2025 Nov 25. OHDSI Community Call: *Early-Stage Researcher Presentations*

ARKE: An Ontology-Driven Framework for Standardizing Radiology Procedure Terminology Using LLMs and RAG

Presenter: *Sumin Lee, Kyulee Jeon*

Leader: *Seng Chan You*

Yonsei University College of Medicine, Seoul, Korea

OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Automated Radiology Knowledge Encoding (ARKE)

ARKE: An Ontology-Driven Framework for Standardizing Radiology Procedure Terminology Using LLMS and RAG

PRESENTER: Sumin Lee, Kyulee Jeon

INTRO

- Radiology procedure codes are fragmented across institutions, embedding modality, anatomy, and contrast inconsistently. This heterogeneity blocks imaging data reuse and phenotype portability in OMOP CDM.
- Generic vocabularies (SNOMED CT, CPT-4, Korea's EDI) cover broad concepts but lack radiology-specific detail. The LOINC/RSNA Playbook provides structured mappings across 18 attributes, yet manual mapping is slow and resource-intensive.
- We developed ARKE, an ontology-driven framework combining knowledge graph retrieval, structured LLM prompting, and RAG for scalable, high-fidelity mapping of local codes to LOINC-RadLex.

METHODS

Data

- 2,126 local procedure terms from Severance Hospital (1,822 training; 304 validation → 237 valid after expert review).

Framework

- ***STEP 1. Knowledge Graph***: Constructed from 2025 LOINC/RSNA Playbook, encoding 18 attributes (e.g, modality, body region, contrast)
- ***STEP 2. Preprocessing***: Local terms translated into Eng, cleaned, and converted into RadLex PartType-based JSON with GPT-4o
- ***STEP 3. Candidate Retrieval***: TOP10 candidates retrieved using graph-based rules (Jaccard, F1, simple overlap, and a weighted match)
- ***STEP 4. Final code selection:*** LLM leverages RAG and CoT prompting to select the optimal mapping among candidates based on attribute consistency.

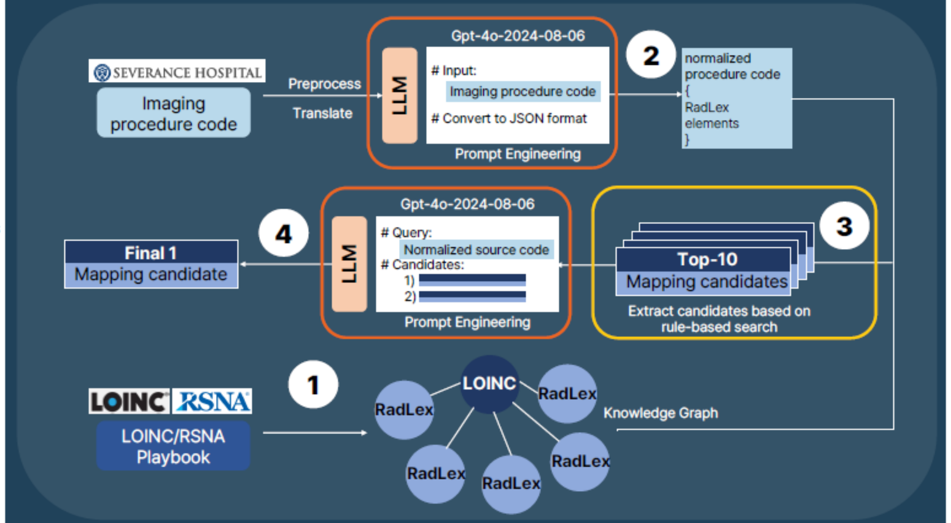
Evaluation

- 4 reviewers created a cross-checked reference.
- Metrics: accuracy, hit rate, MRR, NDCG@10.

RESULTS

- 1:1 mappings for 237 cases.
- Weighted match: accuracy 0.62, Top-5 hit rate 0.68, NDCG@10 0.61.
- Outperformed direct LOINC mapping, and often generated more precise, clinically specific codes than references.

