Vượt qua các LLM tổng quát

Xây dựng hệ thống AI đáng tin cậy, chuyên biệt cho lĩnh vực Tài chính và Pháp lý tại Việt Nam



TS. Lê Xuân Bách

lexuanbach@hcmut.edu.vn

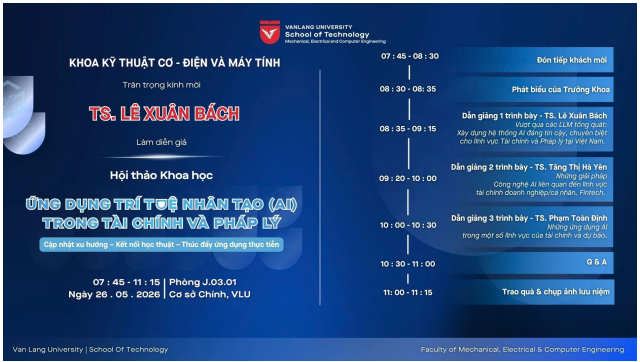
Faculty of Computer Science and Engineering, HCMUT
Vietnam National University Ho Chi Minh City

Van Lang University, School of Technology
26.05.2026, 08:35–09:15, Phòng J.03.01

Nội dung

1. Nhu cầu và cơ hội
2. Bài toán AI đáng tin
3. RAG và AI pháp lý
4. AI tài chính: từ dự báo đến quyết định
5. Ví dụ mở rộng
6. Tổng hợp và kết luận

Bối cảnh buổi nói chuyện



Thông tin từ thư mời

- Hội thảo: **Ứng dụng Trí tuệ Nhân tạo trong Tài chính và Pháp lý**
- Thời gian trình bày: **08:35–09:15**
- Địa điểm: **Phòng J.03.01**, Cơ sở chính VLU

Mục tiêu

Kết nối nghiên cứu gần đây với nhu cầu xây dựng AI có **bằng chứng**, **cấu trúc miền**, và **kiểm soát rủi ro**.

Nhu cầu của nhà trường và cơ hội AI chuyên biệt

Nhu cầu từ các thầy cô

- Có bài toán thật để gắn nghiên cứu, đào tạo và đồ án sinh viên.
- Pháp lý, tài chính, y tế, giáo dục đều cần hệ thống biết ngữ cảnh Việt Nam.
- Cần benchmark, dữ liệu và quy trình đánh giá phù hợp tiếng Việt.



Cơ hội hợp tác

Chuyên gia miền có thể thẩm định dữ liệu, đặt bài toán đúng, và kiểm tra hệ thống trước khi đưa vào vận hành.

Thông điệp mở đầu

Cơ hội không chỉ nằm ở LLM rất lớn, mà ở **dữ liệu miền**, **đánh giá đúng**, và **tích hợp hệ thống** quanh nhu cầu thật.

† Nhận định tổng hợp từ các case studies tiếng Việt/Việt Nam trong deck: SBV-LawGraph, ATM cash management, VSLIM và các survey liên quan.

Lê Xuân Bách: tiểu sử ngắn



TS. Lê Xuân Bách
Faculty of Computer Science and Engineering
HCMUT, VNU-HCM

Hiện tại

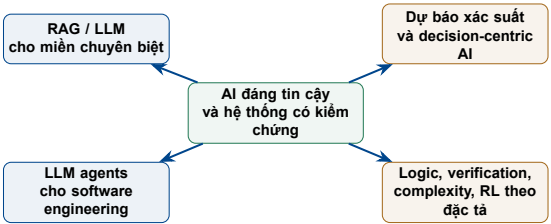
Giảng viên tại Trường Đại học Bách khoa, ĐHQG-HCM; nghiên cứu ở giao điểm giữa **AI đáng tin cậy**, **hệ thống phần mềm**, và **miền ứng dụng rủi ro cao**.

Nền tảng học thuật

- Song bằng **Computer Science** và **Pure Mathematics**, NUS, 2007–2012.
- PhD in Computer Science, NUS, 2013–2017; sau đó làm postdoctoral researcher tại NUS, NTU, Oxford và SMU.
- Trục xuyên suốt: từ logic/verification đến LLM/RAG để xây dựng hệ thống **có thể tin** và **có thể kiểm tra**.

† Tóm tắt từ website cá nhân và CV công khai của Lê Xuân Bách.

Các hướng nghiên cứu chính



Ứng dụng gần đây

Legal AI tiếng Việt; tài chính vận hành với time series forecasting; NLP tiếng Việt cho intent, slot và event extraction.

Nền tảng phương pháp

Formal methods, program verification/synthesis; LLM agents cho static analysis; reinforcement learning với đặc tả logic.

† Tổng hợp từ mục Research Interest trên website cá nhân và các công bố 2024–2026.

