CREWS – Credit Risk Early Warning System

Use of text-analytics methods to increase efficiency in credit analysis

In uncertain times with historically high credit ratings, potential cliff effects induced by IFRS 9 impairment models, and the current recession caused by the COVID-19 pandemic the need to accurately forecast credit events and have timely early warning systems is ever necessary.

Over the last decade the amount of publicly available data has grown significantly and so has the need to systematically analyse it. Credit analysts now need to monitor stock prices, credit spreads, financial statements, newspapers, and a plethora of other indicators in order to identify critical credit exposures and clients. Due to limited resources, there is a compelling need to provide analysts with support tools. Newspapers and articles contain a lot of useful information but are extremely time-consuming to process so text-evaluation algorithms are needed to automate the analysis.

While market data like stock prices or credit spreads and balance sheet data are commonly used by banks within early warning systems, newspaper coverage is rarely integrated as a further source of information. Even though many banks have internal databases with large newspaper coverage that are read by analysts, only few of them have implemented text-mining algorithms to systematically analyse the articles.

This motivated Fintegral to develop a purely news-based early warning system that can be used either as a stand-alone tool to support credit analysts or as a potential extension to existing early warning systems.

CREWS Overview

Fintegral's CREWS (Credit Risk Early Warning System) is an application that implements text-mining techniques on news articles to predict future changes in corporate credit quality. The application has been implemented in a modular fashion to ensure both the extendibility and the exchangeability of each component and therefore can either be used as a stand-alone tool or as component of a larger early warning system.

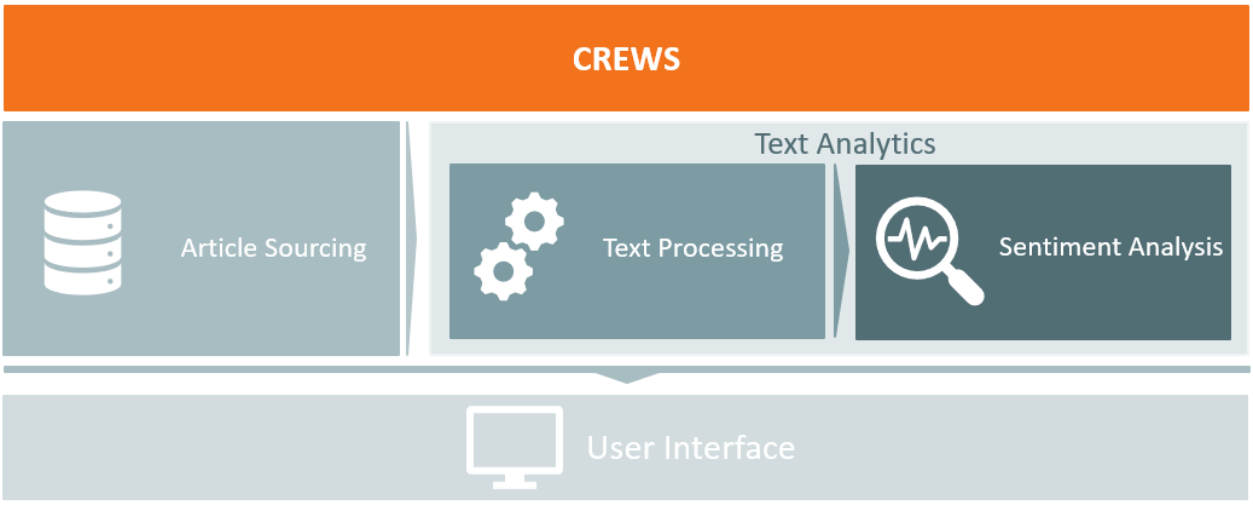

Figure 1: CREWS overview

As shown in Figure 1, the CREWS application consists of three modules and a user interface:

- **Article Sourcing** – This module automatically extracts articles from the web pages of several online journals by analysing the underlying source codes and stores them in a database.
- **Text Processing** – Before applying the sentiment analytics, the underlying data must be processed to maximise the model's accuracy. The text processing module cleans data using a variety of techniques such as lemmatisation and removal of redundant words and characters.
- **Sentiment Analysis** – The sentiment analysis module using text-mining techniques to assign numerical scores to single sentences to calculate an overall sentiment of the article. The scores range from -1 to 1, where the sign of the score signals the tone of the sentence and the magnitude of the score signals the strength of the tone.
- **User Interface** - The web-based user interface allows for the results and summaries to be dynamically viewed in an interactive user-friendly format.

