

UFO

Understanding of non-Financial Objects in Financial Reports

Yasutomo Kimura

Otaru University of Commerce

Takahito Kondo

Kyoto Sangyo University

Kazuma Kadowaki

The Japan Research Institute,
Limited

Hokuto Ototake

Fukuoka University

Makoto P. Kato

University of Tsukuba

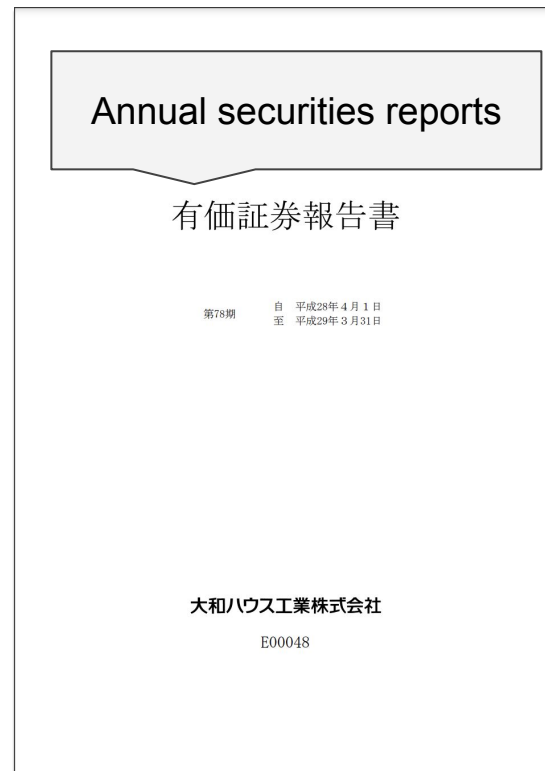
Table of contents

- What is annual securities reports?
 - EDINET
 - XBRL
 - Taxonomy
 - Instance
- What kinds of tables are included in annual securities reports?
- What is the relationship between the text and tables in annual securities reports?
- What is UFO task?
 - Understanding of non-Financial Objects in Financial Reports
 - What is TDE (Table Data Extraction)?
 - What is TTRE(Text-to-Table Relationship Extraction)?

What is annual securities reports?

Annual securities reports are

- reports that listed and other companies are required to prepare.
- disclosed to the outside **every fiscal year**
- electronically submitted to **EDINET**



EDINET (Electronic Disclosure for Investors' NETwork)

- An electronic disclosure system for disclosure documents

金融商品取引法に基づく有価証券報告書等の開示書類に関する電子開示システム English

EDINET Electronic Disclosure for Investors' NETwork

文字の大きさ 小 大

閲覧 有価証券報告書等の開示書類を閲覧するサイト

トップページ 書類検索 公告閲覧 書類比較 ダウンロード

トップメニュー

- お知らせ一覧
- よくある質問 (令和2年3月14日更新)
- 利用環境
- 操作ガイド等
- 御意見・御要望
- 利用規約

書類検索

公告閲覧

新型コロナウイルス感染症に関連する有価証券報告書の提出期限について ⚠️必ず御確認ください。

新型コロナウイルス感染症の感染拡大に伴い、1月7日、新型インフルエンザ等対策特別措置法に基づく緊急事態宣言が発令されたことに関連し、有価証券報告書の提出期限について、以下の通りお知らせいたします。ご質問等ございましたら、ご連絡なく所管の財務(支)局までご連絡ください。

○金融商品取引法に基づく開示書類(有価証券報告書及び内部統制報告書、四半期報告書、半期報告書等)について、今般の新型コロナウイルス感染症の影響に伴い、やむを得ない理由により期限までに提出できない場合は、財務(支)局長の承認により提出期限を延長することが認められていますので、ご連絡なく所管の財務(支)局にご相談ください。

(注)有価証券報告書及び

内部統制報告書の提出期限：事業年度経過後3ヶ月以内
四半期報告書の提出期限：四半期会計期間経過後45日以内
半期報告書の提出期限：中間会計期間経過後3ヶ月以内

○また、臨時報告書についても、新型コロナウイルス感染症の影響により臨時報告書の作成自体が行えない場合には、そのような事情が解消した後、可及的速やかに提出することで、遅滞なく提出したものと取り扱われることとなります。

