sub-categories makes the follow-up tasks easier for the mining tools, and increases the likelihood that these tools will extract the most on-target bits of information from the text. The actual detection of facts within the text is typically performed through information extraction methods.

Information extraction is performed by combining natural language processing tools, lexical resources and semantic constraints. Complementary visualization tools enable the user to explore, check (and correct if required) the results of the process effectively. As a first step in tagging documents, each document is processed to find (extract) Entities and Relationships that are likely to be meaningful and content-bearing. In “Relationships” we refer to Facts or Events involving certain Entities. A possible “Event” may be that a company has entered into a joint venture. A “Fact” may be that a gene causes a certain disease. The extracted information provides more concise and precise data for the mining process than the more naive word-based approaches such as those used for text categorization, and tends to represent concepts and relationships that are more meaningful and relate directly to the examined document’s domain. ,Consequently the information extraction methods allow for mining of the actual information present within

Same **:[1k]Comment** demoib a esu ebyam –comment .hisexampel for t the text, rather than the limited set of tags associated with the documents. Using the information extraction process, the number of different relevant Entities and Relationships on which the data mining is performed is unbounded, typically thousands or even millions, far beyond the number of tags which any automated categorization system could handle.

<u>Tax Mining for Earnings Guidance:</u>

To identify cases of earnings guidance, we use the following rules which search for certain patterns that involve the identification of certain elements in a document. Let us show several examples of such patterns[2]:

Pattern1: ResultsPhrase FinancialQuarter [FinancialAmount] Trend Expectations

Example: "Akorn, Inc. (Nasdaq: AKRN) today announced that earnings for the quarter ending September 30, 2000 will be significantly below expectations"

ResultsPhrase : earnings
FinancialQuarter : the quarter ending September 30, 2000
Trend : will be significantly below
Expectations : expectations

Pattern2: FinancialQuarter ResultsPhrase FinancialAmount Trend Expectations

Example: "INTERLINQ Software Corporation (Nasdaq:INLQ) today announced that it anticipates first-quarter revenue will fall below the company's expectations."

FinancialQuarter : first-quarter
ResultsPhrase : revenue
FinancialAmount :
Trend : will fall below
Expectations : the company's expectations

Example: "Drypers Corporation (Nasdaq: DYPR) today announced that its third quarter revenues and earnings will fall below expectations."

FinancialQuarter : third quarter
ResultsPhrase : revenues and earnings
FinancialAmount :

[2] Text in square brackets indicates an element that may not necessarily appear in the specific document.

Trend : will fall below
Expectations : expectations

Pattern3: Expectations ResultsPhrase FinancialAmount FinancialQuarter [Trend]

Example: "The Company expects earnings per share to be approximately $0.04 per share for the third quarter."

Expectations : The Company expects
ResultsPhrase : earnings
FinancialAmount : $0.04 per share
FinancialQuarter : third quarter

Example: "Featherlite, Inc. (Nasdaq: FTHR), a leading manufacturer and marketer of specialty aluminum trailers and luxury motorcoaches, said today it expects to report a net loss of 10 cents per diluted share for its third quarter ended Sept. 30, which is below previous expectations"

Expectations : it expects
ResultsPhrase : net loss
FinancialAmount : 10 cents per diluted share
FinancialQuarter : third quarter ended Sept. 30
Trend: which is below previous expectations

Pattern4: ResultsPhrase Trend FinancialQuarter FinancialAmount

Example: "Sawtek, which is being acquired by Triquint, also said it sees weaker profit ahead due to a slowdown in the wireless-communications sector"

ResultsPhrase : profit
Trend: weaker
FinancialQuarter : ahead
FinancialAmount :

These and similar rules are used by the text mining software to identify earnings guidance. However, the system may yield announcements that are not necessarily earnings guidance, nor is the software likely to identify all guidance announcements. Thus, we need to impose some additional restrictions to ensure that identified announcements are indeed earnings guidance announcements.

