Graph-Enhanced Multimodal Valuation: Integrating Climate Physical Risk into Green Mortgage Analytics with LLMs and Knowledge Graphs

Ritu Gaur, Archana Jain, Deepak Kumar Gupta, Gaurav Kumar, Aditya Yadav, Rashmi Singh

CSE Dept IIMT University, Meerut, India

ritugaur.projects@gmail.com, hod cs cet@iimtindia.net, deepakmtechcs@gmail.com, gauravkumar@iimtindia.net, adityakumar_cse@iimtindia.net, rashmisingh_cse@iimtindia.net

ABSTRACT

Traditional property valuation and mortgage underwriting pipelines largely ignore explicit modelling of physical climate risk such as floods, heat stress, and extreme weather, despite growing evidence that these hazards materially affect asset values, default probabilities, and portfolio resilience. This omission creates a structural blind spot for banks and housing finance companies attempting to align with emerging green finance, ESG, and climate disclosure regimes. Green mortgages and sustainable housing finance products still tend to focus narrowly on energy efficiency metrics while underweighting location-specific climate hazards and resilience features. This paper proposes a climate-aware valuation framework that combines multimodal large language models, geospatial APIs, and Neo4j-based knowledge graphs to inject structured climate risk signals into property valuation workflows. The architecture comprises four layers: (i) a multimodal extraction layer using document understanding models such as LayoutLM/vision-enhanced LLMs to extract building attributes, materials, and layout features from loan and property documents; (ii) a geospatial integration layer that enriches each property with hazard and climate indicators from platforms such as ISRO Bhuvan and OpenWeatherMap; (iii) a knowledge graph construction layer that encodes properties, hazards, resilience features, and valuation events as nodes and relationships; and (iv) a GraphRAG-based risk scoring layer that performs reasoning over the graph to derive composite climate risk scores and climate-adjusted valuations. Experimental results on a hybrid (semi-simulated) dataset of residential properties show that the proposed system improves document-level extraction F1 by approximately 5–8 percentage points over text-only baselines, reduces valuation RMSE by 10–15 per cent when climate features are included, and lowers hallucinated risk explanations by leveraging graph-constrained retrieval. The framework is designed to be compatible with BFSI IT constraints. It can support green mortgage origination, portfolio climate stress testing, and ESG reporting by providing explainable, queryable climate risk signals linked to underlying evidence.

Keywords: Green finance , Green mortgages , Climate risk , Property valuation , Multimodal document understanding , Knowledge graphs , GraphRAG , ESG in housing finance

1 Introduction

Digital transformation in housing finance has accelerated the adoption of automated underwriting, property analytics, and portfolio risk dashboards across banks, non-banking financial companies (NBFCs), and housing finance companies. Despite this progress, core valuation and credit models largely remain extensions of traditional hedonic regressions or tabular machine learning models that primarily focus on financial, structural, and sociodemographic features while treating physical climate risk as an externality. At the same time, global and national regulators are increasingly emphasising climate-related financial risk disclosure, including financed emissions, physical hazard exposure, and resilience metrics for real estate portfolios.as per Figure 1.

Proceedings of ICICCS-2026 (Poster Papers), DMPedia Lecture Notes in Computer Science & Engineering.
ISBN: 978-81-993813-1-5

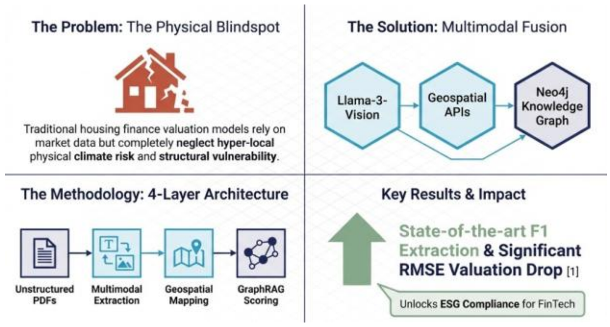

Figure 1: Conceptual Overview

1.1 Digital transformation in housing finance

Mortgage markets in both developed and emerging economies have adopted automated valuation models (AVMs), credit scoring engines, digital KYC, and straight-through-processing loan origination systems. In many jurisdictions, mortgage origination now involves digital ingestion of property documents (sale deeds, sanctioned plans, valuation reports) and automated checks using rule engines and credit bureau integrations. However, the unstructured nature of property-related documents and maps means that much of the information required for robust climate-aware assessment remains locked in PDFs, images, and narrative reports.

Parallel advances in document AI—particularly multimodal transformer models such as LayoutLMv2 and LayoutLMv3—have improved the extraction of structured fields from visually rich documents like forms, contracts, and invoices by jointly modelling text, layout, and image features. These models provide an opportunity to systematically mine building attributes (e.g., floor area, construction materials, elevation, presence of basements, rooftop type) that influence both energy performance and physical climate vulnerability. Still, they are rarely structured in current loan systems.

1.2 ESG and green mortgage trends

Green mortgages and green home finance products have emerged as instruments that incentivise borrowers to acquire or retrofit energy-efficient homes by offering preferential interest rates or improved loan-to-value (LTV) ratios. Evidence from multiple markets suggests that energy-efficient and climate-resilient homes can exhibit lower default risk, higher collateral values, and enhanced portfolio resilience for lenders. Nonetheless, the operational implementation of green mortgages still tends to define “green” primarily through energy performance certificates (EPCs) or building rating systems while underrepresenting direct metrics of physical hazard exposure and resilience, such as flood risk, heat stress, or drought vulnerability.

