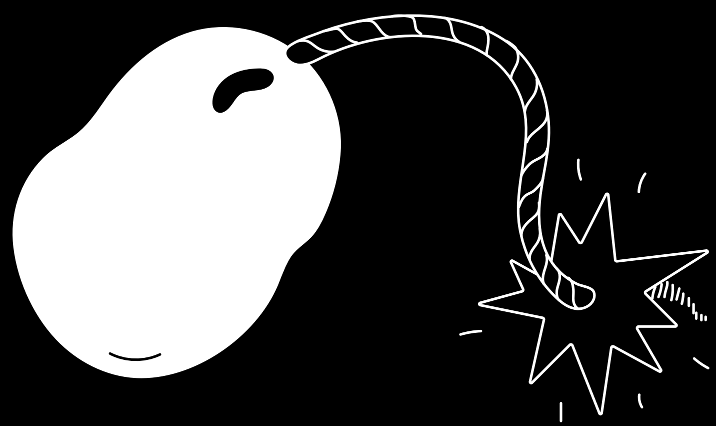
exxeta

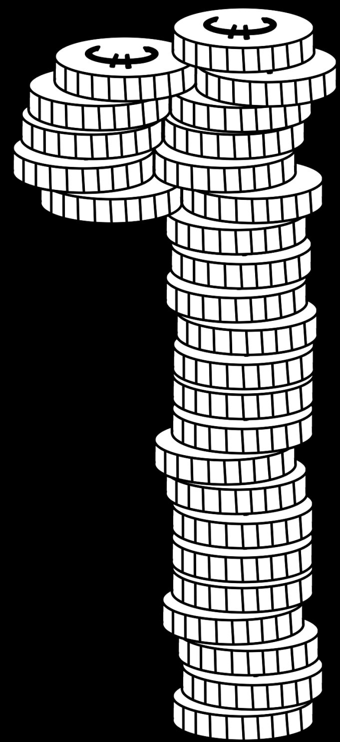
Elevating the Early Warning System and maintaining competitive advantage.

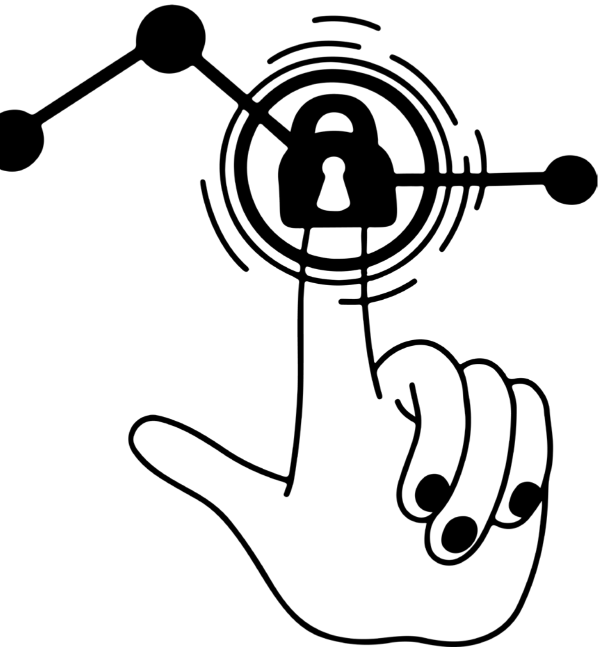
Working Paper on Risk Management and Sustainability.



July 2023

INTRODUCTION





In today's fast-paced and interconnected world, the need for timely anticiptory alerts has become increasingly crucial. These alerts serve as indispensable tools for detecting and predicting potential risks, allowing individuals, organizations, and governments to take proactive measures.

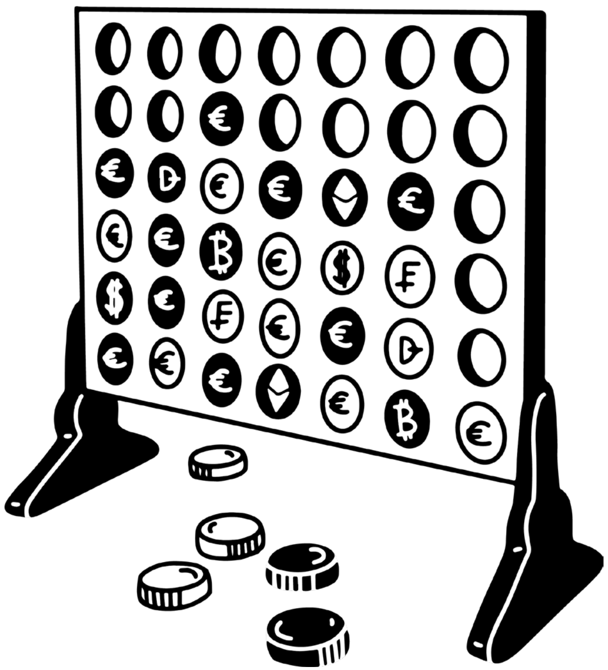
In this Exxeta Working paper on Risk Management we will delve into the enhancement of early warning systems through the incorporation of new risk drivers and the adoption of advanced technology. The importance of digital risk transformation has been highlighted by global CEOs and CROs who express concerns about automating critical operations and relying on automated judgments. Extensive research and discussions have raised questions about the mandate, role, and organization of the risk function in the digital era, as well as the necessary skills, talent, culture, and work methods for a digitized risk function. Additionally, there is a need to understand how to initiate a transformation program and the potential effects of such a transformation.

The impact of data, analytics, and digital tools on risk-related aspects extends to all facets of life, including the banking sector. Initially, banks focused their digitization efforts on enhancing customer-facing experiences such as onboarding and servicing. However, the scope of transformation has expanded, with risk management becoming a vital area of focus. Recognizing the value of digitizing risk, banks are now prioritizing investments in this domain. In fact, digital risk transformations are already in progress in major banks, with over two-thirds of G-SIBs reporting the implementation of such transformations. To keep pace with the rapidly evolving digital landscape, the risk function must accelerate its digitization efforts. Trends like the use of generative AI and widespread data access make it increasingly challenging to maintain analog operations as customer-facing activities and processes transition into the digital age. Banking clients, seeking to maintain a competitive advantage, have raised the demand for early warning systems that provide actionable information well in advance.

As we delve into this paper, we will focus on the enhancement of early warning systems by incorporating new risk drivers and leveraging advanced technology. By gathering and analyzing data from diverse and unconventional sources such as social media platforms (e.g., Twitter, Reddit), business-review ratings, and the entire supply chain, the risk function can enhance its ability to identify emerging risks. Automation of processes and collaboration with other stakeholders will facilitate efficient decision-making and reduce biases. Advanced analytics will further enhance the consistency and accuracy of risk models, while a flexible risk data architecture will enable seamless integration of risk solutions into various banking platforms. Leadership will benefit from self-serve risk explorer dashboards, informed by comprehensive risk capabilities, enabling them to act on risk-driven strategic guidance well in advance.

To fully unlock the potential of the risk function and provide timely, in-depth, and forward-looking insights and guidance, a comprehensive evaluation and redefinition of its mandate, position, organizational structure, culture, talent pool, and operational methods will be crucial. Risk managers widely agree that establishing this digital state within banks will in the near future have significant financial implications, ensuring resilience and success in the dynamic digital era. By undertaking the digitalization of credit procedures, financial institutions can enhance their operations in the face of emerging challenges. The implementation of digital transformation in credit risk management has resulted in increased transparency in risk assessment. It enables financial institutions to adopt a more robust approach to risk management, potentially expanding their operations. This includes implementing risk-based pricing strategies tailored to specific clients, providing expedited customer service without compromising risk thresholds, and improving portfolio management. Automating credit processes and digitizing crucial stages in the credit value chain can lead to cost reductions of up to 50 percent. Moreover, digitizing credit risk offers additional benefits, such as protecting bank revenue and reducing the likelihood of loss by 5 to 10 percent.

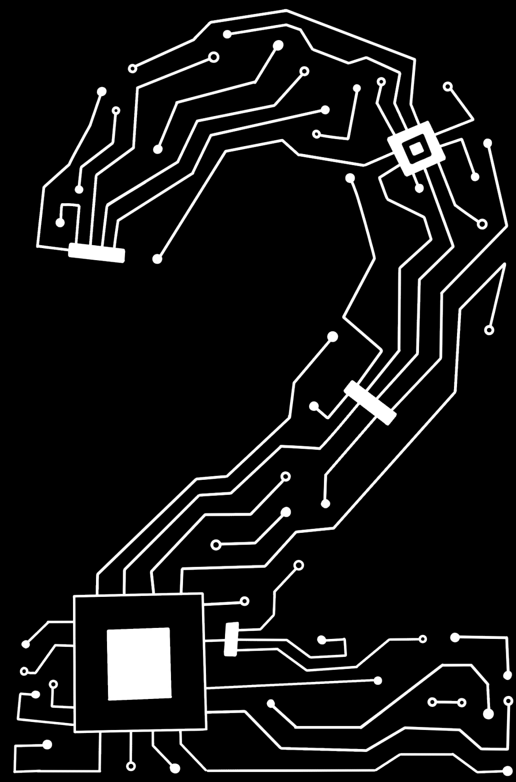
However, several risk-related procedures in financial institutions still lack digital capabilities. While resources have been allocated to develop digital credit risk interfaces, customer demand for seamless digital financial management remains unmet. Enhancing digital capabilities in client-facing operations presents significant opportunities. Financial institutions need to explore the credit risk value chain to identify areas for value generation through digital transformation. A comprehensive mapping and analysis of the credit risk workflow will help



identify and utilize these opportunities systematically. Evaluating the financial benefits of revenue enhancement, cost minimization, and credit risk mitigation at each stage, considering implementation expenses, will help pinpoint the most lucrative domains for digital transformation. The development of a digital credit risk initiative necessitates acknowledging the unique characteristics of the risk function that differentiate it from other functions, such as frontline digital Wealth Management. Regulators are unlikely to accept the traditional characteristic approaches of conventional digital transformations when it comes to risk management. Conducting live launches of „minimum loveable products" for the purpose of assessing and improving in production is generally not a suitable approach for most high-risk endeavors. Most digitization strategies prioritize enhancing the customer experience.

The realm of digital risk encompasses the involvement of external customers in certain domains, such as credit value chain and its delivery. However, the primary emphasis is placed on internal customers, stakeholders, and regulatory bodies in most areas. Furthermore, the management of digital credit risk is an interdependent undertaking that relies on information from all organizational units and operations. The pace of development is constrained by the diligent handling of these interdependencies. Innovative methodologies, such as agile and digital labs, offer efficacious alternatives for gradually executing solutions.

ENHANCEMENT OF CURRENT MONITORING FUNCTION



2.1 UNDERSTANDING EARLY WARNING AND ITS COMPETITIVE ADVANTAGE

The basic principle in early-warning modeling is to inform on various contributors to financial distress. It deals with contributors that **„pave the way"** to performance issues with the purpose of avoiding vulnerabilities all together. As such it lends itself to long-term predictions and less to exact timing of a trigger leading to financial distress. Specifically, in the banking sector these systems should be supported by appropriate IT and data infrastructure to detect increased credit risk in different portfolios, sub-portfolios, industries, geographies, and individual exposures.

The Early Warning Indicators (EWIs) should have predefined trigger levels aligned with credit risk appetite and policies, along with escalation procedures and assigned responsibilities for follow-up actions. Furthermore, the framework should consider the relevance of indicators in relation to transaction characteristics, borrower types, or homogeneous portfolio groups. When an EWI is triggered at the individual exposure, portfolio, sub-portfolio, or borrower group level, institutions should apply more frequent monitoring and consider placing them on a **watch list**. They should also undertake predefined measures and mitigation actions. The risk management function, heads of credit granting functions, and the management body should regularly review specific reports generated from monitoring the watch list. Institutions should tailor their contact and communication with borrowers during payment difficulties based on individual circumstances, following the guidelines on arrears and foreclosure provided by the regulatory body. As part of ongoing **credit risk monitoring**, institutions should consider various signals indicating credit quality deterioration. These signals include negative macroeconomic events, adverse changes in borrowers' financial positions, significant drops in turnover or cash flow, deviations from forecasts or business plans, changes in credit risk of transactions, market volatility, and other factors affecting the borrower's ability to meet debt obligations. Institutions should take immediate action, in accordance with their policies and procedures, when an EWI is triggered. Relevant credit decision-makers should assess the severity of the event and propose suitable actions and follow-up, documented, and communicated to the institution for further implementation.