Additional Filters:

During the period October 2000 through July 2002, we identified 16,026 documents that were classified as earnings guidance announcements of companies that were traded on the NYSE, ASE or NASDAQ. Analysis of a sub-sample of these announcements led us to put some restrictions on results to ensure that the announcements indeed relate to earnings guidance. We have classified an announcement as earnings guidance if it related to a quarter beyond the most recently announced quarter, or to a year beyond the current year. We further classified an announcement as earnings guidance if the text mining rules extracted the trend from the announcement, i.e., whether the item was expected to be "Higher", "Lower", or "Unchanged". Finally, we classified an announcement as earnings guidance if the text mining rules extracted a range for the amount, e.g., $640-$660 million (in revenues). These restrictions resulted in 7,392 announcements that were identified as earnings guidance, although it is possible that the dropped announcements were also in fact earnings guidance. We further deleted any earnings guidance announcement that was within three days of an earlier announcement of earnings guidance for the same company. This reduced the sample to 4,460 announcements, which were further reduced to 3,459 announcements for which we could obtain size-adjusted returns from the price database maintained by Factset Information Services, Inc. This last set of announcements is the sample that is used in the remainder of this study. Note that this is a very large set of earnings guidance announcements. Most prior studies examined much smaller samples.[3]

[3] Roni, Please add the summary about some examples of sample sizes in prior studies.

IV. Results

The sample consists of 3,495 announcements made by 1,788 distinct firms, with mean (median) market capitalization at the end of 2001 of $4,809 ($430) million. Thus, the sample companies are representative of the mid-cap to large-cap sectors of the market, although more than 25% of the announcements are issued by firms with market capitalization of below $100 million on 12/31/2001. Table 1 reports the frequencies of the number of announcements by firms and the median market value as of 12/31/2001. As can be seen in the table, over 50% of the companies in our sample are captured with a single announcement. There is a clear correlation between the number of announcements captured in our sample and the market capitalization of a company; larger firms have typically more appearances in our sample. The largest numbers of announcements in our sample are made by (in descending order) Intel, Microsoft, Yahoo, and Nortel.

(Insert Table 1 about here)

Table 2 reports the distribution of firms in the various 2-digit SIC industries. The table lists only those industries with at least 18 firms (above 1%) of all companies in our sample. As can be seen in the table, there is a reasonably uniform distribution across industries, except for concentrations in the software industry (73, 22.3%), electronics (36, 11.7%), machinery and computer equipment (35, 8%), measuring instruments (38, 5.7%), and chemicals (28, 5.4%). These are likely to be the industries that suffered most in the economic downturn during the sample period, which probably led to many earnings warnings issued by companies.

(Insert Table 2 about here)

Table 3 reports the size-adjusted returns for the three-day window centered on the announcement date. To calculate size-adjusted returns, we first calculate the three-day cumulative return for the announcing firm. Based on its market capitalization of equity (size) at the end of the previous quarter, we assign it to one of ten portfolios ranked on size. We calculate the equally-weighted three-day cumulative return for all companies in the same size decile as the announcing firm, and subtract this return from the announcing firm's return. The table shows that the average preannouncement is associated with a negative size-adjusted return of 2.9%, which is significantly different from zero at a level below 0.001. This is consistent with prior research, intuition, and our knowledge about economic conditions during the sample period, which were deteriorating for most companies, and were associated with negative news during the guidance period. Note, however, that not all announcements were treated by the market with the same force – companies providing guidance during the earnings release period (i.e., within three days of the earnings release date) had a negative average size-adjusted return of 0.5%, which was not significantly different from zero. In contrast, those announcements that occurred between earnings release dates were associated with an average size-adjusted negative return of 4.7%, significantly different from zero at a level below 0.001. Thus, on balance, announcing firms had negative size-adjusted returns, but more so if these announcements were not made during the earnings release period.

(Insert Table 3 about here)

Table 4 reports size-adjusted returns for announcements that indicated trend in earnings (or revenues, or whatever metric was used), and typically did not include any quantitative guidance as well. There were a total of 949 cases where a trend was indicated, with 323 announcements indicating a higher trend, 515 a lower trend, and 111 indicating unchanged trend. The table reports the mean size-adjusted returns for each of these categories. It clearly shows that the market reacted positively and significantly to announcements indicating a higher trend, negatively and significantly to announcements of lower trend, and insignificantly different from zero for those announcements that reported an unchanged trend.

