

GEMR:KG - GLOBAL EMERGING MARKET RISK KNOWLEDGE GRAPH

Adishesh Gonibeed Ravishankar^{*}, Hitesh Kolluru[†], Likitha Sathvik Potineni[‡],
Lalitha Shreya Vanam[§], Hemakshi Pandey[¶] and Srividya Bansal^{||}

*School of Computing and Augmented Intelligence, Arizona State University
Tempe, AZ, USA*

Abstract

A prevailing perception of high risk in emerging markets, often driven by incomplete and fragmented data, leads to capital misallocation and constrains investment. While powerful datasets such as the World Bank’s Global Economic Monitor (GEM) and the Global Emerging Markets (GEMR) Risk Database provide extensive macroeconomic and credit risk information, respectively, they exist in silos, limiting holistic analysis. This paper introduces the Global Emerging Markets Risk Knowledge Graph (GEMR:KG), a novel system that integrates these disparate sources into a unified, semantically rich framework. GEMR:KG models entities including countries, loans, economic indicators, and crisis events, along with their complex interrelationships. This structure enables multi-relational queries that are currently infeasible, such as tracing the geographic contagion of defaults or identifying leading macroeconomic indicators of credit stress. The system features a dual-interface querying capability, supporting both formal SPARQL for quantitative analysts.

KEYWORDS

Knowledge Graphs, Risk, Emerging Markets, World Bank, Ontology Engineering, Macroeconomic Indicator.

1. INTRODUCTION

The analysis of risk in emerging markets is hampered by a fundamental challenge: data fragmentation. On one hand, the World Bank’s Global Economic Monitor (GEM) (World Bank, 2024) provides a rich, high-frequency repository of macroeconomic indicators like GDP, inflation, and trade data. On the other, the Global Emerging Markets (GEMR) Risk Database (World Bank Group, 2025) offers unparalleled, granular data on private sector credit risk, including default and recovery rates over three decades. While invaluable on their own, these critical datasets exist in separate silos.

^{*}adishesh@asu.edu

[†]hkolluru@asu.edu

[‡]lpotinen@asu.edu

[§]lvanam@asu.edu

[¶]hpandey7@asu.edu

^{||}skbansa2@asu.edu

This separation makes it impossible to perform integrated analysis. An investor cannot easily query the relationship between a country’s currency fluctuations (from GEM) and the subsequent default rates in its manufacturing sector (from GEMR). This gap prevents a holistic understanding of risk dynamics and obscures important trends, such as the significant divergence between a country’s sovereign risk rating and the actual, often lower, credit risk of its private sector. To address this, we are developing the Global Emerging Markets Risk Knowledge Graph (GEMR:KG). Our goal is to create a unified system that integrates these two disparate datasets into a single, semantically connected network. To further enhance our analysis of non-economic factors, we have incorporated political risk data from the World Bank’s Governance Indicator Dataset (The World Bank, 2024).

Within this context, this paper addressess four core challenges at the intersection of global finance, risk management, and data science. First, *Inadequate Comprehensive Risk Modeling*: traditional approaches struggle to integrate the multitude of dynamic variables required to assess true country-level risk, including macroeconomic indicators (GDP, CPI, external balances), financial and market data (commodity prices, currency volatility, default histories), crisis-specific behavior (e.g., foreign-exchange shocks), non-economic factors (political instability, regulatory changes), and their temporal dynamics. Second, there is a *poor understanding of systemic contagion*: existing frameworks provide limited support for modeling geographic contagion and tracing the impact of one country’s default or currency shock on its neighbors and trade partners.

Third, stakeholders face *difficulty in extracting actionable insights* from complex datasets. Investors and policymakers lack tools that answer high-level questions directly, such as identifying indicators that best explain a crash in a given market or understanding which countries share similar risk profiles for diversification. Finally, the underlying data remains *inaccessible and siloed*: financial, economic, and political information is distributed across heterogeneous databases and APIs that are difficult for non-technical users to query intuitively.

GEMR:KG is designed to tackle these four challenges by providing a unified, ontology-driven knowledge graph that integrates macroeconomic, governance, and credit-risk data, supports temporal and contagion reasoning, and exposes these capabilities through both SPARQL and an interactive web interface.

2. LITERATURE REVIEW

The intersection of **knowledge graphs (KGs)**, **large language models (LLMs)**, and **financial risk assessment** represents a rapidly evolving research domain with significant implications for emerging market economies. This literature review synthesizes recent advances across multiple disciplines to contextualize our proposed framework for integrated financial risk assessment.