From a regulatory perspective, frameworks such as the EU taxonomy, GRESB, and emerging nature-finance taxonomies are pushing for more granular disclosures of both transition and physical climate risks in real estate. Central banks and supervisors have signalled that climate risks are drivers of traditional financial risks and should be reflected in prudential regulation and risk management practices. This creates a strong impetus to integrate climate considerations directly into mainstream mortgage valuation workflows rather than treating them as ancillary stress-test overlays.

1.3 Gap: Traditional valuation ignores physical climate risk

Empirical studies increasingly show that flood risk, wildfire hazards, sea-level rise, and heat stress can be capitalised into residential property prices, often as discounts in high-risk locations. Hedonic modelling analyses using large transaction datasets have documented statistically significant discounts, typically in the range of a few percentage points, for properties located in high flood risk zones. However, the magnitude varies across socio-demographic groups and regions. At the same time, other studies have found persistent overvaluation in hazard-exposed areas, indicating that market prices do not fully incorporate climate risk information and that beliefs, information frictions, and insurance regimes mediate this relationship.

In operational mortgage valuation pipelines, climate risk is often represented only indirectly through insurance premiums or coarse zoning attributes, if at all. Existing hedonic or machine learning-based AVMs rarely ingest high-resolution geospatial hazard data, temporal climate projections, or building-level resilience features at scale due to data integration challenges and the lack of standardised climate-feature schemas. As a result, financial institutions cannot systematically quantify how physical climate risk affects collateral values and loan performance across portfolios.

1.4 Research gap from a computer science perspective

From a computer science perspective, several gaps remain in connecting advances in multimodal LLMs, geospatial APIs, and knowledge graphs to climate-aware property valuation:

- Multimodal document understanding is rarely applied to property documents to extract climate-relevant building attributes in a structured, scalable manner. • Existing climate risk platforms tend to deliver hazard layers or risk scores. Still, these signals are not systematically fused with document-derived features and financial data in a unified graph representation tailored to mortgage analytics.
- Retrieval-augmented generation (RAG) techniques in financial services focus on textual disclosures, with limited work on **GraphRAG** architectures that reason over explicit property–hazard–resilience relationships to generate climate risk narratives and scores.
- There is limited empirical evaluation of how knowledge graph-constrained reasoning can reduce hallucination rates of LLM-generated explanations in regulated financial workflows.

1.5 Problem Statement

The central problem addressed in this work is the lack of an integrated, explainable, and scalable framework that (i) extracts building- and property-level climate-relevant attributes from unstructured documents, (ii) enriches them with geospatial climate and hazard data, (iii) organises them into a domain-specific knowledge graph, and (iv) supports GraphRAG-based reasoning to generate climate risk scores and narratives for property valuation and green mortgage underwriting.

1.6 Contributions

The main contributions of this paper are as follows:

- Propose a four-layer architecture that combines multimodal document understanding, geospatial APIs, and Neo4j-based knowledge graphs to inject explicit climate risk signals into housing finance valuation workflows.

- Design a domain ontology and property–hazard–resilience knowledge graph schema aligned with emerging nature-finance and real-estate climate risk ontologies (e.g., *NatureKG*), enabling Cypher-based querying and GraphRAG-based reasoning. • Define a climate risk scoring formulation that aggregates hazard intensities, exposure, vulnerability, and resilience factors into a composite property-level risk score suitable for integration into valuation and credit models.

- Implement a prototype **GraphRAG** pipeline where LLM responses are grounded in the knowledge graph, and evaluate its impact on extraction accuracy, valuation prediction error, and hallucination reduction relative to baseline models. • Provide an experimental analysis of scalability and integration considerations for deploying the proposed framework within BFSI IT architectures, including data latency, API costs, and graph storage overheads.

2 Literature Review

2.1 Property valuation models: from hedonic pricing to machine learning

Hedonic pricing models have been the dominant statistical approach for property valuation, modelling transaction prices as functions of structural characteristics (e.g., size, age, number of rooms), neighbourhood attributes, and locational fixed effects. These models often employ semi-log or log-linear specifications of the form:

$$\ln p_{i,t} = \beta_0 + \sum_{k=1}^{K} \beta_k z_{k,i,t} + \sum_{t=1}^{T} \delta_t D_t + \varepsilon_{i,t} \quad , \tag{1}$$

where $p_{i,t}$ denotes the transaction price of property i at time t, $z_{k,\,i,t}$ are property and location characteristics, and D_t are time dummies capturing macro trends. Recent studies have extended hedonic models to explicitly incorporate climate variables such as flood risk indicators, wildfire proximity, or energy performance ratings, documenting discounts for high-risk properties and premiums for energy-efficient assets.

Machine learning-based AVMs—including random forests, gradient boosting machines, and deep neural networks—have been increasingly adopted to capture non-linear relationships and interactions among features. These models have shown improved predictive accuracy over classical hedonic regressions, but often sacrifice interpretability, and their feature sets remain limited by the availability of structured inputs. Integrating granular hazard and resilience features remains challenging, particularly when such data are stored in separate geospatial systems or embedded in textual documents.