(Insert Table 4 about here)

Table 5 is similar to Table 4 but splits the trend announcements into those that were made between earnings releases (stand-alone announcements), and those that were within three days of the earnings release. As can be seen, positive announcements and negative announcements are associated with significant returns in the same direction when these announcements are made outside the usual earnings releases. In contrast, only negative announcements are associated with significant negative returns when the guidance is provided within three days of the earnings release. Positive guidance provided with the earnings announcement causes no significant return. Consistent with Table 4, announcements that indicated unchanged trend did not cause any significant return whether between earnings announcements or with earnings releases.

(Insert Table 5 about here)

Table 6 shows the size-adjusted returns by exchange listing and separately for announcements made between earnings releases and guidance provided with the earnings release. As can be seen from the table, the mean size-adjusted return is typically more negative for NASDAQ-listed companies than the reaction to the announcements made by NYSE or AMEX companies. This may be explained by the richer information environment for NYSE and AMEX companies, which implies that company announcements and guidance may have been anticipated to a greater extent by market participants, and therefore caused a lesser reaction.

(Insert Table 6 about here)

Table 7 reports information about the size-adjusted returns of companies sorted by market capitalization at the end of the previous quarter (size). Panel A reports the average return for all announcements, whereas Panel B reports the returns separately for announcements made between earnings releases and those with earnings releases, after aggregating deciles to obtain sufficient number of observations. As can be seen from the table, typically, larger companies have more negative market reactions to their earnings guidance than smaller firms. This is particularly noticeable for announcements made between earnings releases and deciles 7-10, where the reaction is insignificantly different from zero, but is negative and significant for all other deciles. Thus, the exchange listing results we saw earlier may not be due to size alone but also probably to industry association; companies in high-tech areas (higher proportion on NASDAQ) had more negative market reactions than low-tech companies.

(Insert Table 7 about here)

Table 8 shows the size-adjusted returns for announcements classified by whether they include guidance about revenues or other metrics. For simplicity and ease of exposition, we group all revenue announcements together, even if they also contain guidance about earnings-related items. All other announcements are classified as earnings, even if they refer to earnings per share, operating profits, pretax profits, etc. As can be seen from the table, revenue announcements are typically associated with stronger market reactions; whether the announcements are made with earnings or between earnings release dates. This may be due to the tendency of negative news about deteriorating operations to be framed in terms of revenues rather than earnings.

(Insert Table 8 about here)

V. Summary and Conclusions

Table 1
Distribution of Announcements by Firms

No. of Announcements	No. of Firms	Median Market Capitalization
1	949	337
2	398	392
3	205	667
4	88	670
5	49	814
6	20	1800
7	19	3978
8	6	19600
9	5	16631
10	3	8215
12	1	23975
13	1	10210
14	1	357949
20	1	210401

Table 2
Distribution of Firms by Industries

SIC	No. of firms	% of Total
13	21	1.2%
20	26	1.5%
27	24	1.3%
28	97	5.4%
33	22	1.2%
34	20	1.1%
35	143	8.0%
36	209	11.7%
37	39	2.2%
38	102	5.7%
48	63	3.5%
49	43	2.4%
50	30	1.7%
51	23	1.3%
56	27	1.5%
59	46	2.6%
60	43	2.4%
63	40	2.2%
73	398	22.3%
87	36	2.0%
All Others	336	18.8%
Total	1788	100.0%

Table 3
Size-Adjusted Returns and Timing of Announcement

	Average Size-Adjusted Return	**N**	**Significance**
Guidance between earnings announcements	-4.7%	1957	0.001
Guidance with earnings release	-0.5%	1502	0.112
All guidance announcements	-2.9%	3459	0.001

Notes:

1. Size-adjusted returns are calculated for the three-day window centered on the announcement date. We first calculate the cumulative return for each firm in our sample. We then calculate the equally-weighted mean cumulative return for the same size (market capitalization at the beginning of the quarter) decile. Size-adjusted returns are the return on the company minus the return on the same size decile portfolio.
2. N represents the number of announcements.
3. Significance is the significance level obtained in a t-test that the mean size-adjusted return is zero.
4. Guidance announcements can coincide with preliminary earnings announcements (if they are made within three days of the quarterly earnings announcement), or between earnings announcements.