ARKE: Automated Radiology Knowledge Encoding



Input text: Foot Standing AP Rt
TOP1 candidate: 38815-7 [XR Foot – right Views AP W standing]

Category	Weighted RadLex PartType Name
Modality	Modality. Modality Type
	Modality. Modality SubType
Anatomic Location	Anatomic Location. Region Imaged
	Anatomic Location. Imaging Focus
	Anatomic Location. Laterality.Presence
	Anatomic Location. Laterality
View	View. View Type
Timing	Timing
Pharmaceutical	Pharmaceutical. Substance Given

A: RadLex PartTypes from the input text
B: RadLex PartTypes from the mapping candidate

$$\text{Score} = \sum_{r \in A \cap B} w(r) - \sum_{r \in B-A} w(r) - \sum_{r \in A-B} w(r)$$

$$w(r) = \begin{cases} 1.5 & \text{if } r \text{ is a weighted RadLex PartType} \\ 1 & \text{otherwise} \end{cases}$$

- weighted & match: n=6
- unweighted & match: n=1
- weighted & unmatch: n=1
- unweighted & unmatch: n=1

Score = (1.5 * 6 + 1*1) – (1.5*1 + 1*1) = 7.5

Performance of ARKE Framework

Mapping method		Direct LOINC	LOINC-RadLex			
		cosine	jaccard	simple	weighted	F1
Final	Accuracy	0.536	0.616	0.625	0.625	0.620
TOP1	Accuracy	0.279	0.464	0.481	0.477	0.464
TOP3	Hit Rate	0.502	0.625	0.612	0.620	0.625
TOP5	Hit Rate	0.595	0.662	0.662	0.684	0.662
TOP10	Hit Rate	0.751	0.768	0.755	0.772	0.768
	MRR	0.417	0.556	0.563	0.563	0.556
	NDCG@10	0.496	0.607	0.609	0.612	0.607

Example of radiology procedure codes standardization

			CT Research Liver-LBW (contrast)	CT Abdomen + Pelvis (contrast)	CT Resarch liver + pelvis_PRP (contrast)
Local			CT Research Liver-LBW (contrast)	CT Abdomen + Pelvis (contrast)	CT Resarch liver + pelvis_PRP (contrast)
EDI			RC4018 [Abdominal CT (contrast)]		
SNOMED			429862006 [CT of liver with contrast]	419394006 [CT of abdomen and pelvis]	
LOINC-RadLex	LOINC		24815-3 [CT Liver W contrast IV]	36813-4 [CT Abdomen and Pelvis W contrast IV]	
	RadLex	modality	CT	CT	
		anatomic location	Abdomen	Abdomen	
			Liver	Pelvis	
		Pharmaceutical	Contrast	Contrast	
			IV	IV	
		Timing	W	W	

Example of ARKE framework mapping

Input Text	Silver Reference	Final Mapping Candidate
US Breast Gold insertion	US Guidance for placement of needle in Breast	US Guidance for needle localization of Breast
Rib cage view Both Oblique	XR Chest Right oblique and Left oblique	XR Ribs anterior – bilateral Views
MRI Brain Limited study non contrast	MR Guidance for stereotactic localization of Brain-- WO contrast	MR Brain limited WO contrast
GYN US Routine General Doppler	US.doppler Abdomen and Pelvis	US.doppler Pelvis vessels
Cryosurgical Ablation of Liver	US Guidance for ablation of tissue of Liver	Guidance for cryoablation of Liver
15T MRI Wholespine noncontrast diffusion	MR Spine WO contrast	MR Cervical and thoracic and lumbar spine WO contrast