Nội dung

1 Nhu cầu và cơ hội

2 Bài toán AI đáng tin

3 RAG và AI pháp lý

4 AI tài chính: từ dự báo đến quyết định

5 Ví dụ mở rộng

6 Tổng hợp và kết luận

Một số công trình tiêu biểu

AI ứng dụng và AI chuyên biệt

- **SBV-LawGraph** – Hybrid RAG + Legal Knowledge Graph cho văn bản pháp lý Ngân hàng Nhà nước Việt Nam, **ACIIDS 2026 Oral**.
- **Uncertainty-aware ATM cash management** – dự báo xác suất + tối ưu tồn quỹ ATM tại Việt Nam, **FLINS-ISKE 2026**.
- **VSLIM** – multi-intent detection và slot filling cho tiếng Việt, **ACIIDS 2026**.

Nền tảng tin cậy và kiểm chứng

- **RL với LTL/ω-regular objectives** – optimality-preserving translation sang average rewards, **NeurIPS 2024**.
- **Quantum separation logic** – local reasoning cho quantum programs, **POPL 2022**.
- **Verification / automata / complexity** – rely-guarantee inference, monadic decomposability, tree-share permissions.

Thông điệp cho buổi nói chuyện

Điểm chung: AI hữu ích trong miền rủi ro cao cần kết hợp **mô hình**, **tri thức miền**, **kiểm chứng**, và **đánh giá theo hậu quả quyết định**.

† Nguồn: publications.yml trên website cá nhân; các công trình tiêu biểu tại ACIIDS/FLINS-ISKE 2026, NeurIPS 2024, POPL 2022.

Vấn đề nghiên cứu

Trong pháp lý

Nếu hệ thống nói một quy định còn hiệu lực nhưng văn bản đã bị sửa đổi, câu trả lời **nghe đúng** có thể dẫn đến quyết định sai.

Trong tài chính

Nếu dự báo sai phần đuôi nhu cầu rút tiền, ATM có thể hết tiền đúng lúc khách hàng cần, dù MAE/MASE nhìn rất đẹp.



Khái niệm trung tâm của buổi nói chuyện

AI đáng tin cậy cần **dựa vào bằng chứng**, **hiểu cấu trúc miền**, và **được đo bằng hậu quả quyết định**.

Xu hướng nghiên cứu gần đây

Từ mô hình sang hệ thống

- **Trí tuệ nhân tạo tác tử**: lập kế hoạch, gọi công cụ, tự kiểm tra bước trung gian.
- **Truy xuất–sinh lai với đồ thị**: kết hợp tìm kiếm ngữ nghĩa, từ khóa, đồ thị tri thức và suy luận nhiều bước.
- **Đa phương thức và dữ liệu có cấu trúc**: văn bản, bảng, ảnh, nhật ký, giao dịch, đồ thị pháp lý.

Từ điểm chuẩn sang vận hành

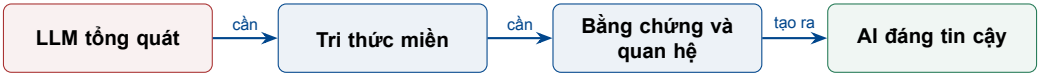
- Đánh giá theo tác vụ, trích dẫn, độ tin cậy, độ trễ và chi phí.
- Thích nghi theo miền thay vì chỉ nhắc lệnh mô hình ngôn ngữ lớn tổng quát.
- Giám sát, nhật ký kiểm toán, chuyên gia trong vòng lặp và chính sách từ chối.

Điểm chung

Nghiên cứu trí tuệ nhân tạo đang dịch từ “mô hình trả lời hay” sang **hệ thống có bằng chứng, công cụ, bộ nhớ miền, và cơ chế kiểm toán**.

† Báo cáo Chỉ số Trí tuệ nhân tạo Stanford HAI 2025; các khảo sát năm 2025 về truy xuất–sinh trên đồ thị, truy xuất–sinh hướng tri thức và truy xuất–sinh tác tử/đa phương thức; các nghiên cứu tình huống trong bộ slide này.

Luận điểm chính



Thông điệp

Trong các lĩnh vực có rủi ro cao, AI không nên chỉ là một mô hình trả lời hay. Nó phải là một **hệ thống ra quyết định** có dữ liệu, cấu trúc, kiểm chứng, và cơ chế từ chối khi thiếu bằng chứng.

Pháp lý

Trích dẫn đúng văn bản, điều khoản, quan hệ sửa đổi/bãi bỏ.

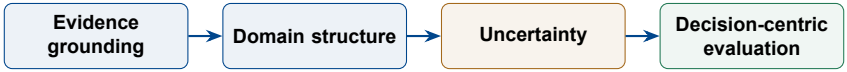
Tài chính

Tối ưu chi phí nhưng vẫn kiểm soát rủi ro vận hành.

Kỹ thuật

Đánh giá bằng tác động cuối, không chỉ bằng benchmark chung.

Điều kiện để tin một hệ thống AI



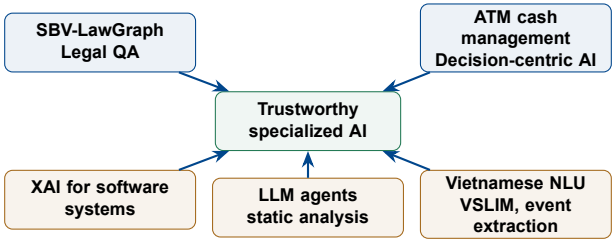
Không đủ

- Câu trả lời nghe hợp lý.
- Điểm retrieval cao nhưng không trích dẫn được.
- Sai số dự báo thấp nhưng quyết định tốn kém.