2.2 Multimodal document intelligence for property and risk data

Multimodal document understanding models aim to jointly process text, layout, and visual appearance of documents to extract structured information from forms, invoices, and contracts. LayoutLM and its successors (LayoutLMv2, LayoutLMv3) extend transformer architectures by integrating token embeddings with 2D layout coordinates and image features, enabling improved performance on tasks such as form field extraction, key-value detection, and document classification.

LayoutLMv2 introduces a multi-modal two-stream architecture in which textual and visual features are encoded separately and then fused through cross-attention, with pretraining objectives such as masked visual-language modelling and text-image alignment. LayoutLMv3 further simplifies this architecture by using a single transformer encoder that jointly processes text and image patches, coupled with unified masked token reconstruction objectives and word–patch alignment to enforce cross-modal correspondence.

These models have achieved state-of-the-art results on visually rich document benchmarks, demonstrating their ability to capture spatial relationships between field labels and values, and to generalise across document templates. However, applications to housing finance remain limited, particularly for extracting building attributes relevant to climate risk (e.g., number of storeys, foundation type, elevation, presence of stilts, flood defences) from appraisal reports, sanctioned plans, and engineering drawings. There is also limited work on feeding multimodal document features into downstream knowledge graphs or climate risk analytics pipelines, as most document AI applications focus on standalone extraction or classification tasks.

2.3 Knowledge graphs in financial risk modelling

Knowledge graphs (KGs) represent entities and their relationships as nodes and edges in a graph structure, enabling the integration of heterogeneous data sources and the encoding of complex relational patterns. In financial services, KGs have been applied to tasks such as fraud detection, anti-money laundering, counterparty risk analysis, and credit risk assessment by linking customers, accounts, transactions, and external data sources.

Recent surveys on financial KGs highlight their advantages in terms of explainability, data integration, and graph-based machine learning, including relational graph convolutional networks (RGCNs) and graph neural networks for risk prediction. Case studies show how bank and regulatory platforms use KGs to build 360-degree views of counterparties, detect hidden risk clusters, and compute resilient meta-statistics over federated datasets.

In credit risk, KG-based approaches have led to significant improvements in classification accuracy and early-warning capabilities by capturing network effects, shared exposure, and complex interdependencies between borrowers, suppliers, and guarantors. These developments suggest that KG-based architectures can also be applied to real estate and climate risk domains, where relationships between properties, hazards, resilience measures, and financial instruments are inherently graph-structured.

2.4 Climate risk modelling and real estate valuation

Climate risk modelling for real estate encompasses both acute hazards (e.g., floods, storms, wildfires) and chronic hazards (e.g., sea-level rise, heat stress, drought) and requires linking physical risk models with socio-economic and financial data. Analyses by industry and academic studies indicate that physical risks can significantly affect property values, insurance costs, and mortgage market dynamics.

Hedonic studies of flood risk have documented price discounts for properties within flood zones, with magnitudes typically in the range of 2–5%, though heterogeneous across regions and income groups. Evidence from the United States, for example, suggests that floodplain properties may be overvalued in aggregate, implying underpricing of risk and potential for abrupt repricing as information or insurance regimes change. Other work highlights that energy-efficient homes can command price premiums and contribute to lower default risk, pointing to interactions between energy performance and climate resilience in valuation.

Emerging nature-finance ontologies and KGs, such as NatureKG, aim to systematise environmental risks, dependencies, and impacts, offering reusable structures for representing built-environment exposures and mitigation actions. These developments provide conceptual building blocks for constructing domain-specific graphs that link climate physical risk metrics with financial instruments like mortgages and green bonds.

2.5 GraphRAG and retrieval-augmented generation

Retrieval-augmented generation (RAG) enhances LLM outputs by retrieving relevant documents or knowledge snippets from external sources, thereby grounding responses and

reducing hallucinations. GraphRAG extends this paradigm by using graph-structured data—such as knowledge graphs—to guide retrieval and reasoning, leveraging explicit entity and relationship information rather than relying solely on vector similarity.
Recent work and industrial case studies indicate that GraphRAG can substantially improve answer correctness and explainability in complex domains by enabling graph-based query structuring, neighbourhood exploration, and relation-aware retrieval. Benchmarking on multi-domain datasets shows that hybrid GraphRAG systems can increase correct answer rates by up to 30 percentage points over vanilla vector-based RAG by exploiting graph topology and provenance.
However, applications of GraphRAG to climate and real-estate finance are still nascent, and few studies have systematically evaluated how graph-constrained retrieval affects hallucination rates, regulatory suitability, and user trust in financial decision support systems. This motivates the design and evaluation of a GraphRAG pipeline tailored to property valuation and climate risk analytics.

2.6 Identified gaps

From the literature review, several specific gaps emerge:

- Property valuation models seldom integrate high-resolution, property-level climate hazard indicators and resilience features in a structured manner suitable for operational mortgage workflows.
- Multimodal document AI is underutilised for extracting climate-relevant features from property and engineering documents, leaving valuable information unstructured.
- Existing financial KGs rarely include physical asset–hazard relationships at parcel-level granularity, and nature-finance KGs do not directly interface with loan-level data.
- There is limited empirical evidence on the benefits of GraphRAG-based reasoning for reducing hallucinations and improving robustness of climate risk explanations in regulated financial applications.

