How Data travel through the market?

——Understand how financial information are used as "data" in the market, and explore how future accounting standard and disclosure should be——

Sharing discussion of "Disclosure data user workshop group" in Japan

Nomura Research Institute
Senior researcher Chie Mitsui

0

About "Disclosure data user workshop group"

We are User group to discuss about issues of corporate disclosure in Japan.

We usually have series of discussions about IFRS financial statements, but this time we picked up special agenda, to discuss "how data are extracted from financial statements, based on XBRL format, how they go around the market, and for what purposes".

Today I would like to share what we recognized of current situations, and what kind of issues are around us.

Date: Jan 10th 2018

Attendees:

Investors (including Pension fund), Information provider/Media/Researcher, Sell-side/ credit /insurance Analyst, CPA, Company side (Accounting Department), Academic, Regulator/Accounting setter, Analyst organization.
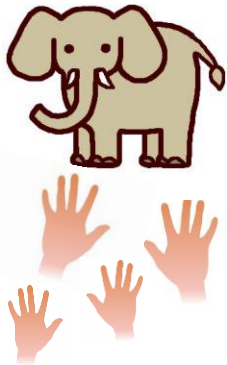
Some foreign organizations, (investor's organization etc) joined from overseas via Phone

The whole discussion result is posted on CFA research website

https://www.arx.cfa/post/How-Data-travel-through-the-market-5065.html

How Data travel through market and its influence

You may imagine that analysts read financial statements prepared by company, evaluate company's value and compare with peer companies….BUT, in reality, financial statements are used in various ways. **There are many quant analysts and passive investors using information as "data", which is standardized in a database (DB),** and **they have become the majority of market participants globally.** In such case, real time data distribution has important role on the impact on the stock price. Also, use of Artificial Intelligence (AI) in analysis and investment has become more common today.

First, financial statements have to be converted into "Data" and then **the data can travel through the market**.

There is a parable about the "blind men and elephant", which resembles using financial statements data. Technological innovation related to data processing is remarkable, it is a truly urgent and critical issue how to organize the data of financial statements in the world in this era. **In other words, there is a risk of penetrating market with biased information as only data processing technology advances**, with the lessons of "blind men and elephant" still to be learnt.

Based on such a sense of crisis, I would like to listen to all of your experiences and explore what we can do as a standard setter in near future.

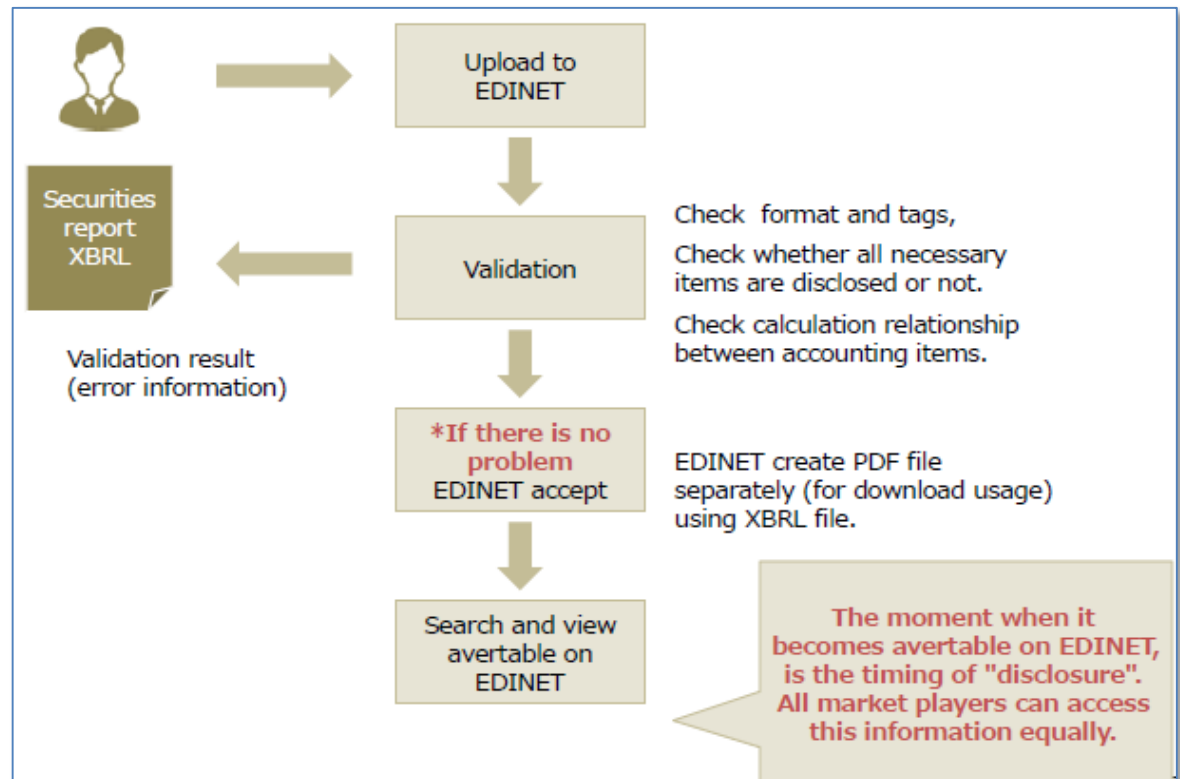One of the attendees, Accounting Setter said in the discussion