2.1 Knowledge Graphs for Financial Applications

Knowledge graphs have emerged as powerful tools for representing complex financial relationships and enhancing analytical capabilities. Mitra et al. (2024) introduced a pioneering framework combining Relational Graph Convolutional Networks (RGCN) with Random Forest algorithms for **credit risk assessment** of Micro, Small, and Medium-sized Enterprises (MSMEs). Their approach achieved 92% balanced accuracy on Indian MSME data by leveraging graph-based representations to capture intricate relationships between borrowers, industries, and temporal features.

Yang et al. (2021) extended graph-based approaches to supply chain contexts, demonstrating how graph mining techniques can identify hidden risk propagation patterns among interconnected

enterprises. Their framework analyzed 97K records, revealing that network topology and relationship dynamics significantly predict financial distress beyond conventional financial ratios.

Yang and Liao (2022) developed a Multi-Network dynamic knowledge graph combining transfer learning, conditional random fields, and BiLSTM architectures to extract enterprise risk signals from news corpora and corporate disclosures. Their **multi-source data fusion** approach achieved superior performance in identifying emerging risks compared to single-source methods, validating the importance of heterogeneous information integration that our framework seeks to operationalize.

2.2 Macroeconomic and Risk Propagation Frameworks

Tilly and Livan (2021) validated that knowledge graphs constructed from news narratives significantly improve macroeconomic forecasting, with disease and economy-related themes showing strong predictive power for industrial production across major economies. Their statistically validated filtering methodology provides a blueprint for extracting parsimonious yet informative graph structures—a principle we extend to financial risk contexts.

Li and Sanna Passino (2024) developed FinDKG, a **dynamic knowledge graph** framework for financial trend detection, integrating temporal dynamics through quadruples (subject, relation, object, timestamp). Their KGTransformer architecture, incorporating heterogeneous graph attention mechanisms, improved link prediction by $\sim 10\%$ on the FinDKG dataset. While focused on thematic investing rather than risk assessment, their temporal modeling techniques inform our approach to capturing evolving risk conditions.

2.3 Financial Contagion and Systemic Risk

Glick and Rose (1999) established the seminal finding that **currency crises propagate primarily through trade linkages** rather than shared macroeconomic fundamentals, with countries tied by international trade experiencing synchronized exchange-market pressure. Their empirical framework provides a canonical model for **geographic contagion** that we adapt for default-risk contexts in emerging markets.

Recent research by Galizia and Lund (2024) challenges conventional risk perceptions of emerging market lending. Analysis of 15,000 loans across 30 years revealed average default rates of 3.6%, comparable to B-rated corporates in advanced economies, with 72% recovery rates substantially higher than emerging market bonds. These findings validate the feasibility of our risk assessment framework while highlighting the importance of nuanced, data-driven approaches for emerging markets.

3. APPROACH

Our approach to building the GEMR:KG system centers on the design and implementation of a comprehensive ontology that serves as the semantic foundation for integrating heterogeneous financial risk data. The ontology engineering process follows established best practices in knowledge representation while addressing the unique requirements of emerging market risk analysis.