Table 4
Size-Adjusted Returns and Indicated Trend

Indicated Trend	Average Size-Adjusted Return	N	Significance
Higher	1.6%	323	0.025
Lower	-7.0%	515	0.001
Unchanged	-0.3%	111	0.804
Total	-3.3%	949	0.001

Notes:

1. Size-adjusted returns are calculated for the three-day window centered on the announcement date. We first calculate the cumulative return for each firm in our sample. We then calculate the equally-weighted mean cumulative return for the same size (market capitalization at the beginning of the quarter) decile. Size-adjusted returns are the return on the company minus the return on the same size decile portfolio.
2. N represents the number of announcements.
3. Significance is the significance level obtained in a t-test that the mean size-adjusted return is zero.
4. The table reports size-adjusted returns for companies that provided guidance about the trend in their earnings (or revenues, or any other metric they used), and typically did not provide any numerical guidance (either an exact number or a range). Only 64 of the trend announcements included any quantitative data as well.

Table 5
Size-Adjusted Returns and Indicated Trend

	Announcement Between Earnings Reports			Guidance Together With Earnings		
Indicated Trend	Average Size-Adjusted Return	N	Significance	Average Size-Adjusted Return	N	Significance
Higher	1.9%	241	0.030	0.9%	82	0.491
Lower	-7.4%	444	0.001	-4.2%	71	0.003
Unchanged	0.6%	80	0.707	-2.5%	31	0.243
Total	-3.7%	765	0.001	-1.6%	184	0.068

Notes:

1. Size-adjusted returns are calculated for the three-day window centered on the announcement date. We first calculate the cumulative return for each firm in our sample. We then calculate the equally-weighted mean cumulative return for the same size (market capitalization at the beginning of the quarter) decile. Size-adjusted returns are the return on the company minus the return on the same size decile portfolio.
2. N represents the number of announcements.
3. Significance is the significance level obtained in a t-test that the mean size-adjusted return is zero.
4. The table reports size-adjusted returns for companies that provided guidance about the trend in their earnings (or revenues, or any other metric they used), and typically did not provide any numerical guidance (either an exact number or a range).
5. The table shows separately the size-adjusted returns for companies that made their guidance announcements between earnings releases (left-most three columns) and those that provided guidance within three days of the earnings disclosure (three right-most columns).

Table 6
Size-Adjusted Returns and Exchange Listing

	Announcement Between Earnings Reports			Guidance Together With Earnings		
Exchange Listing	**Average Size-Adjusted Return**	**N**	**Significance**	**Average Size-Adjusted Return**	**N**	**Significance**
NYSE or AMEX	-2.3%	753	0.001	0.6%	569	0.075
NASDAQ	-6.2%	1204	0.001	-1.2%	933	0.013
Total	-4.7%	1957	0.001	-0.5%	1502	0.112

Notes:

1. Size-adjusted returns are calculated for the three-day window centered on the announcement date. We first calculate the cumulative return for each firm in our sample. We then calculate the equally-weighted mean cumulative return for the same size (market capitalization at the beginning of the quarter) decile. Size-adjusted returns are the return on the company minus the return on the same size decile portfolio.
2. N represents the number of announcements.
3. Significance is the significance level obtained in a t-test that the mean size-adjusted return is zero.
4. The table shows separately the size-adjusted returns for companies that made their guidance announcements between earnings releases (left-most three columns) and those that provided guidance within three days of the earnings disclosure (three right-most columns).