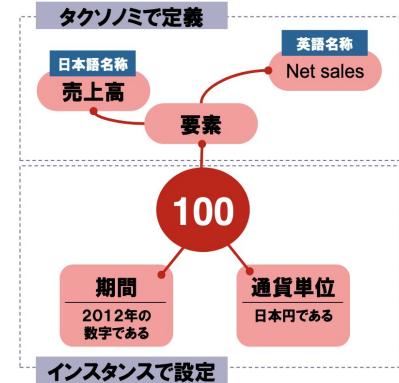
操作ガイド

- EDINET総合ガイドライン
 - EDINET概要書 (令和2年11月22日更新)
 - 提出書類ファイル仕様書 (平成30年3月16日更新)
- 操作ガイド
 - 書類閲覧操作ガイド (令和2年3月14日更新)
 - XBRLからCSVへの変換ツール操作ガイド (令和元年9月6日更新)
 - XBRLからCSVへの変換ツール操作ガイド 変換ルール定義ファイル情報 (令和2年12月26日更新)
- EDINET API関連資料(平成31年3月17日公開)
 - EDINET API利用規約
 - EDINET API仕様書
 - EDINET APIサンプルプログラム (令和2年11月22日更新)
 - EDINETコード集約一覧 (令和元年10月18日更新)
- XBRL関連技術資料
 - フレームワーク設計書 (令和2年11月10日更新)
 - フレームワーク設計書別紙 (平成30年3月16日更新)
 - EDINETタクソノミの設定規約書 (令和2年11月10日更新)
 - EDINETタクソノミの設定規約書別紙 (令和2年11月10日更新)
 - バリデーションガイドライン (令和2年9月19日更新)
 - バリデーションガイドライン添付資料 (令和2年9月19日更新)
 - バリデーションメッセージ一覧 (令和2年9月19日更新)

XBRL eXtensible Business Reporting Language

- A standardized **XML-based language** for making information available for financial reporting
- Link related elements using **taxonomy** and **instance**
 - Taxonomy: Attributes of elements to be linked (name, etc.)
 - Instance: Set period, currency unit, etc. for attribute value

図表 1-1-1 XBRLにおけるタクソノミ及びインスタンス(イメージ)

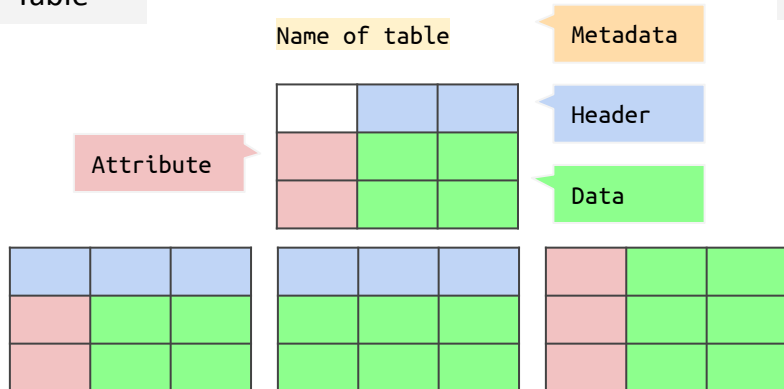


Issues related to annual securities reports

Annual securities reports

- **Different types of tables**
- **Different notation**
- **Non-financial table cells** that is NOT assigned a taxonomy
- **Text without taxonomies**

Table



Text

Ordinary profit for the current consolidated fiscal year was 300.5 billion yen, an increase of 66.9 billion yen from 233.5 billion yen for the previous consolidated fiscal year due to the increase in operating profit.

Reason

UFO Task

Understanding of non-Financial Objects in Financial Reports

Purpose

- We aim to develop a technology for extracting structured information from tabular data and documents.