EXHIBIT: 1 - Understanding Early Warning, from Early, Late to Action of Last Resort

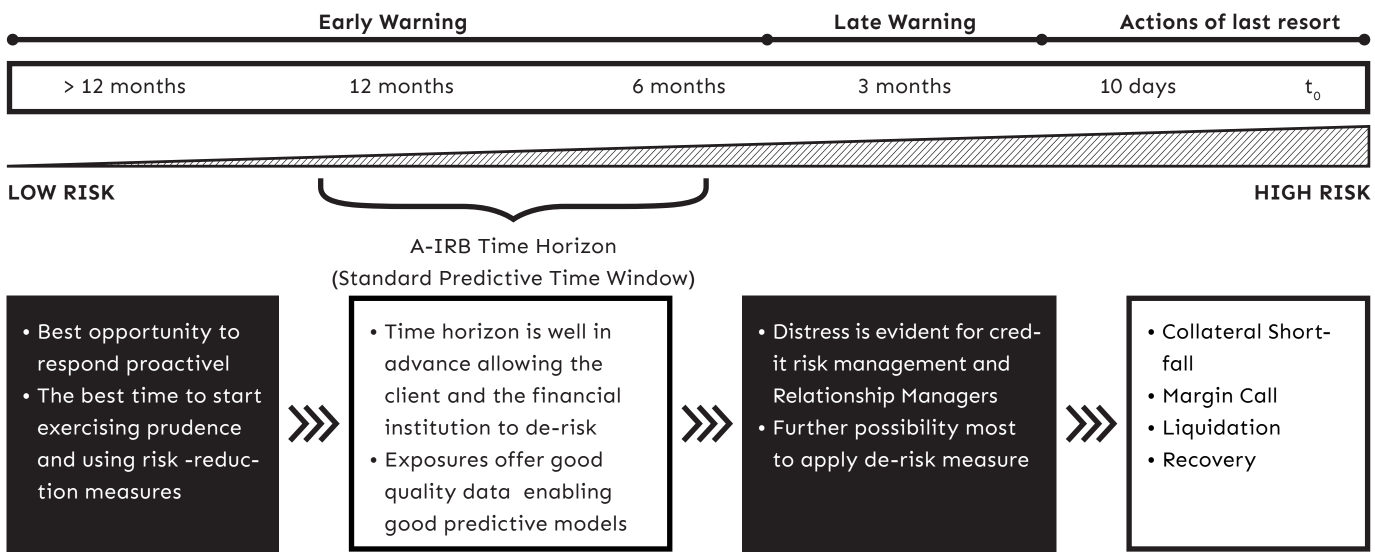
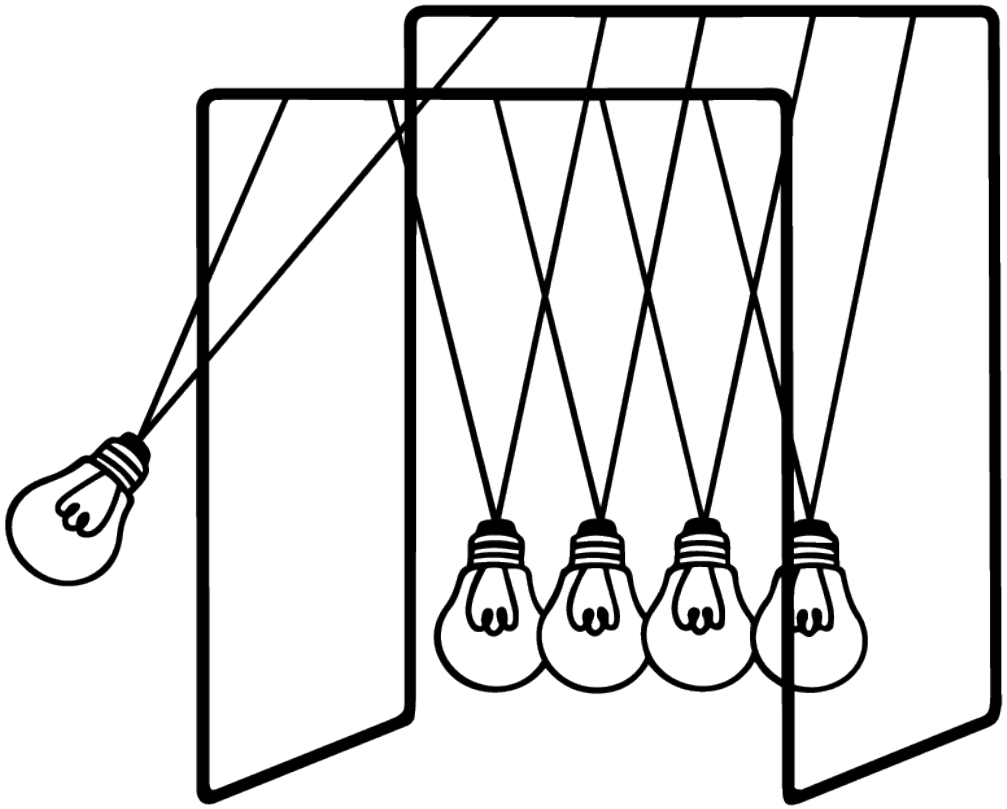


Exhibit 1 depicts a chronological sequence of warning indicators occurring at distinct time intervals, spanning from over 12 months to 10 days. The horizontal axis is demarcated with markings at regular intervals, thereby enabling a sequential representation of time. **Early warning signals with low risk** are displayed at the leftmost end of the timeline, beyond the 12-month mark. This time frame represents the most effective chance to take proactive measures. The view implies that there exists a sufficient duration to recognize potential risk and execute measures aimed at mitigating risks. The initial phase provides a favorable opportunity to exercise caution and implement proactive measures to alleviate any potential hazards. As the time frame approach the present time, it becomes apparent that at the 12-month juncture, there are **early warning indicators that suggest a moderate level of risk**. Within this temporal framework, the horizon of time remains sufficiently distant, affording both the client and the financial institution the opportunity to mitigate their positions' risk. The quality of the available data is sufficient to enable the development of dependable predictive models. This time frame presents a chance to carefully observe the circumstances, make well-informed choices, and adapt investment or lending tactics correspondingly.

As we look towards the near-term future, specifically at the **three-month mark**, we have detected late warning signals that carry a moderate level of risk. During this phase, the manifestation of distress becomes apparent in the realm of credit risk management, prompting heightened involvement from Relationship Managers. The urgency to implement de-risking measures is increasing. Conducting a thorough evaluation of the available alternatives, reassessing one's stance, and implementing appropriate measures to alleviate the potential hazards linked to the detected cautionary indications are of utmost importance. Upon reaching the **10-day threshold**, indicators of elevated risk become apparent. This stage signifies the final option for implementing measures. The imperative to confront the hazards is currently at its apex, and expeditious actions must be taken to alleviate possible damages or unfavorable outcomes. The current era necessitates resolute measures to proficiently handle the hazards, considering the restricted timeframe at hand. The purpose of the chart is to visually depict the chronological sequence of warning indicators, levels of risk, and suggested courses of action. This aids stakeholders in comprehending the ideal timing for proactive responses, the significance of implementing risk reduction measures early on, and the crucial nature of prompt actions in the face of high-risk circumstances. Through the utilization of this chart, both individuals and organizations can proficiently manage and navigate potential risks across varying timeframes.

Early risk detection in the finance industry is greatly aided by the Early Warning System (EWS). The EWS offers a wide range of signals over a period of more than six months. Consolidating these signals and lowering false positive results is difficult, especially when using static rules. It's crucial to remember that signals are constantly suggestive and call for additional research.

Customers are seen to be included in the process as early as possible, ideally before the „Action of Last Resort" is required. The EWS is already being used in the first line of defence, maybe even for the top 100 consumers, according to conversations with Exxeta clients. This emphasizes the significance of early risk identification and management. Customers' participation in the EWS enables proactive risk management and gives them the chance to reduce risk before it's too late. Potential risks can be recognized and swiftly addressed by applying



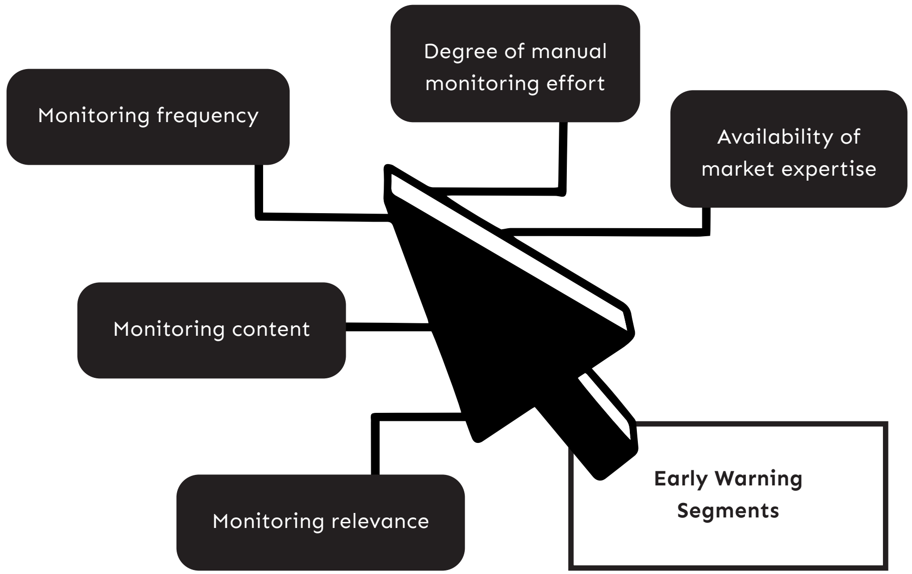
the EWS in the first line, enhancing the credit portfolio's long-term stability.

A comprehensive strategy is necessary for the implementation or adjustment of early warning procedures in the credit industry to satisfy regulatory requirements and maintain the efficiency of early warning systems. It may be advantageous to use automation technologies like Robotic Process Automation (RPA) to generate early warning reports and to use machine learning to increase efficiency. Early warning systems can be tailored to portfolios, allowing monitoring frequency and effort to be adjusted accordingly. Early warning methods should be modified or implemented in an organized manner using tested frameworks. The value contribution of early warning systems should be considered when implementing a comprehensive IT architecture gradually. To secure support and commitment from senior management, it is crucial to measure the results on the profit and loss statement.

2.2 CURRENT EWS

In recent years, especially in the wake of big occurrences like the Subprime Crisis, Euro Crisis, Franks Shock, and the Covid Pandemic, existing Early Warning Systems (EWS) have encountered substantial difficulties and deficiencies. These incidents have brought to light several flaws, including poor models, problems with interpretability, and ineffective notifications such as a high degree of false positives. The historically low investment in risk management, which many banks sometimes overlook or see as secondary, is a significant contributor to these circumstances. Certain automation methods, such as Margin Call Automation, weren't steadily implemented until pivotal times like the Subprime Crisis. However, due to a lack of a strategic approach to take advantage of the advances in technology and techniques, many institutions continue to primarily rely on Excel-based solutions. This goes beyond rule-based systems with infrequent input parameter checking by implementing artificial intelligence techniques such as pattern recognition and automated rule development.

EXHIBIT: 2 - Key criteria for segmenting credit portfolios for the determination of EWI



Process and Structure of an Early Warning System

Nearly all financial organizations have similar structures for their credit operations. The first step is determining the degree to which various credit portfolios should be monitored along with regulatory representatives. It's critical to establish portfolio structures and boundaries in a way that makes it possible to analyze each portfolio effectively and efficiently. Experience has demonstrated that early market side involvement in this process increases participation willingness and, as a result, improves early warning efficacy and efficiency. The portfolios must then be segmented to create so-called early warning segments. Processes and the subject matter to be evaluated are properly designed using early warning segments and levels. Based on this, two-dimensional assignments are made for early warning systems and triggers. The following traits, according to experience, serve as useful segmentation criteria:

EXHIBIT: 3 - Segmentation Criteria

Relevance of Monitoring	This criterion considers the amount of credit being watched. Greater monitoring relevance is indicated by a bigger credit volume.
Monitoring Content	The signals and triggers that are used for monitoring are covered by this criterion. Depending on their features, different portfolios could need various triggers.
Monitoring Frequency	This criterion establishes the regularity of monitoring operations. It depends on the informational availability and detail for each portfolio.
Degree of manual monitoring effort	The complexity and size of the credit business are taken into consideration when determining the degree of manual monitoring effort. Depending on the type of portfolio, it decides how much manual monitoring is necessary.
The availability of market expertise	The availability of market expertise is assessed using this criterion to determine the degree of market knowledge necessary for efficient monitoring. Complex transactions could necessitate the use of specialist knowledge, which is frequently transaction-specific and held by only a few people inside the institution.

Financial institutions can create reliable and personalized early warning systems that enable efficient risk management in their credit portfolios by taking these factors into account. By considering these criteria, financial institutions can establish robust and tailored early warning systems that enable effective risk management in their credit portfolios.

EXHIBIT: 4 - Process for Defining Early Warning Signals

Assessing the validity, development, maintenance, and application costs of the triggers, as well as the potential quantity of corresponding signals that may trigger a subsequent process.

Achieved through expert surveys using the Delphi method, which is particularly suitable for manual triggers that lack automated data analysis capabilities.

Aligning early warning segments and levels to establish a comprehensive overview of existing procedures and triggers, facilitating effective management and monitoring of early warning systems.

Triggers can be essentially differentiated and described based on the following distinguishing factors:

- **Monitoring level** - differentiation based on risk group, individual customer, and account
- **Monitoring segment** - differentiation based on customer group, exposure volume, or complexity level of the exposure

Once the segmentation criteria have been established, the next step is to define early warning signals. The following approach has been established for this purpose (see **Exhibit 4** for reference):

1. A matrix is formed from early warning segments and levels, which can be addressed with dedicated processes and triggers. The matrix provides an optimal overview of the existing early warning procedures as the process continues.

2. The identification of potential triggers and signals is best formulated through expert surveys using the Delphi method. This method is also suitable for potential manual triggers where there is no automated data basis for evaluation.

