

CASE STUDY

Leveraging Machine Readable Filings to Uncover Valuable Investment Insights

THE CLIENT:

A large SWF

USERS:

A newly formed quantitative investment team

As sovereign wealth funds (SWFs) around the world grow in size and number, they are becoming increasingly powerful players on the global financial scene.¹ The investments they make are effective ways for their countries to diversify income streams and become less reliant on a few sources of revenue.

This large SWF invests in all asset classes throughout international markets, including equities, fixed income, real estate, private equity and alternatives. The company had recently put together a new group comprised of quantitative analysts and strategists responsible for extracting important insights from data. The group wanted to expand the information they had available and include additional sources of machine readable annual and interim financial reports to help improve efficiencies and the speed of discovery.

¹ "What is a What is a sovereign wealth fund?", World Economic Forum, www.weforum.org/agenda/2017/10/what-you-need-to-know-about-sovereign-wealth-funds/.



Pain Points

Members of the investment team needed to support both tactical and strategic decision-making at the SWF. They saw the many benefits of using natural language processing (NLP) with large quantities of textual data and had developed an internal solution for obtaining machine readable information on filings for U.S. and global companies. This required an extensive amount of time to gather and maintain the information, however, and team members were concerned about coverage and quality. They wanted to outsource this to a reputable provider, plus find a solution that offered:

- A comprehensive set of **high-quality global information**.
- **Easy access** through a data feed option.
- Hands-on **technical and product support** to address any issues as they arose.

Members of the investment team were looking to find a trusted provider that could download, parse, clean and structure a vast number of filing documents from countries around the world, so they could focus their efforts on higher-value textual analysis.

The team had heard that S&P Global Market Intelligence (“Market Intelligence”) was using a lot of artificial intelligence capabilities to connect the dots across its many datasets and contacted the firm to learn more about its offering for machine readable filings.



The Solution

Market Intelligence discussed its U.S. and Global Machine Readable Filings dataset that provides parsed text of annual and interim reports, broken into the various sections identified by companies, with extraneous information (such as page numbers, images and tables) removed. The data is delivered in a structured format enabling users to perform NLP against it without having to do the document cleanup and structuring themselves. This would enable the team to easily support an analysis of strategic initiatives, earnings, M&A plans, new product possibilities and much more.



Easily screen and evaluate U.S. filings

Machine Readable U.S. Filings provide full coverage of 10K/10Q/8K/6K/20F/40F reports dating back to 2006. The documents are structured by major sections, such as Management Discussions & Analysis, Risks, Competition and Intellectual Property. The document hierarchy is retained, creating a historical baseline for backtesting.



Access data extracted from global filings

Machine Readable Global Filings cover 89% of the MSCI World market cap. Data is gathered from 400+ sources, including company websites, stock exchanges and regulatory websites.



Store data in a centralized warehouse

Xpressfeed™ automates the download and management of filings data to a centralized company warehouse, delivering updates every two hours. This supports easy linking to other Market Intelligence datasets, including financials, estimates and events data.



Key Benefits

Members of the investment team saw many benefits to the offering and subscribed to both the U.S. and Global Machine Readable Filings. In particular, they thought this would help them:

- **Save enormous amounts of time** currently being spent on creating an internal solution for the collection and processing of U.S. and global data.
- **Improve the overall quality of the data** with a tested and scalable solution for text-cleansing, including maintaining consistency across reporting periods regardless of structure changes, reclassifying heading sections for standardization purposes and removing irrelevant elements, such as table headers and page numbers.
- **Access extensive textual information** for companies around the world, which is pre-tagged, structured and organized.
- **Replicate fundamental analyst workflows** across millions of documents to increase the breadth of analysis and identify documents and sections of highest importance.
- **Leverage NLP and data mining techniques** to systematically identify themes, trends and major changes within a company’s reporting of material qualitative information. For example, identify mentions of “ESG”, while filtering out mentions that do not actually apply to this concept.
- **Easily combine the information** with other relevant financial data and market participant actions to establish patterns that could warrant further inquiry.

The investment team is now looking at other machine readable textual datasets provided by Market Intelligence. This includes Transcripts that review data on earnings calls, company conference calls and special calls.

[Click here to explore the datasets mentioned in this case study.](#)

S&P Global

Market Intelligence

Contact Us

Asia-Pacific

+852-2533-3565

Europe, Middle East & Africa

+44-207-176-1234

The Americas

+1 877 863 1306

Copyright © 2022 by S&P Global Market Intelligence, a division of S&P Global Inc. All rights reserved.

These materials have been prepared solely for information purposes based upon information generally available to the public and from sources believed to be reliable. No content (including index data, ratings, credit-related analyses and data, research, model, software or other application or output therefrom) or any part thereof (Content) may be modified, reverse engineered, reproduced or distributed in any form by any means, or stored in a database or retrieval system, without the prior written permission of S&P Global Market Intelligence or its affiliates (collectively, S&P Global). The Content shall not be used for any unlawful or unauthorized purposes. S&P Global and any third-party providers, (collectively S&P Global Parties) do not guarantee the accuracy, completeness, timeliness or availability of the Content. S&P Global Parties are not responsible for any errors or omissions, regardless of the cause, for the results obtained from the use of the Content. THE CONTENT IS PROVIDED ON "AS IS" BASIS. S&P GLOBAL PARTIES DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR USE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE CONTENT'S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE CONTENT WILL OPERATE WITH ANY SOFTWARE OR HARDWARE CONFIGURATION. In no event shall S&P Global Parties be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees, or losses (including, without limitation, lost income or lost profits and opportunity costs or losses caused by negligence) in connection with any use of the Content even if advised of the possibility of such damages.

S&P Global Market Intelligence's opinions, quotes and credit-related and other analyses are statements of opinion as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security. S&P Global Market Intelligence may provide index data. Direct investment in an index is not possible. Exposure to an asset class represented by an index is available through investable instruments based on that index. S&P Global Market Intelligence assumes no obligation to update the Content following publication in any form or format. The Content should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions. S&P Global keeps certain activities of its divisions separate from each other in order to preserve the independence and objectivity of their respective activities. As a result, certain divisions of S&P Global may have information that is not available to other S&P Global divisions. S&P Global has established policies and procedures to maintain the confidentiality of certain non-public information received in connection with each analytical process.

S&P Global may receive compensation for its ratings and certain analyses, normally from issuers or underwriters of securities or from obligors. S&P Global reserves the right to disseminate its opinions and analyses. S&P Global's public ratings and analyses are made available on its Web sites, www.standardandpoors.com (free of charge) and www.ratingsdirect.com (subscription), and may be distributed through other means, including via S&P Global publications and third-party redistributors. Additional information about our ratings fees is available at www.standardandpoors.com/usratingsfees.