Table Data Extraction (TDE)

- TDE is a subtask that extracts data from tables contained in reports for comparison between companies.

Text-to-Table Relationship Extraction (TTRE)

- TTRE is a relationship extraction task that connects the values contained in the table of reports and the related texts.

Why is the task named UFO?

- “有報” stands for “有価証券報告書”
- The pronunciation of 有報 is “júfou”
- The pronunciation of **UFO** is “júfou”, NOT “jú:èfóu” in Japan

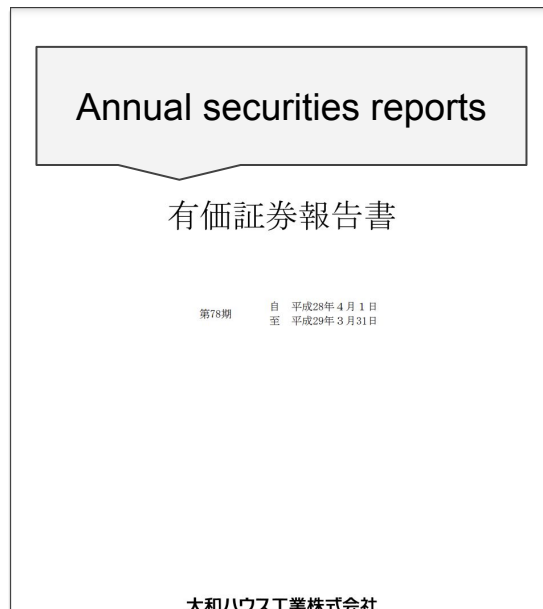


Table Data Extraction (TDE)

Input

- TOPIX100 tables (TABLE in HTML)

Output

- Classification of table cells (**Metadata, Header, Attribute, Data**)
 - For Header, add ontologyID
 - For Data, normalize the value.

Evaluation

- Evaluation by Accuracy

Table Data Extraction (TDE)

Metadata, Header, Attribute and Data

1 【設備投資等の概要】

設備投資の状況をセグメント別に示すと、次のとおりであります。

セグメントの名称	金額 (百万円)	前連結会計年度比 (%)
資源	92,012	689.9
製錬	10,885	△28.5
材料	18,807	△13.1
報告セグメント計	121,704	150.9
その他	727	98.6
調整額	3,519	64.7
合計	125,950	146.9

```
<table class="equipment.per_segment" targetCompany="xxxx">
  <tr>
    <td class="header" ontologyID="segment_name">セグメントの名称</td>
    <td class="header" ontologyID="amount">金額 (百万円)</td>
    <td class="header" ontologyID="">前連結会計年度比 (%)</td>
  </tr>
  <tr>
    <td class="attribute">資源</td>
    <td class="data" valueClass="currency.JPY" normalizedValue="92012000000">92,012</td>
    <td class="data" valueClass="amount.relative" normalizedValue="689.9%">689.9</td>
  </tr>
</table>
```

Metadata

Header

Attribute

Data

Text-to-Table Relationship Extraction (TTRE)

