Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 6, December 2026, pp. 5595~5605
ISSN: 2302-9285, DOI: 10.11591/eei.v15i6.12048

Cross-domain ontology linking user behavior and non-functional requirements

Sali Alas Majapahit[1,2], Purwanto[3], Farikhin[4]
[1]Doctoral Program of Information System, School of Postgraduate Studies, Diponegoro University, Semarang, Indonesia
[2]Department of Informatics Engineering, Faculty of Engineering, Pasundan University, Bandung, Indonesia
[3]Department of Chemical Engineering, Faculty of Engineering, Diponegoro University, Semarang, Indonesia
[4]Department of Mathematics, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia

Article Info

Article history:

Received Dec 6, 2025
Revised Sep 16, 2026
Accepted Oct 4, 2026

Keywords:

Cross-domain ontology
Non-functional requirement
Semantic similarity
User behavior
Web ontology language

ABSTRACT

Integrating user behavioral factors into non-functional requirements (NFRs) analysis remains difficult because NFRs are often abstract and lack explicit links to human factors. This study presents a cross-domain ontology that semantically maps user behavior dimensions to NFR quality attributes, enabling machine-readable cross-domain reasoning for early-stage NFR elicitation and user-aware requirement prioritization. Following the Methontology framework, knowledge was acquired from two sources: a behavioral dataset from a survey of 267 industrial-sector employees in Indonesia (9 dimensions; 36 indicators; validated via partial least squares structural equation modeling (PLS-SEM), average variance extracted (AVE): 0.746–0.807, CR: 0.898–0.926) and an ISO/IEC 25023-derived NFR indicator set (45 indicators across usability, performance efficiency, and security). Semantic similarity was computed using sentence-BERT (SBERT), producing 1,620 pairwise similarity scores (range: 0.20–0.85) instantiated as first-class individuals in a web ontology language-description logic (OWL-DL) ontology. Validation included reasoner-based consistency checking (Pellet: 2,490 ms), SPARQL competency question testing, and expert agreement evaluation by five domain experts (mean satisfaction: 80%; Cohen's Kappa: 0.37–0.58). The resulting ontology provides a reusable semantic backbone for user-aware NFR analysis and supports future integration with embedding-based semantic matching and intelligent requirement prioritization tools.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Sali Alas Majapahit
Doctoral Program of Information System, School of Postgraduate Studies, Diponegoro University
Imam Bardjo Road No. 5, Pleburan, Semarang Sel. District, Semarang 50421, Indonesia
Email: salialasm@students.undip.ac.id

1. INTRODUCTION

Modern software-intensive systems are judged not only by functionality but also by performance, security, and usability—collectively known as non-functional requirements (NFRs) [1]–[3]. NFRs are often poorly specified, ambiguous, and addressed late in development, causing rework and technical debt, especially when rapid delivery prioritizes features [3]–[5].

A key reason is the limited semantic traceability between NFR statements and the human factors that shape system acceptance and perceived quality [3]. User-behavioral signals (e.g., trust, perceived ease of use, risk perception, and habit) often determine whether a system is judged "usable" or "safe," yet these signals are rarely represented in a machine-readable form that can be queried together with NFR models [5], [6].

Recent research has improved NFR analysis through reviews, data-driven natural language processing/machine learning (NLP/ML) classifiers, and ontology-based methods that clarify concepts, standardize vocabularies, and support automated reasoning [7]–[10]. However, most methods only extract or classify NFRs from text, or represent NFR taxonomies and conflicts—they lack a cross-domain semantic bridge connecting user-behavior constructs to NFR measurement indicators in a reusable knowledge model [3], [6]–[9].

This paper addresses this gap. Existing ontology-based approaches—domain ontologies for requirement classification [8] and NFR modeling for specific industries [11]—show the value of formal knowledge representation, yet neither uses empirically grounded behavioral data nor provides a cross-domain semantic bridge to NFR indicators. This work fills that gap with a cross-domain ontology that maps user behavioral patterns to NFR indicators using sentence-BERT (SBERT) similarity scores as first-class individuals, enabling traceable, quarriable reasoning that neither prior NFR ontologies nor general SBERT-based classification approaches achieve together.

We designed and built the ontology in web ontology language (OWL) using a methodology-driven approach, populating it with a behavioral indicator dataset and an International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 25023-based NFR indicator set (author data) [12]–[14]. Cross-domain mapping is enabled by using SBERT similarity scores as first-class individuals, creating traceable links between behavior and NFR indicators [15], [16].

This study makes four contributions: i) a modular OWL ontology connecting user operations and NFR indicators through clear object and data properties; ii) a reproducible mapping pipeline identifying 1,620 behavior–NFR similarity relations; iii) validation via competency-question answering, reasoner consistency checks, and expert review; and iv) a discussion of key limitations, including scalability, similarity bias, and domain-specific dependencies. The research also suggests using ontologies as a foundation for future user-based requirements analysis and semantic matching models [13].

2. METHOD

2.1. Research approach

This research followed the Methontology framework to build the cross-domain ontology through six phases: i) specification and competency questions, ii) knowledge acquisition and dataset construction, iii) cross-domain mapping using SBERT similarity, iv) conceptualization, v) implementation in web ontology language-description logic (OWL-DL), and vi) evaluation and validation, as shown in Figure 1. Each phase is detailed from subsection 2.2 onward, drawing partly on the authors' earlier work.

