

PRIME-RE: A Parallel and Recursive Requirements Elicitation Approach for ML-Based Financial Decision Systems

Cosmin COJOCARU*

National University of Science and Technology Politehnica Bucharest, Bucharest, Romania *Corresponding author, cosmin.cojocaru1906@stud.faima.upb.ro

Sorin IONESCU

National University of Science and Technology Politehnica Bucharest, Bucharest, Romania
sc.ionescu@gmail.com

Abstract. *With the growing use of machine learning models for supporting financial decisions, financial institutions face increasing pressure to provide explanations, traceability and stable risk assessments of their decision-making processes. Current approaches to requirements gathering do not take into account the realities of this domain since they are based on the assumption that decision requirements can be formulated at an early stage and then be refined incrementally. In reality, however, financial decision objectives; data availability; analytical feasibility; and process governance constraints evolve simultaneously and interactively and are often discovered only by way of exploratory analyses or preliminary validation activities. This paper introduces PRIME-RE, a parallel and recursive methodology for eliciting requirements for machine-learning based financial decision support systems. PRIME-RE structures the requirements gathering and elicitation process in four concurrent streams: decision objectives; data feasibility; analytical feasibility; and process-governance constraints, including regulatory constraints. When conflicts arise between streams, the methodology provides for controlled recursive refinement of those streams to resolve the conflicts. The results of the methodology have been validated using two sources: an expert survey with 25 practitioners and prior empirical research on how to structure unstructured data for evaluating risk. In summary, a conceptual architecture is introduced illustrating how non-functional requirements (e.g. stability and interpretability) influence system design and how unstructured information can be integrated as contextual signals for the purpose of informing risk evaluations while preserving core risk assessment capabilities. The results show that parallel and recursive requirements elicitation enables the development of credible, transparent and compliant ML systems.*

Keywords: requirements engineering, financial decision systems, machine learning governance, regulatory compliance, machine learning explainability, unstructured data processing.

Introduction

The last decade has seen a great increase in the use of machine learning (ML) algorithms in the financial sector. This growth is mainly attributable to the faster and more accurate risk assessments enabled by the new algorithms. Financial institutions use ML models to evaluate credit applicants, assess the health of their portfolios, identify potential problem areas, and inform key business decisions. However, there are very specific supervisory guidelines that govern the use of machine learning in finance. These rules emphasize that ML systems need to provide clear explanations for their actions, be auditable, stable, and accountable. For example, the Federal Reserve's SR 11-7 emphasizes that ML system developers should perform model risk management through independent review of their ML models, discipline in developing ML models, validating those models, and continually monitoring them (Board of Governors of the Federal Reserve System, 2011). According to the Basel Committee on Banking Supervision (BCBS), "effective risk data

aggregation and interconnection requires good data management practices, accurate and comprehensive aggregation of risk data, and timely and flexible provision of reports on risk exposure." Moreover, as provided by the EU's Artificial Intelligence Act, "for certain high-risk AI systems, organizations may be required to provide some level of transparency and accountability. For instance, an organization may be required to provide notice that it is using an AI system, to explain the logic used by an AI system to make decisions, and to ensure that there is at least one person or group that can be held accountable for high-risk AI systems." Thus, it is not enough that financial AI is predictive and accurate; it must also be explainable and accountable.

As a result of these challenges, the most typical software engineering solution for handling these issues is to establish a process that involves identifying the requirements for the system at the beginning of the project, defining the functional and non-functional requirements for the system, and refining them step-by-step throughout the development of the system. However, ML-based systems do not fit into this paradigm. Decision objectives may change due to the introduction of new types of data, the presence of data quality issues or sampling bias can invalidate the original assumptions about the input data, and the feasibility of different analytical approaches may not be determinable before they are used. Therefore, requirements such as interpretability and stability typically arise after experimentation with prototypes or early versions of a system. Thus, the process of establishing requirements for ML-based financial decision-making systems is both parallel, with multiple lines of inquiry occurring simultaneously, and recursive, with frequent returns to and revisions of the established requirements.