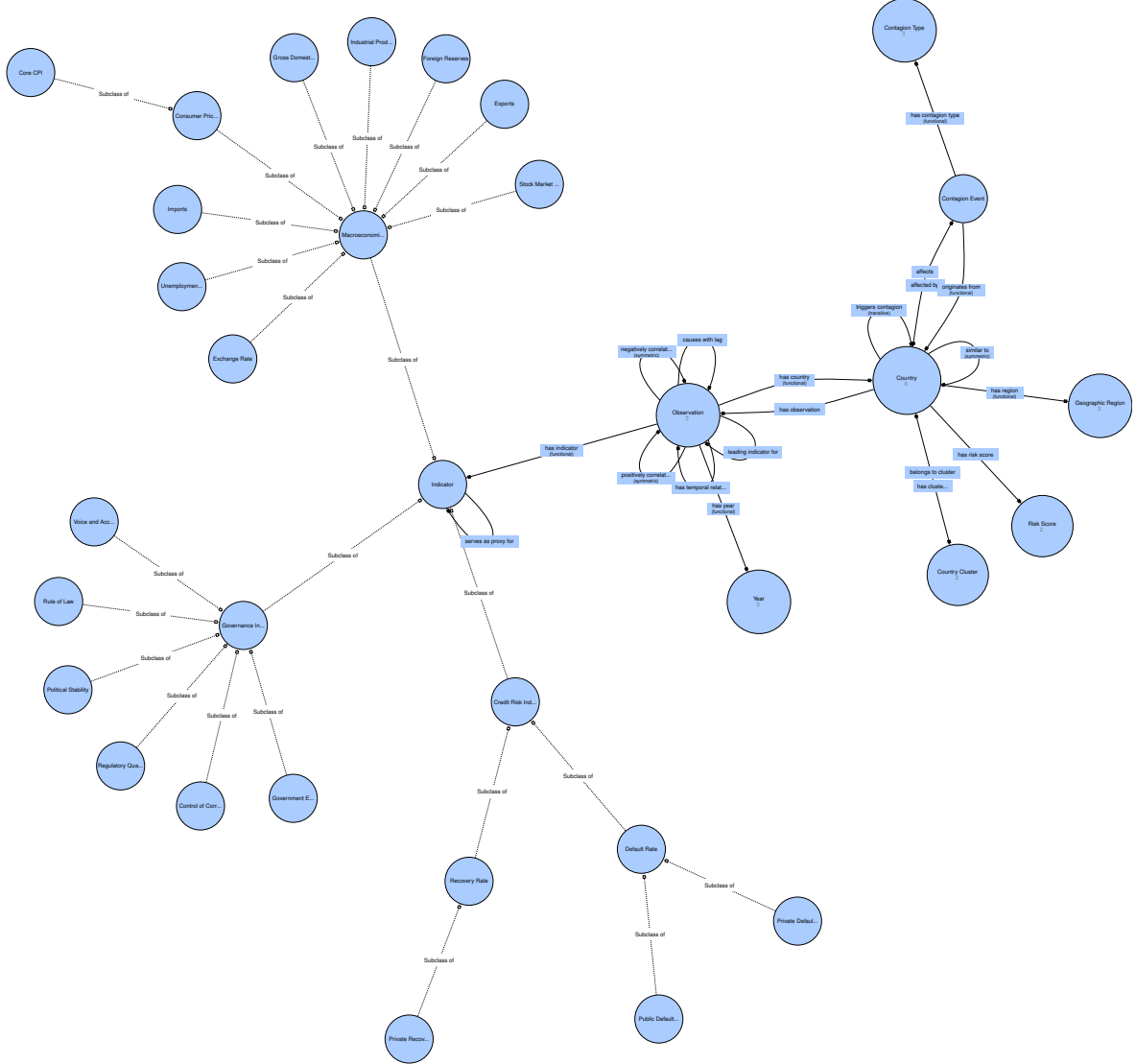


Figure 1. GEMR-KG Ontology Structure

3.1 Ontology Design and Architecture

The GEMR-KG ontology provides a formal specification of the domain concepts, relationships, and constraints necessary for unified risk assessment. The ontology architecture is structured around six primary entity classes that capture the multidimensional nature of emerging market risk:

Core Entity Classes: At the heart of the ontology lies the *Country* class, representing the six emerging markets under analysis (Brazil, China, Mexico, Philippines, Poland, and Thailand). Each country is associated with temporal *Observations* that link specific *Indicators* to particular *Years*, enabling comprehensive time-series analysis across the 2002-2023 period.

Indicator Hierarchy: The ontology models three distinct indicator taxonomies through subclass relationships: (1) *Macroeconomic Indicators* encompassing GDP, inflation (CPI and Core CPI), exchange rates, trade balances (imports/exports), industrial production, stock market indices, unemployment, and foreign reserves; (2) *Governance Indicators* derived from the World Governance Indicators (WGI) framework, including Control of Corruption, Government Effectiveness, Political Stability, Regulatory Quality, Rule of Law, and Voice and Accountability; and (3)

Credit Risk Indicators capturing default rates (private and public) and recovery rates. This hierarchical organization enables both fine-grained queries on specific metrics and aggregate analysis across indicator categories.

Risk Assessment Framework: The *RiskScore* class integrates multiple dimensions of country-level risk through composite scoring across five components: governance quality, economic health, default risk, external vulnerability, and contagion exposure. This multi-factor approach addresses the inadequacy of single-metric risk assessments and provides a holistic view of country risk profiles.

Contagion Modeling: To capture systemic risk propagation, the ontology includes specialized classes for *ContagionEvent* and *ContagionType* (trade-based, currency spillover, political cascade), connected through object properties such as `originatesFrom`, `affects`, and `triggersContagion`. The transitive nature of the `triggersContagion` property enables recursive traversal of contagion chains, facilitating analysis of multi-hop risk transmission across geographic regions.

Geographic and Clustering Constructs: The ontology models regional groupings through *GeographicRegion* (Asia, Latin America, Europe) and similarity-based *CountryCluster* classes. The `similarTo` symmetric property, coupled with quantitative `similarityScore` datatype properties, enables comparative risk analysis and portfolio diversification strategies based on empirically validated country groupings.

Temporal Relationships: Critical to understanding risk dynamics, the ontology encodes temporal dependencies through properties such as `causesWithLag`, `leadingIndicatorFor`, and `negativelyCorrelatedWith`. These relationships, validated through statistical analysis (Pearson correlation coefficients and p-values stored as datatype properties), capture lagged causal effects for instance, the documented 1-year lag between governance improvements (Control of Corruption) and GDP growth in Thailand ($r = 0.563, p = 0.010$).