Cần thêm

- Nguồn chứng cứ có thể kiểm tra.
- Cấu trúc quan hệ của văn bản/quy trình.
- Độ bất định và cơ chế từ chối.
- Đo bằng chi phí, rủi ro, và tác động.

Bản đồ nghiên cứu gần đây



Cách đọc bản đồ

Hai nghiên cứu neo của buổi nói chuyện là **pháp lý ngân hàng** và **quản lý tiền mặt ATM**. Các nhánh còn lại bổ sung cùng một chủ đề: AI phải được thiết kế theo miền, có kiểm chứng và có trách nhiệm.

† Nguồn công bố gần đây: SBV-LawGraph (ACIIDS 2026); ATM cash management (FLINS-ISKE 2026); VSLIM/XAI/LLM agents/event extraction (ACIIDS/IEA-AIE 2026).

Nội dung

1. Nhu cầu và cơ hội
2. Bài toán AI đáng tin
3. RAG và AI pháp lý
4. AI tài chính: từ dự báo đến quyết định
5. Ví dụ mở rộng
6. Tổng hợp và kết luận

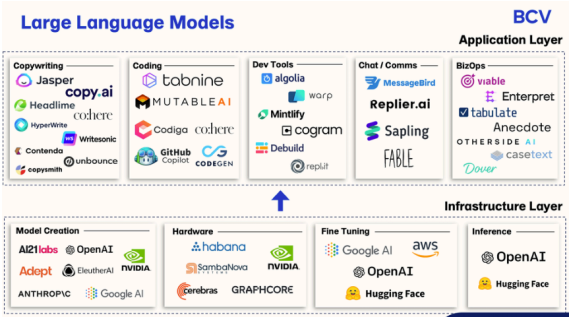
LLM và hallucination

LLM là gì?

- LLM học từ lượng lớn văn bản để dự đoán chuỗi token tiếp theo.
- Nó rất mạnh trong diễn đạt, tóm tắt, suy luận mềm và sinh nội dung.
- Nhưng LLM không tự nhiên là một cơ sở dữ liệu pháp lý hay tài chính.

Hallucination

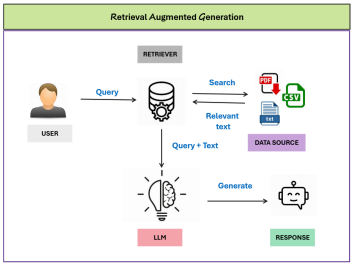
Câu trả lời có thể **trôi chảy** nhưng không có căn cứ, trích dẫn sai, hoặc ghép thông tin không tồn tại.



† Brown et al., NeurIPS 2020; Ji et al., ACM Computing Surveys 2023. Hình minh họa: file trong folder talk.

RAG: đưa bằng chứng vào vòng trả lời





Ý tưởng của RAG

Thay vì yêu cầu LLM nhớ mọi thứ, hệ thống **truy xuất tài liệu liên quan**, đưa bằng chứng vào prompt, rồi buộc câu trả lời bám vào nguồn.

Trong pháp lý

RAG chuẩn là bước đầu; miền pháp lý còn cần quan hệ như sửa đổi, thay thế, bãi bỏ và hiệu lực văn bản.

† Lewis et al., NeurIPS 2020; Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026. Hình minh họa: file trong folder talk.

Vì sao Legal QA tiếng Việt khó?

Đặc thù văn bản pháp lý

- Văn bản có cấu trúc chặt: luật, nghị định, thông tư, điều, khoản.
- Quan hệ liên văn bản: sửa đổi, thay thế, bãi bỏ, hướng dẫn.
- Tính hiệu lực thay đổi theo thời gian.
- Câu trả lời cần trích dẫn và giải thích được.

Thách thức tiếng Việt

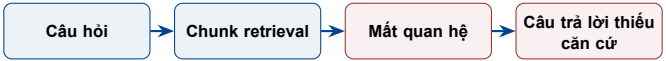
- Tách từ, dấu, cách diễn đạt vùng miền.
- Dữ liệu pháp lý chuyên ngành ít hơn tiếng Anh.
- Một câu hỏi thường cần nhiều văn bản liên quan.
- Sai một căn cứ pháp lý có thể gây hậu quả thật.

Bài toán

Không chỉ tìm đoạn văn giống câu hỏi; hệ thống phải hiểu cả **nội dung** và **quan hệ pháp lý**.

† Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026.

RAG chuẩn thất bại ở đâu?



RAG chuẩn làm tốt

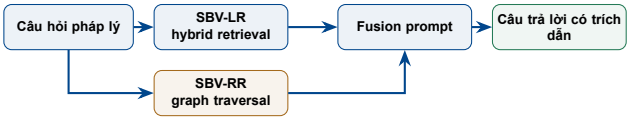
- Tìm đoạn văn liên quan theo từ khóa/ngữ nghĩa.
- Giảm hallucination so với LLM thuần.
- Dễ triển khai bước đầu.