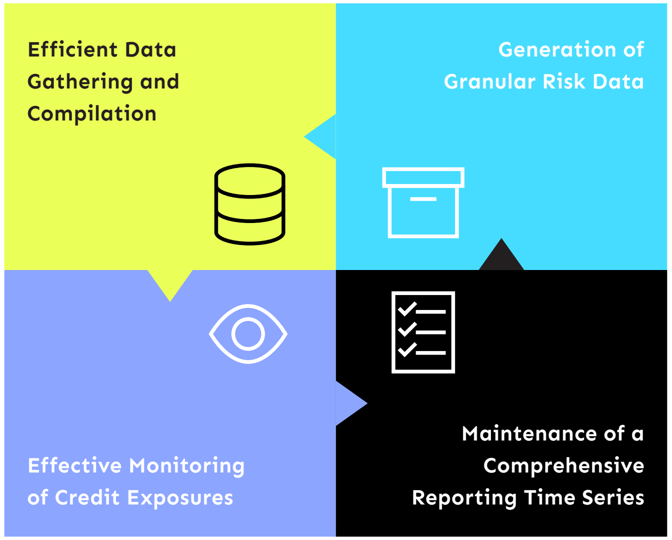
3. This step involves using data mining techniques to identify potential early warning signals in existing data histories. These techniques analyze large datasets to identify frequent patterns that indicate a high likelihood of future defaults. Triggers are classified based on distinguishing characteristics such as **monitoring levels** (risk groups, individual customers, accounts) and **monitoring segments** (customer groups, exposure volume, complexity). Triggers can be qualitative or quantitative in nature, and they can be captured manually or automatically depending on their measurability. The information source also plays a role in trigger classification.

4. At the quantitative level, triggers and signals are evaluated on their individual but also collective merits. Correlations and time alignment with periods of distress or economic downturn are a first step in establishing quality and usefulness. Next steps usually include a multivariate analysis within a set of candidate triggers/signals that can produce sets of possible indicator sets. Finally, as time passes new observations become available and the warning success rate can be observed. If found prudent, as is often the case, recalibration can be performed. Currently, key risk models are reassessed and recalibrated at least once per year.

Aligning to regulatory drivers

EXHIBIT: 5 - Essential Key Focus Areas in a Credit Risk Monitoring Framework

- Enabling automated gathering and compilation of credit risk data, minimizing manual processes.
- Timely and accurate reporting of credit risk information is facilitated.
- Enhances risk management capabilities and decision-making processes.



- Enabling generation of granular risk data.
- Risk data is compatible with institutions' own risk management purposes.
- Risk data meets requirements of competent authorities for prudential and statistical reporting, supervisory stress testing, and crises management.

- Enabling effective monitoring of all credit exposures and collateral.
- Provides a clear overview of the credit decision-making process.
- Ensures transparency and traceability in credit-related activities.

- Ensuring maintenance of an appropriate time series of reporting.
- Captures current exposures, new types of lending, and early warning indicators.
- Covers the credit risk planning horizon.
- Supports tracking and analysis of credit risk trends.
- Facilitates proactive risk management and adaptation of strategies.

Exhibit 5 showcases the effective incorporation of the Loan Origination Early Warning Requirements outlined by the European Banking Authority (EBA) in their Final Report (EBA LOM).

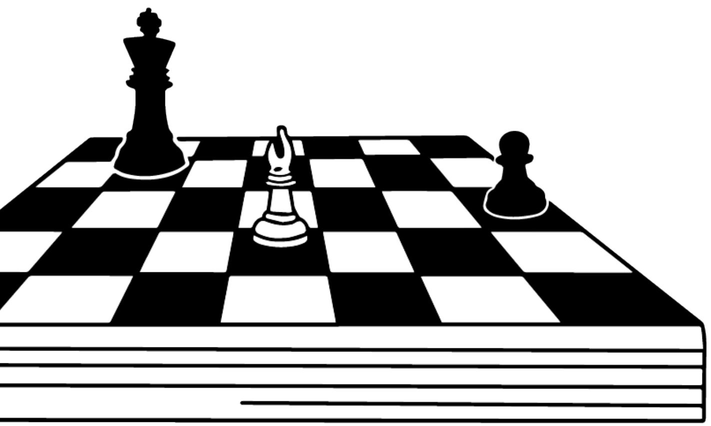
Institutions should think about how to build and apply their framework for credit risk monitoring. These are important things to think about:

Infrastructure for data and the framework: **Automating the collection and compilation of credit risk data** is made possible in large part by the framework and data infrastructure. Institutions can increase the effectiveness and accuracy of data gathering by reducing manual processes. This improves risk management capabilities and supports informed decision-making processes by enabling **rapid and accurate reporting of credit risk information**. Institutions can produce comprehensive and targeted risk data thanks to the framework and data infrastructure. This information complies with the needs of competent authorities as well as their own risk management goals. It guarantees adherence to supervisory stress testing, prudential and statistical reporting, and crisis management. Institutions gain more insight into their risk profiles by having granular risk data. **Effective monitoring of all credit exposures** and collateral is made possible by the framework and data infrastructure, which offer the essential instruments. The openness and traceability of credit-related actions are ensured by this monitoring capability, which provides a comprehensive overview of the credit decision-making process. It assists with risk reduction measures and supports the upkeep of a sound credit portfolio.

Upkeep of an acceptable reporting time series:

The framework maintains an up-to-date reporting time series that includes recent exposures, loan products, and early warning signs. This enables proactive risk management and adaptation to changing market conditions. Institutions can enhance risk management, make informed judgments, and maintain a strong credit portfolio through careful consideration of these variables and implementing a well-designed credit risk monitoring system. Regulators recognize the importance of incorporating Early Warning Signals and customizing the framework to different credit types, ensuring thoroughness and flexibility. This approach enables more effective monitoring and detection of potential hazards, prompt action, and reduces the risk of financial difficulty. It emphasizes the balance between a uniform regulatory framework and flexibility to account for different credit instruments.

2.3 CURRENT ENHANCEMENT CONSIDERATIONS



Banks are aware of the conventional sources and triggers, such as rating deterioration and financial reports, that they monitor. New stressors, such as interruptions in the global supply chain and incidents like the GameStop Short Squeeze, however, demand attention. Even though these factors are acknowledged, many banks continue to use stand-alone Excel-based solutions. It is crucial to update existing procedures and broaden the Early Warning System (EWS) to include more signals to address this. In this development, the watchlist may be enhanced by automation and dynamic control, the integration of Social Media Sentiment Scores with unfavourable news indications, and more. Prioritizing should also be given to aspects like second-order concentrations, supply-chain risks, and sustainability factors including physical and transition risks, disability risks, and reputation risks. Effective risk management in banks depends on modifying the EWS to consider these changing issues and making use of cutting-edge tools and indicators.

A variety of triggers are used by an efficient early warning system to identify possible dangers and problems. These triggers can be broadly divided into two categories: manual expert triggers and technical triggers.

Technical triggers offer objective measurements to signal potential risks and are based on quantitative data. Monitoring overdrafts, examining credit default swap spreads, and keeping watch of changes in credit line drawings above a predetermined level are a few examples of technical triggers. These alerts use data from the bank's credit systems to identify deviations from thresholds or regular trends that might point to a higher risk. However, manual expert triggers rely on the knowledge of credit risk officers and relationship managers. To identify possible threats, these triggers rely on subjective assessment and soft characteristics. Relationship managers might, for instance, view interim financial data or personal issues with business owners as signs of declining financial health. They offer priceless insights that quantitative statistics might not be able to fully capture. Banks can develop a thorough early warning system that detects a variety of potential threats by including both manual and technical expert triggers. These indicators serve as warning signs, prompting banks to take additional steps like tighter scrutiny, more monitoring, or the implementation of risk mitigation techniques. Combining quantitative and qualitative triggers improves banks' capacity to proactively recognize and resolve issues related to their clients.

EXHIBIT: 6 – Exemplary Conceptual View for an EWS in Securities Backed Lending

Social Media Sentiment

- Negative News Sentiment
- Core Triggers
- Days Past Due
- Short-Term Overdraft
- Limit Excess
- Overdue Installments
- Reverse Stress Test Sentiments
- Supply Chain Sentiment
- Sustainability Sentiment
- ...

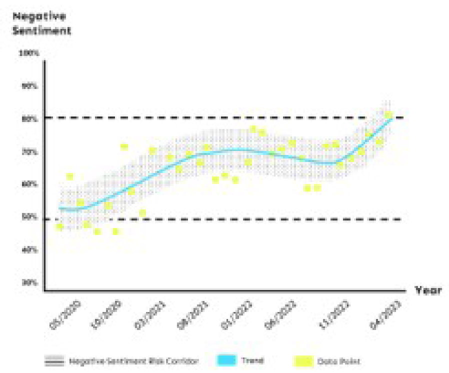
Dimension
- locality
- time horizon
- volume
- variability
- availability
- quality

Scale
- comment
- topic
- subcategory
- category
- Day/Week
- Month /Year

Score (t)

$$\frac{\sum_i^{posts(t)} weight_i(relevance) \cdot Score_i}{\sum_i^{posts(t)} weight_i(relevance)}$$

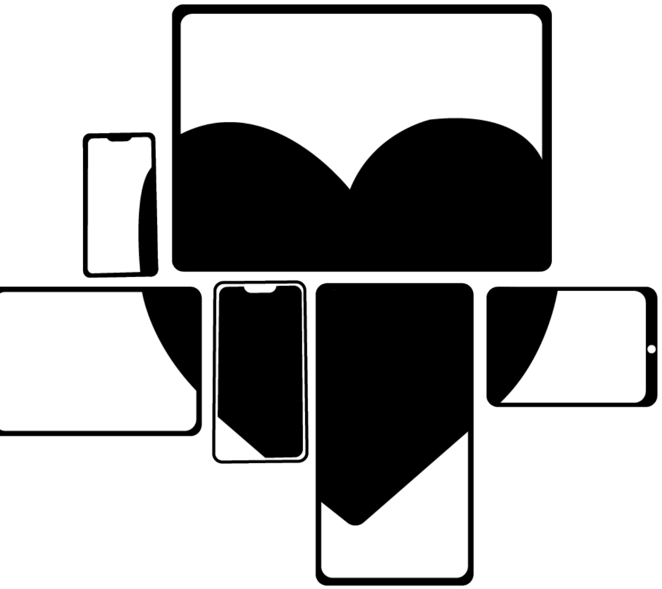
Sentiment Score



- > 50% negative sentiment over 3 months
- 1 Day Change in volume by 50% more
- 1 Day negative sentiment with >80%
- Troll Alert - 5 Day change in „new user"
- Troll Alert – 70% of comments from new users
- ...many other indicators for volume, score, user count, post count, rate change of signaling amongst others

EXHIBIT 6 exemplifies the inclusion of social media risk sentiment as a potential expansion of the Early Warning Signal family. Chapter 3 provides detailed information on this aspect.

CONSTRUCTING A SOCIAL MEDIA RISK SENTIMENT SCORE



Reputation is a valuable intangible asset for any organization. It comes in many forms, whether it is in the eyes of a business partner, a client, or the general public. For an institution like a bank, a good reputation can take years to build and moments to shatter. Among the various relationships upon which an organization establishes its reputation, the one established with the public is by far the most important. A strong reputation in the public's eye is above all else based on trust, transparency, and ethical conduct. A positive reputation enhances an institution's brand value, attracts clients, and helps it maintain a stronghold in the hearts and minds of their customers.

The risk of reputation loss refers to the potential damage to an organization's or individual's reputation that can result from negative public association, loss of trust, or a decline in confidence due to various other factors. This type of risk arises from the conviction that a negative perception can have significant consequences for an institution's everyday operations as it influences the ability to attract and retain clients, maintain business relationships, access capital markets, and ultimately, sustain its financial performance.

As a response to this reality, the inclusion of a Social Media Risk Score within the Early Warning System framework is crucial in today's digital landscape. By automating activities that, until recently, required highly trained professionals, such as sentiment scoring in of social media in the retail world, it is now possible to efficiently process enormous quantities of data from numerous sources. These online communities have a profound impact as they support information exchange, suggest dos and don'ts for retail investors, but also foster swarm intelligence. For that matter, they provide an intimate insight into the users' preferences and thus should be closely monitored.

Notably, sentiment analysis is valuable not only to the retail investors but also to professionals, including hedge funds and short sellers, who actively engage with these platforms. Hedge funds, for instance, closely gauge market sentiment, identify emerging trends, and make informed investment decisions. The GameStop short squeeze demonstrated how sentiment within these communities can have a dramatic impact on stock prices and disrupt traditional market dynamics. Therefore, understanding and accurately interpreting the collective sentiment in these online communities is crucial for hedge funds to assess the potential risks and opportunities, manage their positions, and adapt their strategies accordingly. Due to these circumstances, financial organizations and regulating bodies have started to investigate the possibility of widespread application of advanced machine learning (ML) powered sentiment scoring algorithms in assessing risk.

