International Journal of Graphic Design (IJGD)
Volume 4 No. 1 April 2026
E-ISSN: 2987-9434; P-ISSN : 2988-0343, Page 246-265
DOI: https://doi.org/10.51903/ijgd.v4i1.4001
Available online at: https://journal.stekom.ac.id/index.php/ijgd

From 8-K to Deal Cards: A Source-Grounded Visual Triage Framework for FinTech M&A Filings

Sisi Meng[1], Dingyuan Zhang*[2], Eric Li[3]
[1]Accounting, University of Rochester, New York, USA
[2]Business Analytics, University of Rochester, New York, USA
[3]Applied Analytics, Columbia University, New York, USA
Email Address: jessicazhang0113@gmail.com

***Abstract**. Form 8-K is a timely public channel through which U.S. registrants disclose material events, including acquisition agreements, completed transactions, shareholder communications, and related exhibits. The density of this filing stream creates a visual communication problem: analysts need to identify source-linked deal cues without losing the legal and temporal provenance of each record. This study presents a source-grounded deal-card workflow for FinTech M&A filing triage. The analysis used the official SEC quarterly master indexes and included 16,439 Form 8-K and 8-K/A filings in Q4 2025, together with 276,261 contextual records in the complete Q4 index. Symmetric companion-form windows were protected against quarter-boundary truncation with the last 14 days of Q3 2025 and the first 14 days of Q1 2026. A seven-day window identified 222 rule-defined FinTech M&A targets across 94 company names, or 1.35% of the Q4 corpus. In five-fold cross-validation grouped by CIK, text-only logistic triage reached F1 = 0.224 and M&A-context logistic triage reached F1 = 0.322. Integrated logistic and linear SVM configurations reproduced the rule-defined target exactly (F1 = 1.000). Because the target was constructed from the same FinTech and M&A-neighborhood cues available to those configurations, this exact agreement is interpreted as rule-consistency validation rather than predictive superiority. A structural interface audit found that the deal-card schema exposed 10 documented fields, nine visual anchors, and six configured decision cues. These author-defined attributes describe interface structure; they do not measure comprehension, trust, or review speed. The contribution is a method for organizing SEC filing evidence for subsequent visual review.*

***Keywords:** FinTech; mergers and acquisitions; Form 8-K; information visualization; visual triage*

INTRODUCTION

Mergers and acquisitions in financial technology are information-rich events. They combine corporate strategy, regulatory approvals, financing structure, integration risk, and market interpretation into a disclosure trail that is rarely arranged for quick visual comprehension (Elena & Alexander, 2026). In the United States, Form 8-K is a primary channel for current reporting under the Securities Exchange Act. The official form states that it is used for current reports under Section 13 or 15(d), and its general instructions identify a four-business-day filing framework for many reportable events (U.S. Securities and Exchange Commission [SEC], 2025). For M&A analysis, Item 1.01 may report entry into a material definitive agreement, whereas Item 2.01 applies to a completed acquisition or disposition that meets the relevant significance requirements. Item 9.01 addresses associated financial statements, pro forma financial information, and exhibits (SEC, 2024b). The filing stream is timely, but timeliness alone does not produce comprehension.

Prior accounting research shows that 8-K disclosures carry information value. Lerman and Livnat (2010) found that the expanded 8-K regime shaped the information environment around material events, and McMullin et al. (2019) connected more frequent mandated disclosure to price

Received: Januari 2026; Revised: February 2026; Accepted: March 2026; Published: April 2026
*Corresponding author, jessicazhang0113@gmail.com

formation. At the same time, readability work in financial reporting demonstrates that disclosure value depends on how accessible the text is to users (F. Li, 2008), and domain-specific dictionaries remain necessary because ordinary word lists often misread financial language (Loughran & McDonald, 2011). These findings motivate a design question rather than only a prediction question: how can a filing stream be translated into visual units that expose the event, its evidence, and the next review action?

Large language models (LLMs) provide a possible downstream wording layer for such interfaces. General LLMs show strong few-shot language capabilities (Brown et al., 2020), retrieval-augmented generation conditions language production on external documents (Lewis et al., 2020), and finance-oriented models such as BloombergGPT and FinGPT demonstrate the importance of domain context (S. Wu et al., 2023; Yang et al., 2023). Financial retrieval studies likewise treat evidence links, numeric verification, and hallucination controls as design requirements (C. Li et al., 2023; Q. Wu et al., 2023; Zhang et al., 2023; Zhou et al., 2023). This study evaluates the deterministic filing join, rule-defined triage logic, and deal-card schema. It does not treat language-model output as empirical evidence. The structured payload could constrain a later wording layer, but the reported results concern the source-linked data and interface structure.

The paper contributes three outcomes. First, it reports a complete Q4 2025 processing run on the official SEC quarterly master index, covering 16,439 Form 8-K and 8-K/A filings and 276,261 contextual filing records. Second, it develops an operational FinTech M&A deal-card target based on the intersection of transparent company-name cues and nearby companion-form evidence. Third, it audits how the deal-card schema changes visible fields, visual anchors, and decision cues relative to a raw SEC index row and a keyword-summary baseline. The central claim is that financial disclosure design can recompose filing metadata into a source-linked evidence object without presenting the operational rule as independent deal truth.

Three research questions organize the empirical design. RQ1 asks how EDGAR index fields can be transformed into a source-linked FinTech M&A triage layer. RQ2 asks how the candidate set changes when FinTech cues and event-neighborhood evidence are used separately or together. RQ3 asks how a structured deal card changes field visibility, visual anchors, and decision cues relative to two computational baselines. These questions evaluate triage logic and interface structure rather than deal outcomes or human dashboard effectiveness. FinTech M&A is a useful domain for this test because the events sit at the boundary of financial services, software infrastructure, payment rails, banking regulation, credit products, and data governance (Xin, 2026b). The same filing may matter to investors, designers of financial dashboards, compliance teams, and product strategists. A visually coherent card can therefore serve multiple readers: it can help a finance specialist decide which filing to inspect, while giving a design researcher a repeatable way to study hierarchy, emphasis, and provenance in high-stakes information graphics.

LITERATURE REVIEW

Financial disclosure research provides the foundation for the study. Form 8-K is not a periodic report; it is an event-oriented filing used when a registrant reports current material information. Lerman and Livnat (2010) showed that the expanded 8-K disclosure environment affected how investors receive material-event information, while McMullin et al. (2019) studied the effect of increased mandated disclosure frequency on price formation. These studies explain why an interface that monitors 8-K events has practical value. F. Li (2008) established that financial-report readability relates to earnings and persistence, and Loughran and McDonald (2011) showed that financial language requires specialized textual analysis. Together, this literature suggests that merely placing raw filings in a searchable list does not solve the communication problem.