Nhưng còn thiếu

- Không biết văn bản nào đã bị sửa đổi.
- Không theo được chuỗi viện dẫn.
- Không tự kiểm tra trích dẫn có đủ bằng chứng.

† Lewis et al., NeurIPS 2020; Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026.

SBV-LawGraph: ý tưởng hệ thống



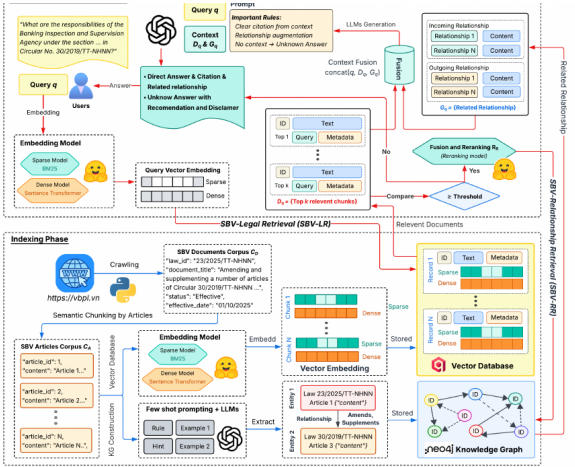


Thiết kế cốt lõi

Kết hợp **sparse–dense reranking** với **Legal Knowledge Graph**: retrieval tìm đúng nội dung, graph tìm đúng quan hệ pháp lý. Khi thiếu bằng chứng, hệ thống trả lời **Unknown Answer** thay vì đoán.

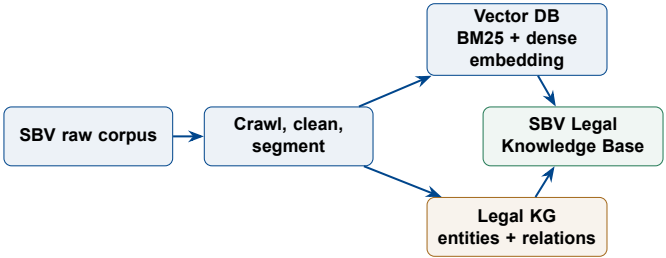
† Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026.

SBV-LawGraph: hình hệ thống từ paper



† Cắt từ Fig. 1 trong Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026.

SBV-LawGraph: offline indexing



Vector side

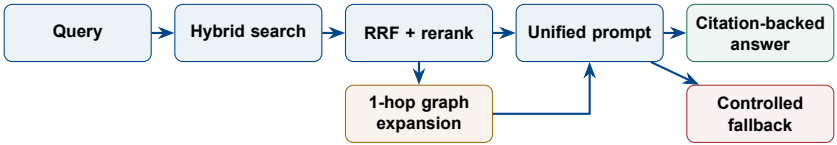
- Article-level chunks.
- BM25 cho từ khóa; legal embeddings cho ngữ nghĩa.

Graph side

- Node: văn bản, điều, thực thể pháp lý.
- Edge: sửa đổi, bãi bỏ, thay thế, hướng dẫn.

† Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026.

SBV-LawGraph: online retrieval và generation

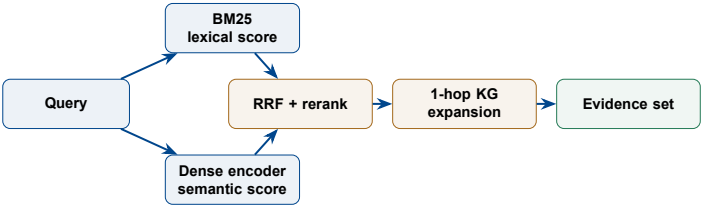


Nguyên tắc sinh câu trả lời

Mỗi khẳng định quan trọng phải được hỗ trợ bởi văn bản hoặc quan hệ trong graph. Nếu câu trả lời không có trích dẫn phù hợp hoặc mâu thuẫn bằng chứng, hệ thống không trả lời bừa.

† Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026.

Kỹ thuật SBV-LawGraph: fusion retrieval



Text signal

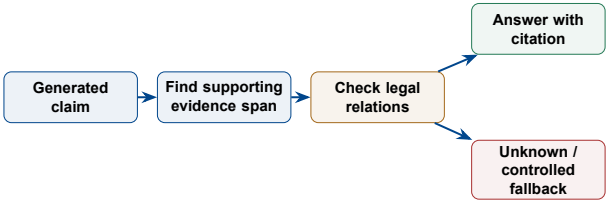
- BM25 giữ matching theo thuật ngữ pháp lý.
- Dense embedding bắt paraphrase và câu hỏi tự nhiên.
- Rerank trộn hai tín hiệu trước prompt.

Graph signal

- Mở rộng theo sửa đổi, thay thế, bãi bỏ, viện dẫn.
- Thêm cấu trúc để kiểm tra lập luận.

† Technical detail tóm tắt từ Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026.

Kỹ thuật SBV-LawGraph: citation audit và refusal



Citation-grounded generation

- Mỗi câu trả lời phải bám vào điều khoản hoặc quan hệ cụ thể.
- Citation không phải trang trí; nó là giao diện kiểm toán.