The CREWS application automatically provides credit analysts with important newspaper coverage and warnings of possible credit downgrades to allow them to focus their attention on mitigating and actioning the risks rather than trying to predict and identify them.

Article Sourcing

A fundamental requirement for any news-based early warning system is an extensive article database as this is needed to calibrate parameters and improve the model's forecasting accuracy. Fintegral's first step when developing CREWS was to create a comprehensive newspaper article database by developing extraction tools as part of the Article Sourcing module.

Most large online journals offer well-structured archives, so we implemented routines to systematically extract the relevant information from several online journals (e.g. Reuters or Handelsblatt) to create a database containing more than 100,000 business- and politics-related articles. Banks that have their own internal article databases or source them from an external provider can easily integrate their solutions into the CREWS application and either combine them with our database or replace it entirely.

Articles contained in the database can be accessed using in-built search functions on the user interface. By simply entering a specific search word (e.g. company name, industry, or topic) the user can retrieve all relevant articles from the database. The interface allows for filtering of articles using date and journal name to help users easily find the specific articles that they want to analyse.

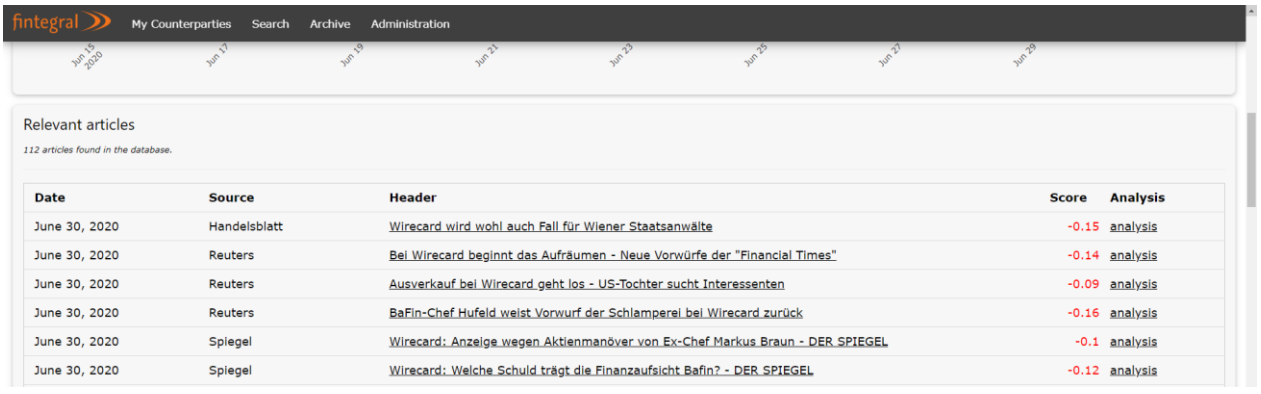

Figure 2: Searching for Wirecard articles

Figure 2 shows the output of the user-interface after searching for the company Wirecard. The results contain both details of the article (date, source and title) as well as the numerical sentiment score that was calculated. As expected most articles from end of June 2020 show a negative sentiment towards Wirecard, this is because they had recently been involved in a balance sheet manipulation scandal.

Text Analytics

The sentiment scores are the heart of the CREWS application and enable the system to identify critical articles which would trigger an early warning signal. The model implemented in CREWS is based on a sentiment analyser which is an established methodology to analyse the general tone of text and has a multitude of applications such as the study of customer reviews, social media posts and newspaper articles.

Text Processing

Before running sentiment analysis on an article, it must first be processed to remove redundant information as natural language contains several elements that do not have a meaningful interpretation and do not affect the overall sentiment. The Text Processing module contains functions to eliminate all unnecessary words (e.g. prepositions, pronouns, special characters and numbers) and leave only words that would potentially affect the overall sentiment of the article.

Afterwards the text is normalised using lemmatisation, this is the process of grouping all words with the same meaning together so they can be analysed as a single item. For example, the word “losses” is derived from “loss” and they both have the same connotation. By bringing all words to their root form, the size of the vocabulary that is needed to design the text-scoring model can be reduced. Figure 3 shows some examples of pre and post processed sentences to show the removal of redundant words.