EXHIBIT: 7 – Use Case Scenarios for Sentiment Scoring in the Financial Sector

Market Oriented Approach

Customer Satisfaction
How - analyze customer feedback from surveys and customer support
Why - identify areas of improvement, address customer concerns, and enhance customer satisfaction

Brand Reputation Management
How - monitor sentiment across internal or external online channels
Why - gauge public opinion about their brand and track any shifts in sentiment over time, address customer concerns, and take proactive measures to manage their brand reputation effectively

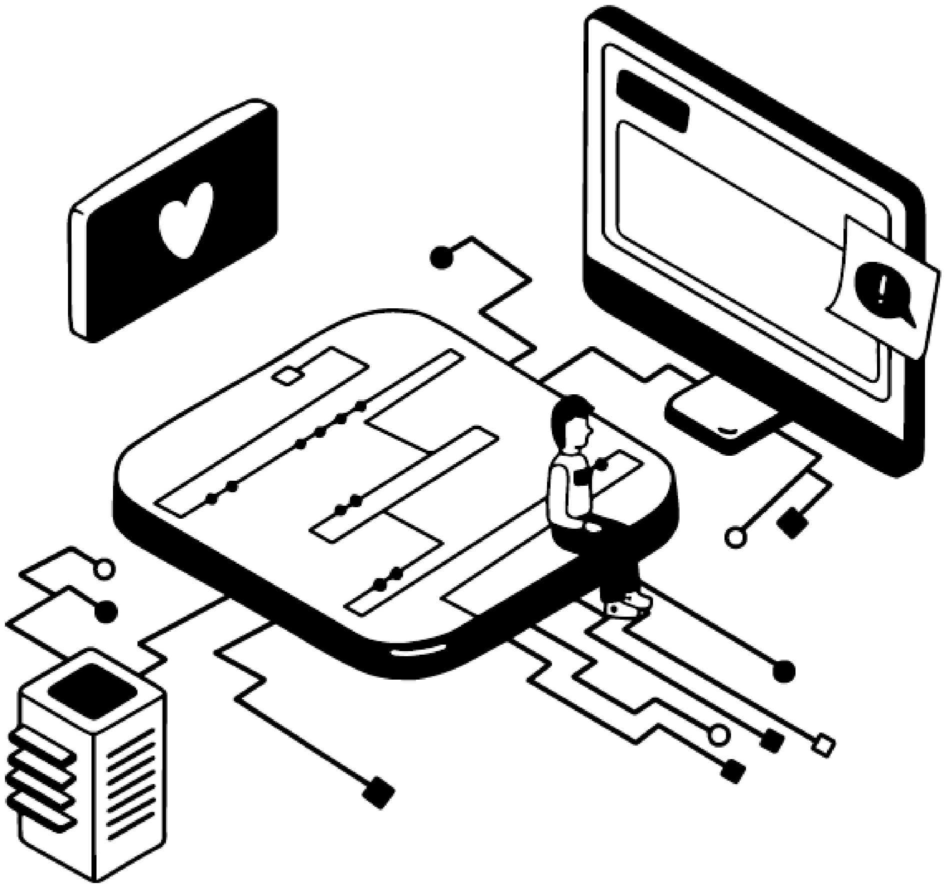
Product Development and Marketing
How - analyze product discussion forums with keyword filtering
Why - calibrate product development efforts, tailor marketing campaigns, and identify opportunities for innovation, gauge sentiments regarding existing products or potential offerings

Customer Oriented Approach

Market Research and Competitive Analysis
How - analyze financial news and associated forums or comment sections
Why - recognize market dynamics, and competitive advantages, assisting in strategic decision-making

Trading and Investment Strategies
How - analyze investment discussion boards, forums, trending/influencer threads
Why - identify sentiment-driven market movements and adjust investment decisions accordingly

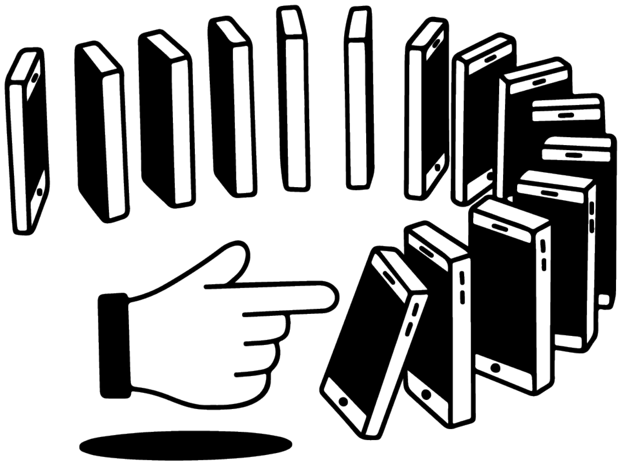
Risk Assessment
How - assess the sentiment of news articles, and key social media posts related to specific companies, industries, or market trends
Why - leverage this information to gauge market sentiment, evaluate higher-risk exposures, and make informed investment decisions



Recognizing the value and potential of sentiment analysis, these entities aim to leverage the insights derived from ML-powered sentiment scoring to enhance risk assessment processes and improve decision-making. This proactive approach allows them to identify potential market disruptions, market manipulation attempts, and systemic risks that could have far-reaching consequences.

As social media platforms continue to dominate communication channels, the influence they wield over financial markets cannot be ignored. Early warning systems enhanced with AI algorithms can identify patterns, detect anomalies, and alert risk managers to emerging social media-driven reputation risks in aggregate but often enough in real-time. To this end, we present a proof-of-concept (PoC) analysis that outlines on one example the effectiveness of ML-powered sentiment scores in predicting distress and its implications for financial organizations and regulatory frameworks.

3.1 MODEL CONSTRUCTION AND CLASSIFICATION RULES



Motivated by recent events where negative public opinion led to the downfall of major corporations, in this analysis we have investigated the feasibility of constructing a lightweight Early Warning Indicator (EWI) to reputation risk. This study considers the aggregated social media sentiment as observed in Open Source Intelligence (OSINT) data for a particular global organization in order to assess the power to signal distress at an early stage.

Standard Risk Modelling Approach

With regards to the current risk management practices, the "social media risk score" can be pictured as an EWI or a risk driver within the normal Internal Ratings-Based (IRB) models. In the simplest terms, EWS could be imagined as a classification problem. If such, then during development of a predictive model the period of distress would be identified on a case-by-case basis, and an appropriate learning model trained using available numeric data. The resulting model output would be considered either a probability, leading to the implementation of linear models like logistic regression, or a "traffic light" flagging, resulting in a decision tree. In any case, it would be an ordinal set that compares contracts and their risk indicators within a target time horizon.

Among the numerous practical reasons why this approach has already not been adopted is the fact that not all exposures, big or small, have a sentiment score that can be computed. For example, there is no large community that is actively commenting on the status of each individual mortgage that a bank has. Secondly, even if there were data available, like client conversations for each exposure, which is not practical in the least, until recently, general sentiment analysis has not risen to the occasion of being applicable on large datasets with a stable output. This lack of good quality data, which cannot support the inference of a long-run averaged (LRA) risk classification regarding observed social media score, motivated risk experts to consider different avenues.

EXHIBIT: 8 - Standard development and application approach to risk modelling within the banking sector

Preparation	**Target Segment** • Identify target exposure class or type of exposure and client pool	**Data** • DQ, pooling, rating and calibration differentiation • external third-party dataset reliance and availability	**Objective** • Internal, behavioral, applications, Advanced/Foundation IRB
Modelling	**Method/Model** • PD, LGD, EAD, EL, RWA • Lin. Regression, Decision tree, NN	**Segments** • Portfolio segmentation • Aggregate, pool, infer	**Objective** • Economic Downturn estimation • Long-Run Average (LRA) • Margin of Conservatism
Application	**Staging** • Software solutions, data acquisition • Cloud, local/internal hardware	**Recalibration** • Material change triggered • Portfolio expansion, re-appropriation • Regulatory mandated re-calibration	**Monitoring** • User-appropriate application • Constant monitoring interface

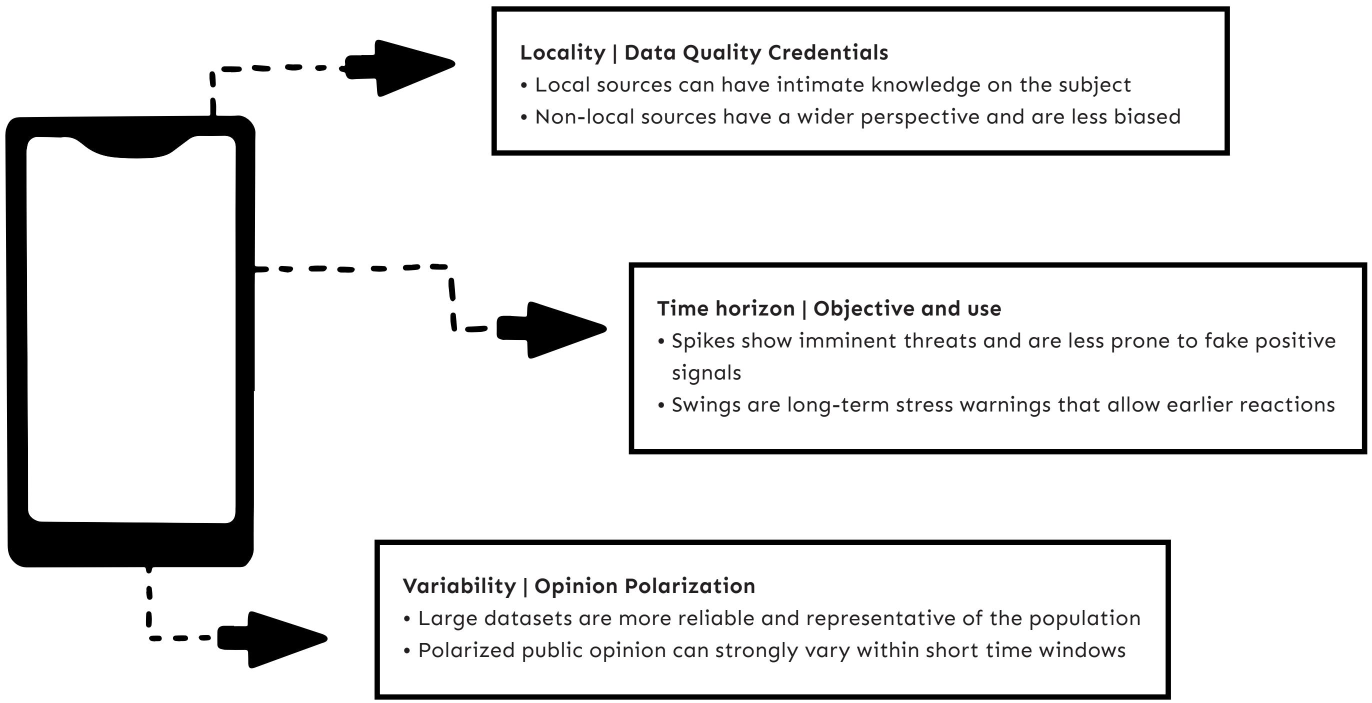
With the aim of still of gaining actionable insights, sentiment scoring has started to be a service provided for those cases where data limitations were under control, and only by those suppliers that had advanced knowledge and expertise to produce good (enough) quality outputs.

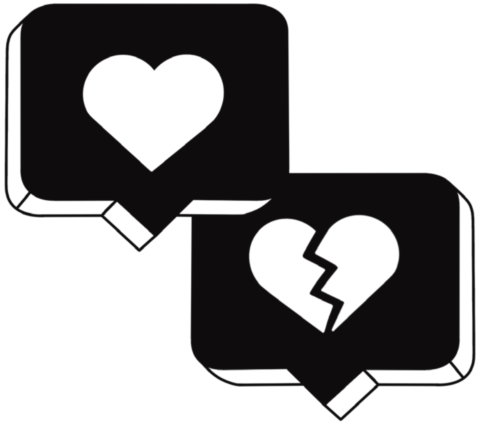
As the time of modern ML algorithms emerged, it has become possible to estimate the positivity/negativity of text messages in a more streamlined manner and thus introduce a new continuous EWI that can be understood on its own merits and observed as a function of time.

Modern Algorithms and Methodology

Most advanced sentiment scoring algorithms, which obtain results that are highly correlated with those created by humans, are all created using Artificial Neural Networks. Specifically, Large Language Models (LLM) like the Generative Pre-trained Transformer (GPT) and Language Model for Dialogue Applications (LaMDA) have started to infiltrate almost every aspect of work in the digital realm. With the many versions of open-source transformer models that were trained on large datasets of high-quality it has become feasible to implement such advanced methodology with household hardware and at high frequency to OSINT data. That is the choice also made in this study.

EXHIBIT: 9 - Key dimensions of sentiment scoring for the purpose of Early Warning flagging





First steps
When aiming to construct a sentiment score that informs on changes in public opinion and thus consequently warns of general risk increase it is important to consider several characteristics of the model and data. Below we outline key considerations for the dimensions of locality, time horizon and volume as crucial first steps. All other degrees of freedom should be based on this initial analysis and understanding of the mentioned general characteristics.

Locality
Information and its availability drive opinions. People's minds are influenced whether they are involved with a specific company, making their involvement a key puzzle piece in interpreting their sentiment. Organizations using general sentiment need to be cautious about overgeneralizing conclusions based solely on local data and consider supplementing it with broader, non-local data sources for a more comprehensive understanding of social media sentiment.

The obvious advantage of using local data for social media sentiment analysis is the relevance and contextual understanding it provides. Local data captures sentiments and opinions from individuals within the specific geographic region, allowing for a more accurate assessment of social media sentiment within that market. It considers cultural nuances, language preferences, and region-specific factors that may influence sentiment towards a company or organization. By using local data, organizations can gain valuable insights that are directly applicable to their target audience. This enables users to tailor their strategies and messaging accordingly.

One limitation is the potential for bias or low representativeness. Local data may not provide a comprehensive view of social media sentiment if it excludes diverse perspectives or fails to capture the opinions of certain demographic groups. Additionally, local data may be influenced by regional events or trends that might not accurately reflect the broader sentiment towards a company or organization.