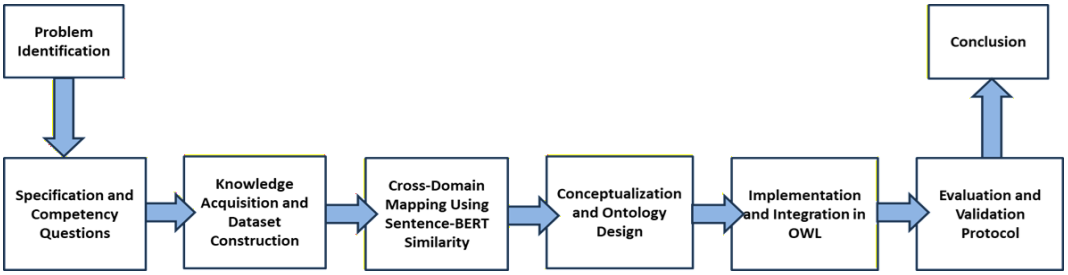


Figure 1. Research phase

2.2. Specification: scope and competency questions

Scope rationale (NFR selection): the NFR domain was scoped to three ISO/IEC 25023-aligned categories—usability, performance efficiency, and security—as they are architecturally significant, commonly shaped by user perception and interaction behavior [2], and frequently discussed in information-systems quality studies [17]. Other categories (e.g., maintainability and portability) were excluded as they require project-specific artifacts not yet grounded in behavioral indicators:

- Ontology goals: (G1) machine-readable behavior–NFR semantics, (G2) cross-domain querying and reasoning, and (G3) reuse in early-stage requirement analysis [18].
- Competency questions (CQs) define what the ontology should support: (CQ1) top behavior indicators linked to a specific NFR indicator; (CQ2) main NFR indicators for a given behavioral dimension; (CQ3) behavior factors affecting several NFR subdimensions; (CQ4) distribution of similarity mappings as weak, moderate, or strong; and (CQ5) derived relations or inverses enabling traceable reasoning [19].

2.3. Knowledge acquisition and dataset construction

Behavioral dataset: we operationalized nine behavioral dimensions into 36 indicators across seven domain clusters—cognitive, social-collective, trust and risk, cultural, digital readiness, economic-pragmatic, and post-adoption. A survey of industrial-sector employees in Indonesia yielded 472 responses; 267 were retained after screening (43.4% excluded), mainly for insufficient daily technology use (<5 hours), non-industrial occupation, or missing respondent data.

Participant demographics: of the 267 valid respondents, 62.2% were male, mostly under 31 years old (29.6%). By position, 58.1% were staff/coordinators/supervisors; 41.9% held a bachelor's degree, and 42.7% had 11–20 years of experience. Nearly all (98.1%) used technology ≥5 hours/day, drawn from seven industry sectors (mainly services 40.1% and manufacturing 21.7%) across ten Indonesian provinces, supporting generalizability of the findings.

Survey instrument validation: the 36 indicators were validated using two-layer partial least squares structural equation modeling (PLS-SEM) (reliability screening, then confirmatory factor analysis): all 29 reflective constructs exceeded thresholds for Cronbach's alpha (0.829–0.880), convergent validity (average variance extracted (AVE) 0.746–0.807), composite reliability (CR 0.898–0.926), and discriminant validity (CR<0.85); all 87 reflective indicators had outer loadings ≥0.70, and one formative construct (infrastructure access) was validated via variance inflation factor (VIF) and outer-weight significance. Each validated indicator was expressed as a short descriptive statement used as SBERT input (subsection 2.4).

NFR dataset: using ISO/IEC 25023 measurement concepts, we compiled 45 NFR indicators covering usability, performance efficiency, and security, each with a name, short description, and keywords for semantic matching [12].

2.4. Cross-domain mapping using sentence–concept resolution similarity

We used the SBERT all-mpnet-base-v2 model to generate 768-dimensional sentence embeddings and computed cosine similarity, $\text{sim}(u,v)=(u\cdot v)/(\|u\|\|v\|)$, for each of the 36×45 behavior–NFR indicator pairs, yielding 1,620 scalar values, categorized post hoc into three bands—weak (0.00–0.39), moderate (0.40–0.69), and strong (0.70–1.00)—following conventions from prior SBERT-matching studies, where scores >0.70 indicate high relatedness and 0.40–0.69 partial overlap [16], [20]. Cross-domain scores are structurally lower than intra-domain ones because behavioral and engineering vocabularies differ, consistent with the observed distribution (range 0.20–0.85; majority 0.30–0.60). Since cosine similarity is deterministic for fixed embeddings, formal confidence intervals do not apply; we instead validated the thresholds via expert agreement (content validity ratio (CVR)=0.99; mean satisfaction=80%; Cohen's Kappa 0.37–0.58). Threshold calibration remains domain-dependent and is flagged as future work.

2.5. Conceptualization and ontology design

We structured the ontology into three modules—behavior.owl, nfr.owl, and relation.owl—with core classes BehaviorDimension, BehaviorIndicator, NFRDimension, NFRSubDimension, NFRIndicator, and SemanticSimilarityRel. Object properties connect dimensions to indicators and relation instances to their behavior/NFR pair, while data properties store identifiers, labels, and similarity scores. For example, CQ1 ('top behavior indicators for a given NFR indicator') is answered because each of the 1,620 pairs is a SemanticSimilarityRel individual carrying identifier, score (hasSimilarityScore), and strength-band properties; a SPARQL query then groups these by NFR indicator and orders by frequency. Encoding links as first-class individuals—rather than bare object properties—lets each relation carry multiple properties at once (subsection 3.1 details the resulting object properties).

Interoperability design: the taxonomy layer can be aligned with simple knowledge organization system (SKOS), and measurement semantics with semanticscience integrated ontology (SIO) patterns; upper ontology mapping descriptive ontology for linguistic and cognitive engineering/ basic formal ontology (DOLCE/BFO) is considered for future extension [21].

2.6. Implementation and integration in web ontology language

We implemented the ontology in OWL-DL and edited it in Protégé, integrating modules via internationalized resource identifiers (IRIs) and owl:imports to preserve modularity [13], [22]. The three OWL modules and SBERT scripts are openly available as reproducibility artifacts (data availability section).