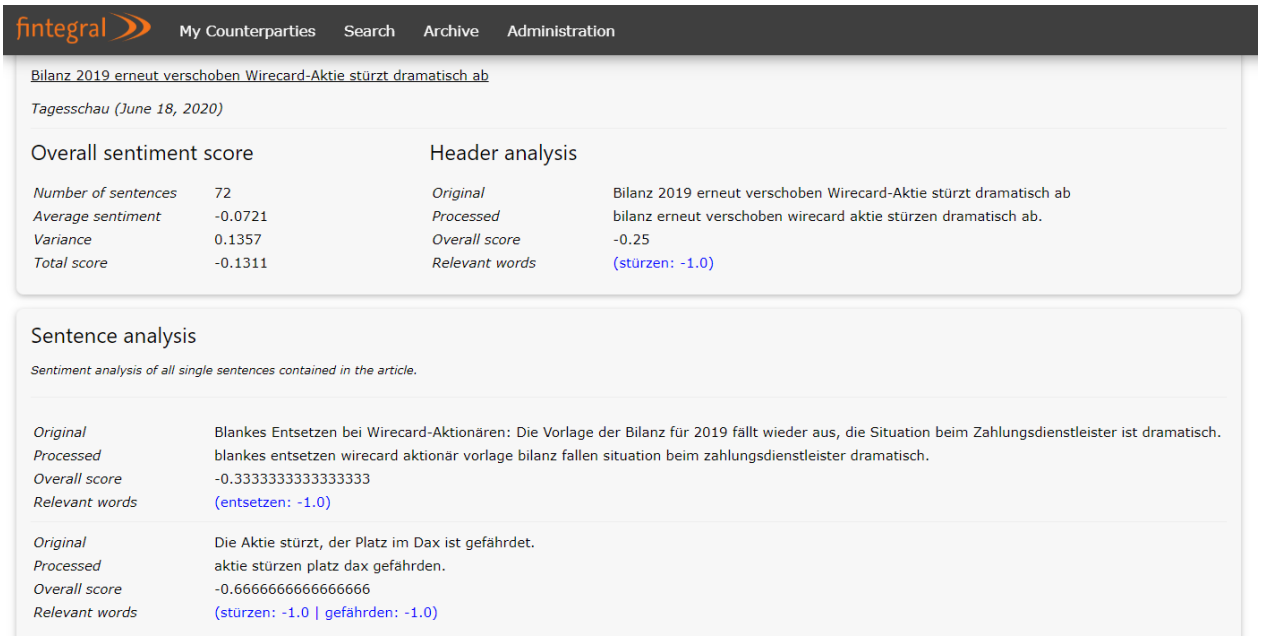

Figure 3: Sentiment analysis drill down

Sentiment Analysis

The overall sentiment of an article is calculated by aggregating the individual sentiment scores of the title and each sentence. The sentiment score of a sentence (or the title) takes a value in the interval [-1; 1] where the sign of the score signals the tone of the sentence and the magnitude of the score signals the strength of the tone (i.e. a score of -0.8 means a strong negative tone and a score of 0.05 signals a weak positive tone). The sentiment of a sentence is calculated from the presence of specific key words that have been pre-determined as predominantly negative or positive.

Negative articles are often written as a series of accusations and justifications so their sentence sentiment scores will often oscillate between negative and positive. To capture this behaviour and better identify negative articles an additional factor is calculated by taking the variance of the individual sentence sentiment scores. In total the overall

sentiment of an article is determined by taking the weighted average of the sentiment of the title, the average sentiment of the text, and the variance factor. Fintegral have already calibrated the weights of these three factors, however users are also able to adjust the weights if they desire.

To ensure maximum transparency and help explain the overall score, CREWS includes a drill down function that allows the user to view the sentence level sentiment derivation. Additionally, CREWS displays: a comparison of the original and processed sentences, the sentiment for each sentence and also lists all pre-determined key words that affect the sentence scores, this is shown in Figure 3.

User Interface

CREWS offers a web-based user interface (UI) build with Django, an established Python web framework, which ensures a user-friendly and intuitive interaction format allowing for flexible enhancements like adding new features and application pages.

The UI not only provides the user with possibilities to search for specific topics and to view detailed derivations of the sentiment scores (as previously discussed) but also allows users to create predefined customer pages and to get an overall portfolio view using the dashboard function.