Khi thiếu bằng chứng

- Không suy diễn từ trí nhớ của LLM.
- Thiếu hoặc mâu thuẫn bằng chứng → Unknown / fallback.
- Từ chối là một hành vi an toàn.

† Refusal/citation-grounding theo thiết kế SBV-LawGraph; Phan, Lê, Quan, ACIIDS 2026.

SBV-LawGraph: kết quả và bài học

Kết quả chính

- Đánh giá trên ALQAC2025 và SBV Legal Questions.
- Vượt BM25, Naive RAG, Advanced RAG trên retrieval metrics.
- Cải thiện correctness, citation inclusion, legal consistency so với LLM tổng quát.

Bài học

- Legal AI cần **retrieval** và **legal structure**.
- Graph không thay thế text; graph bổ sung quan hệ.
- Refusal là một tính năng an toàn, không phải lỗi.

Hệ thống	Text evidence	Legal relations	Controlled fallback
LLM tổng quát	–	–	–
RAG chuẩn	✓	–	hạn chế
SBV-LawGraph	✓	✓	✓

† Kết quả tóm tắt từ Phan, Lê, Quan, SBV-LawGraph, ACIIDS 2026; datasets: ALQAC2025 và SBV Legal Questions.

Nội dung

1 Nhu cầu và cơ hội

2 Bài toán AI đáng tin

3 RAG và AI pháp lý

4 AI tài chính: từ dự báo đến quyết định

5 Ví dụ mở rộng

6 Tổng hợp và kết luận

ATM cash management: bài toán tài chính vận hành

Nếu nạp quá nhiều

- Tiền mặt bị khóa trong ATM.
- Chi phí cơ hội và chi phí giữ tiền tăng.
- Vốn không được dùng cho mục đích tốt hơn.

Nếu nạp quá ít

- ATM hết tiền, mất giao dịch.
- Phát sinh chuyến tiếp quỹ khẩn cấp.
- Ảnh hưởng trải nghiệm và uy tín.

Câu hỏi nghiên cứu

Dự báo tốt hơn có thật sự tạo ra quyết định tiếp quỹ tốt hơn không? Câu trả lời chỉ rõ khi ta đo bằng **chi phí vận hành**, không chỉ MAE hay MASE.

† Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026.

Time series forecasting: dự báo theo thời gian



Time series là gì?

- Dữ liệu được ghi theo thời gian: ngày, tuần, tháng.
- Có xu hướng, mùa vụ, chu kỳ lương, kỳ nghỉ, và nhiễu bất thường.
- Ví dụ: lượng tiền rút hằng ngày tại ATM.

Dự báo tốt

Không chỉ dự báo một đường trung bình, mà cần **dải bất định** để ra quyết định an toàn.

† Hyndman & Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed.; Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026. Hình minh họa: AI-generated bitmap.

Forecast-then-optimize



Điểm quan trọng

Mô hình không chỉ dự báo một con số. Nó dự báo **phân phối nhu cầu**, đặc biệt là phần đuôi trên, để quyết định mức tồn quỹ an toàn.

$$B_{a,t} = Q_{\rho^*}\left(\sum_{i=1}^{R+L} Y_{a,t+i}\right), \qquad Q_{a,t} = \max\{0, B_{a,t} - P_{a,t}\}.$$

† Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026.

Kỹ thuật ATM: cost-aware quantile

Inventory state

- $P_{a,t}$: tiền còn trong ATM a tại ngày t.
- $Y_{a,t+i}$: nhu cầu rút tiền tương lai.
- $R + L$: chu kỳ review cộng lead time.

Quyết định

$$Q_{a,t} = \max(0, B_{a,t} - P_{a,t})$$

Trade-off

- Nạp thấp: stockout, mất giao dịch, tiếp quỹ khẩn.
- Nạp cao: holding cost, idle cash, chi phí cơ hội.
- Vì chi phí thiếu tiền thường lớn, quantile vận hành nằm ở phần đuôi trên.

Ý nghĩa kỹ thuật

Mô hình cần dự báo **distribution tail**, không chỉ điểm trung bình. Sai số trung bình thấp vẫn có thể đặt sai $B_{a,t}$.

† Cost-aware base-stock policy theo Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026.

Dữ liệu và phương pháp

Dữ liệu

- 84 ATM sau lọc chất lượng dữ liệu.
- 3.7 năm giao dịch rút tiền hằng ngày.
- Test horizon: 30 ngày.
- Dữ liệu từ một ngân hàng thương mại Việt Nam, đã ẩn danh.

Benchmark

- 30 cấu hình dự báo.
- Statistical: SARIMA, Exponential Smoothing, Prophet.
- Tree-based: LightGBM, Random Forest.
- Neural: N-BEATS, Temporal Fusion Transformer.

Đánh giá

Mỗi mô hình được đưa qua cùng một chính sách tồn kho. Nhờ vậy, khác biệt chi phí phản ánh chất lượng dự báo và bất định, không phải thay đổi chính sách.

† Dữ liệu và benchmark theo Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026.