Financial NLP and LLM research adds a second foundation. FinBERT demonstrated that pre-trained language models can be adapted to financial sentiment tasks more effectively than general models in the same domain (Araci, 2019). BloombergGPT and FinGPT extended this reasoning to large-scale finance-oriented language modeling and open financial LLM development (S. Wu et al., 2023; Yang et al., 2023). General LLM research shows why prompts can support flexible extraction and summarization (Brown et al., 2020), while RAG highlights the value of coupling language generation with retrievable evidence (Lewis et al., 2020; Su et al., 2025). The risk is hallucination: Ji et al. (2023) documented how natural language generation can produce unsupported content. Recent work on evidence-chain RAG and calibrated abstention similarly emphasizes source attribution, provenance explanation, and refusal when retrieval coverage is insufficient (Jin, 2025a; C. Li et al., 2024; Zhong, Chen, et al., 2025). The present design responds to that risk by making the deal card a source-grounded visual object rather than a free-form summary.

Information visualization and graphic design supply the third foundation. Shneiderman's (1996) visual information-seeking mantra argues for overview, zoom and filter, and details on demand. Card et al. (1999) frame visualization as a way to use vision to think, and Munzner (2014) emphasizes a pipeline from task abstraction to visual encoding. Tufte (2001) and Ware (2012) stress visual clarity and perceptual efficiency; Few (2006) translates those ideas into dashboard practice. Cairo (2016) argues that truthful visual communication must respect evidence, while Heer et al. (2010) show that different chart forms support different analytical questions. Hullman and Diakopoulos (2011) further show that visualization carries rhetorical force, so visual choices must be accountable. Nielsen's (1994) usability heuristics support recognition, feedback, and reduced memory load as design goals. These principles shape the deal-card format: a card should show filing identity, event-neighborhood evidence, display priority, and a source path. Recent studies of AI-enabled financial advice, machine-learning risk assessment, and integrated financial platforms extend this concern to decision support, interpretability, and data integration (Jin, 2025b; Z. Li et al., 2025; Xin, 2026a; Zhang et al., 2025).

A related design issue is the difference between search and sensemaking. Search can locate a filing, but sensemaking requires a visual relation among event identity, timing, evidence, and next action. The 8-K stream is particularly demanding because the same form type can report many kinds of events. A dashboard that treats every row identically leaves the reader to reconstruct the event context. A deal-card design can instead use a compact visual grammar: company identity, filing date, form status, evidence chips, source link, and review action. Financial-search reranking and accounting-aware evidence retrieval studies reinforce this distinction: locating a document is one step before relevance ordering, contextual evidence assembly, and user-facing explanation (Chen et al., 2025; Meng et al., 2026). The present paper links these fields through a graphic design contribution. Instead of treating FinTech M&A as a finance forecasting problem, it treats SEC disclosure as a visual communication medium. The source-grounded card is not a replacement for EDGAR, legal review, or primary-document reading. It is an evidence-preserving interface layer that organizes candidates for subsequent inspection.

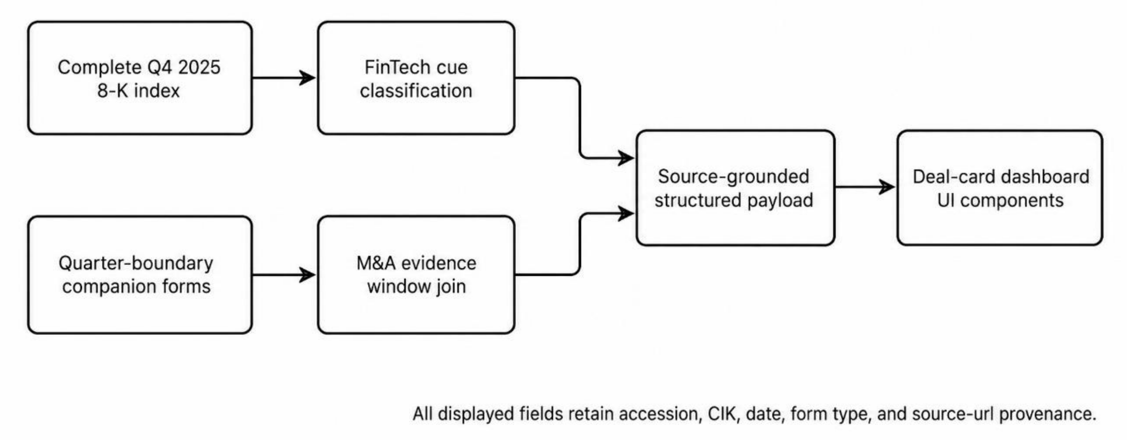

Figure 1. Source-grounded pipeline from official SEC filing indexes to deal-card interface components.

METHODS

The analysis used the official SEC EDGAR quarterly master indexes for 2025 Q3, 2025 Q4, and 2026 Q1 (SEC, 2026). The complete Q4 index contained 276,261 filing records; all

16,439 records with form type 8-K or 8-K/A formed the analysis corpus. The last 14 days of Q3 2025 and the first 14 days of Q1 2026 were used only to protect symmetric companion-form windows at the Q4 boundaries. Each parsed record retained CIK, company name, form type, filing date, SEC filename, accession number, and a filing URL constructed from the SEC archive path. The SEC EDGAR API documentation confirms that public submissions and filing data are available without API keys (SEC, 2024a). The quarterly-index layer provides a compact input for visual triage while retaining a direct path to each primary filing.

Table 1. Dataset profile for the complete Q4 2025 Form 8-K corpus and master index.

Metric	Value
Material-event records	16,439
Unique filing URLs	16,439
Unique CIKs	5,400
Unique company names	5,431
Date range	2025-10-01 to 2025-12-31
Filing dates represented	60
Rows with missing CIK/date/company	0
Repeated company-date-form rows	556
Complete Q4 master-index rows	276,261
Q4 M&A companion-form rows	4,125

The pipeline is shown in Figure 1. It has four stages. The first stage parses the complete Q4 8-K index and validates required fields. The second classifies FinTech cues from company names with a multi-label lexicon covering banking, credit, payments, digital assets, capital markets, and insurance or trust terms. The third joins each Q4 filing to companion forms by CIK and filing date, using quarter-boundary records when necessary. The fourth creates a source-grounded deal-card payload containing company name, CIK, filing date, accession number, form type, source URL, FinTech cue category, companion forms, a rule-derived display priority, and a review action. A deterministic formatter maps these populated fields to the card schema. This separation follows financial disclosure retrieval designs that keep evidence collection, field validation, and explanation as inspectable stages (Chen et al., 2025; Zhou et al., 2023, 2024).