Table 7
Size-Adjusted Returns and Company Size (Market Capitalization)

Panel A:

Decile	Average Size-Adjusted Return	N	Significance
1 (Largest)	-1.9%	957	0.001
2	-3.1%	605	0.001
3	-4.5%	576	0.001
4	-4.2%	464	0.001
5	-3.7%	348	0.001
6	-2.8%	240	0.010
7	-0.1%	142	0.942
8	3.0%	84	0.349
9	-0.5%	34	0.852
10 (Smallest)	2.8%	9	0.805
Total	-2.9%	3459	0.001

Panel B:

	Announcement Between Earnings Reports			**Guidance Together With Earning**		
Decile	**Average Size-Adjusted Return**	**N**	**Significance**	**Average Size-Adjusted Return**	**N**	**Significa**
1-2	-3.5%	947	0.001	-0.6%	615	0.
3-4	-6.9%	565	0.001	-1.3%	475	0.(
5-6	-5.8%	318	0.001	-0.4%	270	0.:
7-10	-0.5%	127	0.778	2.1%	142	0.:
Total	-4.7%	1957	0.001	-0.5%	1502	0.

<u>Notes:</u>

1. Decile 1 consists of companies placed in the top 10% according to market value of equity at the end of the previous quarter. Decile 10 consists of the smallest companies.
2. For other variables, see notes to Table 6.

Table 8
Size-Adjusted Returns and Metric Used in Guidance

Panel A:

Metric	Average Size-Adjusted Return	N	Significance
Earnings	-2.4%	2029	0.001
Revenues	-3.5%	1430	0.001
Total	-2.9%	3459	0.001

Panel B:

	Announcement Between Earnings Reports			Guidance Together With Earning		
Metric	**Average Size-Adjusted Return**	**N**	**Significance**	**Average Size-Adjusted Return**	**N**	**Significa**
Earnings	-4.1%	1175	0.001	-0.2%	854	0.[illegible]
Revenues	-5.6%	782	0.001	-1.0%	648	0.[illegible]
Total	-4.7%	1957	0.001	-0.5%	1502	0.[illegible]

<u>Notes:</u>

1. The metric is what mangers use to guide investors. For simplicity, any time revenues were mentioned, the announcement is classified as revenues, even if it contains some other performance measures (such as earnings, operating profits, etc.). All other metrics are dclassified here as earnings, even if they refer to earnings per share, pretax profits, etc.
2. For other variables, see notes to Table 6.

Appendix

1.1.1 Architecture of Information Extraction Systems

Each Information Extraction system has three to four major components. The first component is Tokenization or zoning, splitting the document into its basic building blocks. The typical building blocks are: words, sentences and paragraphs. In some rare occasion we may have higher building blocks like sections and chapters. Identifying sentence boundaries can be a tricky task since just looking for a "." followed by a capital letter can lead to false sentence boundaries (like the dot in Dr. Ronen Feldman). The second component is the morphological and lexical analysis, the assignment of Part of Speech (POS) tags to the words, creation of basic phrases (like noun phrases and verb phrases) and disambiguating the sense of ambiguous words and phrases. The third component is syntactic analysis, establishing the connection between the different parts of each sentence. This is done by either doing full parsing or shallow parsing. We will elaborate on this component in section 2.1.2. The forth component is the Domain Analysis, where we combine together all the information we have collected from the previous components and create complete frames that describe relationships between entities. The domain analysis contains also an anaphora resolution component (see 2.1.3). Figure 1 show the architecture of a basic IE system. The subcomponents are colored according to their necessity within the full system. We will now elaborate on some of the subcomponents that we mentioned above.

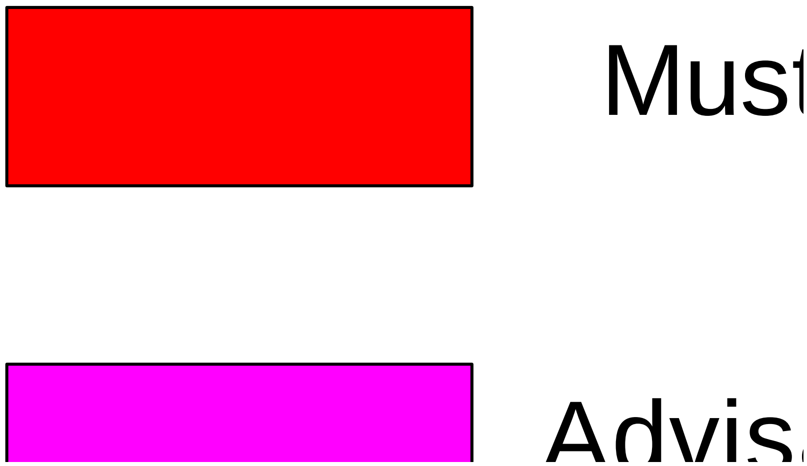

Figure 1 - Architecture of a typical Information Extraction system

DIAL (Declarative Information Analysis Language)