In addition, the user interface offers the functionality to send early warning signals about possible counterparty defaults or downgrades. The warning indicators are sent to the user when a counterparty's sentiment score falls below a pre-determined threshold and triggers the signal.

Customer Pages

CREWS is fully customisable and offers the possibility for users to create their own standardised pages for important clients. For each of clients the user can store relevant search terms (e.g. company name, subsidiaries, etc.) and core facts (e.g. current exposure, sector, region) in the database. Each time the customised page is refreshed, the system automatically analyses all of the relevant recent articles and displays interactive sentiment analysis charts. The page shows a historical time series of the client's sentiment scores as well as the number of articles published each day to allow for users to track the sentiment score and identify if their client is being covered by an unusually high number of articles. An example of the customer page can be seen in Figure 4.

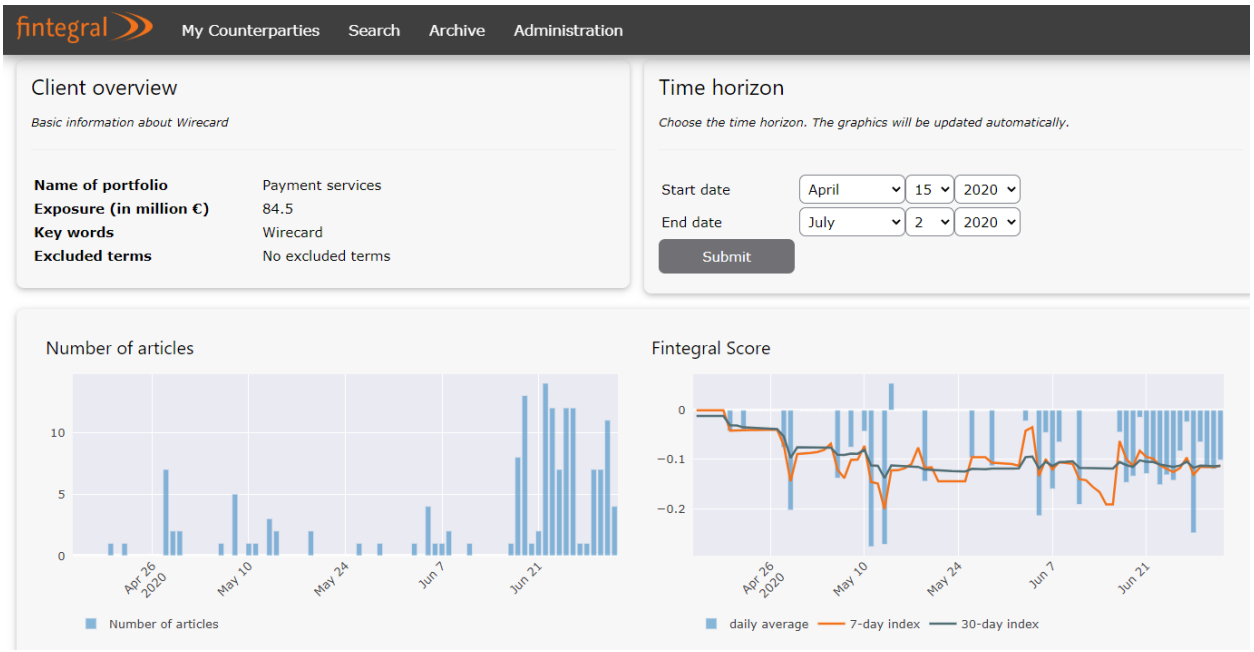

Figure 4: Client overview layout

Dashboard

The dashboard page allows credit analysts to have a holistic view of their portfolio, highlight potentially problematic clients, identify trends and easily compare counterparties. Counterparties are classified into potential risk categories using a simple traffic light (either red, yellow or green) by using their historical average sentiment score. If the warning indicator is red, then this triggers a signal to the user to update them on recent development. The dashboard is an ideal starting point for regular portfolio analysis and provides the user with counterparties that should be investigated in more depth, which they can easily do using the rest of the UI functionality, and example is shown below in Figure 5.