Table 2. Form-type distribution in the complete Q4 2025 corpus.

Form type	Records	Share (%)
8-K	16,082	97.83
8-K/A	357	2.17

Table 1 reports the parsed dataset profile. The corpus contains no missing CIK, filing date, or company-name values. Repeated company-date-form rows are not duplicate filing URLs; they are distinct same-day filings by the same registrant and form type. Table 2 reports the form distribution. Initial 8-K reports dominate the corpus, while amendments remain a visible status cue.

M&A-neighborhood evidence was defined through companion forms that commonly accompany business combinations, exchange offers, tender offers, shareholder communications, or merger registration processes. The companion set included 425, DEFA14A, DEFM14A, PREM14A, S-4, S-4/A, F-4, F-4/A, CB, SC TO-I, SC TO-I/A, SC TO-T, SC TO-T/A, SC 14D9, SC 14D9/A, SC 13E3, and SC 13E3/A. A Q4 filing received an M&A-neighborhood flag when at least one included companion form for the same CIK occurred within the selected symmetric date window. The one-, three-, seven-, and fourteen-day joins used the last 14 days of Q3 and the first 14 days of Q1 when a Q4 filing lay near a quarter boundary. The seven-day window served as the main operational definition because it limited associations to a one-week filing neighborhood while retaining more candidates than the shorter windows.

The accession number was parsed from each filename with a fixed regular expression and retained as a card field. The filing URL was not shortened in the data layer because it functions as a source path, even when the visual card later displays it as an icon or link. The data audit checked record count, unique filing URLs, unique CIKs, missing values, form-type distribution, and repeated company-date-form clusters. These checks are design-relevant: a dashboard that hides repeated same-day filing clusters can make a filing event appear simpler than it is, while a dashboard that exposes every duplicate-looking row without grouping can overload the reader.

The FinTech lexicon was intentionally transparent. It matched words and phrases associated with banking and depository institutions, credit and lending, payments and cards, digital assets, capital markets and asset management, and insurance or trust activity. The lexicon did not attempt to classify every possible business model. Its purpose was to create visible cue chips that can be inspected, accepted, or rejected by the user. In a graphic design context, this matters because a cue chip should show why the interface highlighted a record, not merely that an opaque model assigned a score.

The operational target label was source-derived: a record qualified for the FinTech M&A deal-card view when it had at least one FinTech cue and at least one companion form within seven days. Three rule configurations used the FinTech cue alone, the M&A neighborhood alone, or their structured intersection. Five statistical configurations used logistic regression, linear support vector machines, or multinomial Naive Bayes. Text features were TF-IDF vectors over registrant names; categorical features included form type, companion-form string, and FinTech category; numeric features included companion count, weekday, day of month, and amendment status. Five-fold StratifiedGroupKFold cross-validation grouped observations by CIK so that one registrant did not appear in both training and validation folds. Precision, recall, F1, ROC-AUC, and average precision were computed. Because the integrated feature sets contain the same FinTech and M&A-neighborhood evidence used to construct the label, their agreement with that label is a rule-consistency check rather than an independent predictive test. The implementation used pandas (McKinney, 2010), NumPy (Harris et al., 2020), scikit-learn (Pedregosa et al., 2011), and Matplotlib (Hunter, 2007).

The evaluation therefore separates deterministic eligibility rules from statistical feature-ablation configurations. The single-cue rules describe how broad the candidate set becomes when only one evidence stream is available. The integrated configurations test whether the implementation can reconstruct the operational intersection from its defining features. Without independently adjudicated deal labels, these comparisons do not establish model superiority or external FinTech M&A detection accuracy.

A structural interface audit compared three display conditions. The raw SEC index row presented direct filing fields. The keyword-summary baseline compressed those fields into a deterministic sentence with FinTech and companion-form cues. The source-grounded deal card divided identity, evidence, display priority, source, and review action into separate visual regions (Mu et al., 2026). The audit recorded mean characters per record, documented fields, configured visual anchors, and configured decision cues. These measures describe information quantity and organization; human comprehension, trust, and review speed were outside the scope of the audit.

The card wording was composed with a fixed template that preserved accession numbers and filing URLs and used only populated payload fields. This deterministic design is consistent with evidence-oriented language-system research that emphasizes source attribution, numerical controls, and calibrated abstention before generated text reaches a user-facing surface (C. Li et al., 2023, 2024; Z. Li et al., 2026; Zhong, Chen, et al., 2025). A constrained LLM could later rephrase the same payload, but language-model generation was not part of the reported evaluation.

RESULTS

The complete Q4 corpus covered 60 filing dates from October 1 through December 31, 2025. Table 3 shows a mean daily volume of 273.98 filings and a maximum of 714. Figure 2 shows the day-to-day volume together with the much smaller rule-defined target series. This contrast describes the selection problem addressed by the triage view; it does not measure the effectiveness of the interface for human reviewers.

Table 3. Daily Form 8-K and 8-K/A filing-volume summary.

Statistic	Daily filings
Count	60
Mean	273.98
Standard deviation	117.02
Minimum	99
25th percentile	213
Median	235
75th percentile	281
Maximum	714

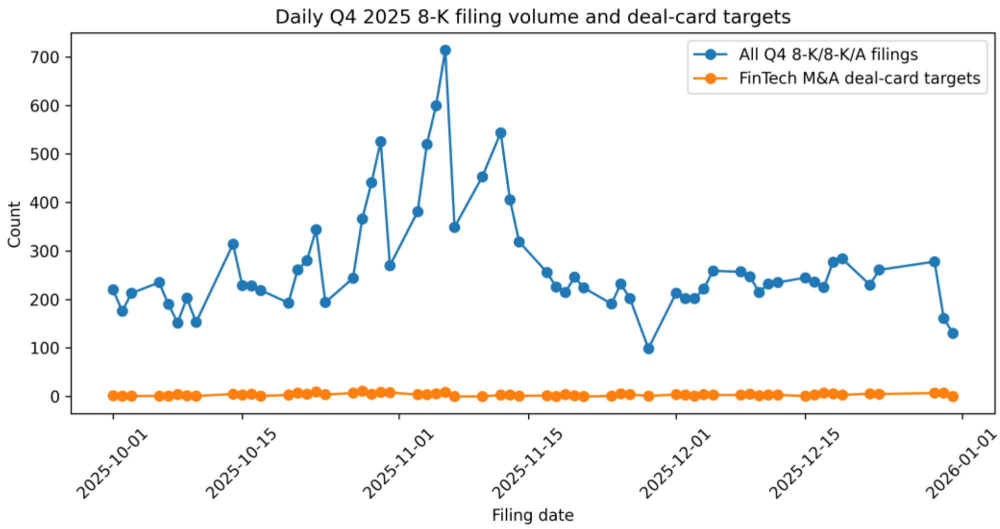

Figure 2. Daily Q4 2025 Form 8-K and 8-K/A volume and rule-defined FinTech M&A targets.