In this subsection we describe DIAL. DIAL is designed specifically for writing IE (information extraction) rules. The complete syntax of DIAL is beyond the scope of this chapter. Here we describe the basic elements of the language.

Basic Elements. The basic elements of the language are syntactic and semantic elements of the text and sequences and patterns thereof. Among these elements the language can identify the following:

- Predefined strings such as “*merger*”

- Word class element: a phrase from a predefined set of phrases that share a common semantic meaning – for example `WC-Countries,` a list of countries.
- Scanner feature (basic characteristic of a token), for example `@Capital` or `@HtmlTag`
- Compound feature: a phrase comprising several basic features. Thus, `Match(@Capital & WCCountries)`, for example, will match a phrase that both belongs to the word class `WCCountries` and starts with a capital letter.
- Part-of-speech tag – for example noun or adjective
- Recursive predicate call – for example Company(C)

Constraints. Constraints carry out on-the-fly Boolean checks for specific attributes. These can be applied to fragments of the original text or to results obtained during processing extraction process.

The marker for a constraint is the word *`verify`*, followed by parentheses containing a specific function, which governs what it is checking for. For example:

```
verify ( StartNotInPredicate ( c , @PersonName ) )
```

ensures that no prefix of the string assigned to variable `c` is a match for the `predicate PersonName`.

An example of a DIAL rule is the following, which is 1 of 10 rules to identify a merger between two companies:

```
FMergerCCM(C1, C2) :-
```

```
Company(Comp1) OptCompanyDetails "and" skip(Company(x),
SkipFail,    10)    Company(Comp2)    OptCompanyDetails
skip(WCMergerVerbs,   SkipFailComp,   20)   WCMergerVerbs
skip(WCMerger, SkipFail, 20) WCMerger
verify(WholeNotInPredicate(Comp1, @PersonName))
verify(WholeNotInPredicate(Comp2, @PersonName))
@% @!
{ C1 = Comp1; C2 = Comp2} ;
```

The rule looks for a company name (carried out by the predicate `Company`, which returns the parameter `Comp1`) followed by an optional phrase describing the company, and then the word *and*. The system then skips up to 10 tokens (within the same sentence, and while not encountering any phrase prescribed by the predicate `SkipFail`) until it finds another company, followed by an optional company description clause. The system then skips up to 20 tokens until it finds a phrase of the word class `WCMergerVerbs`. (This may be something like “approved,” “announced,” etc.). Finally, the system skips up to 10 tokens scanning for a phrase of the word class `WCMerger`. In addition, the rule contains two constraints ensuring that the company names are not names of people.

Each rulebook can contain any number of rules that are used to extract knowledge from documents in a certain domain. The financial rulebook contains 11,500 rules, can identify more than 50 different entity types including company names; people names; organizations; universities; products; positions; locations (cities, countries, states, and addresses); dates, and amounts. In addition it can identify more than 120 different Event types such as: mergers (including a fine-grained distinction between known merger, new merger, rumored merger, planned merger, and cancelled merger); acquisitions (with a similar distinction between acquisition types), joint ventures; takeovers; business relationships; investment relationships; customer-supplier relationships; new product introductions; analyst recommendations for stocks and bonds, associations between companies and people; associations between companies and technologies; associations between companies and products; and many others.