This paper addresses the gap between traditional approaches to Requirements Engineering (RE) and real-world practices related to Machine Learning (ML) in the financial sector. As an extension of the author's previous doctoral work, this paper describes a methodology called PRIME-RE that provides a structured approach to performing Requirements Elicitation otherwise known as eliciting the needs for a financial decision-making process that uses machine learning in a way that is both recursive and parallel. In doing so, PRIME-RE provides a common language framework to enable practitioners to structure the decision objectives for the ML process, assess the feasibility of the available data for use in the ML process, determine the feasibility of various analytical processes, and define the regulatory or compliance-related constraints that will impact the ability to leverage the predictive abilities of ML. This paper expands upon the author's prior research in integrating unstructured data into ML-based financial decision-making, and it demonstrates how contextual information can be used to enhance risk assessments without compromising the integrity of the underlying evaluation layer. Furthermore, this paper includes expert validation and empirically-supported prior research as the foundation for the proposed methodology.

Literature review

Requirements engineering allows us to develop information systems that match the needs of organizations. The traditional definition of requirements elicitation states it as "the process of identifying, evaluating, and recording the expectations and restrictions of all stakeholders related to the project that will be developed." (Kotonya & Sommerville, 1998; Sommerville & Sawyer, 1997) For decision-oriented systems, the requirements go beyond only the functional capabilities and include the accountability, traceability, and justification of decisions taken (Power, 2002). Previous studies indicate that the lack of sufficient detail in the decision requirements leads to systems which may be technologically correct, but wrong in relation to the way of doing things within the organization. The classical methods of interviewing, workshops, and questionnaires

continue to play a central role in this area, although they suffer from a number of limitations due to cognitive biases, the incomplete expression of tacit knowledge and the fact that it is difficult to specify how a system should behave under uncertain conditions (Pohl, 2010). These limitations become even greater where systems use machine learning models.

The process-based approaches can address many of these problems by defining the requirements based on the organization's business processes and control structures. The use of business process modeling to identify the decision points, information flows, and compliance constraints embedded in the organization (Rolland & Prakash, 2001; Dumas et al., 2018) is one example. A goal-oriented requirements engineering approach offers another perspective, relating the requirements of the system to the organization's objectives and constraints and evaluating the trade-offs between conflicting goals (van Lamsweerde, 2001). Although both perspectives have been used in decision support contexts, they typically take place at an abstract level. The translation of high-level goals and process constraints into specific requirements is difficult, particularly when the system's behavior depends on the characteristics of the data and the performance of the model (Horkoff et al., 2016).

Machine-learning lifecycle methodologies, including CRISP-DM and KDD, offer a structured approach to data analysis, and the methodologies define phases such as business understanding, data preparation, modeling and evaluation (Shearer, 2000; Fayyad et al., 1996). The methodologies acknowledge that there are iterations in the process but consider requirements implicitly, and often equate requirements with analytical objectives. Financial methodology extensions have introduced validation and monitoring activities to support the methodologies; however, they do not directly address the elicitation of requirements. In the domain of regulated environments, the lifecycle methodologies are not adequate because they do not take into account explicit governance or regulatory requirements.

These limitations in the requirements engineering literature are further compounded by the regulatory environment in which financial decision systems operate. There are numerous regulatory and supervisory requirements applicable to financial decision systems. The Federal Reserve's SR 11-7 guidance defines model risk as "the potential for adverse consequences to occur as a direct result of model-based decisions that are made using incorrect or misused models." The guidance emphasizes that the model risk increases with the complexity and uncertainty of the model. Model risk is reduced through the disciplines of model development, validation, monitoring, and governance at the board level (Board of Governors of the Federal Reserve System, 2011). The Basel Committee's BCBS 239 principles on risk data aggregation and reporting requires the establishment of strong governance, data architecture, accuracy, completeness, timeliness, and flexibility of risk reports. The BCBS 239 principles also include supervisory review and remedial action (Basel Committee on Banking Supervision, 2013), and establish a foundation for data quality, reporting, and auditability in the context of risk management. The EU's AI Act establishes requirements for the transparency of AI systems, specifically those identified as AI systems that are classified as high risk. The requirements include informing users about the use of AI algorithms in the system and providing explanations in an intelligible form. The focus is on interpretability, accountability, traceability and human oversight. Failure to comply with the requirements set forth in the EU's AI Act will incur severe penalties. SR 11-7, BCBS 239, and the EU's AI Act collectively demonstrate that governance requirements are mandatory. However, none of these regulatory frameworks directly prescribe how requirements elicitation should be organized to ensure compliance of systems development.