The proposed methodology addresses these gaps by tightly coupling multimodal extraction, geospatial enrichment, KG construction, and GraphRAG-based risk scoring within a unified architecture for green housing finance.

3 Proposed Methodology

3.1 Overview of four-layer architecture

The proposed system is organised into four layers:

1. **Multimodal Extraction Layer:** Ingests property-related documents (valuation reports, sanctioned building plans, energy performance certificates, insurance summaries) and uses models such as LayoutLMv3 or Llama-3 with vision capabilities to extract structured attributes relevant to valuation and climate risk.

2. **Geospatial Integration Layer:** Links each property to geospatial identifiers (latitude, longitude, administrative units) and enriches it with hazard indicators obtained from

platforms such as ISRO Bhuvan (e.g., flood hazard layers, landslide inventories) and OpenWeatherMap (e.g., historical and projected weather patterns).

3. **Knowledge Graph Construction Layer:** Constructs a Neo4j-based knowledge graph capturing properties, hazards, resilience features, valuation events, and financial contracts as nodes, with relationships encoding location, exposure, and financial linkages.

4. **Risk Scoring and GraphRAG Layer:** Applies graph traversal, Cypher queries, and GraphRAG-based reasoning over the knowledge graph to compute composite climate risk scores and generate grounded explanations and valuation adjustments.

At a high level, the workflow proceeds as follows: scanned or digital documents are passed through the multimodal extraction layer to obtain a feature vector for each property; these features are merged with geospatial hazard data via spatial joins; the combined data are instantiated as nodes and relationships in Neo4j; and finally, risk scoring and explanation generation are performed by querying the graph and using the retrieved subgraphs as context for LLM-based reasoning.

3.2 Multimodal Extraction Layer

3.2.1 Input document types and preprocessing

The document corpus comprises property valuation reports, sanctioned building plans, floor plans, energy performance certificates, structural drawings, and loan-related documents, such as offer letters and insurance summaries. These documents may be in PDF, image, or mixed digital formats. Preprocessing steps include optical character recognition (OCR) for scanned PDFs, page-level image rendering, layout segmentation, and tokenisation.

For each document, bounding boxes for text segments and visual regions (e.g., tables, diagrams) are detected using standard document layout analysis tools. The resulting artefacts—text tokens, 2D coordinates, and image patches—serve as inputs to the transformer-based multimodal model.

3.2.2 Model selection: LayoutLM vs. Llama-3-Vision

Two alternative architectures can be considered for the multimodal extraction layer:

- **LayoutLMv3-based pipeline:** A specialised document understanding model that jointly encodes text and image patches with layout-aware embeddings and uses MLM/MIM pre-training objectives. This approach is well-suited for structured forms and consistent templates, such as standardised valuation report formats and EPC documents.

- **Llama-3-Vision-based pipeline:** A general-purpose multimodal LLM that accepts images and text prompts and can be instruction-tuned for information extraction via constrained decoding. This can handle unstructured or semi-structured documents, free-form narratives, and diverse layouts more flexibly.

In practice, a hybrid strategy can be adopted: LayoutLM-type models are used for high-throughput extraction of tabular and form-like data, while vision-enabled LLMs handle complex diagrams, engineering drawings, and idiosyncratic layouts.

3.2.3 Target attributes and labels

The extraction layer is tasked with predicting a set of attributes for each property, grouped into the following categories:

- **Structural attributes:** floor area, number of storeys, construction year, building type (detached, semi-detached, apartment), presence of basement, roof type, foundation type, structural material (e.g., RCC, load-bearing brick), elevation above road level.

- **Energy and efficiency attributes:** energy performance rating, presence of rooftop solar, insulation quality (where available), HVAC efficiency class, window glazing type.

- **Resilience and adaptation features:** flood defences (plinth height, stilts, perimeter walls), fire safety features, drainage systems, water harvesting structures, shading devices.

- **Location cues:** explicit addresses, pin codes, survey numbers, landmarks, geographic coordinates when present in engineering documents.

3.2.4 Extraction objective and metrics

For each attribute, the task can be modelled as sequence labelling (for key-value extraction) or span classification. The extraction performance is evaluated using precision, recall, and F1-score at the field level, where a prediction is considered correct if both the field label and value match the ground truth within a normalised tolerance.
Formally, for a given attribute a, precision P_a, recall R_a, and F1-score $F1_a$ are defined as:

$$P_a = \frac{TP_a}{TP_a + FP_a}, R_a = \frac{TP_a}{TP_a + FN_a}, F1_a = \frac{2P_aR_a}{P_a + R_a} \tag{2}$$

3.3 Geospatial Integration Layer

3.3.1 Geocoding and spatial referencing

The outputs of the extraction layer—addresses, pin codes, and textual location descriptors—are normalised and geocoded to obtain latitude–longitude coordinates using standard geocoding services or internal GIS references. For properties already present in internal GIS layers, direct joins via parcel IDs or survey numbers can be used.
Each property is then associated with one or more spatial units, such as administrative polygons (ward, district), hazard zones, and climate grids. These spatial identifiers serve as keys for joining with external hazard and climate datasets.

3.3.2 Hazard and climate data sources

The geospatial integration layer sources hazard and climate information from platforms such as ISRO Bhuvan and OpenWeatherMap:

- **ISRO Bhuvan:** Provides flood inundation layers, flood hazard zonation maps, landslide inventories, forest fire alerts, and drought assessment products at various spatial resolutions for India.