Table 4 reports the full companion-form distribution in the Q4 master index. The 17 included form types accounted for 4,125 records. Form 425 was the largest category, followed by DEFA14A. Figure 3 displays the 12 most frequent companion forms. These forms provide transaction-neighborhood context around an 8-K index row without being treated as independent confirmation of a completed acquisition. This use of nearby filings is consistent with financial retrieval studies that assemble companion evidence and numeric context before producing a risk memo (Q. Wu et al., 2023; Zhang et al., 2023).

Table 4. Included M&A companion-form counts in the complete Q4 2025 master index.

Companion form	Q4 records	Share (%)
425	1,984	48.10
DEFA14A	1,217	29.50
SC TO-I/A	173	4.19
SC TO-I	150	3.64
S-4	122	2.96
S-4/A	84	2.04
SC TO-T/A	81	1.96
DEFM14A	70	1.70
CB	48	1.16
PREM14A	46	1.12
SC 13E3/A	44	1.07
SC 14D9/A	32	0.78
F-4/A	27	0.65
SC TO-T	19	0.46
SC 13E3	10	0.24
F-4	9	0.22
SC 14D9	9	0.22

The distribution has a direct design consequence. A single 8-K index row does not necessarily reveal the surrounding transaction context. The visual schema therefore includes a companion-form region beside the filing identity and source path. This region makes the matched forms visible within the static card without implying that the neighborhood is a legally adjudicated deal label. This arrangement supports financial decision review while keeping the integrated data path visible (Jin, 2025b; Zhang et al., 2025).

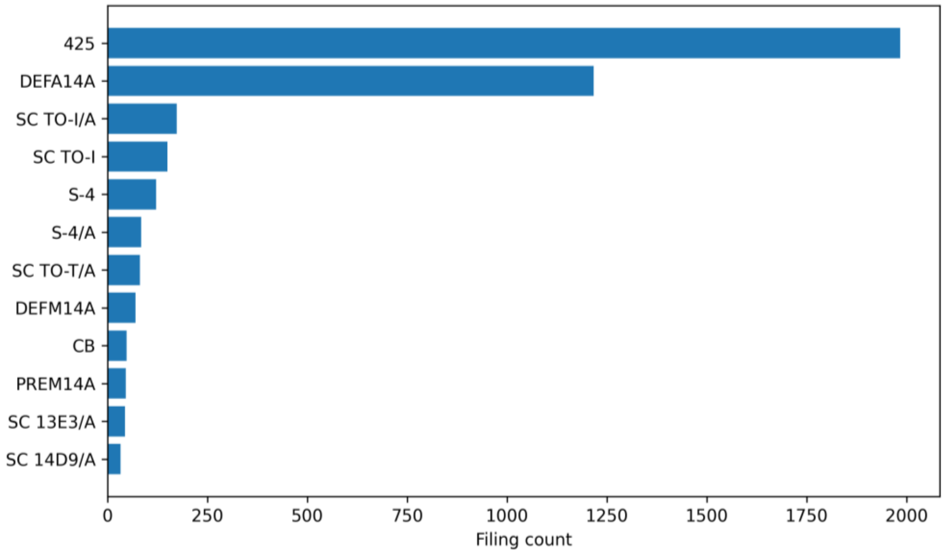

Figure 3. Twelve most frequent M&A companion forms in the complete Q4 2025 master index.

Table 5 reports the company-name FinTech cue categories. The lexicon matched 2,820 filings in the Q4 corpus. Capital-markets and asset-management cues formed the largest category, followed by banking and depository cues, credit and lending cues, and insurance or trust cues. Digital-assets and payments or cards cues were less frequent. Figure 4 shows the multi-label distribution. These categories provide visible reasons for a rule match; they are not comprehensive business-model classifications.

Table 5. FinTech cue categories matched in company names.

FinTech cue category	Records	Unique companies	Share of corpus (%)
Capital markets and asset management	915	293	5.566
Banking and depository	848	202	5.158
Credit and lending	791	179	4.812
Insurance and trust	761	266	4.629
Digital assets	57	19	0.347
Payments and cards	38	19	0.231

Table 6 reports window sensitivity. A one-day window produced 125 FinTech M&A records, the seven-day window produced 222 records across 94 company names, and the fourteen-day window produced 310 records across 110 company names. The seven-day setting was retained as the operational window because it restricted associations to one week while retaining more candidates than the shorter windows; no independent deal truth was available to estimate recall. Table 7 summarizes the seven-day definition. The corpus contained 1,250 M&A-neighborhood filings and 2,820 FinTech-cue filings, with 222 records in their intersection. Figure 5 displays all four cells of this rule matrix.