Data-driven methodologies for gathering system requirements would refine the way requirements are defined as data properties and model behaviors. Vogelsang et al. (2019) found empirical evidence that during the experimentation phase, many feasibility constraints, biases, and limitations in interpretability may become apparent. Thus, the development of system requirements for machine learning systems will need to be viewed as an evolving process of coordination between multiple investigation streams. As a result of the trade-off between predictive ambition and acceptable governance, there has been a call for integrated methodologies that support the parallel exploration and recursive refinement of predictive ambitions and governance acceptability. Recent research has also demonstrated both the potential and the challenges associated with the incorporation of unstructured data into the risk assessment process. Unstructured data can include sentiment analysis, reputation information, and event-based information from news articles, blogs, and social media. Research has indicated the positive impact of incorporating unstructured data into the risk prediction process including the opportunity to identify early warning signs (Cerchiello et al., 2017; Loughran & McDonald, 2011). Due to its heterogeneity, noise, and dynamism, however, unstructured data creates problems regarding the stability, interpretability, and auditability of the predictions. Prior research by the author (Cojocaru & Ionescu, 2025) demonstrates that textual data can be structured in a manner that maintains its relationship to the original source while maintaining traceability. This provides an opportunity to expand the scope of data-driven methodologies for eliciting requirements beyond structured data so that additional types of information can be utilized when defining functional and non-functional requirements such as control, interpretability and decision integration. Nevertheless, the literature does not offer much direction for how to integrate unstructured signals into core risk models without destabilizing them.

There is still a need for integrative methodologies for eliciting requirements which can address the multiple constraints of decision-making (goals), feasibility of data analysis (analytical feasibility), the feasibility of using data (data feasibility) and legal constraints while supporting controlled revisions of decisions. The PRIME-RE methodology addresses this need by modeling the coordination of requirements elicitation as an iterative and recursive process of coordinating decision making goals with the feasibility of using data and performing data analysis and legal constraints on ML-based financial decision systems. The methodology models the elicitation of requirements as a coordination problem and uses coordinated, structured and iterative processes to support the controlled revision of decisions made with financial decision systems that use machine learning.

Methodology

This paper employs Design Science Research (DSR) to create and evaluate PRIME-RE approach to elicit the requirements for machine learning enabled financial decision systems, with emphasis on regulatory constraints, with emerging feasibility and governance requirements.

The research question driving this work is: *How should Requirements Elicitation be organized to create ML-based financial decision systems operating under regulatory constraints, when feasibility and governance requirements emerge through exploration?*

The objective is not to maximize model performance, but rather to understand how to coordinate the emergence of requirements.

DSR is a problem-solving process that includes analysis of the problem, creation of artifacts and reflection on those artifacts. In this case, the artifact is the PRIME-RE methodology, described as a coordination mechanism for running parallel investigations and for initiating recursive

refinement cycles. The requirements elicitation process of PRIME-RE has four streams that run in parallel: decision objectives and context, which deals with the intent and desired outcomes of the system being created; data feasibility, which deals with the availability and quality of both structured and unstructured data; analytical feasibility, which deals with the possible models available to analyze the data, the interpretability of the models, and the stability of the models; and process and governance constraints, which deals with ensuring compliance with regulations. Each of the four streams continuously interact and inform a single backlog of requirements. When there is a conflict, or when there are feasibility issues, or when there are regulatory concerns, a controlled recursive process is initiated to revisit the requirements.

In comparison to traditional software lifecycle methodologies, PRIME-RE does not specify particular analytical techniques or architectures. Rather, it focuses on coordinating decisions and decision-making at the requirements level. By making parallelism explicit, it prevents locking down requirements based solely on first hypothesis assumptions. By providing recursion triggers, it allows findings from exploratory data analysis or regulatory review to feedback into requirement definition. As such, the methodology supports the principle that in high-stakes domains requirements are provisional until proven stable through rigorous challenge, validation and governance review.

PRIME-RE integrates regulatory guidance as described in SR 11-7, through the inclusion of regulations from BCBS 239, and the EU AI Act, with respect to the elicitation streams. Decision objectives stream: clearly defining model purpose & limitations, as required by SR 11-7, directly correlates with the definition of the decision objectives. Data & analytical feasibility streams: a thorough data quality assessment and discipline in the development process, as required by SR 11-7, directly correlates with the data and analytical feasibility streams. Process & governance stream: transparency and timeliness in reporting risks as called for in BCBS 239 is included in the process & governance stream and will impact the auditability and traceability requirements. Interpretability requirements & contextual explanation layer: the EU AI Act's requirements for transparency and human oversight in high-risk systems are included in the interpretability requirements and contextual explanation layer. Compliance therefore is not merely an afterthought, but rather an inherent part of the design process in the PRIME-RE approach.