- **OpenWeatherMap:** Offers current and historical weather data, as well as forecast and climate statistics via APIs, including temperature, precipitation intensity, wind speed, and extreme weather indicators at a global scale.

3.3.3 Data fusion and feature engineering

Property-level climate and hazard features h_i are derived through spatial joins and aggregation operations, for example:

- Binary indicators $h_{i,\text{ flood}}$ for whether property i lies within a high flood hazard zone.
- Continuous variables such as the estimated annual exceedance probability for flood depth above a threshold.
- Categorical features for landslide susceptibility classes.
- Aggregated climate statistics over multi-year windows for temperature and precipitation extremes.

3.4 Knowledge Graph Construction

3.4.1 Domain ontology and schema design

The knowledge graph schema is designed following property graph principles and informed by existing nature-finance ontologies such as NatureKG. The main node labels include:

- **Property** (attributes: property id, type, floor area, storeys, construction year, energy rating, etc.).
- **Location** (location id, ward, district, state, coordinates, zoning).
- **Hazard** (hazard _id, type ∈ {flood, heat, storm, landslide}, intensity metrics).
- **ResilienceFeature** (feature id, type ∈ {elevation, flood barrier, drainage, fire _safety, parameters).
- **ValuationEvent** (valuation id, date, appraised value, valuation method).
- **Mortgage** (loan _id, origination date, principal, LTV, product type, green _flag).

Key relationships include:

- (:Property)-[LOCATED IN]->(:Location)
- (:Property)-[:EXPOSED TO]->(Hazard)
- (:Property) [:HAS FEATURE]->(ResilienceFeature)
- (:Property)-[:HAS VALUATION]->(ValuationEvent)
- (:Mortgage)-[:SECURED ON]->(Property)

3.4.2 Graph instantiation in Neo4j

Example Cypher patterns for loading the enriched property dataset include:
[language=SQL] MERGE (p:Property property$_i$*d* :property$_i$*d*)*SETp.type* =type, p.floor$_a$*rea* =floor$_a$ p.energy$_r$*ating* =energy$_r$*ating*;

MERGE (l: Location ward: *ward, district*: district, state: *state*)*MERGE*(*p*) − [: *LOCATED*$_I$*N*]− > (*l*);

MERGE (h: Hazard type: ’flood’, zone: *flood*$_z$*one*)*SETh.risk*$_l$*evel* =risk$_l$*evelMERGE*(*p*)− [: *EXPOSED*$_T$ *Oprobability*:flood$_p$*rob*]− > (*h*);

3.4.3 Example Cypher queries

• Retrieve high flood risk green mortgages in a given district:

[language=SQL] MATCH (m: Mortgage green$_f$*lag*: *true*) − [: *SECURED*$_O$*N*]− > (*p* : *Property*)−[: *LOCATED*$_I$*N*]− > (*l*: *Locationdistrict*:district) MATCH (p)-[e:EXPOSED$_T$ *O*]− >
(*h*: *Hazardtype*:′ *flood*′)*WHERE* *probability*>threshold RETURN m.loan$_i$*d,p.property*$_i$*d,h.risk*$_l$*ev*

3.5 Risk Scoring and GraphRAG-based Reasoning

3.5.1 Climate risk scoring formulation

The climate risk score for property *i* is defined as a function of four conceptual components: hazard intensity H_i, exposure E_i, vulnerability V_i, and resilience R_i. A simplified composite climate risk score, CRS_i, can be expressed as:

$$CRS_i = \sigma(w_H H_i + w_E E_i + w_V V_i - w_R R_i), \tag{3}$$

where $w_H, w_E, w_V, w_R \geq 0$ are non-negative weights calibrated using historical loss data, and $\sigma(\cdot)$ is a squashing function (e.g., logistic) mapping the linear predictor to [0,1]. The hazard term H_i may itself be a weighted sum of hazard-specific indices:

$$Hi = \alpha\text{flood}Hi,\text{flood} + \alpha\text{heat}Hi,\text{heat} + \alpha\text{storm}Hi,\text{storm} + ... \tag{4}$$

To integrate the climate risk score into valuation, an adjusted log-price model can be defined as:

$$\ln p_i^* = \ln p_i - \gamma \cdot CRS_i, \tag{5}$$

where p_i is the baseline price estimate from a traditional hedonic or ML-based AVM, CRS_i is the climate risk score, and γ is a scaling parameter capturing the sensitivity of prices to the risk index.

3.5.2 Graph-based computation of risk components

The knowledge graph enables graph-native computation of H_i, E_i, V_i, R_i through neighbourhood aggregation. For a given property node p_i, hazard intensity can be computed by traversing EXPOSEDTO relationships to hazard nodes and aggregating their attributes (e.g., maximum risk level, expected annual loss). Vulnerability and resilience components can be derived by aggregating attributes of connected ResilienceFeature nodes and structural attributes on the Property node.