The ontology implements 27 validated cross-dataset relationships spanning macroeconomic, governance, and credit risk dimensions. Functional property constraints ensure data integrity (e.g., each observation has exactly one country, year, and indicator), while cardinality restrictions enforce domain rules such as contagion events originating from precisely one source country while potentially affecting multiple destinations. This formal structure enables SPARQL queries that traverse complex relationship paths answering questions like "Which countries similar to Brazil experienced currency crises that led to manufacturing sector defaults within 18 months?" queries that are impossible with soiled datasets.

3.2 Data Integration Strategy

The ontology functions as a sophisticated integration layer, engineered to semantically harmonize three distinct yet complementary data sources that are fundamental to modern risk assessment. It systematically ingests the World Bank's Global Economic Monitor (GEM), leveraging a robust twenty-two-year longitudinal dataset of high frequency macroeconomic indicators to capture dynamic economic shifts. Parallel to this, it integrates the World Governance Indicators (WGI), which provide critical annual assessments of institutional quality across six foundational dimensions of governance. To address the financial stability of the private sector, the framework further incorporates the IFC GEMs DB, supplying granular, empirical statistics on credit risk and loan performance that are essential for modeling default probabilities and recovery rates. Data harmonization challenges including temporal alignment, unit standardization, and handling sparse coverage in credit risk data are addressed through explicit modeling of data quality attributes (`dataAvailability`, `coverageRatio`) and proxy relationships (`servesAsProxyFor`) within the ontology.

3.3 Ontology Rationale and Provenance

The core rationale for our ontology design is to establish a unified semantic platform that harmonizes heterogeneous data from three distinct silos: the World Bank’s Global Economic Monitor (GEM), World Governance Indicators (WGI), and the GEMS credit risk database. To ensure high architectural trust and interoperability, we extended the Financial Industry Business Ontology (FIBO) , integrating its authoritative standards with our domain-specific requirements for emerging market risk. A central feature of this design is the explicit modeling of temporal dependencies, where the ontology treats time as a critical relational link—transforming isolated statistics into a longitudinal network that supports 22-year time-series analysis. Consequently, the core entities such as Country, Indicator, and Observation are directly mapped from these primary sources, providing a formal structure that can traverse complex relationship paths—such as linking governance shifts to credit defaults across varying time lags.

4. IMPLEMENTATION

The system implementation follows a pipeline designed to transform fragmented raw data into a semantically rich knowledge graph. Data are batch-ingested as raw CSV/XLSX files, undergoing immediate schema validation and range checks to reject anomalous entries.

Subsequently, the data is processed into filtered panels using Python’s Pandas library. This stage involves extensive data cleaning, including normalizing currency units to USD, imputing missing values via linear interpolation, and computing year over year growth metrics. Crucially, the pipeline aligns disparate time axes harmonizing daily financial market data with annual governance scores and programmatically generates temporal lags ($t - 1$, $t - 2$) as illustrated in Fig.?? to facilitate predictive modeling.

In the semantic modeling phase, we constructed the domain ontology by extending the **Financial Industry Business Ontology (FIBO)** to define domain-specific constraints (EDM Council, 2024). The OWL schema was developed and validated within **Protégé**, where we executed reasoning checks to ensure logical consistency and verify relationships between economic indicators and governance data. To populate the graph, we used **Ontotext Refine** to map the processed tabular panels to ontology classes, automatically generating compliant RDF triples from source files.

Finally, the backend architecture is powered by a **Dockerized GraphDB** repository, ensuring a containerized and consistent deployment environment for the semantic store. This layer utilizes OWL 2 RL inference profiles to deduce implicit connections and employs caching to optimize performance. For the user interface, the application is built entirely using **React** and **Vite**. This modern frontend communicates directly with the GraphDB endpoint by executing parameterized SPARQL queries, fetching JSON-formatted results without the overhead of an intermediate middleware layer. Furthermore, this modular architecture facilitates the seamless integration of future visual analytics widgets with significant reconfiguration of the backend logic.

The complete source code, ontology definitions, and documentation for the proposed system are available in this GitHub Repository

5. USER INTERFACE DESIGN

The GEMR:KG web application is designed with a dual-focus interface to accommodate both domain experts and general researchers, who prioritize visual insights. The user interface comprises two primary modules: the Analytical Dashboard and the Knowledge Hub.