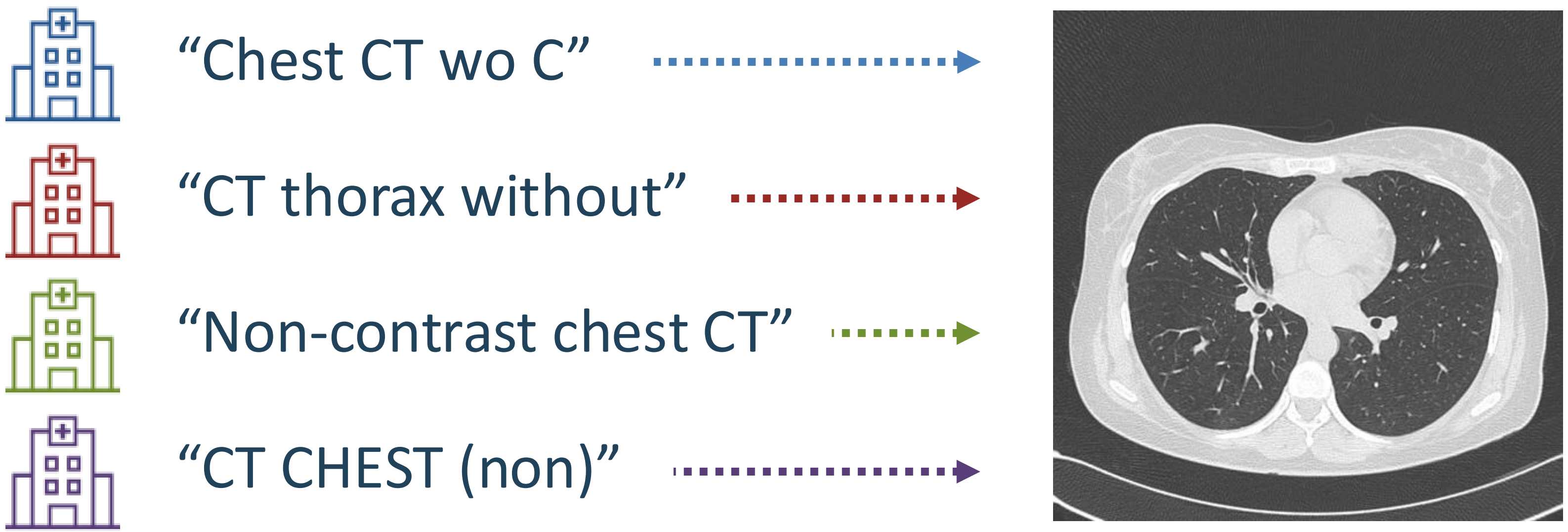
Sumin Lee[1,2*], Kyulee Jeon[1,2*], Yiju Park[1,2], Min Seong Kim[3], Juhyeon Jin[1], Changhoon Han[1], Soonho Yoon[4], Seng Chan You[1,2†]

1. *Department of Biomedical Systems Informatics, Yonsei University College of Medicine*
2. *Yonsei Institute for Digital Health, Yonsei University,*
3. *Department of Material Sciences and Engineering, Yonsei University College of Engineering*
4. *Department of Radiology, Seoul National University College of Medicine*