Kết quả: neural quantile models dẫn đầu

Phát hiện	Tác động	Ý nghĩa
Global neural quantile models đứng trên frontier	Giảm tổng chi phí mô phỏng khoảng một nửa so với baseline cổ điển	Dự báo phân phối tốt hơn chuyển thành lợi ích vận hành
Fill rate xấp xỉ 99%	Ít stockout trong mô phỏng	Kiểm soát rủi ro dịch vụ
Không mô hình nào thắng mọi ATM	Dị biệt theo vị trí và chế độ nhu cầu	Cần giám sát và phân công mô hình động

Bài học cho tài chính

Mô hình vận hành tốt phải đặt xác suất đúng ở vùng ra quyết định: phần đuôi nơi stockout và holding cost được cân bằng.

† Kết quả tóm tắt từ Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026.

Forecast accuracy không đồng nghĩa business value



Dự báo để báo cáo

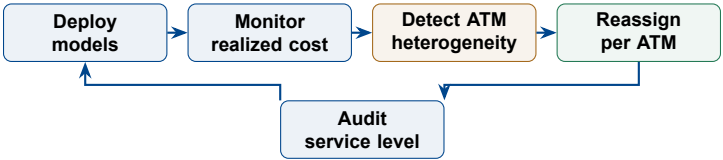
- Sai số trung bình.
- Calibration tổng quát.
- So sánh mô hình theo benchmark.

Dự báo để quyết định

- Chi phí cuối.
- Rủi ro thiếu tiền.
- Tồn quỹ an toàn theo từng ATM.

† Decision-centric evaluation dựa trên Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026.

Triển khai: monitoring-based model assignment



Từ nghiên cứu đến hệ thống

Không cần tin tuyệt đối vào một mô hình duy nhất. Cách thực tế hơn là tạo một danh sách mô hình mạnh, đo chi phí thật theo từng ATM, rồi cập nhật phân công dựa trên hiệu năng vận hành.

† Deployment lesson từ Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026.

Kỹ thuật ATM: model assignment theo monitoring

Online loop

1. Chạy một tập mô hình ứng viên trên từng cụm ATM.
2. Ghi nhận realized cost, stockout, fill rate, holding cost.
3. Ước lượng regret theo ATM hoặc nhóm ATM tương tự.
4. Gán lại mô hình khi chi phí vượt ngưỡng giám sát.

Pseudo-rule

$$m_a^{new} = \arg\min_{m \in \mathcal{M}} \widehat{C}_a(m)$$

- $\mathcal{M}$: danh sách mô hình mạnh.
- $\widehat{C}_a(m)$: chi phí quan sát/ước lượng cho ATM a.

Điểm triển khai

Không cần một mô hình thắng toàn bộ mạng lưới; cần một cơ chế phân công và giám sát được.

† Monitoring-based deployment lesson từ Phan, Lê, Quan, ATM cash management, FLINS-ISKE 2026.

Nội dung

1 Nhu cầu và cơ hội

2 Bài toán AI đáng tin

3 RAG và AI pháp lý

4 AI tài chính: từ dự báo đến quyết định

5 Ví dụ mở rộng

6 Tổng hợp và kết luận

Các nhánh nghiên cứu liên quan

Ngôn ngữ và tri thức tiếng Việt

- **VSLIM**: multi-intent detection và slot filling tiếng Việt.
- **Event extraction survey**: chuyển dịch phương pháp với LLMs.

Giải thích và phần mềm

- **XAI for SDLC**: từ black-box đến glass-box.
- **LLM agents in static analysis**: grounding, validation, reliability.

Sợi chỉ chung

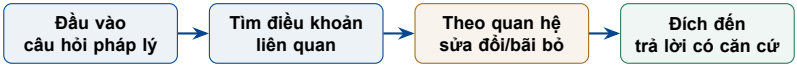
- AI không đứng một mình.
- Mỗi miền cần dữ liệu, cấu trúc và tiêu chí tin cậy riêng.
- Đánh giá phải phản ánh người dùng và quyết định thật.

Cầu nối

Từ pháp lý, tài chính, đến software engineering: vấn đề cốt lõi vẫn là biến LLM thành một thành phần có kiểm soát trong hệ thống.

† Chung et al., VSLIM, ACIIDS 2026; Quang et al., Event Extraction with LLMs, IEA-AIE 2026; Lê & Lê, XAI for SDLC, IEA-AIE 2026; Hua & Lê, LLM Agents in Static Analysis, IEA-AIE 2026.

Ví dụ SBV-LawGraph: từ câu hỏi đến câu trả lời



Đầu bài minh họa

"Quy định A trong văn bản cũ còn hiệu lực không, hay đã bị văn bản mới sửa đổi/thay thế?"

Đích đến mong muốn

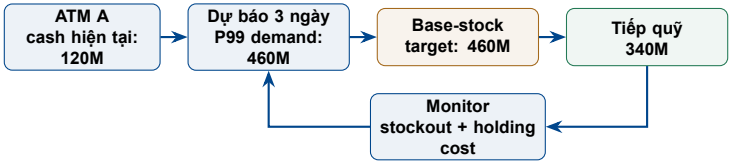
Câu trả lời nêu **còn/không còn hiệu lực**, dẫn ra điều khoản nguồn, và chỉ rõ chuỗi quan hệ pháp lý làm căn cứ.