Input

- **Text and Tables** in the annual securities reports

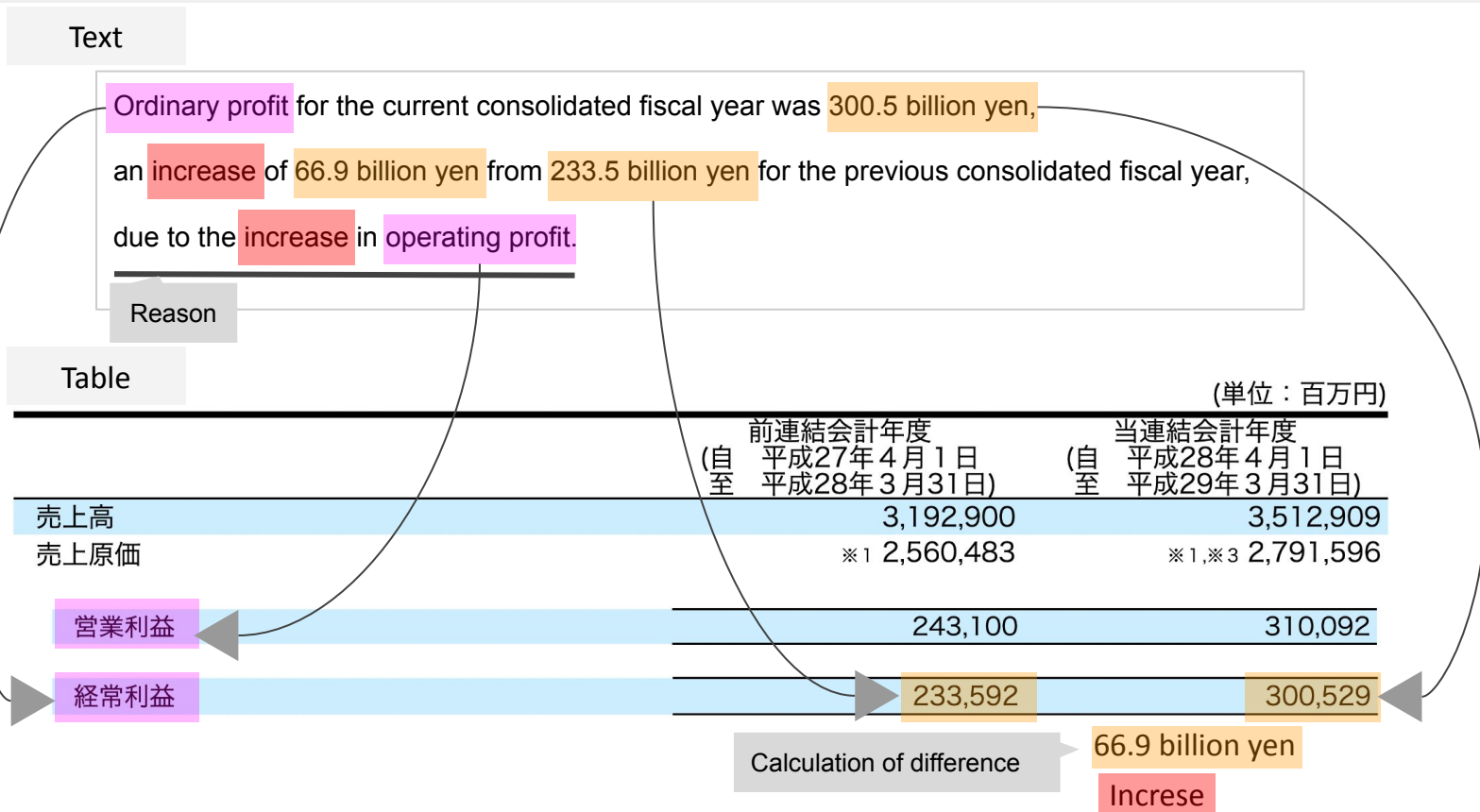
Output

- Text range corresponding to the table
 - Labels for amounts, increases/decreases, references, and reasons
 - Assign IDs for the corresponding tables

Evaluation

- Evaluation by Accuracy

Text-to-Table Relationship Extraction (TTRE)



Schedule

- September 28, 2022 NTCIR-17 Kickoff Event
 - October 8, 2022 JSAI SIG-FIN (Special Interest Group on **Financial Informatics**) in Tokyo
 - Title “UFO: Proposal for an Information Extraction Task for Tables in Annual Securities Reports”
 - November 8, 2022 First table meeting
 - January 2023: Dataset Release*
 - March 2023: Second round table meeting
 - January-June 2023: Dry Run*
 - June 2023: Release of Formal Run dataset
 - June-July 2023: Formal Run*
 - August 1, 2023: Evaluation Result Release
 - August 1, 2023: Draft Task Overview Paper Release
 - September 1, 2023: Draft Participant Paper Submission Due
 - November 1, 2023: All Camera-ready Paper Submission Due
 - December 2023: NTCIR-17 Conference in NII, Tokyo, Japan
- (* indicates that the schedule can be different for different tasks)