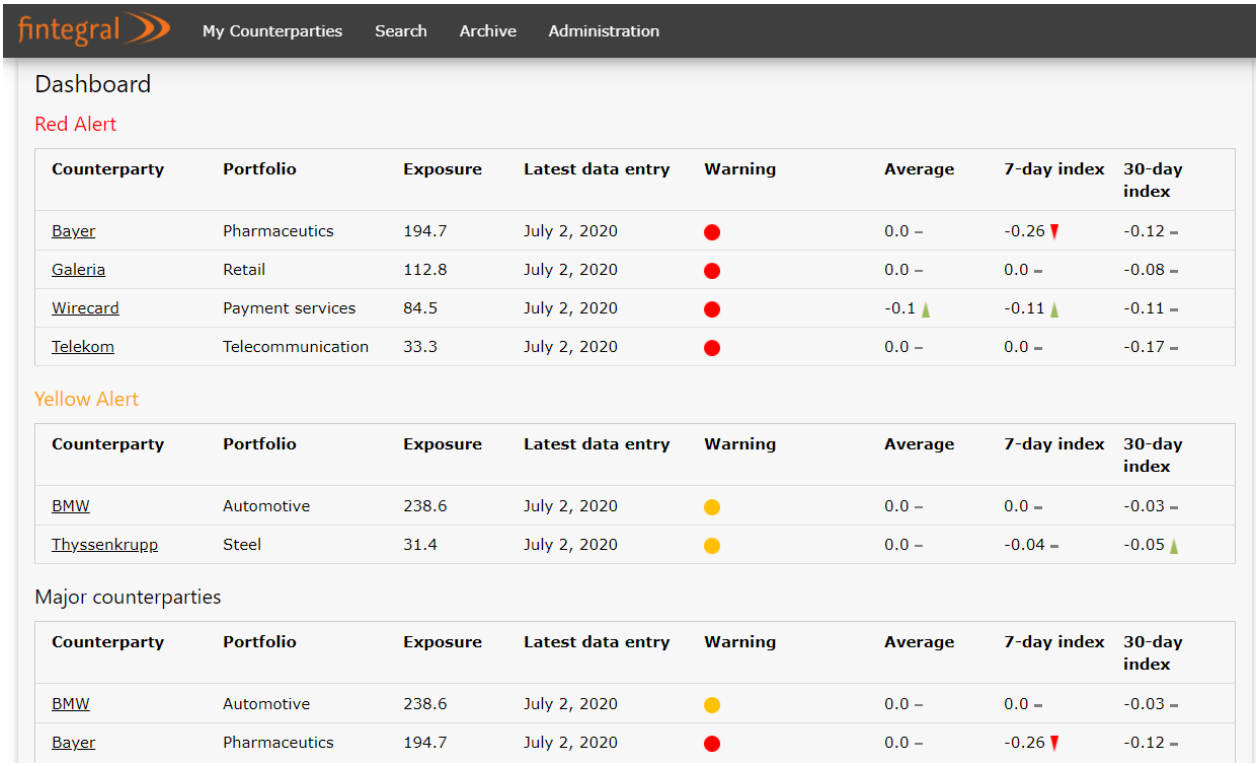

Figure 5: CREWS dashboard

Summary

CREWS is a solely news-based early warning system that calculates the sentiment of articles to help predict changes in corporate credit quality. The application has been developed in a modular fashion to enable clients to use it with flexibility: it can be used standalone or as part of a bigger system, the modules can be easily enhanced to add more functionality, the user-interface is fully customisable to meet all client specific needs.

Although CREWS was originally developed as a tool to increase efficiency for credit analysis it can be used for a wide range of other applications. For example, fund managers face similar problems when identifying critical stock and bond positions and would therefore benefit from the news-based early warning system.

If you have questions regarding CREWS, automated text-scoring, early warning systems or any related topics then please contact us and we will be more than happy to discuss.

Contact us

For more information or if you have any questions please contact:

Dr. Andreas Peter
Managing Partner
+49 160 583 40 66
andreas.peter@fintegral.com

Robert Pullman
Senior Consultant
+44 7784 377 045
robert.pullman@fintegral.com

Tobias Erdweg
Consultant
+49 171 331 86 18
tobias.erdweg@fintegral.com

15+

We have been delivering high quality projects for over a decade

135+

Our clients include global banks, insurers, hedge funds and regulators

700+

We have delivered projects across Europe, North America and Asia

8

We specialise in a core set of risk management topics and techniques

We are a fast-growing specialist risk consultancy with a focus on the finance sector. Our culture of innovation and operational excellence makes us the first choice for our clients and an attractive employer for bright people.