1. Comprehensive Data Retrieval

Fetch complete dataset (all indicators, years, values) for a specific country. Mirrors the IMF “Signal Approach” for aggregating risk factors (International Monetary Fund, 2017).

```
PREFIX gemr: <https://gemr-kg.org/ontology#>
SELECT DISTINCT ?indicator ?year ?value WHERE {
  ?obs a gemr:Observation ; gemr:hasCountry ?c ;
      gemr:hasIndicator ?indicator ;
      gemr:hasYear ?yEnt ; gemr:observationValue ?val .
  ?c gemr:countryName "Brazil" . ?yEnt gemr:yearValue ?
    year .
} ORDER BY ?indicator ?year
```

3. Early Warning: Stock vs Default

Analyzes Stock Market (Year T) vs Private Default Rates ($T + 1$). Tests hypothesis that market sentiment precedes fundamental credit ratings (Wu, 2008).

```
SELECT ?cName ?yInt ?stock (AVG(?defRate) AS ?avgDef)
WHERE {
  ?obsS a gemr:Stock_Market_Index_LCU ; gemr:hasCountry ?c
  ;
      gemr:hasYear ?yE ; gemr:observationValue ?stock .
  ?yE gemr:yearValue ?yLit . BIND(xsd:integer(STR(?yLit))
    AS ?yInt)
  BIND(?yInt + 1 AS ?tYInt)
  ?tYE gemr:yearValue ?tYLit .
  FILTER(xsd:integer(STR(?tYLit)) = ?tYInt)
  ?obsD a gemr:PrivateDefaultRate ; gemr:hasCountry ?c ;
      gemr:hasYear ?tYE ; gemr:observationValue ?defRate
  .
  ?c gemr:countryName ?cName .
} GROUP BY ?cName ?yInt ?stock ?tYInt LIMIT 50
```

5. Default → Recovery Lag

Lag between stock market recovery and real economic (GDP) recovery after default. Asset prices often rebound earlier than real indicators (Kalesnik and Polychronopoulos, 2020).

```
SELECT ?cName ?yT ?yTarget ?gdp (AVG(?stk) AS ?avgStk)
WHERE {
  ?obsS a gemr:Stock_Market_Index_LCU ; gemr:hasCountry ?c
  ;
      gemr:hasYear ?yT_Raw ; gemr:observationValue ?stk
  .
  ?obsG a gemr:GDP_CONST_2010_USD ; gemr:hasCountry ?c ;
      gemr:hasYear ?yTar_Raw ; gemr:observationValue ?
    gdp .
  OPTIONAL { ?yT_Raw gemr:yearValue ?yVT }
  OPTIONAL { ?yTar_Raw gemr:yearValue ?yVTar }
  BIND(COALESCE(xsd:integer(STR(?yVT)), ?yT_Raw) AS ?yT)
  BIND(COALESCE(xsd:integer(STR(?yVTar)), ?yTar_Raw) AS ?
    yTarget)
  FILTER(?yTarget = ?yT + 1) ?c gemr:countryName ?cName .
} GROUP BY ?cName ?yT ?yTarget ?gdp
```

7. Trade-Based Contagion

Traces economic shocks from Source to Dependent Partner via trade. Identifies spillover risks from major trading partners (Van Rijckeghem and Weder, 2003).

```
SELECT ?src ?ptnr ?yT (AVG(?sGro) AS ?aSGr) (AVG(?pExGr)
  AS ?aPEx)
WHERE {
  { ?ptnr gemr:similarTo ?source . } UNION
  { ?ptnr gemr:belongsToCluster ?cl . ?source gemr:
    belongsToCluster ?cl .
    FILTER(?source != ?ptnr) } UNION
  { ?source gemr:countryName ?sN . ?ptnr gemr:countryName
    ?pN .
    FILTER ((STR(?sN)="Brazil" && STR(?pN)="Mexico")) }
  # Growth calcs omitted for brevity
  ?source gemr:countryName ?src . ?ptnr gemr:countryName ?
    ptnr .
} GROUP BY ?src ?ptnr ?yT LIMIT 50
```

2. Comprehensive Risk Profiling

Retrieves total risk score, classification, and component scores. Provides a holistic snapshot of stability vs specific areas of concern.

```
PREFIX gemr: <https://gemr-kg.org/ontology#>
SELECT ?cName ?year ?total ?tier ?gov ?econ ?vuln WHERE {
  BIND("2023"^^xsd:gYear AS ?tYear) BIND("Poland" AS ?
    tCtry)
  ?sObs a gemr:RiskScore ; gemr:hasCountry ?c ;
      gemr:hasYear ?yEnt ; gemr:totalRiskScore ?total .
  ?c gemr:countryName ?cName . ?yEnt gemr:yearValue ?year
  .
  FILTER (?year = ?tYear && STR(?cName) = ?tCtry)
  OPTIONAL { ?sObs gemr:riskTier ?tier }
  OPTIONAL { ?sObs gemr:governanceScore ?gov }
  OPTIONAL { ?sObs gemr:economicHealthScore ?econ }
  OPTIONAL { ?sObs gemr:externalVulnerabilityScore ?vuln
    }
}
```