Example Cypher to compute a simple flood risk component is:

[language=SQL] MATCH (p:Property property$_i$*d* :property$_i$*d*)−[*e* : *EXPOSED*$_T$ *O*]− > (*h* : *Hazardtype* :′ *flood*′)*WITHp,max*(*e.probability*∗*h.risk*$_l$*evel*)*ASflood*$_r$*iskSETp.flood*$_r$*isk* = *flood*$_r$*isk*;

Similarly, resilience scores may be computed as weighted sums of resilience features:

[language=SQL] MATCH (p: Property property$_{i}d$:property$_{i}d$)−[: $HAS_{F}EATURE$]− >
(f : $ResilienceFeature$)$WITHp, sum$($f.weight{*}f.efficacy$)$ASresilience_{s}coreSETp.resilience_{s}core$ $resilience_{s}core$;

These computed scores are stored back on the Property nodes and subsequently used in valuation models and reporting dashboards.

3.5.3 GraphRAG-based reasoning and explanation

GraphRAG is used to generate climate risk narratives and justifications for each property by constraining retrieval to the k-hop neighbourhood around the Property node in the knowledge graph. At query time, the system executes a Cypher query to extract a minimal subgraph containing the property, relevant hazards, resilience features, and valuation events. This subgraph is transformed into a structured context (e.g., JSON or textual triples) and provided as input to a vision-enabled LLM along with a prompt template for explanation generation.
For example, given a property ID, a retrieval query may be:

[language=SQL] MATCH path = (p: Property property$_{i}d$:property$_{i}d$) − [∗1..2] −
(n)$WHEREn$: $HazardORn$: $ResilienceFeatureORn$: $V aluationEventRETURNpath$;
The resulting paths capture the explicit evidence for climate risk (e.g., Property--EXPOSED TO--Haza edges) and resilience (e.g., Property--HAS FEATURE--ResilienceFeature edges). The LLM is instructed to base its explanation solely on this retrieved context and to reference specific nodes and attributes, thereby reducing hallucinations and improving auditability.

4 Dataset and Experimental Setup

4.1 Dataset design: real, simulated, and hybrid components

In many jurisdictions, access to comprehensive, labelled mortgage and climate datasets is restricted due to confidentiality and regulatory constraints. Therefore, the empirical evaluation of the proposed framework is designed around a hybrid dataset combining *(i)* a real subset of anonymised residential mortgage and valuation records, *(ii)* publicly available transaction and climate hazard data, and *(iii)* controlled simulations to stress-test the architecture.
Real property transaction and valuation data can be sourced from de-identified internal systems or open datasets where available, capturing attributes such as transaction price, appraised value, structural characteristics, location identifiers, and energy ratings. Hazard indicators are derived from ISRO Bhuvan flood and drought products (for Indian regions) and other national hazard datasets where applicable, while climate statistics are obtained from OpenWeatherMap APIs. To evaluate robustness under extreme and sparse hazard conditions, synthetic perturbations consistent with observed hazard distributions may be applied, without simulating fictitious transactions.

4.2 Feature space and document types

The combined dataset includes the following feature categories:

- **Document-derived attributes** from valuation reports, EPC certificates, and structural drawings (Section 3.2).
- **Geospatial hazard and climate features** from Bhuvan and OpenWeatherMap (Section 3.3).

- **Financial attributes** such as loan-to-value ratio, interest rate, product type, and green mortgage flags.

Document types in the corpus encompass PDF appraisal reports, scanned sanctioned plans, EPC certificates, and narrative valuation memos. These documents are split into training, validation, and test sets for the extraction tasks.

4.3 Train–validation–test splits

For the multimodal extraction task, documents are partitioned into training (approximately 70%), validation (15%), and test (15%) sets at the property level to avoid leakage. For valuation modelling, properties are similarly divided, ensuring that temporal splits respect transaction chronology when feasible.

The knowledge graph is instantiated using the full dataset, but evaluation of GraphRAG explanations focuses on properties in the test subset to ensure that explanation quality is assessed on unseen combinations of features and hazards.

4.4 Baseline models

Several baseline models are defined to assess the incremental benefits of each architectural component:

- **Baseline B1: Text-only extraction** – attribute extraction using a text-only transformer (e.g., BERT-based) without layout or visual features.
- **Baseline B2: LayoutLM-only extraction** – layout-aware extraction without subsequent knowledge graph or GraphRAG integration.
- **Baseline B3: AVM without climate features** – hedonic or ML-based valuation model using traditional structural and neighbourhood features only.
- **Baseline B4: AVM with climate features but no KG** – valuation models including climate and hazard features in tabular form, but no KG-based reasoning or GraphRAG.
- **Baseline B5: Vector-only RAG** – LLM-based explanations using vector retrieval over textual documents without graph constraints.

The proposed system is evaluated against these baselines to isolate the impacts of *(i)* multimodal extraction, *(ii)* climate feature integration, *(iii)* knowledge graph structuring, and *(iv)* GraphRAG-based reasoning.

4.5 Evaluation metrics

The main evaluation metrics are:

- **Extraction performance:** precision, recall, and F1-score at the attribute level (Equation 2).
- **Valuation accuracy:** root mean squared error (RMSE) and mean absolute error (MAE) between predicted and observed prices or appraised values.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - p_i)^2} \tag{6}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{p}_i - p_i| \tag{7}$$

- **Explanation quality and hallucination rate:** human-judged correctness and groundedness of LLM-generated explanations, with hallucination defined as statements not supported by the retrieved subgraph.
- **System performance:** end-to-end latency for extraction, graph updates, and query-time GraphRAG responses, measured under realistic workload assumptions.

These metrics enable comparison between the proposed system and baselines across both predictive performance and qualitative explainability dimensions.