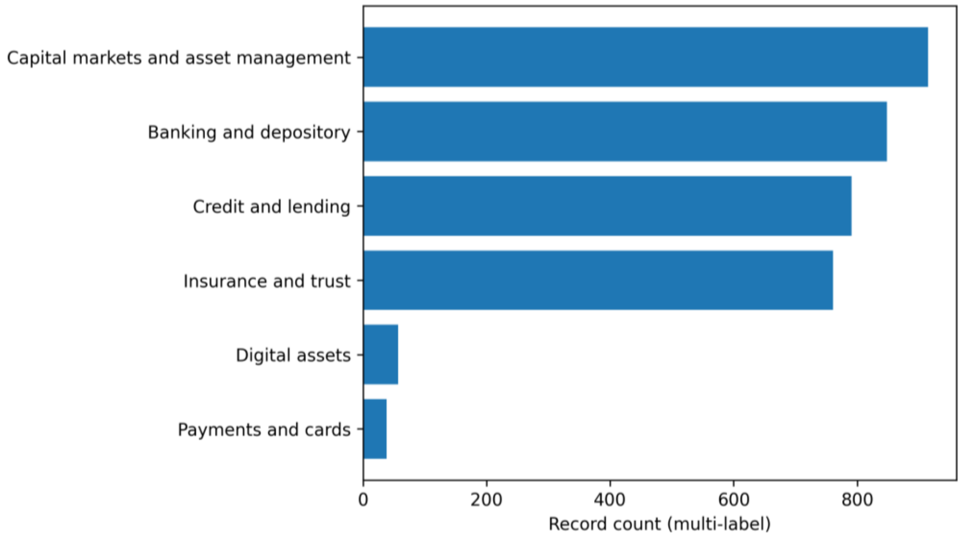

Figure 4. Company-name FinTech cue categories in complete Q4 2025 Form 8-K filings.

Table 8 reports agreement with the rule-defined target. The keyword-only rule retrieved every target by construction but had precision of 0.079 and F1 of 0.146 because it selected a much broader set. The M&A-neighborhood rule likewise had recall of 1.000, with precision of 0.178 and F1 of 0.302. Under CIK-grouped cross-validation, text-only logistic regression reached F1 = 0.224 and M&A-context logistic regression reached F1 = 0.322. Integrated logistic and linear SVM configurations reproduced the rule-defined labels exactly (F1 = 1.000), while integrated multinomial Naive Bayes reached F1 = 0.872. Figure 6 visualizes F1 and average precision against the operational label.

Table 6. Sensitivity of rule-defined targets to companion-form window size.

Window (days)	All 8-Ks with companion	FinTech 8-Ks with companion	Unique target companies	Target share (%)
1	811	125	65	0.76
3	956	151	82	0.92
7	1,250	222	94	1.35
14	1,670	310	110	1.89

The perfect score of the structured intersection rule is definitional because the rule is evaluated against its own output. The same caution applies to integrated statistical configurations that receive the two label-defining evidence streams. Their exact agreement confirms implementation consistency, not independent predictive superiority. The partial-evidence configurations are useful only for showing how candidate scope changes when one cue is absent.

Table 8. Rule-consistency and feature-ablation audit for the operational FinTech M&A target.

Configuration	Precision	Recall	F1	Average precision
Keyword-only rule	0.079	1.000	0.146	0.079
M&A-neighborhood rule	0.178	1.000	0.302	0.178
Structured intersection rule	1.000	1.000	1.000	1.000
Text-only logistic	0.139	0.591	0.224	0.174
M&A-context logistic	0.192	1.000	0.322	0.514
Integrated logistic	1.000	1.000	1.000	1.000
Integrated linear SVM	1.000	1.000	1.000	1.000
Integrated multinomial NB	1.000	0.780	0.872	0.988

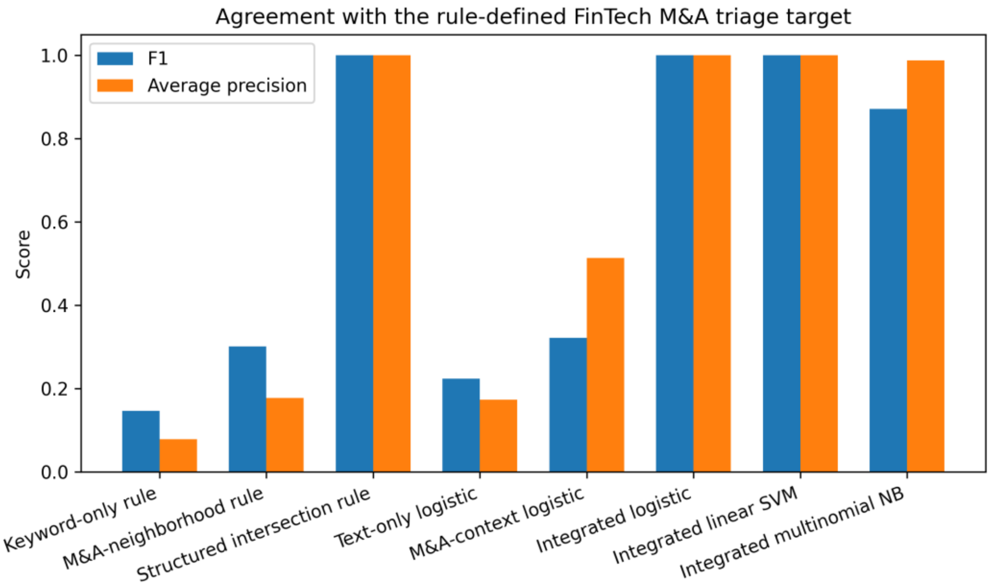

Figure 6. F1 and average precision against the rule-defined target across cue and feature configurations.

Table 9. Structural interface inventory for the rule-defined target records.

Interface condition	Mean characters	Documented fields	Visual anchors	Decision cues
Raw SEC index row	241.2	8	2	0
Keyword-summary baseline	97.9	5	4	2
Source-grounded deal card	218.4	10	9	6

Table 9 reports the structural interface audit for the 222 target records. The raw SEC index row averaged 241.2 characters and contained eight documented fields, two visual anchors, and no configured decision cues. The keyword-summary baseline averaged 97.9 characters, with five documented fields, four anchors, and two cues. The source-grounded deal card averaged 218.4 characters and was configured with 10 documented fields, nine anchors, and six cues. The card therefore exposes a different information structure rather than simply minimizing text. Whether that structure improves comprehension or review speed requires user testing (Jin, 2025b; Z. Li et al., 2025).

Table 10. Target-card field population and provenance classification.

Deal-card field	Population rate (%)	Source or derivation
CIK	100.0	direct SEC index field
Company name	100.0	direct SEC index field
Filing date	100.0	direct SEC index field
Form type	100.0	direct SEC index field
Accession number	100.0	parsed from SEC filename
Filing URL	100.0	constructed from SEC archive path
FinTech segment	100.0	company-name lexicon (derived)
M&A companion forms	100.0	SEC index window join (derived)
Confidence tier	100.0	cue intersection rule (derived)
Review action	100.0	interface rule (derived, not a filing item)

Table 10 reports field population and provenance for the 222 target cards. By definition, every target contains both a FinTech cue and companion-form evidence, so those two fields are populated for all target cards. CIK, company name, filing date, form type, and the archive path are direct or directly constructed SEC index fields; accession numbers are parsed from SEC filenames; segment, display priority, and review action are documented rule-derived fields. The review action is an interface instruction, not an actual Form 8-K item. Figure 7 shows a target card with the EDGAR source path visibly included. This distinction between source fields and derived fields follows evidence-chain and numerical-guardrail approaches that make transformations inspectable (C. Li et al., 2024; Z. Li et al., 2026; Zhang et al., 2023).