In contrast, utilizing non-local data for social media sentiment analysis is also beneficial due to its broader perspective. Non-local data incorporates sentiments and opinions from a wider range of sources, potentially spanning multiple regions, countries, or even global audiences. This broader view can reveal trends, patterns, or sentiments that might not be evident when analyzing only local data. It allows organizations to gain insights into how their brand or organization is perceived on a larger scale, helping them identify opportunities for expansion or detect potential risks across different markets. However, non-local data may not capture specific regional or cultural factors that influence sentiment within a particular market. Analyzing sentiments without considering local context can lead to misinterpretations or misalignment in strategies and messaging. Additionally, language and translation issues may arise when dealing with non-local data, as accurate translation and interpretation are crucial for understanding sentiments accurately.

To obtain a comprehensive view of social media sentiment, organizations should consider a balanced approach by combining both local and non-local data sources. This allows them to benefit from localized insights while also gaining a broader understanding of sentiments across different markets.

Time horizons

Early warning signaling aims to detect signs of potential distress at an early stage. Sentiment scores are advantageous because they measure time dependent public opinion in a numerical form without calibration to previously observed distress. Depending on the type of sentiment time behavior it is possible to recognize early signs of distress, as a 6-month plateau of bad sentiment, or imminent risk, if for example the negative score hits an overwhelmingly bad grade. Depending on the aimed time horizon it is crucial to recognize trigger types.

Standard credit risk models consider periods from 6 to 12 months as adequate time horizons. With this aim an EWI would be appropriate if it could be shown to correlate with historical downturn periods, while not correlating with other financial KPIs. These models generally use the Through-the-cycle (TTC) approach and capture a robust estimation of risk over the entire business cycle. As the horizon is well in the future, this approach measures gradual evolution of risk drivers and allows proactive intervention before problem escalation but can suffer from false positives and requires additional efforts and engagement.

On the other hand, in moments of large and statistically significant deviation in public opinion, last-minute analyses that focus on the point-in-time (PIT) approach signal evident problems leaving little room for ambiguity.

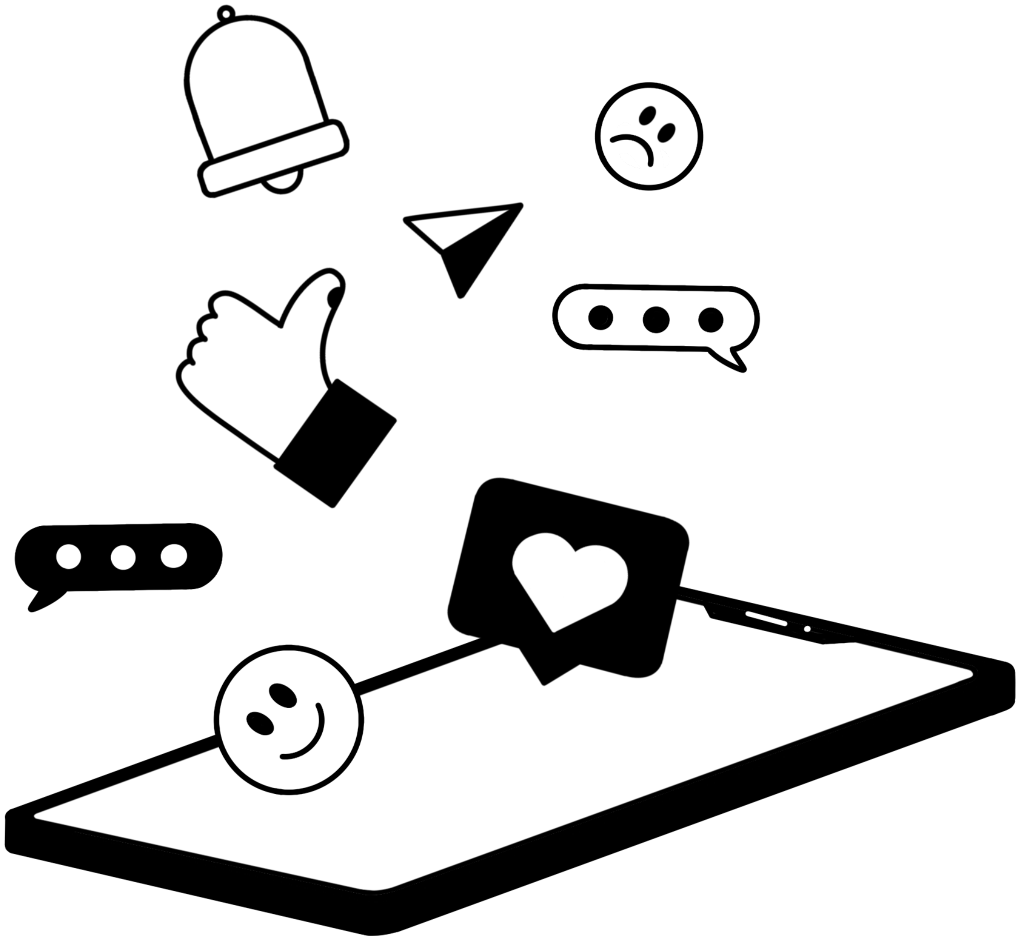
The triggers here would be for example, a signal outside previously observed behavior at the level of several standard deviations. In this case the warnings come late and are limited of scope because often enough the public is informed using other news sources that are readily available in an earlier stage thus making proactive mitigation procedures late to act. When developing a sentiment score often it is not possible to observe long periods of time or large numbers of companies thus learning "from experience" and developing a robust distress predictor. To this end, it is important to develop support indicators or triggers that will not be able to offer an absolute PD score but can indicate change in sentiment and thus quantify the size of the risk via the size of the change.

Volume and variability

Not all companies have the same visibility or are equally in the public spotlight. For this reason, the use of large enough data sets is indispensable when estimating social media sentiment for several reasons. Firstly, large data sets provide a more representative sample of the population's sentiment. With a larger data set, there is a higher likelihood of capturing a diverse range of opinions and perspectives, ensuring a more accurate estimation of social media sentiment. This helps to mitigate the risk of bias and ensures that the analysis reflects the true sentiment of a larger segment of the population.

Secondly, large data sets allow for more robust analysis and statistical modeling. With a sufficient volume of data points, it becomes possible to apply advanced techniques and algorithms to extract meaningful and actionable insights. These techniques can include sentiment analysis, natural language processing, and machine learning algorithms that can uncover patterns, trends, and sentiment shifts at a granular level. Large data sets enable the identification of subtle nuances and sentiment variations across different demographics, regions, or time periods, providing a more comprehensive understanding of social media sentiment.

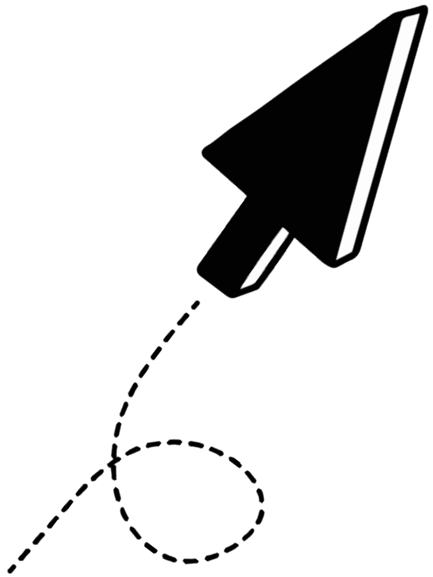
Finally, a large dataset might stem from the fact that the data source has a high frequency of incoming observations. Such a condition can allow finer time granularity and perhaps even real-time measurement of sentiment (live chat/comment sections).



The ramifications of such detailed look are immense as they provide the capability to quickly react and realign before larger damage is done.

Overall, using large data sets to estimate social media sentiment increases the accuracy, reliability, and generalizability of the analysis. It supports organizations and researchers in making informed decisions, design effective strategies, and tailor their communication approaches based on a solid foundation of comprehensive and representative social media sentiment data.

3.2 A CASE STUDY FOR A SWISS G-SIB



In the banking sector, using sentiment scores holds significant business implications, particularly in monitoring and managing their public perception. Tracking sentiment allows banks to understand customer sentiment trends, in-time reactions during critical events or product launches, identify general areas of concern or dissatisfaction, and proactively address them.

Positive sentiment indicates customer satisfaction, which can contribute to increased brand loyalty and customer retention. On the other hand, negative sentiment signals potential issues that need immediate attention to prevent reputational damage and customer attrition. As a POC a Social Media Risk Sentiment has been implemented in an efficient manner, in extreme cases the Social media Risk Sentiment could on time signal the possibility of such adverse events like a bank runs or company boycotts.

The cost-to-benefit ratio of employing social media sentiment analysis is generally regarded as highly favorable. The reason why is that insights derived from sentiment scores have a high actionability and thus very effectively contribute to improved customer experiences, overall better marketing campaigns, and both credit and reputation risk management. In general, the net effect of using sentiment scores is positive and thus beneficial. Furthermore, non-aggregated sentiment analysis empowers banks to stay connected with their clients, respond promptly to feedback, and ultimately drive their long-term success in a competitive marketplace.

The analysis

The purpose of this study was to explore the applicability of Neural Network (NN) algorithms to (OSINT) within the Early Warning System (EWS) domain. Specifically, this PoC focuses on the time evolution of user-provided sentiment and its predictive capabilities to identifying vulnerable states prior to financial crises.

A Swiss based Global Systemically Important Bank has been selected as the subject of investigation. The choice was made due to the public discourse. The dataset under analysis consisted of relevant social media platforms entries collected over a three-year period commencing on May 1st, 2020. The social media risk sentiment score was calculated by analyzing the language used in individual social media posts and classifying it as positive, negative, or neutral.

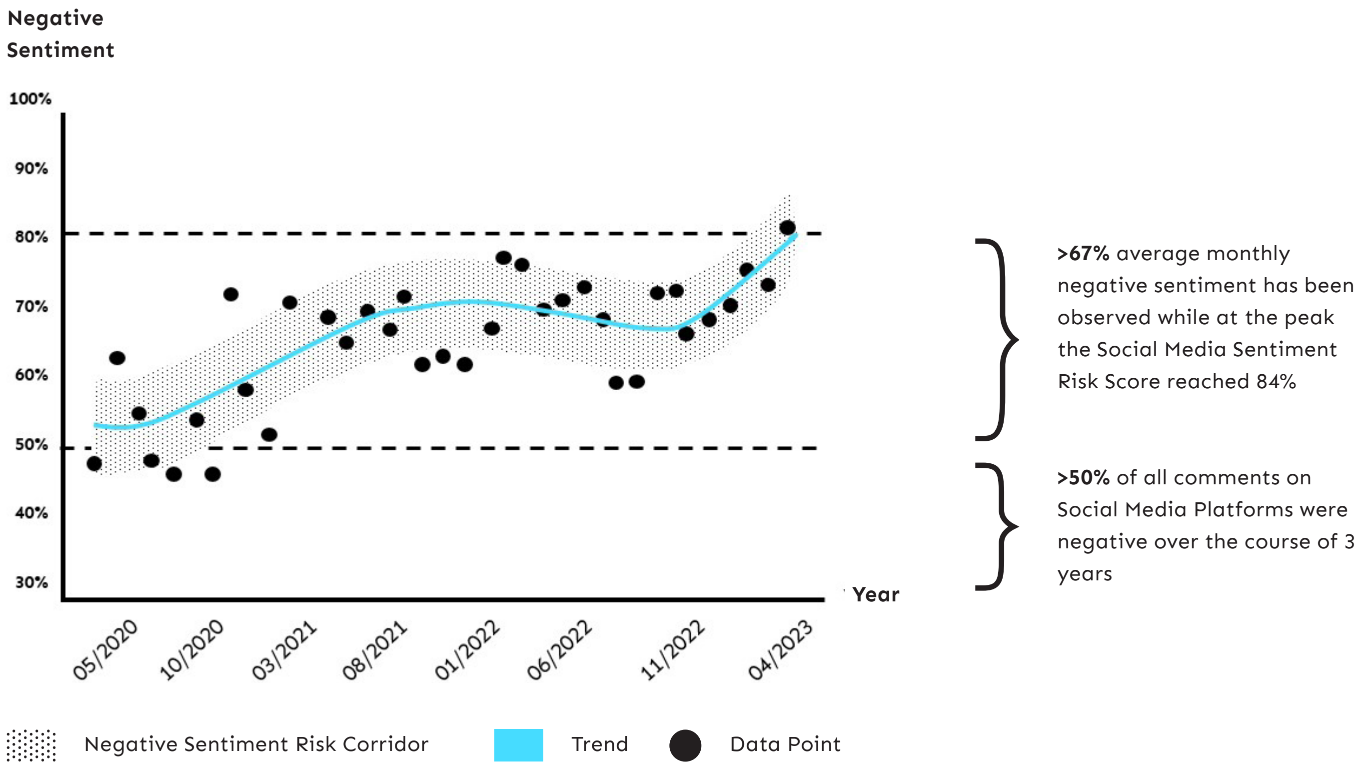
By aggregating the categorized social media comments over time, such as days, weeks, or months, a way to recognize and investigate trends and recognize the overall sentiment trajectory was developed.

To construct the social media risk sentiment score, multiple data streams have been identified, each presenting its own unique set of challenges. To avoid known data quality pitfalls, after initial investigations, appropriate data filtering methods were developed so that irrelevant entries such as spam or commercials did not play a large role. The scores were assigned on posts on a whole, but also on a sentence level, from which post sentiments were calculated. Each entry was considered as a sum of statements with different weight of relevance to the considered topic, namely the Bank. The results were integrated over a selection of time periods and a polynomial trend line was obtained using standard statistical analysis methods. Different approaches to filtering, aggregation, scoring and fitting have been considered, at both post and sentence level, and used to estimate systemic uncertainty due to method selection.