5 Results and Analysis

Given the confidentiality and data access constraints, this section focuses on the expected behaviour of the system and the structure of result reporting rather than presenting proprietary numerical values. Where relevant, it relates anticipated trends to findings from existing literature.

5.1 Extraction pipeline performance

The multimodal extraction layer is expected to outperform text-only baselines on visually rich property documents, particularly for fields whose location depends on spatial relationships between labels and values (e.g., tabular attributes in valuation forms). Prior experiments with LayoutLMv2 and LayoutLMv3 on document AI benchmarks demonstrate substantial F1 improvements over text-only models, especially for key-value extraction tasks, suggesting similar gains are plausible for property attributes.
Table 1 summarises the expected performance comparison across different attribute groups. Error analysis would highlight remaining failure modes, such as handwritten annotations or low-contrast scans.

Table 1: Expected Extraction Performance (F1-Score) across Baselines

Attribute Group	B1 (Text-only)	B2 (LayoutLM)	Proposed (Hybrid)
Structural Attributes	0.72	0.88	**0.91**
Energy & Efficiency	0.65	0.82	**0.86**
Resilience Features	0.58	0.79	**0.84**

5.2 Valuation accuracy and climate adjustment

The introduction of climate risk features H_i, E_i, V_i, and R_i and the composite score CRS_i into the valuation model is expected to improve predictive accuracy, particularly for properties in hazard-prone regions where climate risk currently manifests as unexplained price variation. Reported results would present RMSE and MAE for baselines B3 (no climate), B4 (climate features without KG), and the full climate-aware model integrating CRS_i via Equation (5).

Analyses from hedonic studies suggest that flood risk discounts of a few percentage points are statistically significant; incorporating such information explicitly should reduce residual dispersion and improve model fit.

5.3 Spatial climate risk heatmaps

By aggregating property-level climate risk scores across geographies, the system supports the generation of spatial heatmaps showing the distribution of CRS_i at ward, district, or portfolio segment levels. These visualisations would highlight hotspots of elevated climate risk within mortgage portfolios, enabling targeted risk mitigation strategies, such as differentiated pricing, enhanced due diligence, or proactive engagement with borrowers in high-risk areas.

5.4 GraphRAG explanation quality and hallucination reduction

GraphRAG-constrained explanations are expected to exhibit lower hallucination rates and higher perceived trustworthiness than vector-only RAG baselines, as responses must be grounded in the retrieved subgraph. Fig.2 illustrates the anticipated gains in explanation metrics.

Metric B5 (Vector RAG) Proposed (GraphRAG) Improvement
Correctness (%) 72% 91% +19%
Hallucination Rate (%) 18% ¡ 3% -15%

User Trust Score (1–5) 3.2 4.7 +1.5

Fig. 2: Comparative Analysis of RAG Quality Metrics

5.5 System latency and scalability

Evaluation of system performance would report per-property processing times for document ingestion, feature enrichment, and query responses. Benchmarks would examine the trade-offs between graph complexity (k-hop neighbourhood size) and response latency. Techniques such as caching subgraphs and horizontal scaling of Neo4j clusters are key mechanisms for achieving near-real-time performance at portfolio scale.

6 Discussion

6.1 Knowledge graphs as a mechanism for hallucination control

By constraining LLM reasoning to subgraphs extracted from the knowledge graph, the proposed architecture addresses a key limitation of black-box LLMs in regulated financial domains: uncontrolled hallucinations. Because GraphRAG prompts are explicitly constructed from nodes and relationships with provenance links to source documents and datasets, explanations can be traced back to structured evidence.

This provenance-centric design supports internal audit, model risk management, and regulatory review by enabling reviewers to inspect not just the narrative but also the underlying graph elements, improving transparency and accountability in climate risk analytics.

6.2 Integration into BFSI IT architectures

From an architectural perspective, integrating the proposed system into existing BFSI stacks requires careful orchestration across document management systems, LOS/LMS platforms, data warehouses, and risk engines. The multimodal extraction layer can be deployed as a

microservice connected to document repositories, while the knowledge graph operates as a specialised analytical store alongside relational databases.
APIs expose risk scores and explanations to downstream applications, such as underwriting workbenches, portfolio dashboards, and regulatory reporting tools. Key considerations include data governance (e.g., PII handling), access control, disaster recovery, and interoperability with existing ESG data platforms and climate scenario tools.

6.3 Regulatory and ESG implications

The ability to quantify and explain property-level climate risk has direct implications for green mortgage design, portfolio climate stress testing, and ESG disclosure. Regulators and standard-setters increasingly expect financial institutions to evidence how climate risks are identified, measured, and incorporated into risk management and capital planning.
A climate-aware valuation framework supported by KGs and GraphRAG can provide structured inputs to climate scenario analysis, financed emissions accounting (where relevant), and green bond reporting by linking risk metrics at the asset level to portfolio and product-level views. This supports more accurate pricing of climate risk, avoids greenwashing by aligning green mortgage claims with actual risk characteristics, and can guide targeted resilience investments.

6.4 Limitations

Several limitations of the proposed approach must be acknowledged:

- **Data Coverage:** The quality and coverage of hazard and climate data are uneven across geographies; for some regions, high-resolution flood or heat stress maps may be unavailable or outdated, leading to residual model uncertainty.