2.7. Evaluation and validation protocol

We validated the ontology through: i) consistency checking with description logic (DL) reasoners, ii) competency questions (CQ) answering via SPARQL, iii) expert review of conceptual adequacy and domain relevance, and iv) structural quality metrics using ontology evaluation and pitfall-scanning frameworks [23], [24], plus runtime performance (Pellet reasoning time; SPARQL execution time for CQ1–CQ5 via RDFLib), reported in subsubsection 3.3.4.

3. RESULTS AND DISCUSSION

3.1. Ontology model overview

The ontology comprises three interoperable modules: behavior.owl (behavioral dimensions/indicators), nfr.owl (NFR dimensions/subdimensions/indicators), and relation.owl (instantiated behavior–NFR similarity relations with explicit scores). This modularity supports maintainability and partial reuse in other RE pipelines [21], [25]; Figure 2 shows the class hierarchy.

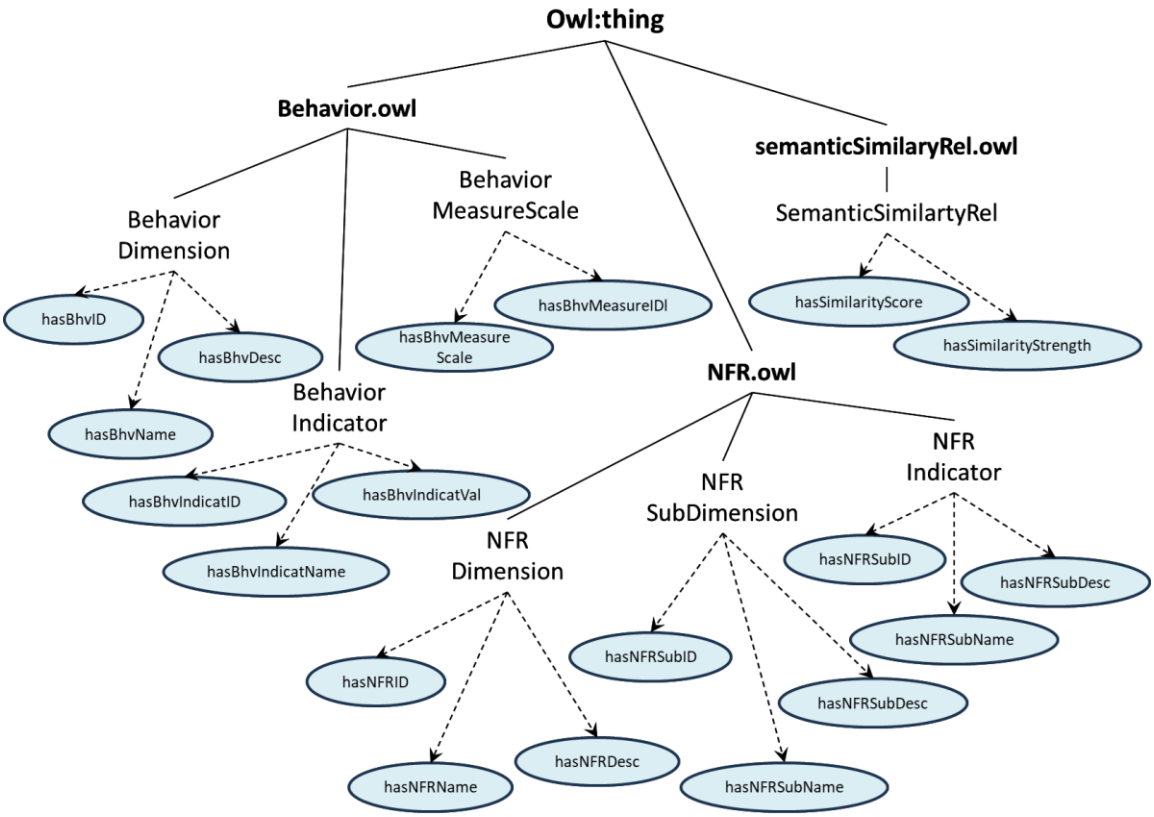


Figure 2. Class hierarchy diagram

To support reasoning regarding behavior–NFR semantics, the following components were created:

a. Object properties

Object properties connect classes/entities across domains, defining relationships between individuals, their applicable domains, and special characteristics such as inverses where relevant, see Table 1.

Table 1. Object properties definition

Object properties	Relationship type
RBDimensionToIndicator	BehaviorDimension to BehaviorIndicator; Cardinality 1 to n
RBIndicatorToMeasure	BehaviorIndicator to BehaviorMeasurementScale; Cardinality n to 1.
RNDimensionToSub	NFRDimension to NFRSubDimension; Cardinality 1 to n.
RNIndicatorToMeasure	NFRIndicator to NFRMeasurementScale; Cardinality n to 1.
RNSubToIndicator	NFRSubDimension to NFRIndicator; Cardinality 1 to n.
relatedToBehavior	SemanticSimilarityRel to BehaviorIndicator; Cardinality n to m.
relatedToNFR	SemanticSimilarityRel to NFRIndicator; Cardinality n to m.

b. Data properties

Data properties such as hasScore and hasStrengthLevel hold literal values of semantic relations; 25 primary data properties were designed, each described in detail (see data availability).

c. Instances

Each instance represents a real-world entity from the behavior, NFR, or semantic relationship domain, derived from the processing results as follows: i) Domain_Behavior: BehaviorDimension (9),

BehaviorIndicator (36), and BehaviorMeasurement (11) individuals; ii) Domain_NFR: NFRDimension (3), NFRSubDimension (13), NFRIndicator (45), and NFRMeasurement (10) individuals; iii) SemanticSimilarity: SemanticSimilarityRel (1,620) individuals; and iv) a complete list of instances is provided in the supplementary materials (see data availability).

