From Workshop Transcripts to Requirements Models
A Heuristic and RAG-Based Approach for Reusing Elicitation Data

Roxana L. Quintanilla Portugal[1][0000-0001-7693-5353] , Julio C. S. P. Leite[0000-0002-0355-0265], Lyrene Silva[0000-0003-1772-6062] and Edgar Sarmiento [0000-0002-0956-3091]

[1] Departamento Académico de Informática, UNSAAC, Cusco, Peru
[2] Escola Politécnica da USP, São Paulo, Brazil
[3] Universidade Federal do Rio Grande do Norte, UFRN, Natal, Brazil
[4] Universidad Nacional de San Agustín de Arequipa, Peru

roxana.quintanilla@unsaac.edu.pe,julioleite@usp.br, lyrene.silva@ufrn.br, esarmientoca@unsa.edu.pe

Abstract. Elicitation activities such as workshops generate rich qualitative data; however their reuse in Requirements Engineering (RE) remains limited due to the lack of explicit structure and traceable interpretation. While modeling approaches such as i* and the NFR Framework define modeling constructs, they do not make explicit how model elements are derived from elicitation data. In this paper, we present an industrial case based on workshop transcripts and conduct an exploratory exercise of how RE engineers derive goals, tasks, and quality concerns from the same textual input. From a limited but representative sample, we identify recurring patterns in their interpretive reasoning and derive an initial set of heuristics that make this reasoning explicit. We further illustrate how these heuristics can be operationalized within a Retrieval-Augmented Generation (RAG) pipeline, enabling semantic and semi-automated processing grounded in RE engineers' interpretations while preserving traceability. Our results suggest that, in this exploratory setting, making interpretative reasoning explicit supports the reuse of elicitation data through requirements models.

Keywords: Elicitation Data, Requirements Modeling, Interpretative Reasoning, Data Reuse, Qualitative Data, Semantic Analysis.

1 Introduction

Elicitation activities such as workshops and interviews are widely used in Requirements Engineering (RE) to capture stakeholders' perspectives. These activities generate rich qualitative data, often in the form of transcripts or notes. However, such data is often difficult to reuse in practice due to its lack of explicit structure and the need for significant interpretation [1][2]. One key reason is that transforming elicitation data into requirements models depends heavily on the RE engineers' interpretation, which is typically implicit and difficult to trace.

This challenge becomes particularly relevant in the context of FAIR principles [3], where data reuse requires not only access to data but also transparency in how it is

Proceedings of the 29th Workshop on Requirements Engineering (WER2026),
August 11-13, 2026, La Plata, Argentina.
https://doi.org/10.29327/1884517.29-30

interpreted. In many cases, elicitation datasets contain sensitive information, making controlled and traceable reuse even more important.

In practice, different RE engineers may construct different models from the same data, deciding what constitutes a goal, a task, or a quality concern, and how these elements relate. While modeling frameworks such as i* [4] and the NFR Framework [5] define these concepts and relationships, they do not make explicit how RE engineers should interpret discourse to produce modeling elements. This process is often supported by heuristic reasoning [6] or linguistic cues [7][8], but these approaches do not capture the diversity of RE engineers' viewpoints.

Approaches such as the LEL [9] emphasize preserving domain language and links to the original discourse. However, intentional modelling still require transforming textual expressions into modeling constructs. Thus, traceability is needed across this transformation, so that model elements remain grounded in their source evidence. IRES [9] structure reasoning about requirements, but does not capture the diversity of RE engineers' interpretations observed in elicitation data.

Recent NLP4RE and LLM4RE work has explored embeddings, semantic similarity, text generation, and modelling support for requirements-related text [10][11]. However, elicitation documents and RAG-driven approaches remain comparatively underexplored in LLM4RE. Our focus is on making explicit the interpretive reasoning that links workshop evidence to intentional modelling elements.

In this paper, we address this challenge through an industrial case involving workshop transcripts collected in the context of a project on responsible AI in journalism [12]. We used a short excerpt independently interpreted by multiple RE engineers. By comparing these interpretations, we identify recurring reasoning patterns used to transform textual evidence into requirements concepts.

Based on this task, we propose an initial set of heuristics that make interpretive reasoning explicit. These heuristics capture how RE engineers move from textual expressions, such as dissatisfaction, perceived limitations, or user experiences, to modeling elements such as goals, tasks, and quality concerns. We further illustrate how these heuristics can be used within a Retrieval-Augmented Generation (RAG) pipeline to support traceable and semi-automated reuse of elicitation data.