4. Political Stability Impact Tracker

Tracks negative Political Stability (T) impact on GDP ($T + 1$). Validates the ‘Governance-Growth’ transmission channel (Aisen and Veiga, 2011).

```
SELECT ?cName ?yInt ?gdp (AVG(?polStab) AS ?avgPol) WHERE
  {
    ?obsP a gemr:PoliticalStability ; gemr:hasCountry ?c ;
      gemr:hasYear ?yE ; gemr:observationValue ?polStab
    .
    FILTER(?polStab < 0)
    ?yE gemr:yearValue ?yLit . BIND(xsd:integer(STR(?yLit))
      AS ?yInt)
    BIND(?yInt + 1 AS ?tYInt)
    ?obsG a gemr:GDP_CONST_2010_USD ; gemr:hasCountry ?c ;
      gemr:hasYear ?gYRaw ; gemr:observationValue ?gdp .
    OPTIONAL { ?gYRaw gemr:yearValue ?gYVal }
    BIND(COALESCE(xsd:integer(STR(?gYVal)), ?gYRaw) AS ?gYInt
      )
    FILTER(?gYInt = ?tYInt) ?c gemr:countryName ?cName .
  } GROUP BY ?cName ?yInt ?tYInt ?gdp LIMIT 50
```

6. GDP Growth Tracker

Calculates annual GDP growth percentage. Provides a standardized metric for comparing economic performance across economies.

```
SELECT ?cName ?year ?gdpCurr ?growthPct WHERE {
  ?obsG_T a gemr:GDP_CONST_2010_USD ; gemr:hasCountry ?c ;
      gemr:hasYear ?yT_R ; gemr:observationValue ?
    gdpCurr .
  ?obsG_P a gemr:GDP_CONST_2010_USD ; gemr:hasCountry ?c ;
      gemr:hasYear ?yP_R ; gemr:observationValue ?
    gdpPrev .
  OPTIONAL { ?yT_R gemr:yearValue ?yVT }
  OPTIONAL { ?yP_R gemr:yearValue ?yVP }
  BIND(COALESCE(xsd:integer(STR(?yVT)), ?yT_R) AS ?year)
  BIND(COALESCE(xsd:integer(STR(?yVP)), ?yP_R) AS ?yPrev)
  FILTER(?yPrev = ?year - 1)
  BIND(((?gdpCurr - ?gdpPrev)/?gdpPrev)*100 AS ?growthPct)
  ?c gemr:countryName ?cName .
} LIMIT 100
```

8. Currency Crisis Early Warning

Detects regional currency contagion based on “Neighborhood Effects” in emerging markets (Basu, 2001).

```
SELECT ?src ?tgt ?yr ?sDep ?tDep ?risk WHERE {
  { SELECT ?src ?tgt ?yr (AVG(?sD_P) AS ?sDep) (AVG(?tD_P)
    AS ?tDep)
    WHERE {
      ?obsS a gemr:Exchange_rate_new_LCU_per_USD ;
      gemr:hasCountry ?s ; gemr:observationValue ?sV
      .
      # Calculation logic omitted
      VALUES (?sN ?tN) { ("Brazil" "Mexico") ("Thailand" "
        Philippines") }
      ?s gemr:countryName ?sN . ?t gemr:countryName ?tN .
      ?s gemr:countryName ?src . ?t gemr:countryName ?tgt
      .
    } GROUP BY ?src ?tgt ?yr }
  BIND(IF(?tDep > 10, "HIGH", IF(?tDep > 5, "MOD", "LOW"))
    AS ?risk)
}
```