1. Public filing systems in Japan

■ Fundamental environment, Regulatory filing system and digital reports (XBRL)

In Japan companies must file their regulatory reports to EDINET, which is provided by the Japan FSA. The public cannot access this information before companies have filed it, so EDINET provides us (users) with the quickest possible access to this information. The regulatory reports are edited in XBRL format, it means, there are no argument whether to see or not to see XBRL, if we want to use data immediately.

In Japan, company also disclose preliminary reports, known as "the earning digests", that are not audited. These documents also must be filed with another disclosure system (TDnet; Timely disclosure network) provided by the Tokyo Stock Exchange.

As with regulatory reports, the pubic cannot access this information before it is filed. For investors to obtain the information immediately after it is disclosed, they must access TDnet.
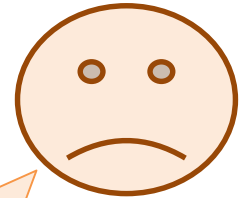
Securities report XBRL

Validation result (error information)

Upload to EDINET

Validation

Check format and tags,
Check whether all necessary items are disclosed or not.
Check calculation relationship between accounting items.

*If there is no problem
EDINET accept

EDINET create PDF file separately (for download usage) using XBRL file.

Search and view avertable on EDINET

The moment when it becomes avertable on EDINET, is the timing of "disclosure". All market players can access this information equally.

2. Use of financial data on a real-time basis

■ Digital reportings are expected to be able to use in real time.
Basically Market have wanted to do so, waited until digital report came.
But data users faced problems…

We provide data terminal services to market players. We provide real-time data and news relating to financial results to customers who instantly make deals based on the information in earning digests, changes in management forecasts, or other event information.
Our customers do not have time to closely follow every company's announcement, especially when many companies are publishing earning digests at the same time. This explains why many customers follow the headlines provided by the information terminal. Stock prices react to the news instantly, we think it most important to provide data on a real-time basis.
We automatically produce "News" (ie articles) from XBRL based data submitted to the TSE. Since some companies misunderstand that "XBRL data is not required for change in management forecast", there are cases where XBRL data is missing from the filings. This means that the data is not always complete and correct because some companies fail to include XBRL filing. This results in us having to manually check the data to avoid distributing incorrect information. Currently XBRL can't utilize without visual check....

Real time
News media

Why did it happen? Lack of compliance? Quality of standards or rules?

These are the two most important assumptions when using XBRL in real time.
1. PDF and XBRL contents are always same
2. PDF and XBRL submission timings are always same
In addition, it is important regulators (accounting standard setters) provide clear rules for utilizing XBRL in real-time.

Global information provider

Current rules are not clear

3. Changing data-use in today's market

■Increasing Passive investment
They have many shares and need screening. So Data should be comparable and stored in DB.