The main contributions of this paper are: a) the identification of a gap in making interpretive reasoning explicit in requirements modeling; b) a set of heuristics derived from RE engineers' interpretations of elicitation data; c) the operationalization of these heuristics through structured examples; and d) and illustration of how their use in a RAG-based pipeline for traceable and reusable analysis.

The remainder of the paper is structured as follows. Section 2 describes the industrial case and dataset. Section 3 presents the proposed heuristics and illustrates their use in a RAG-based pipeline. Section 5 discusses conclusion and lessons learned.

2 Industrial Case

This work is grounded in an industrial case conducted in the context of the project "Towards responsible AI in local journalism" [12]. As part of this project, participatory co-creation workshops were organized with journalists to explore their needs, challenges, and possible AI-based support. The sessions were audio-recorded and transcribed, resulting in a corpus of qualitative data capturing stakeholders' perspectives in their own words.

Due to the sensitive nature of the data, publication requires multi-level anonymization. This limits the possibility of openly sharing the full corpus and motivates controlled reuse mechanisms. For this reason, we consider a RAG-based setting in which the data remains within a protected environment and only selected evidence is retrieved for analysis.

The complete dataset comprises approximately 10 hours of workshop transcripts; this paper focuses on a short translated excerpt selected for its rich expressions of user experiences, perceived limitations, and expectations regarding AI-supported journalism. The results presented in this paper is exploratory in nature and focuses on a limited portion of the available material.

Beyond privacy constraints, an additional motivation for this work is the need to support the reuse of elicitation data. In the the project, i* and the NFR models were used to structure discussions and summarize results [13][14][15]. These models provided initial evidence of how elicitation data can be abstracted into reusable representation, a capability also explored in prior work on reusable goal models[16x]. However, the workshops also revealed a limitation: models alone do not capture how RE engineers identified, selected, and justified abstraction from stakeholders' discourse. This motivates our focus on traceable interpretative reasoning.

3 Solution and Results

To illustrate the interpretative task performed by requirement engineers when transforming elicitation data into modeling elements, one of the authors selected an excerpt from the workshop transcript. The excerpt was chosen because it contained explicit expressions of needs, problems, user reactions, and limitations of current practices. It was presented to senior RE engineers with experience in intentional modelling for independent interpretation.

RE.eng.1 (see Table 1), initially provided a focused example of a possible modeling insight derived from the excerpt. This example was intended to illustrate the nature of the task rather than to produce an exhaustive analysis. The remaining participants developed their own independent interpretations, with varying levels of detail and abstraction. To provide a common starting point while preserving interpretive freedom, participants were guided through four steps:

1. Identifying relevant statements in the discourse, focusing on expressions that indicate problems, user reactions, or limitations of current practices;
2. Interpreting the meaning of these statements by identifying the concerns or issues expressed by participants;
3. Mapping the identified elements to requirements concepts, such as actors, goals, tasks, and quality concerns; and
4. Making explicit the interpretive rule used to transform textual expressions into modeling elements.

These steps are grounded in the practical experience of one of the co-authors in modeling elicitation data. They do not constitute a formal or validated method; rather, they serve as a reference point to support and compare how different RE engineers reasoned about the same textual input. The excerpt used in this exercise is presented below:

“We realized that simply translating our articles is not enough. Many migrant readers still feel that the stories are not written for them. The tone often feels distant, and as a result they do not engage with the content.”

3.1. RE Engineers´ Interpretations

The excerpt introduced in the previous subsection was independently evaluated by four RE engineers. Table 1 summarizes these interpretations, highlighting the progression from textual evidence to modeling elements across four levels: relevant statements, interpretation, requirements concepts, and applied heuristics.

The comparison reveals several differences. First, RE engineers focused on different aspects of the text: some empathize expressions of dissatisfaction to identify inefficient operationalizations, while others identify contextual elements, such as actors and objectives. Second, the level of abstraction vary, ranging from task-level elements to softgoals. Third, the reasoning strategy differs, with some interpretations relying on linguistic cues and others on more conceptual approaches. Fourth, most RE engineers explicitly stated the rule used to derive the model elements. When this rule is absent, the reasoning remains tacit.