d. Axioms

An axiom is a formal statement that defines rules, constraints, or logical relationships (e.g., subclass, domain–range, disjointness, functional, and semantic web rule language (SWRL) rules) that ensure the ontology reasons consistently. Each axiom type is briefly described below:

- TBox axioms: these axioms define class hierarchies, equivalence, and disjointness, forming the foundation for formal class-relatedness reasoning: 14 total (5 behavior, 2 NFR, 4 SemanticSimilarity, and 3 generalization).
- RBox axioms: RBox axioms define the relationship structure among classes—domain/range, object properties, and inverses—supporting valid cross-domain reasoning: 7 total (2 behavior, 3 NFR, and 2 SemanticSimilarity).
- ABox axioms: ABox axioms define concrete facts about instances across the behavior, NFR, and similarity domains, supporting reasoning validation of individuals, classes, relations, and literals: class assertion (1,747), object property assertion (176), and data property assertion (2,503).
- Restriction axioms: restriction axioms establish cardinality constraints, value constraints, and SWRL-based inference rules, enabling detection of logical inconsistencies and automatic classification: cardinality (7), value constraints (6), and SWRL rules (3).

A summary of all axioms is provided in the supplementary materials (see data availability).

e. Uniform resource identifier customization

We standardized the uniform resource identifier (URI) and global namespace (prefix http://example.org/ontology/) to keep resource description framework (RDF) entity URIs consistent and avoid redundancy or incompatibility between modules, ensuring compatible RDF structures and formal reasoning, see Table 2.

Table 2. URI component and global namespace

Component	Value
Ontology IRI	http://www.example.org/ontologi/BhvNFR
Default namespace	http://www.example.org/ontologi/BhvNFR#
Prefix	bhvNFR:
Resource examples	bhvNFR:BehaviorIndicator

Explanation of Table 2: i) ontology IRI: global address identifying the behavior–NFR ontology; ii) namespace: used for all ontology entities (class, property and individual); iii) prefix: shortened namespace for resource URIs; iv) the domain www.example.org can be customized; v) initialization occurs in the OWL file header (rdf:RDF declaration or ontology block); and vi) the header file is provided in the supplementary materials (see data availability).

3.2. Similarity landscape and behavioral implications

The SBERT-based similarity mapping was implemented in Python using pandas 1.5.3, numpy 1.23.5, nltk 3.8.1, scikit-learn 1.2.2, sentence-transformers 2.2.2, matplotlib 3.7.1, and seaborn 0.12.2, see Table 3

Table 3. Use of python libraries

Steps	Matching technique	Python library/module	Functions/transformer used
Preprocessing	All techniques	pandas, re, string	extract(), split(), token filtering
SBERT-semantic similarity	Sentence+concept	sentence-transformers	SentenceTransformer ('all-mpnet-base-v2'), util.cos_sim()
Analysis and summary	All techniques	pandas, groupby(), melt()	sum(),mean(),pivot(), manipulasikolom
Visualization	All techniques	matplotli, seaborn	barplot(), tight_layout(), savefig()

Scores ranged from 0.20–0.85, with most mappings falling in the moderate range, suggesting behavioral factors can serve as elicitation prompts and prioritization cues rather than direct proxies for NFR statements (complete results are available in the supplementary materials; see data availability).

3.3. Ontology evaluation summary

3.3.1. Web ontology language structure validation

We validated the ontology by counting classes, object properties, data properties, and individuals, confirming that all entities loaded without errors, the modular structure was unified, and URIs were standardized; reasoning was further validated in Protégé as follows: i) check ontology consistency using the RDFLib or Protégé validator and ii) validation covers classes, properties, and individuals. OWL validation results: number of classes: 7, number of object properties: 5, number of data properties: 2, and number of individuals: 1620. Notes: i) all entities were successfully loaded without RDF/extensible markup language (XML) parsing errors, ii) the modular structure was successfully merged, and the URI was standardized, and iii) for advanced reasoning validation, Protégé (HermiT/Pellet reasoner) is recommended.

3.3.2. Semantic consistency test

The SBERT-based similarity scores ranged from 0.20 to 0.61, with most mappings falling within the moderate band, indicating that behavioral and NFR descriptions are semantically related but not interchangeable—expected, since the constructs belong to different conceptual domains. In practice, moderate similarity suggests behavioral factors can serve as early elicitation prompts rather than direct replacements for NFR statements. Previous research has performed semantic similarity calculations. Complete artifacts are available in the supplementary materials (see data availability).

3.3.3. Competency validation

We tested the competency questions from the specification phase using SPARQL; we successfully answered all five questions (subsubsection 3.3.4). Although there were initial execution failures with modules loaded via `owl:imports`, Protégé—using HermiT/Pellet—was ultimately able to handle `owl:imports` correctly. We tested reasoning and inference, as well as semantic similarity and relational queries. Table 4 presents examples and the corresponding results. These queries demonstrate that the ontology can retrieve the most behavior-relevant NFR indicators, support threshold-based reasoning for prioritization, and reveal coverage gaps that guide further knowledge acquisition.