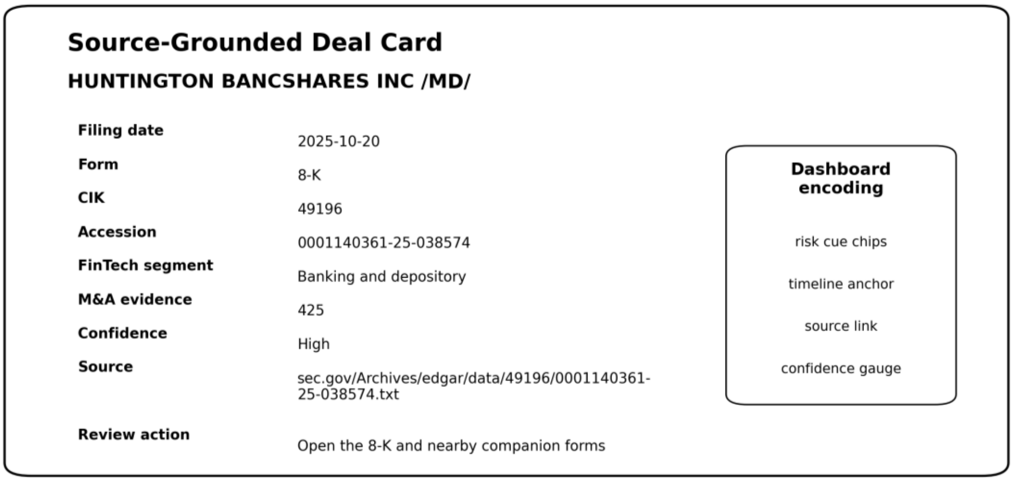

Figure 7. Source-grounded deal-card layout for a rule-defined target record.

DISCUSSION

The results establish the feasibility of transforming a complete quarterly SEC index into a structured visual triage object. The raw index row preserves filing metadata but offers little visual differentiation. The deterministic summary compresses the row but removes spatial separation among identity, evidence, priority, and source. The deal card retains those elements in distinct regions. This organization follows Shneiderman's (1996) overview-and-details principle and Few's (2006) dashboard design logic, and it is consistent with financial-platform studies that emphasize interpretable risk information and integrated data (Z. Li et al., 2025; Zhang et al., 2025). The present structural audit does not establish that users read the card faster or trust it more.

The rule comparison clarifies the scope of each cue. FinTech terms alone select many financial-sector filings that are unrelated to M&A, while companion forms alone include transaction contexts outside the FinTech cue set. The integrated configurations reproduce the operational target because they contain both variables used to construct it. Their perfect scores therefore confirm implementation consistency rather than independent detection performance. In design terms, the intersection links a sector cue with a transaction-neighborhood cue and shows both reasons on the card.

The source-grounded payload also addresses a recurring challenge in language-model interfaces. Brown et al. (2020) and Lewis et al. (2020) explain why LLMs and retrieval-augmented generation are attractive for transforming text, while Ji et al. (2023) documents the risk of unsupported generation. Financial retrieval studies extend this concern to corporate memo and disclosure settings, where numerical errors or missing evidence can alter the meaning of a risk communication (Q. Wu et al., 2023; Zhou et al., 2023). The present card avoids that issue by using deterministic wording over a fixed payload. A future language layer could be evaluated against the same source fields, but it would require direct output auditing before any claims about generated explanations (Zhong et al., 2026).

The interface schema therefore precedes any optional language formatter. Source fields populate the schema, derived fields are labeled as rules, and the static renderer controls card-ready wording. This ordering makes the visual system inspectable and reduces the chance that fluent prose will be mistaken for verified disclosure content. Design-rationale and evidence-card research similarly treats the model as a formatter or critic over a structured brief rather than as the source of truth (Jin, 2025b; Y. Li et al., 2025; Q. Wu et al., 2025).

The graphic design contribution is the visual grammar formed by identity anchors, evidence chips, a rule-derived display priority, and a source path. Tufte (2001), Ware (2012), and Cairo (2016) emphasize that visualizations should communicate without distorting evidence. The card follows that principle by separating observed index fields from derived interface cues (Zhong, Wu, et al., 2025). The display priority indicates whether the two operational signals are present;

it is not a probability, a measure of deal success, or a legal assessment. The review action directs the reader to the source filing and matched forms; it is not a claim about the filing's actual item numbers. Risk-card research in adjacent LLM-agent interfaces likewise separates labels, evidence, and human review actions (Su et al., 2026).

Nielsen's (1994) recognition principle and Munzner's (2014) abstraction framework provide a rationale for making transaction-neighborhood terms visible rather than leaving them implicit in a filing list. The relevant analytical sequence is to notice a candidate, inspect its evidence, and decide whether to open the primary filing. The card maps those tasks to visible regions. This mapping is a design proposition; the study did not test whether people perform the sequence more accurately or quickly.

Algorithmic triage and graphic design address different parts of the workflow. A classifier can rank records, while the interface determines which evidence accompanies that ranking. In this study, the integrated score is not independent because it reconstructs a label defined by the same cues. The design value lies in exposing the FinTech cue, companion-form context, filing date, and source path instead of displaying a score alone. Financial-search reranking and disclosure-review-card studies support this auditability stance by exposing ranking cues and evidence links (Meng et al., 2026; Zhang et al., 2025).

The structural metrics complement character count by documenting fields, visual anchors, and configured decision cues. They describe differences among the three representations but do not establish that one condition is easier, faster, more useful, or more trustworthy for users. A user study is necessary before the static card can be described as an effective dashboard.

LIMITATION

The study evaluates a filing-index and companion-form layer. It does not estimate transaction value, market reaction, or post-merger performance, and it does not replace legal interpretation of a filing. The M&A-neighborhood flag is based on included forms around the same CIK and date window; it is not a manually adjudicated deal database. The rule is suitable for constructing a transparent review queue, but it should not be interpreted as a definitive catalog of completed acquisitions. The exact F1 obtained by the structured intersection and integrated feature sets is likewise definitional rather than evidence of external predictive validity. This distinction is consistent with evidence-calibrated and numerical-guardrail research that avoids turning a triage signal into a legal or market outcome (Z. Li et al., 2026; Zhong, Chen, et al., 2025).