Table 1. Comparison of RE engineers’ interpretations

Step	RE.Eng.1	RE.Eng.2	RE.Eng.3	RE.Eng.4
Relevant statements	“translating is not enough” → ineffective operationalization		Migrant readers don't engage with content; Writers realized that translation is not enough; Translations ignore cultural aspects	Migrant readers do not engage with the content; Writers realized that translation is not enough.
Interpretation	Ineffective operationalization affecting engagement		Audience not reached; readers feel distant/disconnected	Translation process does not consider migrant readers' profiles; Migrant readers do not engage with the content due to the tone.
Requirements concepts	Task: ineffective operationalization; Softgoal: engagement	Actors: editorial, translator, sociologists, readers; Goals: The translation should be as good as possible ; Softgoals: cultural context, engagement, readability, inclusiveness; Tasks: translate, improve translation, text adaptation, text use, news production, read news, comment news, improve adaptation; Resources: text, translated text, adapted text, comment	Actors: writers, readers; Goal: engagement; Tasks: change tone; Softgoal: sense of belonging	Actors: writers/readers; Goal: increase readership; Task: adapt tone; consider migrant profiles; Softgoals: engagement, readability, awareness, enjoyability
Heuristics	dissatisfaction with operationalization → unmet goal		Decompose into subject/problem/cause/solution → map to RE concepts	Subjects/nouns → actors; verbs → tasks; adjective/adverbs → softgoals

3.2. From Interpretations to Heuristics

The comparison presented in the previous section showed that different requirements engineers derived distinct modeling abstractions from the same textual evidence. While interpretations vary, they exhibit recurring and comparable reasoning patterns. The heuristics were consolidated by comparing recurring mappings across interpretations, rather than by measuring agreement quantitatively.

Table 2 presents the heuristic catalogue derived from the comparison. Each row describes one heuristic, its type of reasoning, the textual trigger that may activate it, the transformation logic used to move from text to modelling abstraction, the corresponding RE mapping, and its relation to IRES concepts. The IRES column is used as a conceptual anchor for the heuristic; the arrows indicate conceptual links or progressions between elements, rather than strictly temporal or causal sequences.

Table 2. Heuristic catalogue derived from the comparison

Heuristic Name	Type of heuristic	Trigger in text	Transformation Logic	RE Mapping	IRES
Actor Identification	Basic/Linguistic	Mentions of groups (users, readers, stakeholders)	Identify explicitly mentioned actors	Actor	Action → performs → Actor
Goal Inference from Dissatisfaction	Interpretive	not enough, fail, problem	dissatisfaction suggests an unmet goal or deficient practice	Goal/unmet goal	Necessity → Motivation → Goal
Softgoal from User Experience	Interpretive	feels, experience, distant, engage	subjective perceptions suggest quality concerns	Softgoal	Action → affects/achieve → Goal
Task from Problem Cause	Structural	cause of issue, action needed	infer an action that addresses the identified cause	Task	Goal → conducts → Action; Necessity (problem) → performs → Action
Verb-based Task Identification	Basic/Linguistic	action verbs	map explicit actions to candidate tasks	Task	Action
Concept Abstraction	Interpretive	specific issue descriptions	generalize from concrete expression to abstract quality concern	Softgoal	State of Affairs
Narrative Decomposition	Structural	problem–cause–effect pattern	split discourse into actor, issue, and response	Multiple (actor, goal, task)	Full IRES
Implicit Domain Inference	Interpretive	gaps between problem and explanation	infer missing contextual elements	Goal/Softgoal/Task	Action → Evolves → State of Affairs; State of Affairs →shapes → Neccesity

Table 3 provides an operational view of the heuristics by showing how concrete textual triggers can be transformed into candidate RE elements. The table preserves the correspondence with the heuristic catalogue in Table 2, but focuses on their practical application over the analyzed excerpt. Some mappings are close to the textual evidence, while others require additional interpretation, such as inferring a task from a perceived problem or from domain knowledge. The arrows indicate interpretive transformations rather than deterministic rules.

Table 3. Operational view of the heuristics

Trigger example	Heuristic applied	Candidate Inference	RE element
"many migrant readers"	stakeholder mention → actor	migrant readers	Actor
"translation is not enough"	dissatisfaction with operationalization → unmet goal	improve translation	Goal
"the tone feels distant"	user perception → quality concern	belonging	Softgoal
"they do not engage with the content"	user perception → quality concern	engagement	Softgoal
"translating our articles"	Verb-based Task Identification	translating articles	Task
"the tone feels distant"	problem cause → task (inferred corrective action)	adapt tone	Task
"[translating our articles is not enough]. [Many migrant readers still feel that the stories are not written for them]. [The tone often feels distant]"	problem → cause → solution (task/goal/softgoal)	adapt tone	Task
"stories are not written for them"; "Migrant readers do not engage with the content due to the tone."	Implicit Domain Inference	adapt content to migrant readers /lack of user profiling	Task/Goal

Rather than imposing a single correct interpretation, this heuristic catalogue and it operational view organize plausible reasoning observed across experts. This improves the traceability of the analytical process and provides an initial basis for reusing these heuristics in RAG-assisted contexts.