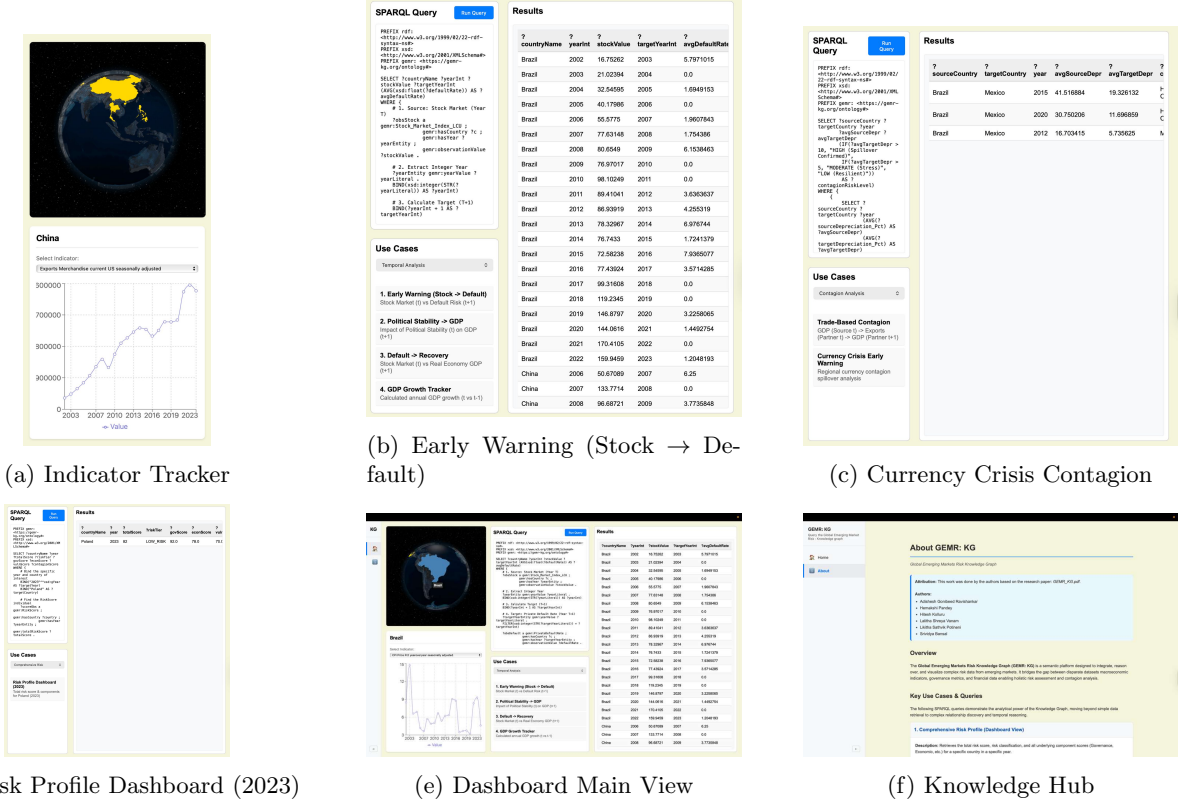


Figure 2. Analytical visualizations and user interface views generated by GEMR:KG.

5.1 Analytical Dashboard

The core interaction layer integrates visual exploration with semantic querying. A **geospatial selection module** implemented as an interactive 3D globe allows users to select target emerging market countries (e.g., Brazil), automatically filtering the downstream data context. A **dynamic visualization pane** renders time-series charts for selected indicators (e.g., “CPI Price Pct”) over the 2002–2023 period, highlighting trends and volatility spikes. The **SPARQL editor** at the center displays the active query and offers pre-defined templates (such as “Early Warning (Stock → Default)” or “Political Stability → GDP”), while the **tabular results view** in the right-hand panel presents query outputs in an exportable tabular format.

5.2 Knowledge Hub

To ensure transparency and understanding some concepts, the application includes a dedicated “About” module. This section documents our paper’s attribution and provides detailed textual explanations for the supported SPARQL patterns. By listing key use cases such as the “Comprehensive Risk Profile” alongside their descriptions, the interface bridges the gap between raw ontology classes and practical financial concept explanation.

6. QUERYING WITH SPARQL

To validate the ontology’s analytical depth, we executed a series of queries ranging from basic data retrieval to complex predictive modeling. We specifically focus on validating contagion mechanisms,

where financial distress spreads across borders. The grid below demonstrates the system’s ability to handle temporal lags, risk profiling, and contagion mechanics through concise SPARQL queries, and Figure 2 shows example visual outputs and interface views generated by these queries.

7. CHALLENGES AND LIMITATIONS

Developing GEMR:KG exposed two main limitations. (i) **data scarcity and continuity**: Many high-quality financial datasets are paywalled or incomplete. The public subsets we use contain missing years and mixed reporting frequencies (daily vs. annual). We mitigate this with interpolation and unit harmonization, but sparse coverage still limits the granularity of our risk analysis compared with commercial systems. (ii) **restricted temporal scope**: After cleaning and alignment, our panel is effectively constrained to 2002–2023. As a result, GEMR:KG cannot yet model earlier crisis episodes such as the 1997 Asian Financial Crisis, which restricts long-run validation of contagion patterns.

8. CONCLUSIONS AND FUTURE WORK

The Global Emerging Markets Risk Knowledge Graph (GEMR:KG) addresses the critical challenge of data fragmentation in financial risk assessment by creating a unified semantic framework. Through the execution of advanced SPARQL queries, the system validated its ability to trace temporal dependencies, such as the lag between political instability and GDP decline, and to identify systemic vulnerabilities through trade-based contagion. Ultimately, GEMR:KG empowers investors and policymakers to move beyond isolated statistics, offering a tool to visualize the intricate web of causal relationships that drive stability and risk in the global economy.