Bibliography

1. Allen J. (1995). “Natural Language Understanding”, Addison-Wesley.
2. Aumann Y., Feldman R., Ben Yehuda Y., Landau D., Lipshtat O., Schler Y. “Circle Graphs: New Visualization Tools for Text-Mining”. PKDD 1999: 277-282
3. Cardie C. (1997). “Empirical Methods in Information Extraction”, AI Magazine, 18, #4, pp. 65-80.
4. Charniak E. (1993). “Statistical Language Learning”, MIT Press.
5. Cohen W. and Singer Y., 1996. Context Sensitive Learning Methods for Text categorization. In Proceedings of SIGIR’96.
6. Cohen. W., "Compiling Prior Knowledge into an Explicit Bias". Working notes of the 1992 AAAI spring symposium on knowledge assimilation. Stanford, CA, March 1992.
7. Cowie J., and Lehnert W., "Information Extraction," Communications of the Association of Computing Machinery, vol. 39 (1), pp. 80-91.
8. Daille B., Gaussier E. and Lange J.M., 1994. Towards Automatic Extraction of Monolingual and Bilingual Terminology, In Proceedings of the International Conference on Computational Linguistics, COLING’94, pages 515-521.
9. Dumais, S. T., Platt, J., Heckerman, D. and Sahami, M. Inductive learning algorithms and representations for text categorization. Proceedings of the Seventh International Conference on Information and Knowledge Management (CIKM’98), 148-155, 1998.
10. Feldman R. and Dagan I., 1995. KDT – Knowledge Discovery in Texts. In Proceedings of the First International Conference on Knowledge Discovery, KDD-95.
11. Feldman R., and Hirsh H., 1996. Exploiting Background Information in Knowledge Discovery from Text. Journal of Intelligent Information Systems. 1996.
12. Feldman R., Aumann Y., Amir A., Klösgen W. and Zilberstien A., 1997. Maximal Association Rules: a New Tool for Mining for Keyword co-occurrences in Document Collections, In Proceedings of the 3rd International Conference on Knowledge Discovery, KDD-97, Newport Beach, CA.

13. Feldman R., Aumann Y., Finkelstein-Landau M., Hurvitz E., Regev Y., Yaroshevich. “A Comparative Study of Information Extraction Strategies”. CICLing 2002: 349-359
14. Feldman R., Aumann Y., Liberzon Y., Ankori K., Schler J., Rosenfeld B. “A Domain Independent Environment for Creating Information Extraction Modules”. CIKM 2001: 586-588
15. Feldman R., Rosenfeld B., Stoppi J., Liberzon Y. and Schler, J., 2000. “A Framework for Specifying Explicit Bias for Revision of Approximate Information Extraction Rules”. KDD 2000: 189-199.
16. Fisher D., Soderland S., McCarthy J., Feng F. and Lehnert W., "Description of the UMass Systems as Used for MUC-6," in Proceedings of the 6th Message Understanding Conference, November, 1995, pp. 127-140.
17. Frantzi T.K., 1997. Incorporating Context Information for the Extraction of Terms. In Preceedings of ACL-EACL’97 .
18. Frawley W. J., Piatetsky-Shapiro G., and Matheus C. J., 1991. Knowledge Discovery in Databases: an Overview. In G. Piatetsky-Shapiro and W. J. Frawley, editors, Knowledge Discovery in Databases, pages 1-27, MIT Press.
19. Grishman R., The role of syntax in Information Extraction, In: Advances in Text Processing: Tipster Program Phase II, Morgan Kaufmann, 1996.
20. Hayes P. 1992. Intelligent High-Volume Processing Using Shallow, Domain-Specific Techniques. Text-BasedIntelligent Systems: Current Research and Practice in Information Extraction and Retrieval. New Jersey, P.227-242.
21. Hayes, P.J. and Weinstein, S.P. CONSTRUE: A System for Content-Based Indexing of a Database of News Stories. Second Annual Conference on Innovative Applications of Artificial Intelligence, 1990.
22. Hobbs, J. (1986), Resolving Pronoun References, In B. J. Grosz, K. Sparck Jones, & B. L. Webber (Eds.), Readings in Natural Language Processing (pp. 339-352), Los Altos, CA: Morgan Kaufmann Publishers, Inc.
23. Hofmann T. (1999). “The Cluster-Abstraction Model: Unsupervised Learning of Topic Hierarchies from Text Data”, Proc. of IJCAI’99, pp. 682-687.
24. Joachims, T. Text categorization with support vector machines: Learning with many relevant features. Proceedings of European Conference on Machine Learning (ECML’98), 1998
25. Larkey, L. S. and Croft, W. B. 1996. Combining classifiers in text categorization. In Proceedings of SIGIR-96, 19th ACMInternational Conference on Research and Development in Information Retrieval (Zurich, CH, 1996), pp. 289–297.
26. Leek T. R. Information extraction using hidden Markov models. M.Sc. thesis, Dept. of Computer Science, University of Califonia, San Diego, 1997.
27. Lehnert W., Cardie C., Fisher D., Riloff E., and Williams R. “Description of the CIRCUS System as Used for MUC-3”, Proceedings, Third Message Understanding Conference (MUC-3), San Diego, California, pp. 223-233 .
28. Lent B., Agrawal R. and Srikant R., 1997. Discovering Trends in Text Databases. In Proceedings of the 3rd International Conference on Knowledge Discovery, KDD-97, Newport Beach, CA.
29. Lewis, D. D. 1995a. Evaluating and optmizing autonomous text classification systems. In Proceedings of SIGIR-95, 18th ACMInternational Conference on