Table 4. SPARQL testing result examples

No.	Scenario	Description	Expected	C	Status
1	Inferensi automatic classification Query example: SELECT ?indikatorBehavior ?indikatorNFR ?score WHERE { ?rel rdf:type ns1:SemanticSimilarityRel ; ns1:hasSimilarityScore ?score ; ns1:hasBhvIndicat ?indikatorBehavior ; ns1:hasNFRIndicat ?indikatorNFR . FILTER (?score >= 0.8) }	If SimilarityScore≥0.80, then the instance must be in "HighSimilarityMapping."	BI_12, BI_14	Fit	Fit
2	Inferential relationship chaining Query example: SELECT ?indikatorBehavior ?indikatorNFR ?score WHERE { ?rel rdf:type ns1:SemanticSimilarityRel ; ns1:hasBhvIndicat ?indikatorBehavior ; ns1:hasNFRIndicat ?indikatorNFR ; ns1:hasSimilarityScore ?score . }	If BehaviorIndicator is related to TaskDifficulty and TaskDifficulty influences NFRPerformance, then BehaviorIndicator influences Performance.	Reasoner adds relations automatically	Fit	Fit
3	Query to verify the type of match (SBERT) Query example: SELECT ?indikatorNFR (COUNT(?indikatorB) AS ?jumlahIndikator) WHERE { ?rel rdf:type ns1:SemanticSimilarityRel ; ns1:hasBhvIndicat ?indikatorB ; ns1:hasNFRIndicat ?indikatorNFR.} GROUP BY ?indikatorNFR ORDER BY DESC(?jumlahIndikator)	SPARQL check MatchingType property	Distribution according to the preprocessing table	Fit	Fit

3.3.4. Runtime performance

On the full merged ontology (1,620 individuals), Pellet reasoning in Protégé completed in 2,490 ms. SPARQL execution for CQ1–CQ5 via RDFLib averaged 84.3 ms (range 0.41–244.86 ms), with full retrieval of all 1,620 individuals (CQ5) at 224.3 ms and threshold-filtered retrieval (CQ3, ≥0.50, 266 mappings) at 108.9 ms—both sub-second, serving as a reproducibility baseline for future benchmarking at larger scales (5,000–10,000 individuals).

3.3.5. Evaluation using expert

We assessed the conceptual adequacy of the SBERT-based categorization through an expert agreement approach [26], [27]: five domain experts—requirements engineers and a human–computer interaction (HCI) specialist—rated their agreement with the SBERT-assigned similarity categories, directly validating computational categorization against expert judgment.

Instrument and sampling: a stratified random sample of 86 pairs was drawn from the full 4,050-pair space (90 behavioral×45 NFR indicators), proportionally covering four bands: strong (sim >0.50, n=6, census), moderate (0.30–0.50, n=30), weak (0.10–0.30, n=30), and not relevant (≤0.10, n=20), split into two 43-pair sessions completed by all five experts, who were blinded to SBERT scores and rated each pair on a 5-point Likert scale across four dimensions: semantic relatedness, band appropriateness, cross-domain clarity, and RE domain relevance.

Statistical results: Shapiro–Wilk testing showed largely non-normal distributions, consistent with ordinal Likert data. Cronbach's Alpha showed good internal consistency (α>0.70); Spearman correlation between expert pairs was high (r_s>0.60); and receiver operating characteristic/ area under the curve (ROC/AUC) showed experts effectively differentiated similarity categories (AUC>0.70). CVR [26], with five experts and p<0.05, requires a critical CVR of 0.99 (≥4 of 5 experts rating ≥4/5); the resulting mean expert satisfaction of 80% confirms the SBERT categorization aligns well with expert judgment. Cohen's Kappa between expert pairs ranged 0.37–0.58 (mean 0.468), a moderate agreement [28], [29] acceptable for subjective ordinal evaluation across diverse disciplinary backgrounds [30] (see data availability).

3.4. Discussion

These results confirm that Methontology is effective for building complex cross-domain ontologies such as this behavior–NFR model, consistent with Zhou *et al*. [31], who adapted Methontology for real-time ontologies integrating multiple knowledge domains, and Liyanage *et al*. [32], who applied it to formalize a legal ontology in Sri Lanka, highlighting its flexibility in complex domains. Conceptually, Methontology's systematic process from requirements definition to evaluation supports knowledge traceability and semantic consistency—echoed by Junior *et al*. [21], whose oil-industry ontology validation using Methontology matched expert manual verification results [21].

From an implementation standpoint, OWL enables automated reasoning and semantic interoperability across systems, aligning with Medinacelli *et al*. [33], who showed OWL integration into model-based systems engineering (MBSE) supports automated consistency checking and cross-domain model sharing. Integrating SBERT further enhances the ontology's ability to detect ML-based semantic similarity across new contexts, consistent with Chu *et al*. [34] refined SBERT, which maps embeddings into manifold space for more stable similarity detection, and Bekamiri *et al*. [35], who used SBERT to dynamically measure patent semantic distance for adaptive classification and retrieval.

This use of SBERT for cross-domain mapping is consistent with its broader RE applications: Alhoshan *et al*. [36] achieved F1=0.72–0.80 for zero-shot NFR classification without task-specific training, and Das *et al*. [37] found SBERT outperforms word-level embeddings for requirement-similarity detection. This study extends that work by using SBERT not to classify requirements within one domain, but to bridge user behavioral and NFR indicators across domains—a use case prior RE applications have not addressed.

The predominantly moderate scores (0.30–0.60) reflect genuine semantic distance between the experiential language of behavioral indicators (e.g., 'perceived ease of use') and the engineering vocabulary of NFR indicators (e.g., 'task completion rate'), not a model limitation—a high cosine similarity (>0.70) between such divergent registers would paradoxically suggest insufficient domain specificity. Moderate-strength mappings are sufficient for the intended use: elicitation prompts and prioritization cues in early-stage NFR analysis, without claiming direct semantic equivalence.

As a usage illustration, a requirements analyst eliciting NFRs for an industrial IoT dashboard queries CQ1 for 'error recovery efficiency' (a usability sub-indicator) and retrieves three moderately similar behavioral indicators—'perceived difficulty recovering from system errors' (sim=0.58), 'habit of workaround behavior when system fails' (sim=0.51), and 'trust reduction after repeated task failures' (sim=0.47)—used as elicitation prompts rather than equivalences: testing error-recovery flow with low-self-efficacy users, watching for workarounds that mask failures, and addressing trust erosion through security and reliability NFRs. This turns a single NFR indicator into a behavioral risk map that neither a plain taxonomy nor a general SBERT classifier provides on its own. Thus, this model functions not only as a static knowledge representation but also as a semantic bridge between user behavior and system quality attributes.