Observations

The data presented in EXHIBIT: 10 shows a consistent negative social media platform based public perception of the Bank, with over 50% of the sentiment being negative throughout the considered timeline. This sustained negative sentiment raises concerns about the Bank's reputation and public perception. Furthermore, it is notable that the negative sentiment reached a plateau in the second half of 2021 and persisted for more than a year. This extended period of negativity suggests deep-rooted and continued issues that were of significant concern to the public. The subsequent steep rise in negative social media sentiment observed by the end of 2022, consistently increasing month-over-month, highlights a worsening situation.

EXHIBIT: 10 - Social Media Sentiment Risk Score



This indicates that there were ongoing discussions and concerns among the public, potentially indicating further deep nested problems faced by the financial institution. The presence of a Negative Sentiment Risk Corridor, with one standard deviation of variability between the predicted and observed sentiment, adds to the indicate the significance of the observation. This risk corridor underscores the predictive power of social media sentiment analysis in signaling aggregated risks such as reputational, operational, and credit risks. Credit Risk from a collateral management point of view considering that a permanent negative perception leading to a steadily decrease of the share-price, which also strongly correlated in the observation period. A continued decrease in the share price would have potential impacts from a credit monitoring point of you considering a financial security backed loan (e.g. Lombard) having shares of that particular institution in the collateral portfolio.

It can be supposed that the maintained plateau of negative sentiment culminated with the early 2023 worsening. In that sense, the late 2020 increase in negativity can be recognized as an early warning trigger, pointing to distress to come in about 18 months. If the sentiment went down, then the plateau would be no more, thus signaling a healing of sort and an increase in the positive outlook.

Ex ante, it is challenging to define a statistically motivated plateau length that if achieved can serve as a hard trigger. However, dynamic thresholds can be set by senior management and for instance including public relations as organizational experts determining the trigger for negative public opinion and making it a candidate at early stage for the monitoring watchlist.

Overall, these findings prove that the public discourse can be of short memory, while in contrast a sentiment score could still have the power to point to a shift in outlook and thus aid in addressing possible distress well in advance. Finally, it highlights the importance of considering various potential risks and implications associated with such sentiment patterns, including their impact on the Bank's reputation, operational effectiveness, and creditworthiness.

3.3 A CRITICAL VIEW ON THE SOCIAL MEDIA SENTIMENT RISK SCORE

It is crucial to acknowledge that the integration of AI, particularly the social media sentiment score, into the traditional rules based early warning system could have played a pivotal role in preventing losses and increasing preparedness and readiness. By leveraging responsible AI practices and ensuring high data quality, the early warning system can capture and analyze sentiment from social media, providing valuable insights into public perception and potential risks. This enhancement allows financial institutions to proactively identify emerging trends and sentiment shifts, enabling them to take appropriate measures to mitigate potential risks and optimize their risk-return profile.

Still, with advances being published on a weekly basis, there is no single best modeling technique. The approach of choice should be influenced to a large extent by the purpose that the model is supposed to serve and how it will be used considering current regulatory minimum requirements and leading practice. Furthermore, there is an unavoidable variability in sentiment scoring due to its subjective nature. It is known that some sentiments can be subtle, context-dependent, or rely on understanding metaphors or sarcasm. This can be challenging to humans as well as automated systems, including modern Neural Network algorithms like transformer models. Moreover, sentiment analysis models can be sensitive to biases present in the training data, potentially leading to biased predictions. To address this concern, multiple score aggregation strategies must be considered, enabling consistency checks, and providing an estimate of the inherent systematic uncertainty arising from model selection.

With regards to the model inputs, first efforts in initiating any largescale and future-proof scoring system must be attributed to an efficient data storage for facilitating data access and analysis and complying with the BCBS239 principles. The aim should be a streamlined data processing and visualization for effective information communication. If setup is made with foresight, then extracting important information and predicting user content and sentiment requires low maintenance and can itself be automated. Data preprocessing pipelines and aggregation procedures should incorporate modular architecture so that any additions to the procedure are “plug-n-play” while reversal to older versions is equally simple. This approach would also support auditing efforts as change log creating is effortless.

It is beneficial to differentiate OSINT sources regarding their type and thus the **kind of influence exerted on their readership**. In general, data from reputable news sources is in many ways more uniform and easier to interpret. There is a strong incentive for clear communication and double meaning is almost nonexistent. On the other hand, blogs, forums, or chatrooms offer a more direct way of measuring the "pulse of the people" but suffer a tradeoff in quality as they often serve as echo chambers to the extreme point of view, contain flight of fancy opinions, or off-topic conversation that can confuse the algorithms, or in the most recent cases, contain bot created non-opinions used more to influence or confuse the readership.

These are of course not the only sources of opinion that can be scanned and considered. For example, valuable insights can be obtained from internal chat rooms or customer complaints contact points. Such sources do not drive social media sentiment, as there is no feedback loop, but do measure a segment of the population, usually the more extremely opinionated one. Due to such issues, it is essential to develop prudent data pre-processing and filtering steps to minimize the effects of vacuous or disingenuous text entries.

Finally, as is the case in this POC, assigning a number that informs on the amount of positivity or negativity that is expressed in a sentence is at this moment not possible. Two sentences can both be positive but still exude very different effects on the readership ("OK" vs. "the best"). With that observation it is rather easy to conclude that any approach that needs absolute values, which hold a generally interpretable quality, is challenging.

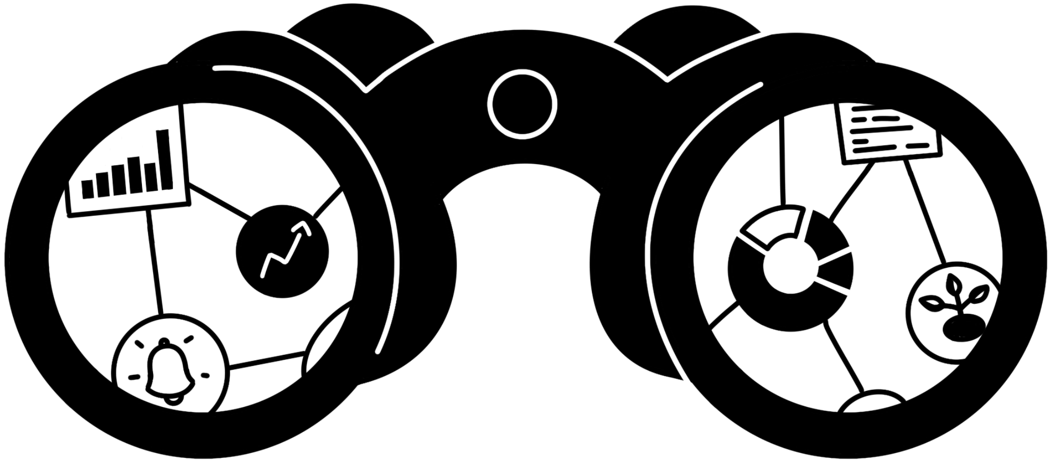
To this end, in this POC demonstrated that it is not **important to measure absolute sentiment**, but to in a uniform way **observe a change**. Hard triggeres might be possible to set, but would still require expert setting as a starting point and overtime, a What-Is and What-If based approach including automated calibration.

WHY THE SUPPLY CHAIN MATTERS

4.1 THE WATCH LIST

Within the domain of credit risk management, financial institutions encounter the task of proficiently overseeing and alleviating risks linked to their credit exposures. The utilization of a Watchlist has surfaced as a valuable mechanism to tackle this challenge. A Watchlist refers to a specific group of financial instruments, such as securities or loans, that are subjected to heightened scrutiny within a broader range of assets. Watchlists are as such an integral component of the Credit Risk Monitoring framework within Credit Portfolio Management. The objective of these efforts is two-pronged, namely, to augment risk management strategies and to potentially recognize favorable prospects. Comprehensive policies and procedures are established by institutions to monitor credit exposures and borrowers who exhibit heightened risk levels. These policies entail the ongoing surveillance of EWI and the subsequent discernment of entities that are to be incorporated into the watchlist. To ensure a proactive approach to risk mitigation, effective watchlist monitoring entails the consideration of various aspects. The process includes the examination of adverse occurrences or patterns that could potentially jeopardize a cohort of debtors, encompassing factors such as economic, demographic, or technological elements. The borrower's credit score or rating is evaluated, enabling more frequent communication and the gathering of supplementary details such as financial forecasts. The monitoring process includes evaluating credit limits and considering potential adjustments or cancellations of underutilized limits. Regular review of specific reports generated from monitoring this watchlist is conducted by key stakeholders, such as the head of the risk management function, heads of credit-granting functions, credit portfolio management and joint management bodies between first line and second line of defense. The adoption of all-encompassing watchlist protocols and meticulous surveillance of entities enlisted on the watchlist can enhance the risk mitigation strategies of financial establishments. The adoption of a proactive approach can be efficacious in addressing credit risks and safeguarding portfolios against possible unfavorable occurrences. This holistic approach to credit risk management also includes considering signals from the supply chain view. Institutions recognize that disruptions and challenges in the supply chain can have a significant impact on the creditworthiness of borrowers.

Economic, demographic, or technological threats that affect a group of borrowers can



originate from disruptions within their supply chains. effective watchlist monitoring involves considering various aspects, including climate-related risks. Physical risks, such as extreme weather events and rising sea levels, as well as transition risks associated with the shift to a low-carbon economy, can have a material impact on borrowers' credit profiles. Therefore, the monitoring of the watchlist also involves analyzing signals related to the supply chain view and climate change in order to cope with second order effects. This includes assessing the stability and resilience of suppliers, the presence of concentration risks, or vulnerabilities to events such as natural disasters or geopolitical shifts.

4.2 PREVENTING MANAGEABLE CREDIT RISK – WHAT IS AND WHAT IF

When considering the supply chain view in credit risk management, it is crucial to assess the individual borrower but also to delve in detail into their broader securities portfolio (collateral).
One aspect of particular importance is screening for second order concentration.
Second order concentration arises when a client possesses a seemingly diversified securities portfolio, but certain holdings within that portfolio are interconnected through the supply chain.

EXHIBIT: 11 - Second Order View of Portfolios

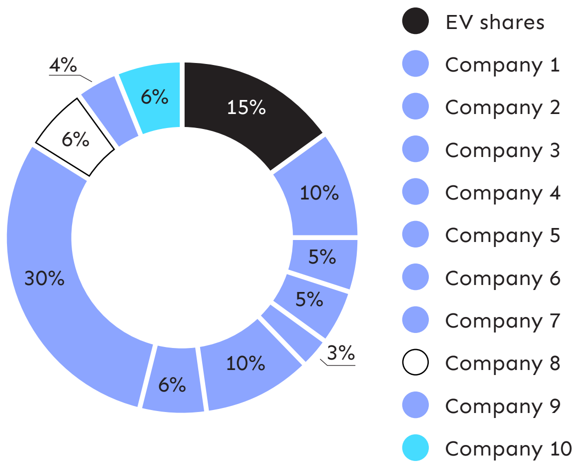


Figure 1: Diversified Client Portfolio - First Order

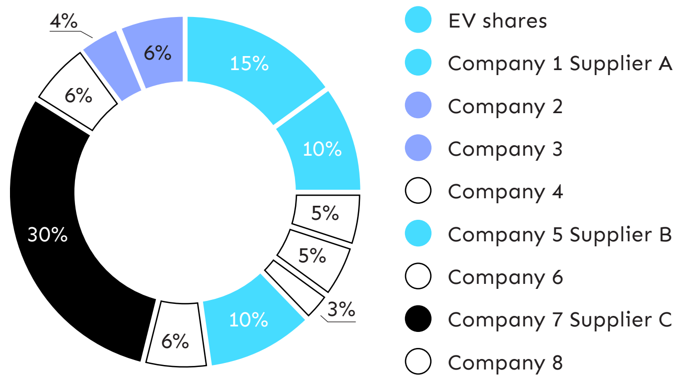
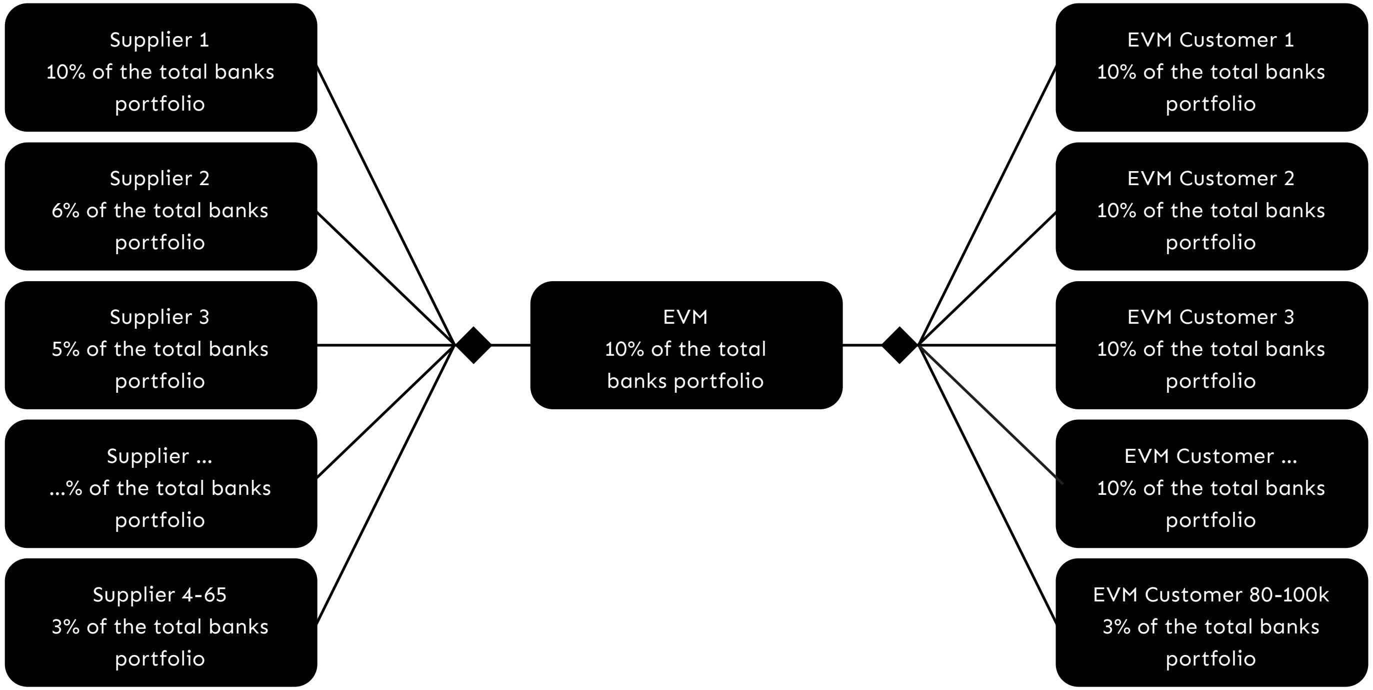