The FinTech cue model is based on company-name terms rather than SIC codes or business descriptions. This makes each cue visible and reproducible, but it can include broad financial-services entities and miss firms whose names do not reveal payment, lending, banking, or digital-asset activity. The card preserves the matched term so that readers can see why a record entered the view. Industry codes, entity profiles, and primary-document extraction could add context while retaining the same provenance distinction (Chen et al., 2025; Zhou et al., 2024). The interface audit is computational. Character length, field population, visual anchors, and decision cues are properties assigned or counted by the authors. They do not measure comprehension, trust, task accuracy, cognitive load, or review time. Accordingly, the results support only the structural description of the data-to-card transformation and cannot establish real dashboard effectiveness.

CONCLUSION

This study developed a source-grounded visual triage framework that converts complete Q4 2025 SEC Form 8-K index records and nearby companion filings into structured FinTech deal cards. The corpus contained 16,439 filings, including 2,820 company-name FinTech matches, 1,250 seven-day M&A-neighborhood records, and 222 records in their intersection across 94 company names. Integrated configurations reproduced the rule-defined target exactly because they used the same FinTech and companion-form cues from which that target was constructed; F1 = 1.000 is therefore a rule-consistency result rather than evidence of predictive superiority. The structural audit showed that the card schema contained more documented fields, visual anchors, and configured decision cues than the two computational baselines, but these counts do not establish gains in comprehension, trust, or review speed. The contribution is a transparent method for organizing filing-stream evidence for visual triage, not a validated AI detector or a user-tested dashboard.

FUTURE RESEARCH

Future research can test the framework against independently adjudicated FinTech M&A cases and evaluate the interface with analysts, investors, and disclosure professionals. A within-participant comparison of the raw row, deterministic summary, and deal card could measure task accuracy, completion time, source verification, trust, and workload. Filing-level SEC items and primary-document passages could also replace the current review-action rule. If a language model is added, its outputs should be assessed separately for supported claims, omissions, contradictions, numerical fidelity, and source attribution (Chen & Xu, 2026).

REFERENCES

Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1908.10063

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), Advances in neural information processing systems (Vol. 33, pp. 1877–1901). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

Cairo, A. (2016). The truthful art: Data, charts, and maps for communication. New Riders.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (Eds.). (1999). Readings in information visualization: Using vision to think. Morgan Kaufmann.

Chen, Y., & Xu, H. (2026). Trust-calibrated multilingual RAG for humanitarian information platforms: Empirical evaluation on OMoS-QA for migration information access. International Journal of Graphic Design, 4(1), 141–164. https://doi.org/10.51903/ijgd.v4i1.3552

Chen, Y., Zhou, S., & Lin, E. (2025). Accounting-aware evidence retrieval for institutional due diligence of tokenized trade receivable RWA. Journal of Technology Informatics and Engineering, 4(3), 649–663. https://doi.org/10.51903/jtie.v4i3.542

Elena, S., & Alexander, L. (2026). Simulating Sustainable Color-Form Decisions in AI-Driven Eco-Brand Identity Design. International Journal of Graphic Design, 4(1), 88–104. https://doi.org/10.51903/IJGD.V4I1.3313

Few, S. (2006). Information dashboard design: The effective visual communication of data. O'Reilly Media.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585, 357–362. https://doi.org/10.1038/s41586-020-2649-2

Heer, J., Bostock, M., & Ogievetsky, V. (2010). A tour through the visualization zoo. Communications of the ACM, 53(6), 59–67. https://doi.org/10.1145/1743546.1743567

Hullman, J., & Diakopoulos, N. (2011). Visualization rhetoric: Framing effects in narrative visualization. IEEE Transactions on Visualization and Computer Graphics, 17(12), 2231–2240. https://doi.org/10.1109/TVCG.2011.255

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), Article 248. https://doi.org/10.1145/3571730

Jin, J. (2025a). Evidence-chain reliable RAG: Hallucination detection, source attribution, and deterministic provenance explanations. Journal of Technology Informatics and Engineering, 4(2), 520–533. https://doi.org/10.51903/jtie.v4i2.535

Jin, J. (2025b). LLM-style evidence cards for scientific search interfaces: A UI/UX design framework for retrieval transparency, ranking trust, and visual evidence hierarchy. International Journal of Graphic Design, 3(2), 397–414. https://doi.org/10.51903/ijgd.v3i2.3698

Lerman, A., & Livnat, J. (2010). The new Form 8-K disclosures. Review of Accounting Studies, 15(4), 752–778. https://doi.org/10.1007/s11142-009-9114-7

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), Advances in neural information processing systems (Vol. 33, pp. 9459–9474). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html

Li, C., Bai, J., & Wang, S. (2024). Evidence-chain reliable RAG: Word-level hallucination detection, source attribution, and provenance explanation for LLM applications. Journal of Advanced Computing Systems, 4(2), 76–92. https://doi.org/10.69987/JACS.2024.40207

Li, C., Su, W., & Zhang, E. (2023). Lightweight hallucination firewall for enterprise LLM applications: Evidence consistency, self-checking, and small-model detection on TruthfulQA. Journal of Advanced Computing Systems, 3(1), 49–65. https://doi.org/10.69987/JACS.2023.30104

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics, 45(2–3), 221–247. https://doi.org/10.1016/j.jacceco.2008.02.003

Li, Y., Lu, S., & Zhao, L. (2025). LLM-as-design-critic: Aligning AI-generated UI feedback with human graphic design judgment. International Journal of Graphic Design, 3(1), 196–215. https://doi.org/10.51903/ijgd.v3i1.3661

Li, Z., Zhang, K., & Wong, A. (2026). Numerical-reasoning guardrails for a quant research assistant: A compact reproducible benchmark using SEC and FRED data. Journal of Technology Informatics and Engineering, 5(2), 75–90. https://doi.org/10.51903/jtie.v5i2.541

Li, Z., Zhou, S., & Zhou, Z. (2025). Financial risk dashboard design for institutional RWA investors: Visual hierarchy, chart comprehension, and explainability in FinChart-Bench. International Journal of Graphic Design, 3(1), 196–210. https://doi.org/10.51903/ijgd.v3i1.3715