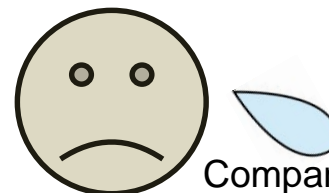Representative of an investor group in Japan

In passive investment portfolio management, the list of stocks included in the index and the index-weights are the only necessary information. Financial statement information may be used to select shares included in the index, but it is not used by investors. Thought, indices are classified based on market capitalization size or industries, the TOPIX index is one of the commonly used benchmarks; it includes all of the shares listed on TSE 1st section (approx. 2000 shares).

Passive investor's engagement is expected under the Japan FSA stewardship code (introduced 3 years ago). Fundamental information such as capital efficiency, policy on shareholder returns are required for engagement purposes. It is unrealistic to engage with all of the companies included in the index, therefore we need comparable data on companies to perform screening, and ultimately to decide on the target companies for engagement.

This upward trend in passive investing is also seen in Europe and USA (33% of assets in EU, 47% in US). In the US, it is expected that the allocation to passive funds will overtake active funds in 2019.

These days we have more interest in how can we get into index, not how we should disclose well for investors.
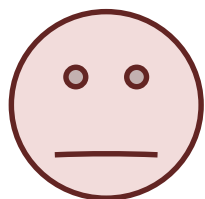
Company IR person

3. Changing data-use in today's market (cont)

■Work of quant analysts
Quant analysts predict which factor will be the most effective on stock price performance during the next month. They usually do not care so much about the original data, but rely on third-party data base.
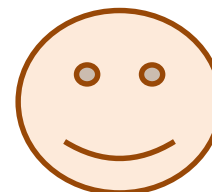
Most of Quants analysts also do not use detailed information on financial statements. In addition, we do not usually use historical results, we often use the forecasts – especially net profit estimates.
When calculating the effectiveness of factors, it does not matter if some companies are missing from the dataset. This is because we have plenty of alternatives to substitute for companies that do not disclose the forecast data.
Also, comparability problems due to differences in accounting standards do not matter to our approach. In fact, many Quants analysts do not know the difference between IFRS and J - GAAP.
Most of Quants analysts are not concerned about the company's fundamental value, instead we want to understand how other market participants view the company.

Quants analyst

So we always extract data from XBRL immediately after they are disclosed, then store them in DB and provide to the end users.

Information provider

I do not need to care about comparability because it is guaranteed by information providers, I don't care which company adopt IFRS.

■Data supply-chain is existing.

4. Data should be more consistent and comparable

■Mission of Data Analysts of Banks
Even if not relying on the third-party vendors, still someone need to make data comparable, even with XBRL...This is the issue of original financial statements.
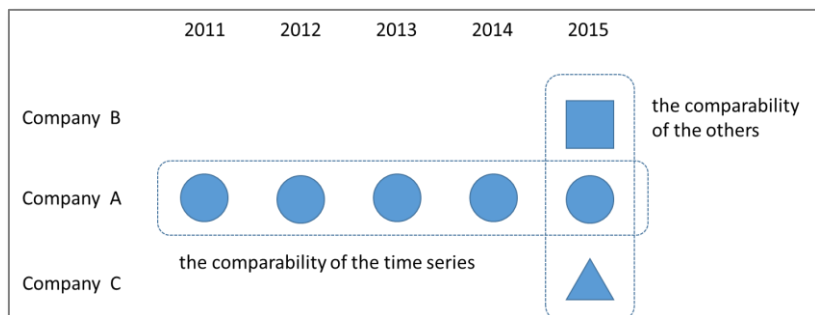
I am in charge of editing and planning databases used for investment and financing at a bank. It is necessary to process as "data", not directly from the disclosed information.
In the example of company A, it is necessary to ensure that its data has continuity over time (from 2011 to 2015) and comparability with other companies (company B and C). If these conditions are not met, the disclosed information will not reach users. However, accounting standard changes in the middle of this period, then the definition of the disclosed items would change.
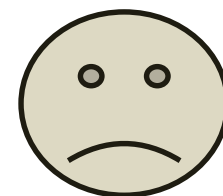
Also, If a group of companies that disclose account items related to products as various names, users may take the basic common item and so we will only store the item, for example "goods" in the database. Then, various items such as "products", "work in progress", "raw materials" are gathered as "inventory assets", and items such as "inventory assets", "cash", "sales receivables" , "Liquid assets", etc, are categorized at the discretion of the database making side (ie criteria of data and editing).