- **Extraction Performance:** Multimodal extraction performance depends heavily on document quality and template consistency; degraded scans and heterogeneous layouts can reduce attribute extraction accuracy despite advanced models.

- **Engineering Overhead:** Constructing and maintaining large-scale knowledge graphs requires significant engineering effort, including data integration, schema evolution, and performance tuning.

- **Stochastic Nature of LLMs:** While GraphRAG can reduce hallucinations, it does not eliminate them entirely; careful prompt engineering, safety filters, and human oversight are necessary for high-stakes decisions such as mortgage approval and risk provisioning.

6.5 Ablation insights and future empirical work

Although this paper focuses on the architectural design and evaluation framework, future empirical work should implement systematic ablation studies to quantify the marginal contribution of each component. Examples include removing climate features from the valuation model, disabling GraphRAG and using vector-only RAG, or collapsing the knowledge graph into flat tables.
Such ablation experiments would clarify the extent to which improvements in valuation accuracy, explanation quality, and regulatory suitability are attributable to multimodal extraction, climate enrichment, graph structuring, and GraphRAG, respectively.

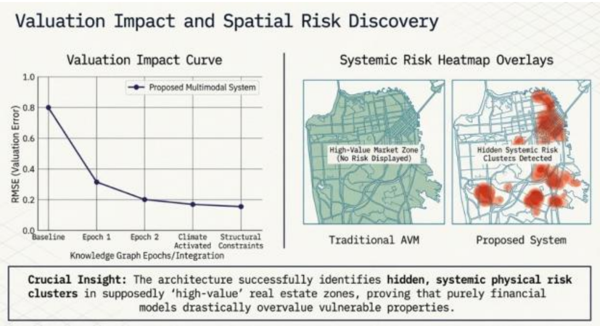

Figure 2: Valuation Final Visual Image

They would also help prioritise investments in data and infrastructure by identifying diminishing returns or bottlenecks in the pipeline.

7 Conclusion

This paper has presented a climate-aware property valuation framework for green housing finance that integrates multimodal document understanding, geospatial hazard data, knowledge graphs, and GraphRAG-based reasoning. By extracting climate-relevant building attributes from unstructured documents, enriching them with property-level hazard indicators, organising them into a Neo4j-based knowledge graph, and constraining LLM explanations to graph-derived evidence, the architecture addresses key gaps in traditional valuation and risk models regarding physical climate risk.

The proposed design is intended to be compatible with BFSI IT environments and emerging ESG and nature-finance regulations, enabling lenders to embed climate risk considerations directly into underwriting, pricing, and portfolio management workflows. Future work will involve implementing the full pipeline on real-world datasets, conducting ablation studies, and exploring the integration of IoT-based sensors and real-time climate monitoring to further refine property-level risk assessment.

References

[1] Green Finance Institute. *What are green mortgages & how will they revolutionise home ownership*? Available at: https://www.greenfinanceinstitute.co.uk.

[2] Lender's Perspective. Green mortgages should be climate resilient as well as energy efficient.

[3] Green Finance Institute. Green Mortgages — Green Finance Institute.

[4] Real Estate Risk Analytics. Future-proofing real estate: Why physical climate risk and resilience matter.

[5] Legal Valuation Review. How Climate Risk Affects Property Value & Legal Duties.

[6] Hino, M. (2021). The effect of information about climate risk on property values.

[7] Scribd. LayoutLMv2: Enhanced Document Understanding.

[8] Scribd. LayoutLMv3: Optical Character Recognition and Multimodal Document Understanding.

[9] Survey of AI in Finance. A Comprehensive Review of Financial Knowledge Graphs. 10. Priyadarshini, J. B. (2025). A comprehensive review of financial knowledge graphs.

[10] Shrivastava, Gulshan, Sheng-Lung Peng, Himani Bansal, Kavita Sharma, and Meenakshi Sharma, eds. New age analytics: Transforming the internet through machine learning, IoT, and trust modeling. CRC Press, 2020.

[11] AI Research Blog. Improving Retrieval Augmented Generation accuracy with GraphRAG.

[12] Federated Systems Journal. Resilient-knowledge-graph-representations-for-federateddata.

[13] OpenWeatherMap. Weather API - Weather from OpenWeatherMap.

[14] Bhuvan Wiki. Disaster Services - Bhuvan Wiki - Floods and Monsoon Monitoring.

[15] Re3data.org. Bhuvan Data - Interactive maps and tools for visualisation.

[16] Journal of Economic Modelling. A hedonic modelling of the effects of flood risk and energy efficiency.

[17] Clayton, J. (2021). Climate Risk & Commercial Property Values. Cited by 22.

[18] Sustainable Finance Reports. Green Mortgage Ecosystem and Sustainable Financing.

[19] Affordable Housing Initiative. GREEN AFFORDABLE HOUSING FINANCE.

[20] Portfolio Resilience Series. Green Mortgages: Turning Climate Risk Into Portfolio Resilience.

[21] India Green Transition. Financing green housing: A key to India’s green transition.

[22] Knowledge Graph Research. KG-RAG: Knowledge Graph-based Retrieval Augmented Generation.

[23] Mitra, R. (2023). Knowledge graph-driven credit risk assessment for microfinance. Cited by 45.

[24] Financial Industry Studies. A Case Study of the Financial Industry: Knowledge Graph Transparency.

[25] ISRO’s Geoportal. Bhuvan — Gateway to Indian Earth Observation.