Compared with existing RE ontologies, this work fills a gap left open by prior models: Alrumaih *et al*. [8] built a Methontology/OWL requirements-classification ontology—the same foundation used here—without incorporating behavioral semantics; Monaco *et al*. [11] modeled OWL 2 DL-based industrial-energy NFRs but left behavioral factors descriptive and empirically ungrounded. This study extends both by introducing a data-driven semantic similarity layer—from a 472-respondent behavioral survey and ISO/IEC 25023-derived NFR indicators—as first-class OWL-DL individuals, constituting the model's primary structural novelty.

As reported in subsubsection 3.3.5, inter-rater Kappa (mean 0.468) falls in the moderate range [28], [29], acceptable given experts' diverse disciplinary backgrounds [30]. Notably, the lowest pairwise agreement occurred for the security subdomain, which also showed lower SBERT similarity scores—

suggesting security-related behavioral indicators may need further domain-specific refinement. This direction aligns with emerging work on ontology learning from large language models [38].

Regarding the behavioral dimension, many constructs are drawn from technology acceptance model (TAM) and unified theory of acceptance and use of technology (UTAUT). However, certain constructs implying user behavior—such as emotion and motivation—could be considered for inclusion in the behavioral dimension [39].

4. CONCLUSION

This study presented a cross-domain ontology bridging user behavior semantics and NFRs by representing behavior indicators, NFR indicators, and their semantic similarity relations in a modular OWL-DL model, enabling cross-domain querying and early-stage, user-aware NFR analysis through competency-question-driven retrieval and threshold-based reasoning. The current version is limited by its focus on three NFR categories, reliance on embedding-based similarity heuristics, and the need for broader real-world empirical validation. Future work will expand ontology scope, publish reproducibility artifacts (OWL files, indicator lists, similarity matrices), benchmark reasoning/query performance at larger scales, and strengthen validation through comparative ontology evaluation.

This research opens up several opportunities for further development, including: i) ontology integration with adaptive inference systems: developing a hybrid ML-based reasoning system for adaptive inference as the knowledge domain evolves, following an ontology-learning-from-data direction; ii) empirical evaluation of ontology-based systems: testing this ontology in real-world contexts (e.g., in system requirements analysis tools or user behavior analytics platforms) will provide external validation of its practical benefits; iii) domain ontology expansion and framework reusability: developing an upper ontology usable across domains such as e-commerce, education, and health will enhance reuse and semantic interoperability between systems; iv) integration with user cognitive and affective models: future research could incorporate user emotional and motivational dimensions, to enrich behavioral representations with cognitive and affective aspects; and v) to support discoverability and reuse, the ontology artifacts—the three OWL modules, behavioral and NFR indicator sets, SBERT similarity matrices, and SPARQL competency-question scripts—are available via Pasundan University's institutional repository and a supplementary Google Drive archive (data availability). Individual modules can be imported via owl:imports for partial adoption: behavior.owl alone can represent user behavioral dimensions in other industrial technology-adoption studies, while nfr.owl can serve as a standalone NFR taxonomy aligned with ISO/IEC 25023. Extensibility follows the modular design: new NFR categories can extend nfr.owl, new behavioral dimensions can extend behavior.owl, and new similarity relations can be instantiated in relation.owl by rerunning the SBERT pipeline, with the CQ1–CQ5 questions serving as a reuse contract any extension should keep answerable. Future studies can benchmark this ontology against existing NFR ontologies using established quality frameworks (e.g., class richness, inheritance depth, and relationship consistency). Planned alignment with upper ontologies (DOLCE, BFO) or vocabularies (SKOS, SIO) will further improve interoperability and enable automated ontology matching in the requirements engineering (RE) domain.

The theoretical implications of this research lie in extending the ontology-based requirements engineering (OBRE) approach by incorporating human behavioral dimensions into the NFR model. The practical implications include a formal ontology model that can be integrated into requirements management tools or semantic search engines to help analyze NFRs based on user behavior.

ACKNOWLEDGMENTS

The authors thank experts in requirements engineering and human-computer interaction who contributed to and validated the developed ontology model.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors' institutional affiliations at Diponegoro University and Pasundan University supported the work through internal research facilities and academic supervision.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sali Alas Majapahit	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓
Purwanto		✓			✓	✓		✓	✓	✓	✓	✓		
Farikhin	✓		✓	✓			✓			✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Reproducibility artifacts—including the three OWL ontology module files (behavior.owl, nfr.owl, relation.owl), SBERT similarity computation scripts (Python, all-mpnet-base-v2 configuration), the full 1,620-pair similarity matrix, SPARQL competency question scripts (CQ1–CQ5), and Appendices A to G — are openly available at: https://drive.google.com/file/d/1SxR3c0KF6lx0k1mgCMGScrrh2TC10h8e/view?usp=sharing. Researchers who want to reuse or extend the ontology can import individual modules via owl:imports declarations, as described in subsection 2.6.

REFERENCES

[1] M. A. Gondal, N. A. Qureshi, H. Mukhtar, and H. F. Ahmed, "An engineering approach to integrate Non-Functional Requirements (NFR) to achieve high quality software process," in *ICEIS 2020 - Proceedings of the 22nd International Conference on Enterprise Information Systems*, SCITEPRESS - Science and Technology Publications, 2020, pp. 377–384, doi: 10.5220/0009568503770384.

[2] ISO/IEC, "Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Product quality model," ISO, 2023, [Online]. Available: https://www.iso.org/standard/78176.html, (Accessed: Mar. 31, 2026).