Data analyst of Bank

	2011	2012	2013	2014	2015	
Company B					■	the comparability of the others
Company A	●	●	●	●	●	
Company C		the comparability of the time series			▲	

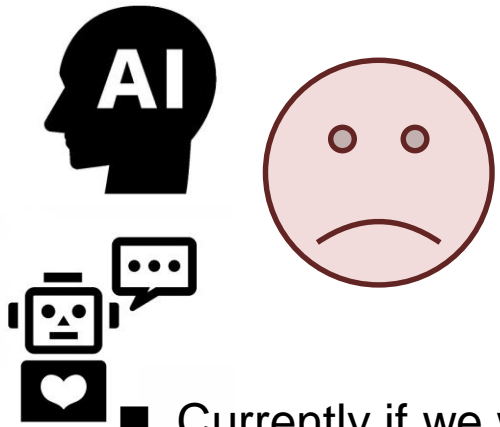Should all disclosure, including non-financial information, be considered as well?

Company IR person

5. To have computer system "read" data in place of human

■ "Writing" articles about financial results by AI, its possibilities and challenges

We use AI and prevent necessary documents from being buried and overlooked.
We are now,
1. Automatically creating news headlines that let users know the contents of the story/results.
2. Extracting the important parts from a huge amount of information and summarizing it automatically.
 We started these efforts because all of EDINET's submission became XBRL, and TDnet requires companies to submit financial part in XBRL. Even without tags, XHTML is more helpful than simple PDF files. When reading data from PDFs, there can be garbled characters or collapsed tables . We work on the automatic creation of articles, there is a part which can not be read accurately in text part of filings, we check numbers from XBRL data and complement / correct them.
One of the problems we face occurs in cases where data in PDF filing is correct, but data in XBRL filing is incorrect. We focus on speed and process data by AI system, but we later confirm it by making manual comparisons with PDF.

Information provider

Under current situation, **AI can write a news article?**
In one case, a Company wrote a lot of criticism about a previous board member in the earnings digest. Current AI can find which part of story is possibly important in the whole article. So the **AI article picked up this part of the earnings digest and sent out the article automatically.** That was really strange, because the article mentioned only criticism of previous management, when article has to be on the earning digest.
AI cannot understand the meaning the same way a human can.

■ Currently if we want to use AI, we need more clear and correct data before it.

Who can take the cost of capturing data from digital report

■ Systems are developed for capturing data, but still need manual check. Besides, we need to always change this system.
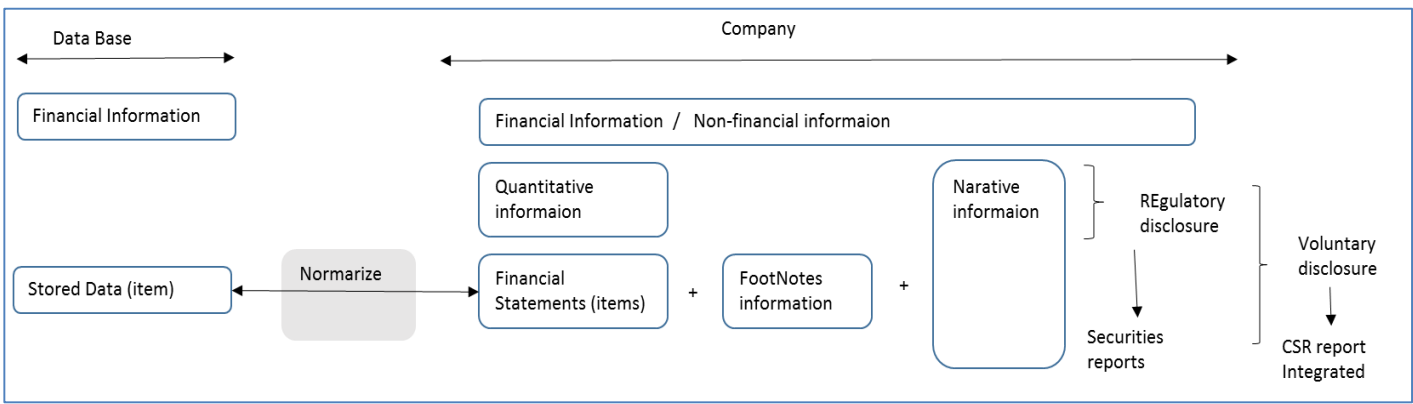
Data analyst of Bank