YONSEI UNIVERSITY
SNUH
OHDSI

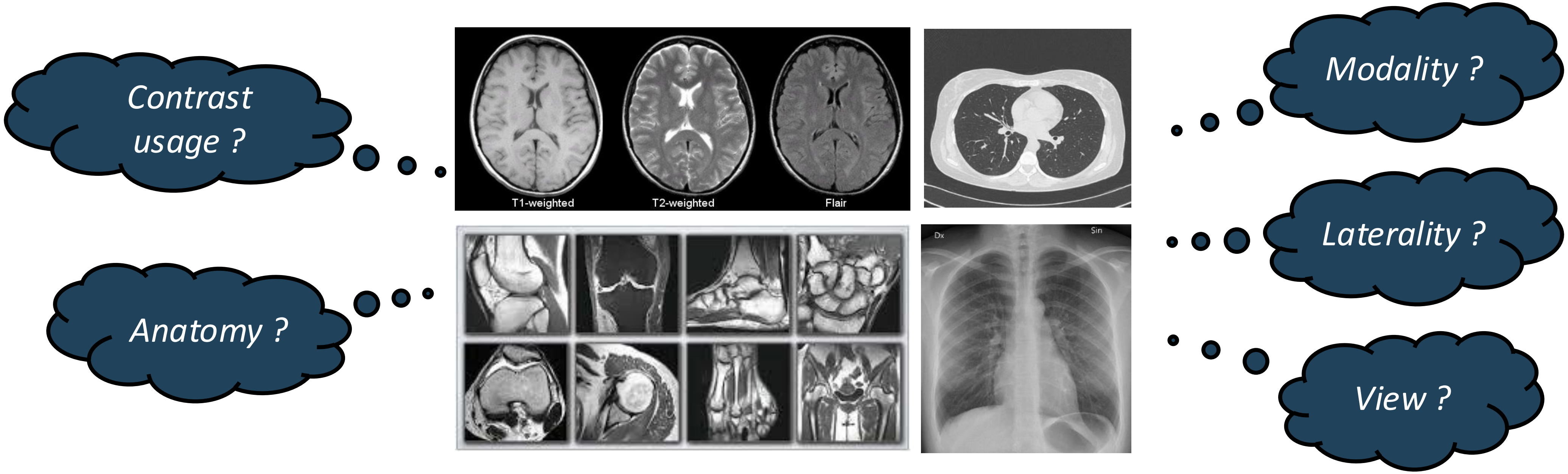
Motivation

- Imaging research integrates data across institutions and health systems.
- Before multi-site analyses can occur, **the fundamental question—"Was the same exam performed?" is often ambiguous**.



Existing Standards… But Not Radiology-Specific

- Multiple coding systems exist, yet none were designed to capture the **compositional nature** of imaging exams
  - Imaging exams are built from components like **modality, contrast, and anatomical focus**
  - A component-based encoding is needed **to group, route, and match** similar exams



LOINC–RadLex Playbook

- To address this gap, the **LOINC–RadLex Playbook**, an ontology explicitly optimized for imaging procedures was developed.

Local radiology term	Billing Code System e.g., CPT-4	Typical Standard e.g., SNOMED	LOINC/RSNA Radiology Playbook		
			LOINC	RadLex Playbook	RadLex
CT Research Liver - LBW (contrast)	74160 [Abdomen with Contrast]	429862006 [CT of liver with contrast]	24815-3 [CT Liver W contrast IV]	LP200067-9 [Modality]	RID10321 [CT]
				LP199933-5 [Anatomy]	RID58 [Liver]
				LP200088-5 [Timing]	RID49853 [W]
				LP200085-1 [Substance]	RID11582 [Contrast]
CT Abdomen + Pelvis (contrast)	74117 [CT Abdomen & Pelvis W/Contrast]	419394006 [CT of abdomen and pelvis]	36813-4 [CT Abdomen and Pelvis W contrast IV]	LP200067-9 [Modality]	RID10321 [CT]
				LP199933-5 [Anatomy]	RID56 [Abdomen], RID2507 [Pelvis]
CT Research liver + pelvis_RPP (contrast)				LP200088-5 [Timing]	RID49853 [W]
				LP200085-1 [Substance]	RID11582 [Contrast]