Expert elicitation was carried out to validate the requirements, prioritize them, and further enhance the methodology. An expert questionnaire was performed during the author's Ph.D. research as part of the methodological evaluation. The questionnaire was distributed online to 25 experts selected intentionally to represent key roles in credit risk, ML engineering, and managerial decision-making. The questionnaire had both closed-ended questions (choice among single-choice, multiple-choice, Likert-scale), and some open-ended questions to identify potential emerging issues. Experts provided their opinion about what they consider to be relevant in terms of risk appetite, and whether they would like to use a risk evaluation system for different aspects of loan underwriting, credit monitoring, and managerial decision-making. They could also rate the importance of different data sources, non-functional properties, together with the use of unstructured information.

Experts participating in the study had an average of over ten years of experience, which is important since the requirements collected will represent the real-world situation. The results showed that the system utility is in all aspects of loan underwriting, credit monitoring, and managerial decision-making. Structured and verifiable information such as legal public records and financial statements received the highest utility rating. The need for good quality data was emphasized. The volatility and the contextual relevance of unstructured online information were

identified as major contributors to the evaluation. Non-functional requirements such as interpretability, stability, and traceability were rated higher than the correct functioning of the algorithm. All respondents agreed on the distinction between the main evaluation and the signal, which was used to develop the layered approach.

Table 1. Summary of the expert ratings for different categories of public information sources used in risk assessment

Information Category	Examples	Mean Rating (1–5)	Interpretation
Legal records	Court records, insolvency announcements	4.68	Highest utility; core to risk evaluation
Financial statements	Balance sheets, income statements	4.52	Essential structured data
Data quality	Timeliness, completeness, reliability	4.48	Critical for credibility
Firmographics	Company age, size, sector	4.00	Contextual but not determinative
Macroeconomic indicators	GDP growth, interest rates	4.00	Contextual calibration
Online information	News, social media, web pages	3.08	Volatile; used for alerts and context

Source: Author's survey of 25 experts. Higher scores indicate greater perceived utility.

Overall, the methodology integrates theoretical synthesis, regulatory mapping, expert elicitation and application in a real financial decision environment. The recursive nature of PRIME-RE reflects the fact that requirements for ML systems are discovered, negotiated and stabilized through investigation. The following section will discuss the results of using PRIME-RE, and explain the conceptual and architectural implications.

Results and discussions

In the context of the risk evaluation, the PRIME-RE process identified the need for a fixed, simple-to-understand core risk score. Additionally, the approach identified the need for a decision-support tool to enable executives to understand why their firm is a high/low-risk entity, and to question the answers of the decision-support tool as new evidence comes to light. The data feasibility stream identified the need for quality structured data; and the unstructured data required for the main evaluation task was viewed somewhat negatively. The analytical feasibility stream identified the trade-off between prediction capability and interpretability of the output as being important; and the process governance stream identified the need for auditability and documentation of the system's purpose and the ability to update the system's models. These streams interacted continually and new requirements emerged as conflict and/or feasibility limitations appeared.

Figure 1 illustrates the PRIME-RE elicitation process. The diagram illustrates the four investigation streams contributing to the common requirements backlog and causing recursive refinement cycles as conflicts or feasibility limits arise. The diagram also illustrates the role of regulatory mapping as a constraint for each stream.