Figure 2: Concentrated Client Portfolio - Second Order

Let's think about an example to demonstrate this idea. Let's say a client's portfolio includes shares of a well-known electric vehicle manufacturer (EVM). Even though this may appear to be a diversified investment and solid collateral for a Lombard facility, it raises questions when we consider that the EVM supply chain is also included in the same portfolio. In this case, any obstacles or disruptions to EVM's supply chain could have a major effect on the client's total portfolio, raising concerns about concentration and risk. The client securities portfolio illustration makes it abundantly evident that, in the first instance, the credit risk officer can draw the conclusion that the portfolio appears to be well-diversified and the risk is in on acceptable level considering the potential assigned Standard Loan-to-Value (LtV) ratios. With over 60% of the shares belonging to the same value chain, a closer examination reveals that in the second order perspective, supply chain concentration is fairly significant. In a perfect world, the LtV calculation engine of the bank would take into account supply chain data when allocating the appropriate LtV to the assets. In spite of the possible lack of supply chain resilience demonstrated by the COVID-19 pandemic, experience with client systems has indicated that this has not yet been explored or is only seen as a nice-to-have feature.

Identifying and addressing second order concentration is critical in credit risk management. By thoroughly analyzing the interconnections within a client's portfolio, financial institutions can better understand the potential ripple effects of disruptions or vulnerabilities in the supply chain. This knowledge allows them to assess the overall risk exposure accurately and implement appropriate risk mitigation measures, such as diversification strategies or alternative sourcing arrangements. Identifying and addressing second order concentration is critical in credit risk management. By thoroughly analyzing the interconnections within a client's portfolio, financial institutions can better understand the potential ripple effects of disruptions or vulnerabilities in the supply chain. This knowledge allows them to assess the overall risk exposure accurately and implement appropriate risk mitigation measures, such as diversification strategies or alternative sourcing arrangements.This entails investigating the interconnectedness of client portfolios, taking into account joint exposures, supplier linkages, or widespread supply chain weaknesses. With the use of such analysis, the institution is able to evaluate the potential knock-on consequences of disruptions and take preemptive steps to reduce systemic risks. Additionally, credit portfolio managers may make the best use of this data to maximize their resource and risk appetite allocation.

EXHIBIT: 12 - A Total Bank Second Order Asset Concentration View



They can decide on risk mitigation techniques including reducing credit limits, diversifying client contacts, or entering into alternative sourcing agreements by determining areas of overexposure or excessive concentration.

When discussing early warning systems and the integration of reverse stress, it is crucial to take into account the regulatory frameworks, namely IFRS 9 and CECL. The mentioned prospective regulations were formulated in reaction to the financial crisis that occurred between 2008 and 2009, with the objective of detecting potential risks in banks' financial records at an earlier stage. The IFRS 9 places emphasis on the concept of Expected Credit Loss whereas the Current Expected Credit Loss methodology is utilized within the framework of the Generally Accepted Accounting Principles in the United States.

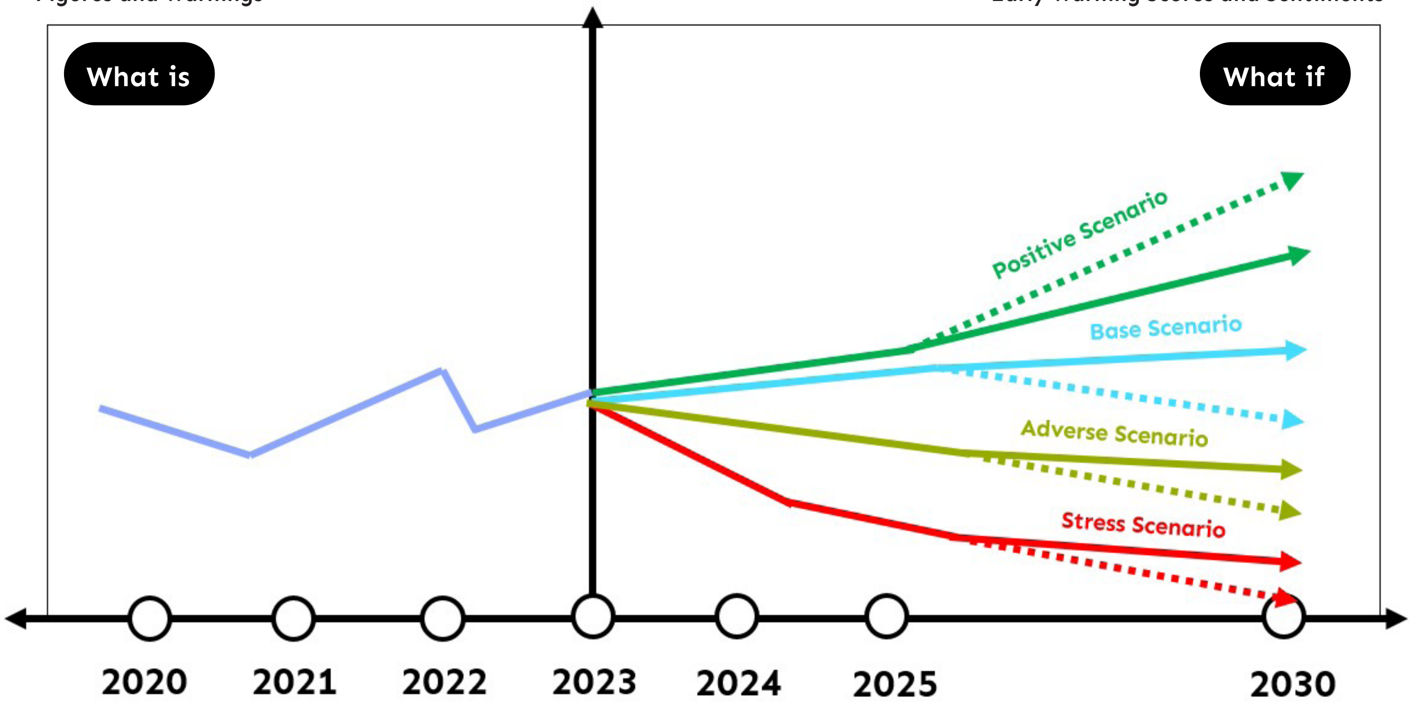
IFRS 9 and CECL necessitate that financial institutions evaluate and integrate prospective data into their credit risk models. This entails the contemplation of multiple scenarios and conducting stress tests on their portfolios to assess the possible ramifications of unfavorable occurrences.

EXHIBIT: 13 - What Is and What If

Actual Performance Indicators
Figures and Warnings

Forward Looking Performance Indicators
Early Warning Scores and Sentiments

- Expected Credit Loss
- Profit, Capital
- Liquidity
- Credit Risk
- Climate Risk
- Supply Chain Risk
- Geopolitical Risk

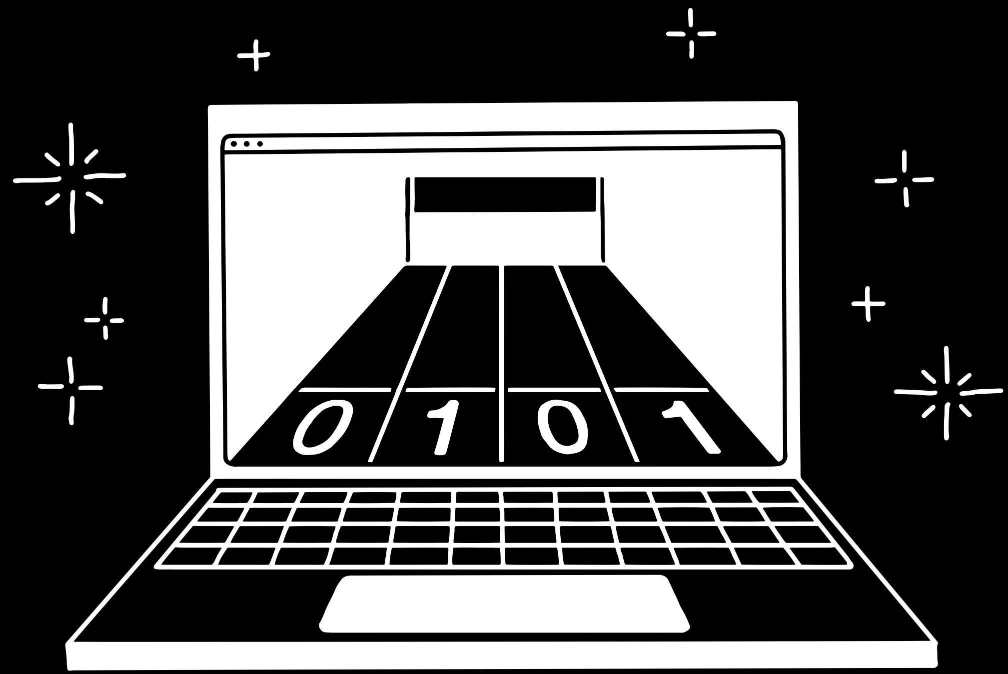


The integration of reverse stress testing enables financial institutions to detect potential weaknesses and evaluate their capacity to withstand adverse circumstances. This enables individuals or organizations to anticipate and tackle potential hazards by implementing suitable measures to minimize their impact. The adoption of these standards has resulted in an increased level of intricacy and computational burden in the process of impairment calculations, necessitating substantial investment of both time and resources. It is imperative that financial institutions possess a comprehensive comprehension and ownership of their credit loss computations. This will allow for enhanced management of their balance sheets and profit and loss statements, as well as a holistic understanding of potential future developments. Recent events, such as the COVID-19 pandemic, have underscored the significance of robust model governance and model risk management procedures. It is key for banks to ensure that their models are statistically sound, relevant, and flexible, particularly when confronted with unforeseen circumstances. This entails the evaluation of the effects of climate change and the incorporation of climate-related risk elements into their models and stress testing endeavors.

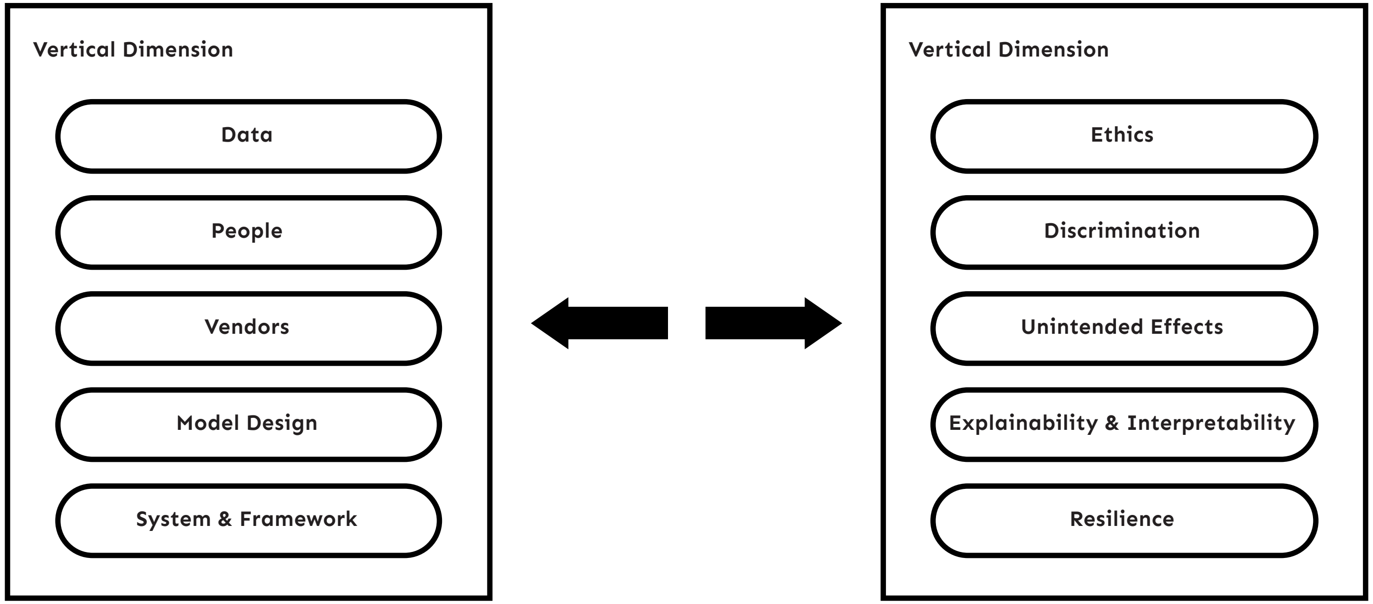
The regulatory focus on climate change underscores the necessity for prospective analysis and evaluations based on scenarios to identify possible risks.