Điểm cần thấy

Khác biệt nằm ở bước giữa: hệ thống không chỉ lấy một đoạn giống câu hỏi, mà phải nối được văn bản hiện hành, văn bản sửa đổi, và phần bằng chứng dùng để sinh câu trả lời.

† Ví dụ minh họa theo pipeline SBV-LawGraph; không phải tư vấn pháp lý. Phan, Lê, Quan, ACIIDS 2026.

Ví dụ ATM: từ dự báo đến quyết định tiếp quỹ



Đầu vào

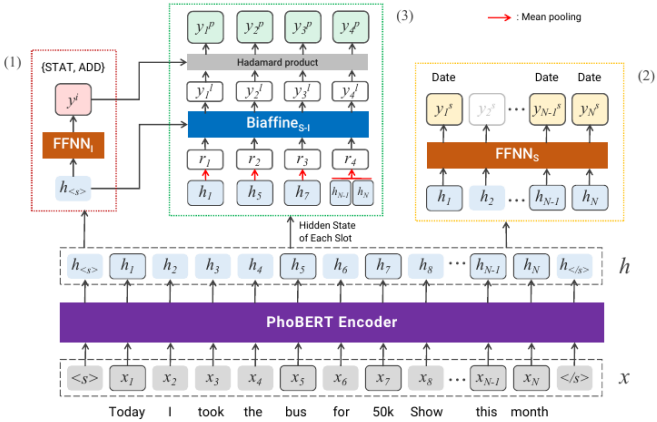
Lịch sử rút tiền, ngày lương/kỳ nghỉ, lịch tiếp quỹ, số dư hiện tại và chi phí thiếu tiền/giữ tiền.

Đích đến

Không phải "dự báo đẹp", mà là quyết định tiếp quỹ làm giảm chi phí nhưng vẫn giữ fill rate cao.

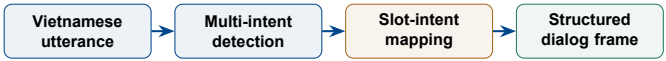
† Các số trong ví dụ là minh họa cho công thức base-stock; pipeline dựa trên Phan, Lê, Quan, FLINS-ISKE 2026.

VSLIM: hình hệ thống từ paper



† Cắt từ Fig. 2 trong Chung, Le-Minh, Lê, Quan, VSLIM, ACIIDS 2026.

Ví dụ VSLIM: multi-intent + slot filling



Ví dụ câu nói

"Tìm chuyến bay từ TP.HCM ra Hà Nội sáng mai và thêm hành lý 20kg."

Intent

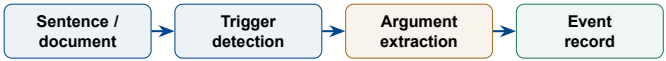
`search_flight` + `add_baggage`

Slot gắn với intent

Slot	Intent
from=TP.HCM	search_flight
to=Hà Nội	search_flight
date=sáng mai	search_flight
weight=20kg	add_baggage

† Chung, Le-Minh, Lê, Quan, VSLIM, ACIIDS 2026. Ví dụ chỉ minh họa dạng tác vụ.

Ví dụ Event Extraction: từ câu sang event record



Ví dụ câu

"Ngân hàng X ký hợp tác tín dụng với công ty Y tại TP.HCM ngày 12/5."

LLM có thể tham gia

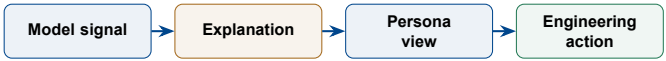
Data augmentation, prompting, fine-tuning, hoặc chuẩn hóa schema đa ngôn ngữ.

Event record

field	value
type	cooperation/signing
trigger	ký hợp tác
org1	Ngân hàng X
org2	công ty Y
place/date	TP.HCM / 12-5

† Quang, Vo, Lê, Quan, Event Extraction with LLMs, IEA-AIE 2026. Ví dụ chỉ minh hoạ dạng tác vụ.

Ví dụ XAI for SDLC: giải thích như artifact kiểm toán



Ví dụ

Mô hình đánh dấu `PaymentService` là module rủi ro cao trước release.

Đích đến

Reviewer ưu tiên test payment flow, yêu cầu bằng chứng bổ sung, hoặc trì hoãn release nếu explanation không đủ tin cậy.

Bằng chứng giải thích

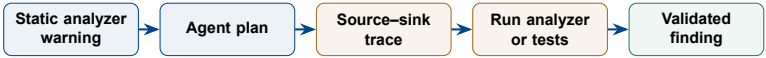
Churn gần đây, coverage thấp, dependency quan trọng, lịch sử lỗi, và thay đổi API.

Rủi ro

Giải thích nghe hợp lý nhưng không faithful có thể làm tăng over-trust.

† Lê & Lê, XAI for SDLC, IEA-AIE 2026. Ví dụ chỉ minh họa dạng workflow.