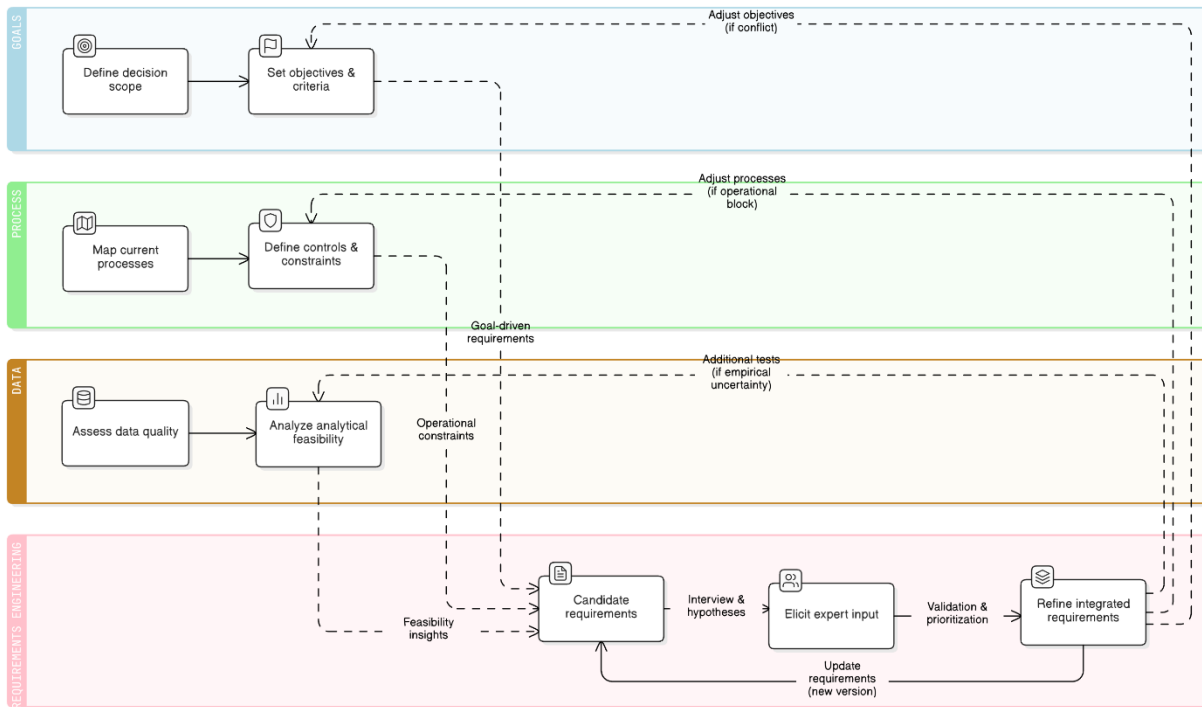


Figure 1. PRIME-RE parallel and recursive requirements elicitation process

Source: Author's own research.

As the elicitation streams continued and the unstructured data was being included into the risk assessment model, a number of issues and the subsequent recursion of the refinement process took place. The primary reason that triggered the recursive refinement cycle was related to the treatment of unstructured data. While it was possible for the sentiment of the news and social media to be used as an early warning system for the risk score, experts indicated that the inclusion of the unstructured data would negatively affect the stability of the evaluation. To address this issue, the decision objectives were re-evaluated, and the unstructured data was reclassified from an analytic source to contextual information. Therefore, while the contextual information would assist in providing early warnings with respect to trending analysis, it would not influence the core risk rating. As a result of the recursive refinement cycle, the functional requirements of "include all available data" was modified to "include structured data within the core risk score, and include unstructured data as contextual and alert information." The same modification was done with respect to update frequency, model complexity, and documentation. Requirements were rewritten based upon the regulatory guidelines and expert advice.

The recursive refinement cycle also illustrated the significance of non-functional requirements. Non-functional requirements identify how a system behaves and under what conditions the results of the system may be accepted. The respondents viewed interpretability and stability as being much more significant than small increases in accuracy. The respondents also identified the importance of traceability which permits the tracking of each input and transformation made within the system. The importance of non-functional properties is consistent with regulatory guidance. Regulatory guidance does not view non-functional properties as constraints, but rather as decision-making factors.

The stabilized requirements used as a basis for developing a conceptual layered architecture for the risk evaluation system. As shown in figure 2, there are three layers in the architecture; the first is an explanatory structural evaluation layer, which is a layer that calculates the risk score using structured, verifiable data and models; the second is a contextual layer that uses signals from recent unstructured data sources; and finally, the third is an explanation and governance layer that spans all layers and enables the provision of explanations for interpretations, documentation of the process and auditability of decisions made within the system. In addition, a presentation layer is provided to allow the results to be presented to end-users in formats that meet their specific needs.

The separation of these layers addresses the conflicting priorities that were identified during the elicitation process. Separating the core risk score from the other elements of the system will provide the system with both stability and comparability. Routing unstructured signals to a contextual layer will provide the ability to adapt to changes in the environment in a timely manner without impacting the primary assessment. Finally, implementing the governance layer will facilitate regulatory compliance through the tracking of decisions, documentation of the limitations of the models used in the system and independent validation of the system. The presentation layer will satisfy the expert need for different displays for executive and analytical users. Therefore, this architecture demonstrates how non-functional requirements can affect the design of systems and also demonstrates the application of PRIME-RE in practice.