3.3. Applying the Heuristics in a RAG-based Pipeline

Fig. 1 structures the proposal as three SADT activities. First, RE engineers transform selected transcript excerpts into structured expert examples, linking evidence, heuristic, RE mapping, and inferred element. Second, the retrieval layer uses these examples as semantic seeds to retrieve related transcript chunks. Third, the heuristic-guided component combines retrieved evidence, stored heuristics, thresholds, and RE concepts to generate candidate RE elements with traceability links.

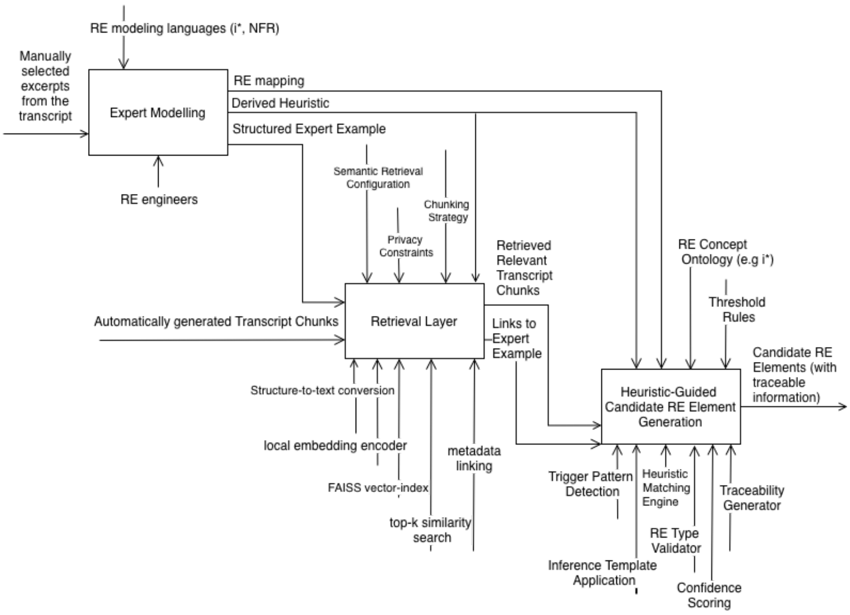

Fig. 1. SADT representation of a heuristic-guided pipeline supporting elicitation data reuse through traceable RE abstractions generation

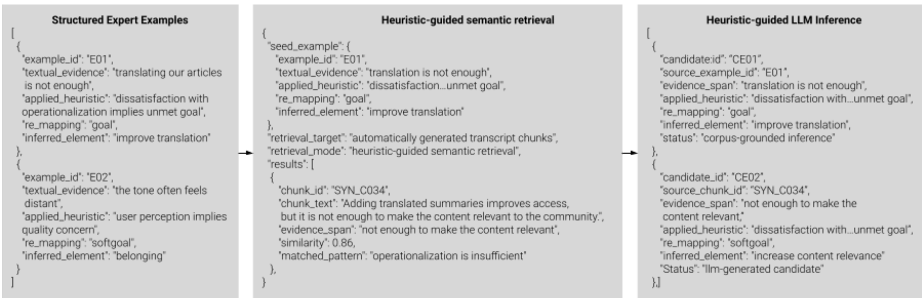

Fig. 2. Example of outputs of each stage: a structured expert example, retrieved transcript chunks, and the heuristic-guided LLM inference for candidate RE elements.

To illustrate this process, Fig. 2 shows the outputs of each activity in the SADT model. A structured expert example is not limited to textual evidence; it also includes the applied heuristic, the RE mapping, and the inferred element. This structured example is converted into a textual representation and used as a semantic seed to retrieve transcript chunks that express similar reasoning patterns, rather than merely similar wording. Finally, the retrieved chunks are processed by a local LLM to generate candidate RE elements subject to expert review, preserving links to the source example and chunk.

Since this is an initial architecture, retrieval quality, chunking strategy, confidence thresholds, and validation of retrieved evidence remain open issues for future implementation and evaluation.

4 Conclusion

This paper explored how recent AI technologies can support the structuring and systematic reuse of qualitative elicitation corpora. In particular, semantic retrieval combined with local generative AI deployment may enable the modeling processing of sensitive transcript data while respecting GDPR constraints and supporting FAIR-oriented reuse.