Research and Development in Information Retrieval (Seattle, US, 1995), pp. 246–254.
30. Lewis, D. D. and Hayes, P. J. 1994. Guest editorial for the special issue on text categorization. ACMT ransactions on Information Systems 12, 3, 231.
31. Lewis, D. D. and Ringuette, M. 1994. A comparison of two learning algorithms for text categorization. In Proceedings of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval (Las Vegas, US, 1994), pp. 81–93.
32. Lewis, D. D., Schapire, R. E., Callan, J. P., and Papka, R. 1996. Training algorithms for linear text classifiers. In Proceedings of SIGIR-96, 19th ACMInternational Conference on Research and Development in Information Retrieval (Z¨urich, CH, 1996), pp. 298–306.
33. Manning C.D. and Schutze H. (1999). “Foundations of Statistical Natural Language Processing”, MIT Press.
34. Rajman M. and Besançon R., 1997. Text Mining: Natural Language Techniques and Text Mining Applications. In Proceedings of the seventh IFIP 2.6 Working Conference on Database Semantics (DS-7), Chapam & Hall IFIP Proceedings serie. Leysin, Switzerland, Oct 7-10, 1997.
35. Riloff E. and Lehnert W. Information Extraction as a Basis for High-Precision Text Classification, ACM Transactions on Information Systems (special issue on text categorization) or also Umass-TE-24, 1994.
36. Salton G. (1989). “Automatic Text Processing: The Transformation, Analysis and Retrieval of Information by Computer”, Addison-Wesley.
37. Sebastiani F. Machine learning in automated text categorization, ACM Computing Surveys, Vol. 34, number 1, pages 1-47, 2002.
38. Sheila Tejada, Craig A. Knoblock and Steven Minton, Learning Object Identification Rules For Information Integration, Information Systems Vol. 26, No. 8, 2001, pp. 607-633 .
39. Soderland S., Fisher D., Aseltine J., and Lehnert W., "Issues in Inductive Learning of Domain-Specific Text Extraction Rules," Proceedings of the Workshop on New Approaches to Learning for Natural Language Processing at the Fourteenth International Joint Conference on Artificial Intelligence, 1995.
40. Vapnik, V., Estimation of Dependencies Based on Data [in Russian], Nauka, Moscow, 1979. (English translation: Springer Verlag, 1982(.
41. Vapnik, V., The Nature of Statistical Learning Theory, Springer-Verlag, 1995.
42. Yang, Y. and Chute, C. G. 1994. An example-based mapping method for text categorization and retrieval. ACMT ransactions on Information Systems 12, 3, 252–277.
43. Yang, Y. and Liu, X. 1999. A re-examination of text categorization methods. In Proceedings of SIGIR-99, 22nd ACM International Conference on Research and Development in Information Retrieval (Berkeley, US, 1999), pp. 42–49.