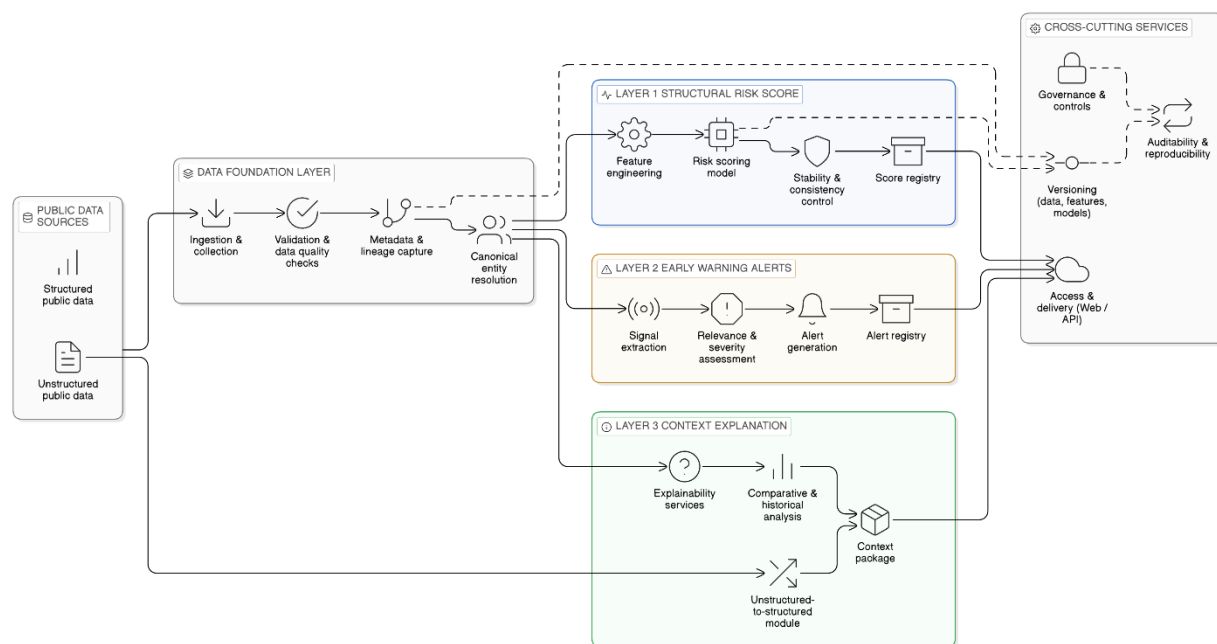


Figure 2. Conceptual architecture for machine learning based financial decision systems

Source: Author's own research.

Figure 2 also illustrates the expert opinion that a risk evaluation system must provide a complete set of related outputs. The respondents stated that the score should be supported by explanations, current signals and contextual information to facilitate a comprehensive decision. Therefore, the architecture includes explicit modules for explanation and alerting.

To operationalize the contextual layer, prior work developed an unstructured document processing pipeline that extracts ad-hoc metadata from textual sources while preserving traceability.