LOINC–RadLex Playbook

LOINC Radiology		LOINC/RSNA Radiology Playbook				RadLex	
LoincNumber	LongCommonName	PartNumber	PartTypeName	PartName	PartSequence Order	RID	PreferredName
100349-0	US.doppler Penis vessels W vasodilator IV	LP207608-3	Rad.Modality.Modality Type	US	A	RID10326	Ultrasound
		LP207609-1	Rad.Modality.Modality Subtype	Doppler	A	RID10375	Doppler
		LP199998-8	Rad.Anatomic Location.Region Imaged	Pelvis	A	RID2507	Pelvis
		LP208051-5	Rad.Anatomic Location.Imaging Focus	Penis vessels	A	RID49825	set of penis vessels
		LP200088-5	Rad.Timing	W	A	RID49853	with
		LP432424-2	Rad.Pharmaceutical.Substance Given	Vasodilator	A	RID50652	Vasodilator
		LP200078-6	Rad.Pharmaceutical.Route	IV	A	RID11160	intravenous
100369-8	NM Thyroid gland Views W Tl-201 IV	LP208891-4	Rad.Modality.Modality Type	NM	A	RID10330	nuclear medicine imaging
		LP200078-6	Rad.Pharmaceutical.Route	IV	A	RID11160	intravenous
		LP208673-6	Rad.Pharmaceutical.Substance Given	Tl-201	A	RID11753	Thallium-201
		LP221404-9	Rad.View.Aggregation	Views	A	RID49567	multiple views
		LP200088-5	Rad.Timing	W	A	RID49853	with
		LP199995-4	Rad.Anatomic Location.Region Imaged	Neck	A	RID7488	neck
		LP206648-0	Rad.Anatomic Location.Imaging Focus	Thyroid gland	A	RID7578	thyroid gland
100760-8	XR Radius and Ulna - left Single view	LP212285-3	Rad.Modality.Modality Type	XR	A	RID10345	projection radiography
		LP200031-5	Rad.Anatomic Location.Region Imaged	Upper extremity	A	RID1850	Upper extremity
		LP222059-0	Rad.View.Aggregation	View	A	RID43399	1 view
		LP199954-1	Rad.Anatomic Location.Laterality.Presence	TRUE	A	RID49844	Laterality present
		LP199952-5	Rad.Anatomic Location.Laterality	Left	A	RID5824	LEFT
		LP221452-8	Rad.Anatomic Location.Imaging Focus	Radius	A.A	RID2109	Radius
		LP221453-6	Rad.Anatomic Location.Imaging Focus	Ulna	A.B	RID2119	Ulna

Real-World Gap Between Standard and Practice

- Despite its strengths, the Playbook is **not widely adopted** in practice.
- Mapping local terms to Playbook concepts remains challenging due to:
  - **Time-consuming and labor-intensive**,
  - **Difficult to perform consistently even for radiologists**, and
  - Complicated by the **frequent updates to RadLex–LOINC Playbook**,

  making it challenging for institutions to maintain up-to-date standardized mappings.