Future work will focus on three directions: (i) adding an **LLM-based query layer** that translates natural-language questions into SPARQL, lowering the barrier for non-technical users; (ii) **expand geographic coverage** by ingesting additional GEM and GEMs data, extending the knowledge graph beyond the current six countries; (iii) **enhance visual analytics** for multi-hop queries, including network views of contagion paths and correlation heatmaps, so that complex graph patterns are accessible.

8.1 Strategic Significance and Analytical Utility of GEMR:KG

The Global Emerging Markets Risk Knowledge Graph (GEMR:KG) is designed to resolve the critical challenge of data fragmentation by integrating the World Bank’s Global Economic Monitor, the World Governance Indicators, and the Global Emerging Markets Risk Database into a unified, semantically connected network. This research serves a dual-purpose audience, providing quantitative analysts with formal SPARQL querying capabilities for multi-relational analysis while offering policymakers and general researchers an interactive web interface for geospatial and visual risk exploration. By modeling the complex interrelationships between macroeconomic indicators, governance quality, and credit performance, the system generates actionable intelligence regarding systemic contagion paths and temporal dependencies such as identifying how currency volatility in one region triggers lagged credit stress in another. Ultimately, this framework allows stakeholders to move beyond the limitations of isolated statistics to visualize the intricate web of causal relationships that drive global economic stability, specifically capturing the nuanced divergence between a country’s sovereign rating and the actual credit risk of its private sector.

ACKNOWLEDGMENTS

The authors thank the World Bank for providing the datasets used in this study. They also used Google’s Gemini LLM to assist with language editing, LaTeX formatting, and syntax checking of SPARQL queries; all analyses and final text were reviewed and approved by the authors, who take full responsibility for the work.

References

- Aisen, A. and Veiga, F. J. (2011). How does political instability affect economic growth? IMF Working Paper WP/11/12, International Monetary Fund.
- Basu, T. G. (2001). What drives contagion: Trade, neighborhood, or financial links? IMF Working Paper WP/01/29, International Monetary Fund.
- EDM Council (2024). Financial industry business ontology (fibo). [Online]. Available: <https://spec.edmcouncil.org/fibo/ontology>.
- Galizia, F. and Lund, S. (2024). Reassessing risk in emerging market lending: Insights from GEMs consortium statistics. Ifc research note, International Finance Corporation.
- Glick, R. and Rose, A. K. (1999). Contagion and trade: Why are currency crises regional? *Journal of International Money and Finance*, 18(4):603–617.
- International Monetary Fund (2017). Assessing country risk: Selected approaches. Technical note, International Monetary Fund, Washington, D.C.
- Kalesnik, V. and Polychronopoulos, A. (2020). Value in recessions and recoveries. Technical report, Research Affiliates.
- Li, X. and Sanna Passino, M. (2024). FinDKG: Dynamic knowledge graphs with large language models for detecting global trends in financial markets. *arXiv preprint arXiv:2407.10909*.
- Mitra, R., Dongre, A., Dangare, P., Goswami, A., and Tiwari, M. K. (2024). Knowledge graph driven credit risk assessment for micro, small and medium-sized enterprises. *International Journal of Production Research*, 62(12):4273–4289.
- The World Bank (2024). Worldwide governance indicators. Available: <https://www.worldbank.org/en/publication/worldwide-governance-indicators>.
- Tilly, S. and Livan, G. (2021). Macroeconomic forecasting with statistically validated knowledge graphs. *Expert Systems with Applications*, 186:115765.
- Van Rijckeghem, C. and Weder, B. (2003). Spillovers through banking centers: A panel data analysis of bank flows. *Journal of International Money and Finance*, 22(4):483–509. See also IMF Working Paper 00/55.
- World Bank (2024). Global economic monitor — data catalog. Available: <https://datacatalog.worldbank.org/search/dataset/0037798/Global-Economic-Monitor>.
- World Bank Group (2025). Global emerging markets risk database (gems) consortium. Available: <https://www.worldbank.org/en/about/unit/brief/gems>.
- Wu, T. (2008). The dynamic pricing of sovereign risk in emerging markets: Fundamentals and sentiment. *Journal of International Money and Finance*, 27(7):1171–1188.
- Yang, B. and Liao, Y. M. (2022). Research on enterprise risk knowledge graph based on multi-source data fusion. *Neural Computing and Applications*, 34(4):2569–2582.
- Yang, S., Zhang, Z., Zhou, J., Wang, Y., Sun, W., Zhong, X., Fang, Y., Yu, Q., and Qi, Y. (2021). Financial risk analysis for SMEs with graph-based supply chain mining. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, pages 4661–4667.