[3] T. Olsson, S. Sentilles, and E. Papatheocharous, "A systematic literature review of empirical research on quality requirements," *Requirements Engineering*, vol. 27, no. 2, pp. 249–271, Jun. 2022, doi: 10.1007/s00766-022-00373-9.

[4] A. Melo, R. Fagundes, V. Lenarduzzi, and W. B. Santos, "Identification and measurement of Requirements Technical Debt in software development: A systematic literature review," *Journal of Systems and Software*, vol. 194, p. 111483, Dec. 2022, doi: 10.1016/j.jss.2022.111483.

[5] A. Nemeth and A. Bekmukhambetova, "Achieving Usability: Looking for Connections between User-Centred Design Practices and Resultant Usability Metrics in Agile Software Development," *Periodica Polytechnica Social and Management Sciences*, vol. 31, no. 2, pp. 135–143, Jun. 2023, doi: 10.3311/PPso.20512.

[6] Y. Hao and X. Yu, "Ontology-based Software Trustworthy Requirements and Behavior Modeling," in *Proceedings - 2021 International Conference on Networking, Communications and Information Technology, NetCIT 2021*, IEEE, Dec. 2021, pp. 460–466, doi: 10.1109/NetCIT54147.2021.00097.

[7] T. Wang, P. Liang, and M. Lu, "What Aspects Do Non-Functional Requirements in App User Reviews Describe? An Exploratory and Comparative Study," in *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*, IEEE, Dec. 2018, pp. 494–503, doi: 10.1109/APSEC.2018.00064.

[8] H. Alrumaih, A. Mirza, and H. Alsalamah, "Domain Ontology for Requirements Classification in Requirements Engineering Context," *IEEE Access*, vol. 8, pp. 89899–89908, 2020, doi: 10.1109/ACCESS.2020.2993838.

[9] J. Dąbrowski, E. Letier, A. Perini, and A. Susi, "Analysing app reviews for software engineering: a systematic literature review," *Empirical Software Engineering*, vol. 27, no. 2, p. 43, 2022, doi: 10.1007/s10664-021-10065-7.

[10] S. C. Necula, F. Dumitriu, and V. Greavu-Șerban, "A Systematic Literature Review on Using Natural Language Processing in Software Requirements Engineering," *Electronics (Switzerland)*, vol. 13, no. 11, p. 2055, 2024, doi: 10.3390/electronics13112055.

[11] R. Monaco, X. Liu, T. Murino, X. Cheng, and P. S. Nielsen, "A non-functional requirements-based ontology for supporting the development of industrial energy management systems," *Journal of Cleaner Production*, vol. 414, p. 137614, 2023, doi: 10.1016/j.jclepro.2023.137614.

[12] ISO/IEC, "Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of system and software product quality," ISO, 2016, [Online]. Available: https://www.iso.org/standard/35747.html, (Accessed: Mar. 31, 2026).

[13] W3C, "OWL 2 Web Ontology Language Document Overview (Second Edition)," [Online]. Available: http://www.w3.org/TR/owl2-overview/, (Accessed: Mar. 31, 2026).

[14] M. Fernandez, A. Gómez-Pérez, and N. Juristo, "Methontology: from ontological art towards ontological engineering," in *Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering*, 1997, pp. 33–40.

[15] B. Li, H. Zhou, J. He, M. Wang, Y. Yang, and L. Li, "On the Sentence Embeddings from BERT for Semantic Textual Similarity," *arXiv preprint*, 2020, doi: 10.48550/arXiv.2011.05864.

[16] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using siamese BERT-networks," in *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, Hong Kong, China, 2019, pp. 3982–3992, doi: 10.18653/v1/D19-1410.