The industrial case showed that requirements modelling from elicitation data is inherently interpretive. Different RE engineers derived distinct yet valuable abstractions from the same textual evidence, confirming the need to preserve not only the resulting model elements, but also the reasoning that led to them.

Based on this observation, we proposed an initial SADT architecture in which expert-derived heuristics are represented as structured examples and used to guide semantic retrieval and candidate RE generation. This can help preserve expert reasoning, scale scarce modelling expertise, and make intentional modelling more feasible in industrial contexts..

However, the proposed architecture remains exploratory. Future work should provide synthetic examples and traceability schemas to support inspection and reuse. The approach does not remove analyst judgement: incorrect retrievals, biased heuristics, and overgeneralized LLM inferences remain risks requiring human validation.

Acknowledgments. Portugal thanks the Volkswagen Foundation for enabling access to data from the research project "Towards Responsible AI in Local Journalism" (grants 9B761 and 9B390). Silva thanks the PROPESQ-UFRN for financial support.

References

1. Diamantopoulos, T., Symeonidis, A.L.: Mining software requirements. In: Mining Software Engineering Data for Software Reuse, pp. 75–97. Springer, Cham (2020)
2. Feinberg, M., Sutherland, W., Nelson, S.B., Jarrahi, M.H., Rajasekar, A.: The new reality of reproducibility: The role of data work in scientific research. Proc. ACM Hum.-Comput. Interact. 4(CSCW1), Article 35, 1–22 (2020)
3. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., et al.: The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* 3(1), 1–9 (2016)
4. Franch, X., do Prado Leite, J.C.S., Mussbacher, G., Mylopoulos, J., Perini, A. (eds.): Social modeling using the i framework: Essays in honour of Eric Yu. Springer, Cham (2024)
5. Chung, L., Nixon, B.A., Yu, E., Mylopoulos, J.: Non-functional requirements in software engineering. Springer, Boston (2012)
6. Gonzalez-Baixauli, B., do Prado Leite, J.C.S., Laguna, M.A.: Eliciting non-functional requirements interactions using the personal construct theory. In: Proc. 14th IEEE Int. Requirements Engineering Conf. (RE'06), pp. 347–348. IEEE (2006)
7. Portugal, R.L.Q., do Prado Leite, J.C.S.: Usability related qualities through sentiment analysis. In: Proc. 1st Int. Workshop on Affective Computing for Requirements Engineering (AffectRE), pp. 20–26. IEEE (2018)
8. Younas, M., Abang Jawawi, D.N., Shah, M.A., Mustafa, A., Awais, M., Kamran Ishfaq, M., Wakil, K.: Elicitation of nonfunctional requirements in agile development using cloud computing environment. IEEE Access 8, 209153–209162 (2020)

9. Oliveira, A.P.A., Werneck, V.M.B., Cysneiros, L.M., do Prado Leite, J.C.S.: The philosophy behind IRES, an intentional requirements engineering strategy. In: Proc. Workshop on Requirements Engineering (WER) (2021)
10. Sabetzadeh, M., Arora, C.: Practical guidelines for the selection and evaluation of natural language processing techniques in requirements engineering. In: Ferrari, A., Ginde, G. (eds.) Handbook on Natural Language Processing for Requirements Engineering. Springer, Cham (2025)
11. Zadenoori, M.A., Dąbrowski, J., Alhoshan, W., Zhao, L., Ferrari, A.: Large language models (LLMs) for requirements engineering (RE): A systematic literature review.
12. Towards Responsible AI for local journalism. Available at: https://www.algorithmic.news (accessed April 2026)
13. Portugal, R.L.Q, Wilczek, B., Eder, M., Thurman, N., Haim, M.: Design thinking for journalism in the AI age: Towards an innovation process for responsible AI applications (2023)
14. Portugal, R.L.Q., Delle Ville, J., Antonelli, L.: Implementing accuracy quality for responsible AI in newsrooms. In: Proc. Workshop on Requirements Engineering (WER) (2024)
15. Portugal, R.L.Q., Silva, L.F., Sousa, H.P., Almentero, E.: Utilizing LLMs for early-stage qualities elicitation: A study on responsible fact-checking in journalism. In: *Proc. Workshop on Requirements Engineering (WER)* (2025)
16. Duran, M.B.: Reusable goal models. In: Proceedings of the 25th IEEE International Requirements Engineering Conference (RE), pp. 532–537. IEEE (2017)