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

McKinney, W. (2010). Data structures for statistical computing in Python. In S. van der Walt & J. Millman (Eds.), Proceedings of the 9th Python in Science Conference (pp. 56–61). https://doi.org/10.25080/Majora-92bf1922-00a

McMullin, J. L., Miller, B. P., & Twedt, B. J. (2019). Increased mandated disclosure frequency and price formation: Evidence from the 8-K expansion regulation. Review of Accounting Studies, 24(1), 1–33. https://doi.org/10.1007/s11142-018-9462-2

Meng, S., Chen, J., & Zheng, I. (2026). LLM-inspired offline reranking for financial search: Query rewriting, hybrid retrieval, and listwise relevance ranking on FiQA. Journal of Technology Informatics and Engineering, 5(1), 361–378. https://doi.org/10.51903/jtie.v5i1.537

Mu, J., Lu, Y., & Hwang, E. (2026). Structured visual brief interfaces for advertising design: A UI/UX framework for turning creative intentions into designer-editable graphic design cards. International Journal of Graphic Design, 4(1), 192–208. https://doi.org/10.51903/ijgd.v4i1.3702

Munzner, T. (2014). Visualization analysis and design. CRC Press. https://doi.org/10.1201/b17511

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In B. Adelson, S. Dumais, & J. Olson (Eds.), CHI '94: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 152–158). Association for Computing Machinery. https://doi.org/10.1145/191666.191729

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(85), 2825–2830. https://jmlr.org/papers/v12/pedregosa11a.html

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In Proceedings 1996 IEEE Symposium on Visual Languages (pp. 336–343). IEEE Computer Society Press. https://doi.org/10.1109/VL.1996.545307

Su, W., Chen, S., & Zhao, C. (2025). Budgeted multi-hop retrieval agent for compositional question answering: A retrieval-policy evaluation on the official MultiHop-RAG benchmark. Journal of Technology Informatics and Engineering, 4(3), 649–662. https://doi.org/10.51903/jtie.v4i3.543

Su, W., Rao, H., & Ma, E. (2026). Privacy and data-integrity risk cards for LLM agents: A UI/UX design framework for secure human oversight under prompt-injection attacks. International Journal of Graphic Design, 4(1), 186–191. https://doi.org/10.51903/ijgd.v4i1.3699

Tufte, E. R. (2001). The visual display of quantitative information (2nd ed.). Graphics Press.

U.S. Securities and Exchange Commission. (2026). EDGAR full-index archive [Data set]. Retrieved July 23, 2026, from https://www.sec.gov/Archives/edgar/full-index/

U.S. Securities and Exchange Commission. (2024a, June 6). EDGAR application programming interfaces (APIs). https://www.sec.gov/search-filings/edgar-application-programming-interfaces

U.S. Securities and Exchange Commission. (2024b, June 24). Exchange Act Form 8-K. https://www.sec.gov/rules-regulations/staff-guidance/compliance-disclosure-interpretations/exchange-act-form-8-k

U.S. Securities and Exchange Commission. (2025). Form 8-K. https://www.sec.gov/files/form8-k.pdf

Ware, C. (2012). Information visualization: Perception for design (3rd ed.). Morgan Kaufmann.

Wu, Q., Bai, J., & Zhou, X. (2023). Evidence-grounded financial RAG: Reducing numerical hallucination in LLM-generated corporate risk memos. Journal of Advanced Computing Systems, 3(3), 65–84. https://doi.org/10.69987/JACS.2023.30306

Wu, Q., Meng, S., & Zhao, J. (2025). Text-grounded LLM-assisted design rationale interfaces: Turning advertising layout metadata into explainable UI/UX decision cards. International Journal of Graphic Design, 3(1), 216–240. https://doi.org/10.51903/ijgd.v3i1.3713

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). BloombergGPT: A large language model for finance [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2303.17564

Xin, Q. (2026a). Explainable and fair credit risk scoring with counterfactual explanations: A reproducible evaluation on the German Credit dataset (HELOC-motivated). J-INTECH (Journal of Information and Technology), 14(2), 215–231. https://doi.org/10.32664/j-intech.v14i02.2228

Xin, Q. (2026b). Self-supervised customer representation learning for segmentation and next-purchase prediction on UCI Online Retail. J-INTECH (Journal of Information and Technology), 14(1), 20–37. https://doi.org/10.32664/j-intech.v14i01.2229

Yang, H., Liu, X.-Y., & Wang, C. D. (2023). FinGPT: Open-source financial large language models [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2306.06031

Zhang, K., Chen, Y., & Qian, A. (2025). Evidence-grounded accounting disclosure review cards: A visual communication framework for LLM-style explanations over SEC financial statements and notes. International Journal of Graphic Design, 3(2), 395. https://doi.org/10.51903/ijgd.v3i2.3710

Zhang, K., Meng, S., & Zhou, E. (2023). Evidence-grounded trading desk risk memos over SEC filings: Retrieval-augmented generation with XBRL numeric verification. Journal of Advanced Computing Systems, 3(2), 60–76. https://doi.org/10.69987/JACS.2023.30205

Zhong, Z. S., Chen, J., Zhong, E., & Sun, X. (2025). Evidence-calibrated RAG for unanswerable question answering: Retrieval coverage, abstention calibration, and hallucination-proxy analysis on SQuAD 2.0. Journal of Technology Informatics and Engineering, 4(2), 502–520. https://doi.org/10.51903/jtie.v4i2.536

Zhong, Z. S., Li, C., & Rao, H. (2026). Trajectory reliability prediction for generalist AI agents: Tool-use failure analysis and success forecasting on ZClawBench. Journal of Technology Informatics and Engineering, 5(1), 341–360. https://doi.org/10.51903/jtie.v5i1.539

Zhong, Z. S., Wu, Q., & Mi, G. (2025). Uncertainty-aware medical image explanation cards: LLM-generated visual explanations for AI-assisted radiology interfaces. International Journal of Graphic Design, 3(2), 415–436. https://doi.org/10.51903/ijgd.v3i2.3616

Zhou, S., Li, Z., & Wang, E. (2023). Evidence-grounded RAG for tokenized trade receivable disclosure QA under U.S. capital market standards. Journal of Advanced Computing Systems, 3(7), 41–57. https://doi.org/10.69987/JACS.2023.30704

Zhou, S., Li, Z., & Wang, E. (2024). Long-document RAG for contractual and insurance clause analysis in receivables RWA structures. Journal of Advanced Computing Systems, 4(8), 88–104. https://doi.org/10.69987/JACS.2024.40810