[17] D. Mairiza, D. Zowghi, and N. Nurmuliani, "An investigation into the notion of non-functional requirements," in *Proceedings of the ACM Symposium on Applied Computing*, Sierre, Switzerland, 2010, pp. 311–317, doi: 10.1145/1774088.1774153.
[18] K. Mokos, T. Nestoridis, P. Katsaros, and N. Bassiliades, "Semantic Modeling and Analysis of Natural Language System Requirements," *IEEE Access*, vol. 10, pp. 84094–84119, 2022, doi: 10.1109/ACCESS.2022.3197281.
[19] C. M. Keet and Z. C. Khan, "Characterising Competency Questions for Ontologies," in *CEUR Workshop Proceedings*, Catania, Italy, 2025.
[20] Z. Liu, J. Zhu, X. Cheng, and Q. Lu, "Optimized Algorithm Design for Text similarity Detection Based on Artificial Intelligence and Natural Language Processing," *Procedia Computer Science*, vol. 228, pp. 195–202, 2023, doi: 10.1016/j.procs.2023.11.023.
[21] M. R. N. M. Junior, E. M. N. Gonçalves, S. S. da C. Botelho, and E. da S. D. Estrada, "Data Digitalization and Conformity Verification in Oil and Gas Industry Databooks Using Semantic Model Based on Ontology," in *Proceedings of the International Conference on Informatics in Control, Automation and Robotics*, 2023, pp. 180–187, doi: 10.5220/0012176700003543.
[22] B. Glimm, I. Horrocks, B. Motik, G. Stoilos, and Z. Wang, "HermiT: An OWL 2 Reasoner," *Journal of Automated Reasoning*, vol. 53, no. 3, pp. 245–269, 2014, doi: 10.1007/s10817-014-9305-1.
[23] M. Poveda-Villalón, A. Gómez-Pérez, and M. C. Suárez-Figueroa, "OOPS! (OntOlogy Pitfall Scanner!): An on-line tool for ontology evaluation," *International Journal on Semantic Web and Information Systems*, vol. 10, no. 2, pp. 7–34, 2014, doi: 10.4018/ijswis.2014040102.
[24] A. Duque-Ramos, J. T. Fernández-Breis, R. Stevens, and N. Aussenac-Gilles, "OQuaRE: A square-based approach for evaluating the quality of ontologies," *Journal of Research and Practice in Information Technology*, vol. 43, no. 2, pp. 159–176, 2011.
[25] M. Buranarach, S. R. Bhatta, and B. R. Ghimire, "Ontology-based knowledge management in income tax of Nepal," *International Journal of Metadata, Semantics and Ontologies*, vol. 15, no. 2, p. 144, 2021, doi: 10.1504/ijmso.2021.10044071.
[26] C. H. Lawshe, "A quantitative approach to content validity," *Personnel Psychology*, vol. 28, no. 4, pp. 563–575, 1975, doi: 10.1111/j.1744-6570.1975.tb01393.x.
[27] N. F. Noy and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," Stanford Knowledge Systems Laboratory Technical Report KSL-01-05, Stanford, CA, USA, 2001. [Online]. Available: https://protege.stanford.edu/publications/ontology_development/ontology101.pdf.
[28] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.
[29] W. Tang, J. Hu, H. Zhang, P. Wu, and H. He, "Kappa coefficient: a popular measure of rater agreement," *Shanghai Archives of Psychiatry*, vol. 27, no. 1, pp. 62–67, 2015, doi: 10.11919/j.issn.1002-0829.215010.
[30] M. L. McHugh, "Interrater reliability: the kappa statistic," *Biochemia Medica*, vol. 3, no. 22, pp. 276–282, 2012.
[31] J. Zhou, X. Song, Y. Li, Y. Gao, and X. Zhang, "Building Real-Time Ontology Based on Adaptive Filter for Multi-Domain Knowledge Organization," *IEEE Access*, vol. 9, pp. 66486–66497, 2021, doi: 10.1109/ACCESS.2021.3076833.
[32] C. Liyanage, T. Ilayperuma, J. J. Archchige, and F. A. Bukhsh, "Formalising Legal Knowledge of Sri Lankan Civil Appellate High Court Domain from Ontological Perspective," in *Advances in Conceptual Modeling (ER 2023)*, Cham, Switzerland: Springer, 2023, vol. 14319, pp. 216–226, doi: 10.1007/978-3-031-47112-4_18.
[33] L. P. Medinacelli, F. Noyrit, and C. Mraidha, "A Methodology for Knowledge Integration and Acquisition in Model-Based Systems Engineering," in *International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K - Proceedings*, 2023, pp. 264–271, doi: 10.5220/0012233900003598.
[34] Y. Chu, H. Cao, Y. Diao, and H. Lin, "Refined SBERT: Representing sentence BERT in manifold space," *Neurocomputing*, vol. 555, p. 126453, 2023, doi: 10.1016/j.neucom.2023.126453.
[35] H. Bekamiri, D. S. Hain, and R. Jurowetzki, "PatentSBERTa: A deep NLP based hybrid model for patent distance and classification using augmented SBERT," *Technological Forecasting and Social Change*, vol. 206, p. 123536, Sep. 2024, doi: 10.1016/j.techfore.2024.123536.
[36] W. Alhoshan, A. Ferrari, and L. Zhao, "Zero-shot learning for requirements classification: An exploratory study," *Information and Software Technology*, vol. 159, p. 107202, 2023, doi: 10.1016/j.infsof.2023.107202.
[37] S. Das, N. Deb, A. Cortesi, and N. Chaki, "Sentence Embedding Models for Similarity Detection of Software Requirements," *SN Computer Science*, vol. 2, p. 69, Apr. 2021, doi: 10.1007/s42979-020-00427-1.
[38] A. S. Lippolis *et al.*, "Ontology Generation Using Large Language Models," in *The Semantic Web. ESWC 2025*, vol. 15718 LNCS, E. Curry, M. Acosta, M. Poveda-Villalon, M. van Erp, A. Ojo, K. Hose, C. Shimizu, and P. Lisena, Eds., Cham, Switzerland: Springer, 2025, pp. 321–341, doi: 10.1007/978-3-031-94575-5_18.
[39] J. M. López-Gil, R. Gil, and R. García, "Web ontologies to categorialy structure reality: Representations of human emotional, cognitive, and motivational processes," *Frontiers in Psychology*, vol. 7, p. 551, 2016, doi: 10.3389/fpsyg.2016.00551.

BIOGRAPHIES OF AUTHORS



Sali Alas Majapahit settled in Bandung, Indonesia, was born in 1967, received his B.S. degree from the Department of Informatics Engineering, STT-Indonesia, Bandung, Indonesia, in 1996, and his M.S. degree from the Department of Information System Engineering, STMIK-Likmi, Bandung, Indonesia, in 2013. He is currently pursuing his Ph.D. degree in the Doctoral Program of Information Systems at the School of Postgraduate Studies, Diponegoro University, Semarang, Indonesia. His current research interests include IS development, web-based applications, and smart systems. He can be contacted at email: salialasm@students.undip.ac.id.



Purwanto currently works at the Department of Chemical Engineering, Faculty of Engineering, Diponegoro University in Semarang, Indonesia. His graduated with a bachelor's degree in Chemical Engineering from Universitas Diponegoro and a master's and a doctorate in Process Engineering from INPL Toulouse, France. His current projects include hazardous waste management, waste-to-product, electrochemical processes (electro-Fenton, electrocoagulation), green chemistry, eco-industrial estate, and information systems. He can be contacted at email: purwanto.profundip@gmail.com.



Farikhin earned his master's degree in Mathematics in 2003 and obtained a Ph.D. in Mathematical Computation in 2011, from UGM-Indonesia and UMT-Malaysia, respectively. His research interests include mathematical analysis-based computing, approximation methods/theory, and optimization/soft computing for data mining. He can be contacted at email: farikhin@lecturer.undip.ac.id.