Ví dụ LLM agents: kiểm chứng static analysis



Ví dụ

Analyzer báo nguy cơ SQL injection tại hàm đăng nhập.

Agent làm gì?

Truy vết input, kiểm tra sanitizer, đề xuất parameterized query, rồi chạy lại analyzer/test.

Đích đến

Cảnh báo được xác nhận hoặc bác bỏ bằng artifact thật: code path, test output, analyzer log.

Guardrail

Không chấp nhận lời giải thích thuần văn bản nếu thiếu bước kiểm chứng.

† Hua & Lê, LLM Agents in Static Analysis, IEA-AIE 2026. Ví dụ chỉ minh hoạ dạng workflow.

Nội dung

1. Nhu cầu và cơ hội
2. Bài toán AI đáng tin
3. RAG và AI pháp lý
4. AI tài chính: từ dự báo đến quyết định
5. Ví dụ mở rộng
6. Tổng hợp và kết luận

Blueprint cho AI chuyên biệt ở Việt Nam



Pháp lý

Văn bản, điều khoản, hiệu lực, quan hệ sửa đổi.

Tài chính

Dự báo bất định, chi phí, service level, rủi ro.

Tổ chức

Quy trình kiểm tra, giám sát, audit, human-in-the-loop.

Thách thức cần xử lý

Dữ liệu và đánh giá

- Dữ liệu phân tán, nhiễu, thiếu nhãn, khó chia sẻ vì riêng tư/bảo mật.
- Ground truth trong pháp lý/tài chính thường cần chuyên gia thẩm định.
- Benchmark chung không đo được rủi ro vận hành thật.

Hệ thống và trách nhiệm

- Hallucination, citation sai, drift dữ liệu và prompt injection.
- Chi phí inference, latency, logging, truy vết và bảo trì tri thức.
- Người dùng có thể quá tin hoặc không tin hệ thống nếu thiếu giải thích.

Cách nhìn

Thách thức không chỉ là thuật toán; đó là bài toán **data governance**, **evaluation**, **UX tin cậy**, và **vận hành dài hạn**.

† Stanford HAI AI Index Report 2025; Ji et al., hallucination survey, ACM CSUR 2023; RAG robustness/evaluation surveys 2025.

Lời khuyên: từ demo đến hệ thống đáng tin

Cho nghiên cứu

- Workflow hẹp, rủi ro thật.
- Eval protocol rõ.
- Metric theo quyết định cuối.

Cho triển khai

- Evidence log trước.
- Refusal + human review.
- Monitor cost, latency, drift.

Cho sinh viên

- Học domain.
- Biết đánh giá.
- Prototype kiểm tra được.

Nguyên tắc ngắn

Đừng hỏi “dùng model nào?” trước; hãy hỏi: quyết định nào cần cải thiện, bằng chứng nào chứng minh hệ thống đúng, và ai chịu trách nhiệm khi nó sai?

† Lời khuyên tổng hợp từ các bài học triển khai SBV-LawGraph, ATM cash management và xu hướng trustworthy/domain-specialized AI.

Research agenda và cơ hội hợp tác

Hướng nghiên cứu

- Legal KG chất lượng cao cho văn bản Việt Nam.
- RAG có kiểm chứng, citation audit, và refusal policy.
- Decision-focused learning cho bài toán tài chính.
- Benchmark gắn với tác động vận hành.

Hợp tác cần thiết

- Dữ liệu miền và chuyên gia thẩm định.
- Bài toán có chi phí/rủi ro rõ ràng.
- Môi trường thử nghiệm có giám sát.
- Chuẩn đánh giá phù hợp Việt Nam.

Quan điểm

AI chuyên biệt là đồng thiết kế giữa thuật toán, dữ liệu, chuyên gia miền và quy trình vận hành.

† Blueprint tổng hợp từ hai case studies chính: SBV-LawGraph (ACIIDS 2026) và ATM cash management (FLINS-ISKE 2026).

Ba thông điệp chính

1. **LLM tổng quát là điểm khởi đầu, không phải hệ thống hoàn chỉnh.**
   Trong tài chính và pháp lý, cần grounding, cấu trúc miền, và kiểm soát rủi ro.
2. **Tin cậy phải được thiết kế, không thể chỉ được tuyên bố.**
   SBV-LawGraph dùng retrieval kép, Legal Knowledge Graph, citation-grounded generation, và fallback.
3. **Đánh giá phải đi đến quyết định thật.**
   Trong ATM cash management, mô hình tốt là mô hình giảm chi phí và giữ service level, không chỉ giảm sai số dự báo.

Kết luận

AI đáng tin cậy cho Việt Nam sẽ đến từ các hệ thống chuyên biệt, có cấu trúc, có bằng chứng, và được đo bằng tác động thực tế.

† Ba thông điệp tổng hợp từ SBV-LawGraph (ACIIDS 2026) và ATM cash management (FLINS-ISKE 2026).

Q & A

Xin cảm ơn!

TS. Lê Xuân Bách

lexuanbach@hcmut.edu.vn

https://lexuanbach.github.io