Early warning systems, particularly in the credit value chain, require the necessary attention, complexity management, and continuous enhancement to ensure competitiveness and maintain a comprehensive view of risks. By adopting a holistic and forward-looking approach while including “What Is” with “What If” and the technology, methodology and people-skill overhaul, financial institutions can better anticipate and respond to emerging challenges in the ever-evolving financial landscape and maintain competitive advantage.

APPENDIX - IMPORTANT DIMENSIONS FOR RESPONSIBLE AI



In the construction of AI models, it is vital to prioritize responsible AI practices to mitigate risks effectively. Responsible AI encompasses various dimensions that are critical for model development and deployment. Firstly, ethics play a crucial role in ensuring that AI systems operate within ethical boundaries, respecting human values, privacy, and societal norms. This involves considering the potential ethical implications of AI algorithms and ensuring transparency and accountability in decision-making processes



Discrimination is another dimension that needs careful attention. AI models should be designed and trained to avoid bias and discrimination, ensuring fair treatment and equal opportunities for all individuals, regardless of their characteristics or background. Unintended effects are also significant concerns. AI systems should be thoroughly assessed to identify and address any unintended consequences or negative impacts that may arise from their use. This involves anticipating and mitigating potential risks associated with the deployment of AI models. Explainability and interpretability are essential for building trust and understanding in AI systems. Models should be designed to provide clear explanations and justifications for their decisions, enabling stakeholders to comprehend how and why specific outcomes are reached. This fosters transparency and accountability in AI applications. Lastly, resilience is a crucial aspect of responsible AI. AI models should be designed to be robust and resilient, capable of adapting to changing conditions and withstanding potential attacks or vulnerabilities. This involves considering potential threats, ensuring system integrity, and implementing measures to enhance model resilience.

Considering these dimensions – ethics, discrimination, unintended effects, explainability & interpretability, and resilience – is vital throughout the model construction process. By adhering to responsible AI practices, organizations can develop AI models that are trustworthy, fair, and aligned with societal values, thereby mitigating risks and promoting positive outcomes in the AI ecosystem.

Ethics

Data
- Ethical sourcing and use of data
- Protecting user privacy and consent
- Ensuring data quality and representativeness

Model Design
- Ensuring ethical design and use of AI models
- Preventing malicious uses of sentiment analysis
- Avoiding amplification of harmful content or sentiment

People
- Promoting ethical use of sentiment analysis by individuals
- Educating users on potential biases and limitations of sentiment scores
- Encouraging responsible sharing and interpretation of sentiment analysis results

System & Framework
- Ethical deployment and use of sentiment analysis systems
- Protecting against algorithmic manipulation or abuse
- Ensuring user consent and control over data usage

Vendors
- Ethical development and deployment of sentiment analysis tools
- Responsible sharing of user data with vendors
- Evaluating vendor practices for bias mitigation and ethical considerations

Discrimination

Data
- Identifying and addressing biases in training data
- Ensuring diverse and inclusive data representation
- Mitigating the risk of reinforcing stereotypes or discrimination

Model Design
- Bias mitigation in model training and validation
- Fair representation of different user groups
- Addressing algorithmic discrimination or prejudice

People
- Addressing user concerns and feedback regarding biased or unfair sentiment analysis results
- Ensuring user trust in sentiment analysis systems
- Fostering user empowerment and control over system outputs

System & Framework
- Ensuring equitable access to sentiment analysis tools
- Avoiding discriminatory or biased system behavior
- Addressing potential harms or unintended consequences

Vendors
- Vendor accountability for biased or unfair sentiment analysis outcomes
- Ensuring diversity and inclusion in vendor teams and practices
- Addressing potential vendor lock-in or monopolistic behavior

Unintended Effects

Data

- Unforeseen and unintended consequences of data usage
- Potential for malicious data manipulation
- Misinterpretation of user-generated content

Model Design

- Uncertain and unintended model behavior
- Potential for model-generated content to spread misinformation or harmful sentiment

People

- Potential for misuse or misinterpretation of sentiment analysis results by individuals
- Emotional impact on individuals from system-generated sentiment scores

System & Framework

- System errors or malfunctions leading to incorrect sentiment scores
- Potential for system vulnerabilities or exploitation

Vendors

- Reliance on vendor expertise and proprietary algorithms for sentiment analysis
- Potential vendor limitations or biases in sentiment analysis models

Explainability & Interpretability

Data
- Understanding data selection criteria and weighting
- Transparency in data collection methods and pre-processing steps
- Mitigating the risk of data manipulation or tampering

Model Design
- Interpretability and Transparency of model decisions
- Ability to explain the factors contributing to sentiment scores
- Identifying and addressing model limitations or biases

People
- Promoting transparency in sentiment analysis processes and limitations
- Encouraging critical thinking and contextual interpretation of sentiment scores
- Supporting user feedback and accountability mechanisms

System & Framework
- Ensuring transparency in system design and functionality
- Auditing and accountability mechanisms for system behavior
- Addressing user concerns and feedback regarding system performance

Vendors
- Transparency in vendor practices and model training methodologies
- Vendor cooperation in addressing biases, limitations, or concerns
- Supporting user choice and interoperability between sentiment analysis vendors

Resilience

Data
- Monitoring data sources for relevance and accuracy
- Adapting to evolving user behaviors and communication patterns
- Ensuring data currency and relevance over time

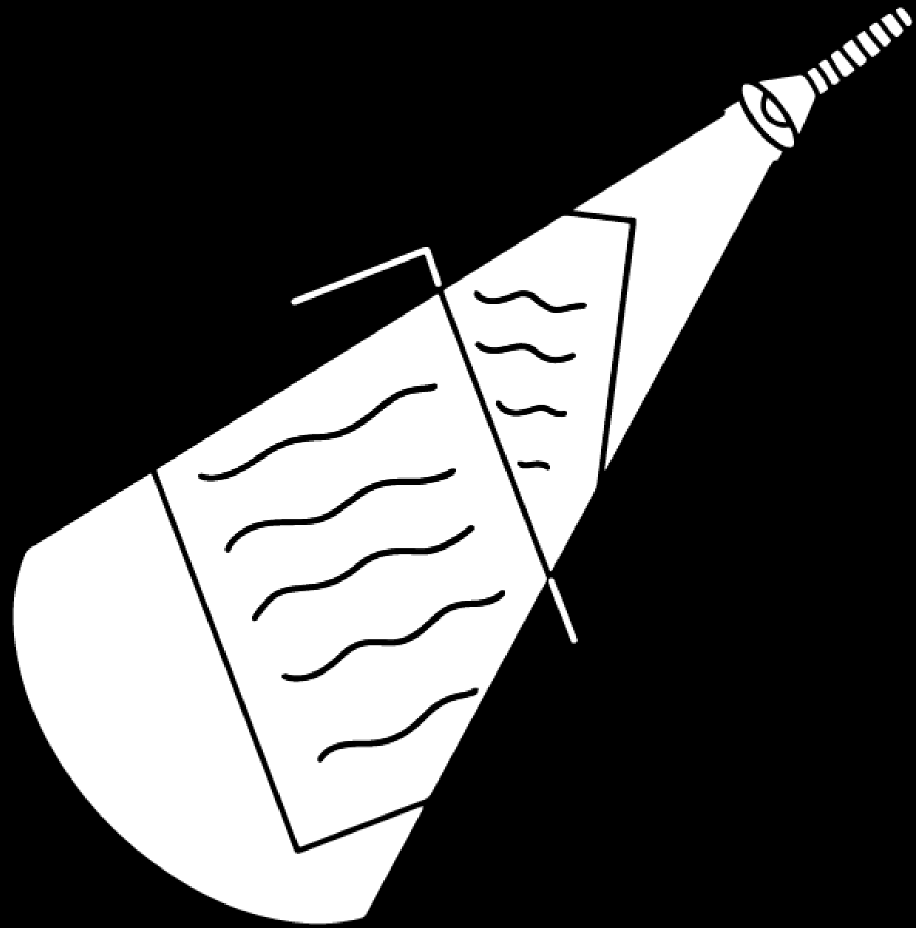
Model Design
- Continuous model monitoring and performance evaluation
- Regular model updates and retraining to adapt to evolving language and user behavior
- Mitigating model degradation or drift over time

People
- Building user awareness and understanding of sentiment analysis technology
- Adapting to changing user expectations and needs
- Empowering users to provide input on system improvements

System & Framework
- Continuous system monitoring and risk assessment
- Regular system updates and patches to address vulnerabilities
- Adapting to changing regulations and standards

Vendors
- Regular vendor evaluations and risk assessments
- Diversifying vendor partnerships to avoid over-reliance
- Adapting to changes in vendor offerings or market dynamics

END NOTES



BIBLIOGRAPHY

Cambria, E., Das, D., Bandyopadhyay, S. & Feraco, A., 2017. A Practical Guide to Sentiment Analysis. Berlin: Springer.

EBA, 2020. Guidelines on Loan Origination and Monitoring, s.l.: s.n.

EBA-GL-2017-16, 2018. Guidelines on PD and LGD estimation, s.l.: s.n.

ECB, 2018. ECB guide to internal models, s.l.: s.n.

J. H. Lang, T. P. P. S., 2018. A framework for early-warning modeling with an application to banks. ECB Working Paper Series No. 2182.

Jan Hannes Lang, T. A. P. P. S., 2018. A framework for early-warning modeling with an application to banks, s.l.: ECB.

Jarmulska, B., 2020. ECB. [Online]
Available at: https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp2408~aa6b05aed7.en.pdf?9551c7c6e8e8fdbd35e-5512b5afcf097

K. Bluwstein, M. B. A. J. S. K. Ö. Ş., 2021. Credit growth, the yield curve and financial crisis prediction: evidence from a machine learning approach. ECB Working Paper Series No 2614.

K. Bluwstein, P. B. E. C. M. L., 2021. Asset encumbrance in euro area banks: analysing trends, drivers and prediction properties for individual bank crises.. ECB Working Paper Series No 261.

Keenan, M. J. S., 2020. Advanced Positioning, Flow, and Sentiment Analysis in Commodity Markets: Bridging Fundamental and Technical Analysis, 2nd Edition. s.l.:Wiley.

Loki, E., 2019. Managing preventable risk. The opportunity for machine learning in credit risk early warning system, s.l.: Professional Risk Managers' International Association: In Intelligent Risk Report.

Loki, E. K. T. B. B., 2022. The Future Risk Imperative.. Exxeta Working Paper on Risk and Sustainability.

My T. Thai, W. W. H. X., 2017. Big Data in Complex and Social Networks. s.l.:Chapman & Hall.

Powell, J., 2021. FT - GameStop can't stop going up. [Online] Available at: https://www.ft.com/content/7aa60aa1-484f-4747-9136-cd0a560dd2d8

Sohangir, S. W. D. P. A. e. a., 2018. Big Data: Deep Learning for financial sentiment analysis. J Big Data 5.

Theo Notteboom, T. P. J.-P. R., 2021. Disruptions and resilience in global container shipping and ports: the COVID-19 pandemic versus the 2008-2009 financial crisis. In: Mariticm Exonomics & Logistics.

ABOUT THE AUTHORS

Elias Loki

Elias Loki is the practice's leader in risk management and sustainability in Switzerland.

He is a specialist in risk data management, credit risk management, and sustainability. For many years, he has provided systemically important banks with advice and guidance in implementing improvements to risk resiliency.

His experience in mechanical engineering, business economics, and an MBA has given him a comprehensive foundation from which to approach a variety of problems in the field of risk digitalization.

Goran Simatović

Data scientist in Risk Management Dr. Goran Simatović has spent more than ten years analysing and modelling statistical data.

His primary research focus is on creating data-driven software solutions for real-world credit risk management issues.

Goran has a broad background from which to approach numerous topics in the field of Financial Risk Management and Data Science thanks to his prior employment at ABN Amro and CERN Switzerland.

Sarah Salem

With more than 10 years of experience as a skilled financial analyst and consultant, Sarah Salem is a seasoned expert in the field of wealth and asset management.

She has had the privilege of working with major financial institutions and leading asset management change throughout her career. One of the most significant facets of Sarah's job is renewing the front organization, which she excels at. She deftly strikes a balance between the need to revitalize front-end operations and the need to keep an all-encompassing perspective that includes the complete front-to-back structure.

She is an invaluable addition to the industry because of her ability to handle these complexities.

Exxeta AG
Albert-Nestler-Straße 19
76131 Karlsruhe
Tel. +49 721 50994-5000
Fax. +49 721 50994-5299
E-Mail info@exxeta.com

HRB 702566
Amtsgericht Mannheim
USt-IdNr. DE813026499

Exxeta Schweiz
Talacker 41
8001 Zürich
Tel. +41 79 102 99 12

Vertretungsberechtigter Vorstand
Dr. Peter Heine
Erwin Kiefer
Achim Kirchgässner
Andreas Ritter

Aufsichtsratsvorsitzender
Ulrich A. Götz

exxeta.com