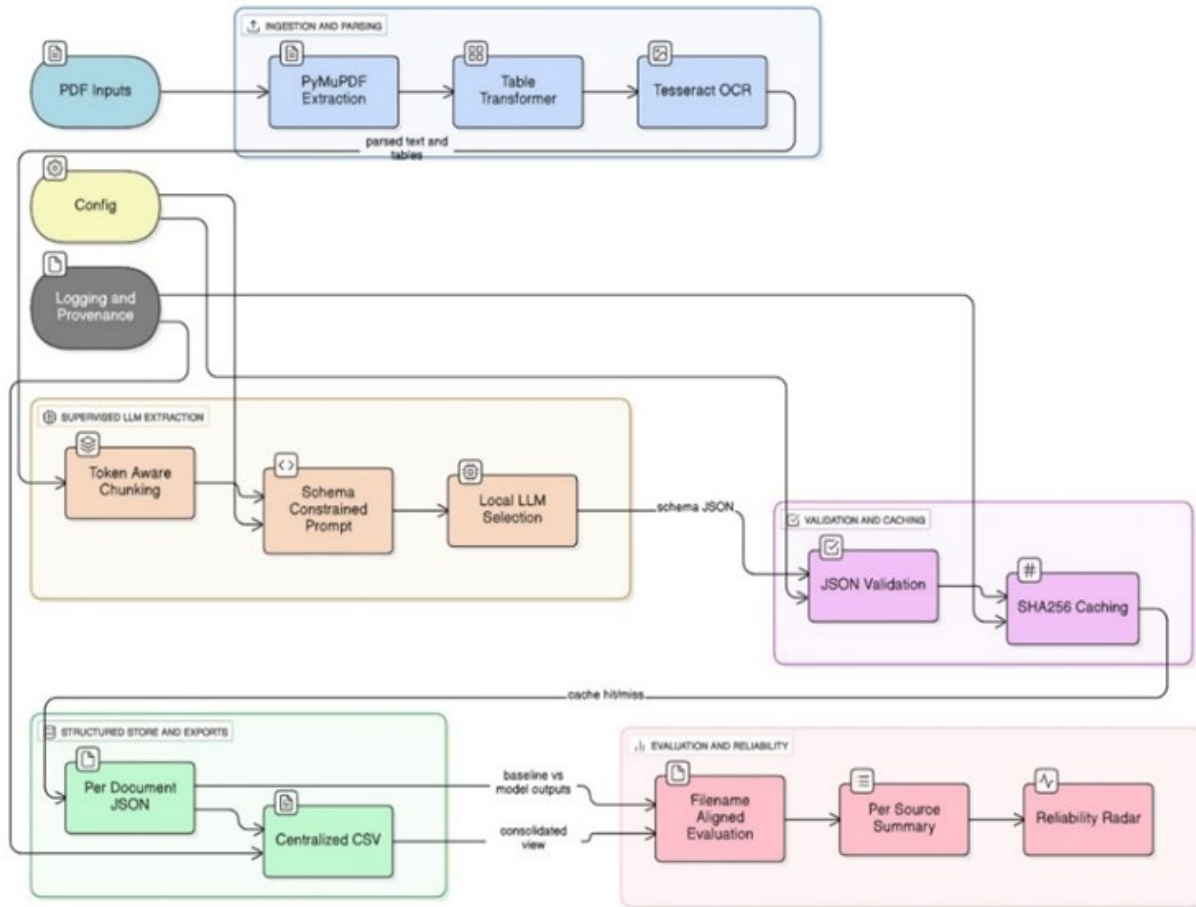


Figure 3. Unstructured document processing architecture used for contextual signals

Source: Cojocaru & Ionescu, 2025.

The figure 3 illustrates the pipeline implemented in the author's earlier research and reused in this work for the contextual layer. Information will be collected either by performing a web crawl or via API and stored into an object storage system. A large language model will be employed to perform named entity recognition (NER) and determine the time validity of news articles. The NER output and time validity of each news article are then mapped to the source of the news item and are stored in a structured format. News items concerning the risk exposure of companies are evaluated as being positive or negative. Finally, the signal will be aggregated and transmitted to the contextual layer for early warning purposes and to provide explanations available for users.

This architecture provides the ability to utilize unstructured information to enhance the decision context in a manner that is both stable and traceable for the regulators. In addition, this process exemplifies the application of concepts related to data provenance and the application of processing metadata. Furthermore, this process exemplifies the controlled use of recent technological developments such as large language models for a specific, well-defined purpose.

The relationship between key requirement categories identified through the application of PRIME-RE and the layers of the resultant architecture is summarized in Table 2. The mapping explicitly identifies the relationships between functional requirements and non-functional requirements and the architectural decisions they influence.

Table 2. Mapping requirements to proposed architecture

Requirement Category	Description	Example Requirements	Mapped Layer
Decision objectives	Defines purpose and acceptable outcomes of the risk evaluation.	Provide a stable risk score comparable across firms.	Structural evaluation layer
Data and sources	Covers data quality, provenance and availability for structured and unstructured data.	Use audited financial statements; use unstructured data only as contextual signals.	Structural evaluation layer (structured) and contextual layer (unstructured)
Analytical feasibility	Concerns choice of modelling approaches and interpretability.	Employ interpretable models such as gradient boosting; document model limitations.	Structural evaluation layer and governance layer
Governance and regulation	Reflects compliance with SR 11-7, BCBS 239 and the EU AI Act.	Maintain audit trails; implement human oversight; ensure explainability and traceability.	Governance layer and presentation layer
Presentation and user roles	Addresses differentiation of outputs by role and scenario.	Provide executive summary view and detailed analytics view.	Presentation layer

Source: Author's analysis based on PRIME-RE application.