It is difficult to say that the convenience for users would improve only because various data become available in XBRL. To incorporate data into the DB it must be checked by visual inspection, and becomes additional cost. Furthermore, if the XBRL specification changes, then this requires a system revision. The figure on the right is part of the table about renovation of the DB when EDINET (XBRL) was revised in 2013. Currently, such upgrades are increasing. Although Narrative

procedure	What DBJ has to do
(1)Dimension	Checking dimensions
find way to detect data	Find way to extract data from Dimension and store to DB
Count target items	Categorize target items and counting number of items
(2) Form /Capturing facts	Checking Primary Financial Statements of the securities report
Find way to detect data	Examin difference between Third generation EDINET taxonomy and Current generation EDINET taxonomy. Find way how new data can store DB
Examin items	Categorize target items and counting number of items
Mapping items	Survey on correlation between account items and DB (standardized) items
How to transfee current mapping to next mapping for new generation EDINET	Find way of transission from Current EDINETtaxonomy mapping to Third generation EDINET taxonomy.
(3) Other procefure for this upgrade to Third generation EDINET	Examine difference between Third generation EDINET and Current EDINET. (Regarding with other specifications)
Changing the hierarchy of taxonomies	Investigation of program and database upgrade accompanying change of EDINET taxonomy hierarchy
Change extension link role	Investigation of cost of program and database upgrade accompanying change of extention link role
Integration of item schema	Investigation of cost of program and database upgrade accompanying change of items schema from separated industries type to integrated type.
(4) Compatibility from Current EDINET	With Consideration of current EDINET compliant XBRL data coming to the DB after third generation EDINET has launched, Need to find way to handle new data as same way of data for current EDINET.
Find Compatible functions	Find way to handle with data for Third generation EDINET while maintaining the current functions

disclosures are increasing, it is difficult to determine the criteria for continuity and comparability when capturing it in a DB. The figure below is the overall picture of information disclosed by companies. The right side of the figure shows what kind of items are recorded for each category such as financial, non-financial, statutory disclosure, voluntary disclosure. The left side of the figure shows the connection between the items contained in the DB and the information disclosed by the company. Regarding financial information, the items disclosed by companies are recorded with processing of standardization. But, the information contained in the DB captures a smaller portion of whole disclosure.

Data Base

Company

Financial Information

Financial Information / Non-financial informaion

Quantitative informaion

Narative informaion

REgulatory disclosure

Voluntary disclosure

Stored Data (item)

Normarize

Financial Statements (items) + FootNotes information +

Securities reports

CSR report Integrated

Where are we heading? What should we do?

In the current situation, what we can do, what we should do?

- ✓ Increasing number of passive investors. They need comparable data on a large number of companies across markets. They do not have the time to read them carefully, just need screening systematically.
- ✓ Quants analyst challenges AI. AI needs to read data correctly first.

- ✓ But digital report (XBRL) is currently not perfect. Rules are not clear, allowing companies to make mistakes. Data usually needs manually checking.
- ✓ However, issues that we recognize on digital reporting today have happened on paper based disclosure already.

This situation has a risk actually like "blind men and elephant". We might need to understand what is going on in the market and tighten rules if it is necessary. But the most important is….

Think and work together !

Thank you for listening!

Questions and comments
c-mitsui@nri.co.jp