Conclusion

The approach creates a parallel process to the identification of financial decision system requirements and incorporates decision objectives, data feasibility, analytical feasibility, and governance. Additionally, the approach offers a recursive way to manage conflicting requirements. Further, the approach includes regulatory guidelines that relate to financial regulation so that the financial regulation and governance elements can be integrated into the financial decision system and not added on at a later date.

The results demonstrate that the use of explicit parallelism and recursion in requirements elicitation provides a mechanism for practitioners to identify problems that would have otherwise gone unnoticed. The emphasis on non-functional requirements (interpretability, stability, traceability, and governance) in high-stakes environments indicates the architectural and elicitation process requirements. The separation of the core score and contextual signals allow for the mitigation of volatility concerns and to meet supervisory requirements. The direct relationship between the requirements and architecture creates a tangible reference for the systems designer, and reemphasizes the importance of governance.

PRIME-RE provides assurance that regulatory requirements will not be forgotten during the requirements elicitation process, and are therefore important to incorporate into the requirements elicitation process as soon as possible to avoid costly redesigns and to create confidence. Expert validation provides additional assurance that experts place greater value on stability and interpretability over slightly better performance. Therefore, trustworthiness must be established as a-priori condition for the deployment of financial AI systems.

The processing of unstructured documents demonstrates the potential for the responsible use of emerging data sources, and is used to mediate the tension between obtaining a large quantity of information and achieving stable decision-making. The results present a solution to a significant problem in current risk analysis where the need for a large quantity of information typically conflicts with the need for stable decision-making.

Although there are some limitations to the study, such as applying the method in only one application area and having a limited number of experts validate the methodology, these are consistent with the Design Science Research approach employed, which focuses on the construction and initial validation of a conceptual artifact rather than its empirical testing across multiple projects. Future

research is needed to further test and evaluate the applicability of PRIME-RE in other fields of finance and in various system development settings, in order to compare the results against established frameworks such as CRISP-DM and traditional requirements engineering approaches.

In spite of these restrictions, the research work undertaken contributes to the improvement of the decision systems development process in data driven, machine learning environments.

References

- Basel Committee on Banking Supervision. (2013). Principles for effective risk data aggregation and risk reporting. Bank for International Settlements.
- Board of Governors of the Federal Reserve System. (2011). Supervisory guidance on model risk management (SR 11-7). Federal Reserve, Washington, DC.
- Cerchiello, P., Giudici, P., & Nicola, G. (2017). Twitter data models for bank risk contagion. *Neurocomputing*, 264, 50–56.
- Cojocaru, C., & Ionescu, S. (2025). A Supervised Framework for Document Processing at Scale with Large Language Models in Credit-Risk Research. *International Conference of Management and Industrial Engineering*, 12. <https://doi.org/10.56177/12icmie2025.196>.
- Doshi-Velez, F., et al. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.
- Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). *Fundamentals of business process management* (2nd ed.). Springer.
- European Parliament and Council. (2024). Artificial Intelligence Act. Proposed regulation, Brussels.
- Fayyad, U., et al. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37–54.
- Guidotti, R., et al. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), 93.
- Horkoff, J., et al. (2016). Goal-oriented requirements engineering: An extended systematic mapping study. *Requirements Engineering*, 21(2), 133–160.
- Kotonya, G., & Sommerville, I. (1998). *Requirements engineering: Processes and techniques*. John Wiley & Sons.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65.
- Pohl, K. (2010). *Requirements engineering: Fundamentals, principles, and techniques*. Springer.
- Power, D. J. (2002). *Decision support systems: Concepts and resources for managers*. Greenwood Publishing Group.
- Rolland, C., & Prakash, N. (2001). Matching ERP system functionality to customer requirements. *Proceedings of the Fifth IEEE International Enterprise Distributed Object Computing Conference*, 66–75.
- Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing*, 5(4), 13–22.
- Sommerville, I., & Sawyer, P. (1997). *Requirements engineering: A good practice guide*. John Wiley & Sons.
- van Lamsweerde, A. (2001). Goal-oriented requirements engineering: A guided tour. *Proceedings of the Fifth IEEE International Symposium on Requirements Engineering*, 249–262.
- Vogelsang, A., et al. (2019). Requirements engineering for machine learning: Perspectives from data scientists. *Proceedings of the IEEE/ACM International Workshop on Requirements Engineering for Machine Learning*, 1–8.