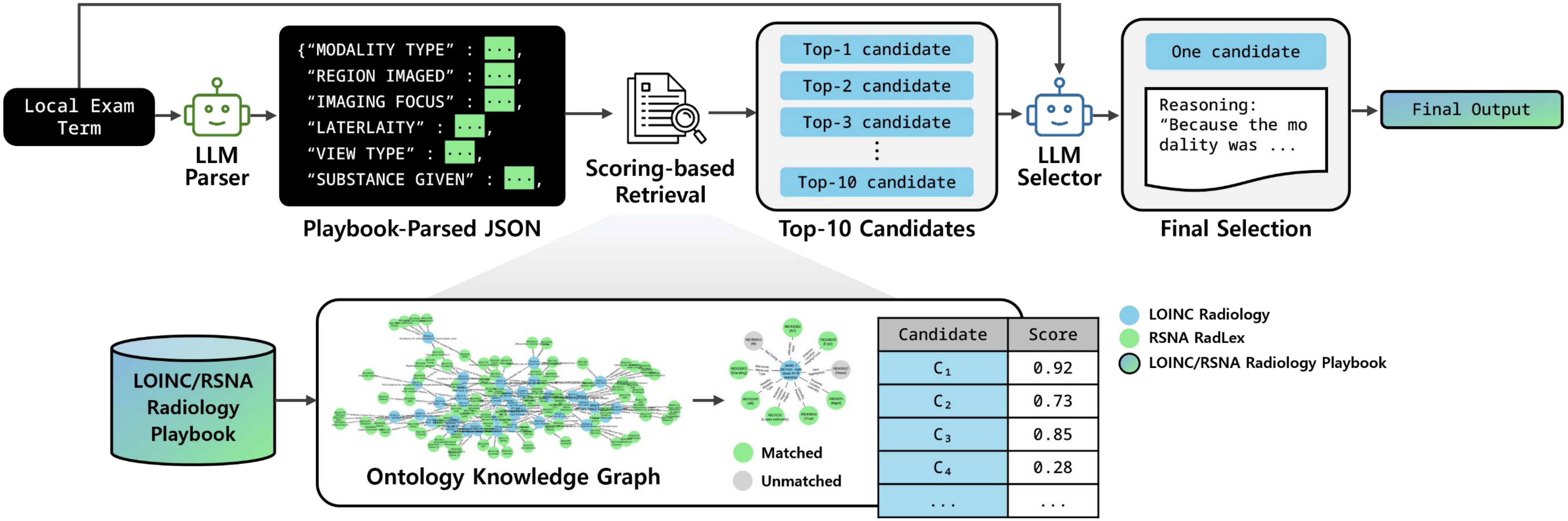
Limitations of Direct LLM-Based Mapping

- Recent attempts to use large language models (LLM) for terminology mapping demonstrate notable limitations:
  - **Hallucinated attributes**, adding or changing contrast or anatomy
  - **Incomplete interpretations**, losing key attributes
  - **Inconsistency**, with variable outputs for the same input

ARKE: A Knowledge-Encoding Approach

- We introduce a **knowledge-encoding** framework for standardizing imaging exam names. ARKE integrates two core principles:
  1. **Ontology-Based Weighted Matching**
     - Utilizes the RadLex to assign weights to modality, anatomy, and other key attributes.
     - Retrieves Top-N candidate terms from the ontology, preventing hallucination.
  2. **Constrained LLM Reasoning**
     - LLM usage is restricted to two structured steps:
       1) Converting the local exam term into a RadLex–based JSON structure
       2) Selecting one final term from the ontology-derived Top-10 candidates
     - The LLM does not generate or extend codes independently.

Overview



Local Exam Term
LLM Parser
{"MODALITY TYPE" : ...,
"REGION IMAGED" : ...,
"IMAGING FOCUS" : ...,
"LATERLAITY" : ...,
"VIEW TYPE" : ...,
"SUBSTANCE GIVEN" : ...,
Playbook-Parsed JSON
Scoring-based Retrieval
Top-1 candidate
Top-2 candidate
Top-3 candidate
Top-10 candidate
Top-10 Candidates
LLM Selector
One candidate
Reasoning:
"Because the mo
dality was ...
Final Selection
Final Output
LOINC/RSNA Radiology Playbook
Ontology Knowledge Graph
Matched
Unmatched
Candidate
Score
C1
0.92
C2
0.73
C3
0.85
C4
0.28
...
...
LOINC Radiology
RSNA RadLex
LOINC/RSNA Radiology Playbook

Data Source

- **Local radiology terms** collected from two tertiary hospitals in South Korea (2024-12)
  - Yonsei University Hospital System (3,036)
    - Random splitting: Training, Validation, Internal Test Set
  - Seoul National University Hospital (290)
    - 2,900 → 10% used for External Test Set
- **The LOINC/RSNA Playbook** from LOINC 2.80 (2025-02-26 released)

Our proposed Framework: ARKE

1. Preprocessing
   – **RadLex PartType Structurization**: LLM for parser

2. Framework
   – **Ontology-Based Weighted Matching:**
     - Knowledge Graph
     - Rule-based Scoring
     - RAG

   – **Final Selection**: LLM with reasoning

Data Preprocessing

- **Normalize local terms** into RadLex–based JSON structure

Local Radiology Terms

Foot Standing AP Rt

LLM as **Parser**

Structured Local Radiology Terms


```
{
  "Name": "Foot Standing AP Right",
  "Parts": {
    "Rad.Modality.Modality Type": "XR",
    "Rad.Modality.Modality Subtype": "",
    "Rad.Anatomic Location.Region Imaged": "Lower extremity",
    "Rad.Anatomic Location.Imaging Focus": "Foot",
    "Rad.Anatomic Location.Laterality.Presence": "True",
    "Rad.Anatomic Location.Laterality": "Right",
    "Rad.View.Aggregation": "",
    "Rad.View.View Type": "AP",
    "Rad.Timing": "",
    "Rad.Maneuver.Maneuver Type": "Standing",
    "Rad.Pharmaceutical.Substance Given": "",
    "Rad.Pharmaceutical.Route": "",
    "Rad.Reason for Exam": "",
    "Rad.Guidance for.Presence": "",
    "Rad.Guidance for.Approach": "",
    "Rad.Guidance for.Action": "",
    "Rad.Guidance for.Object": "",
    "Rad.Subject": ""
  }
}
```

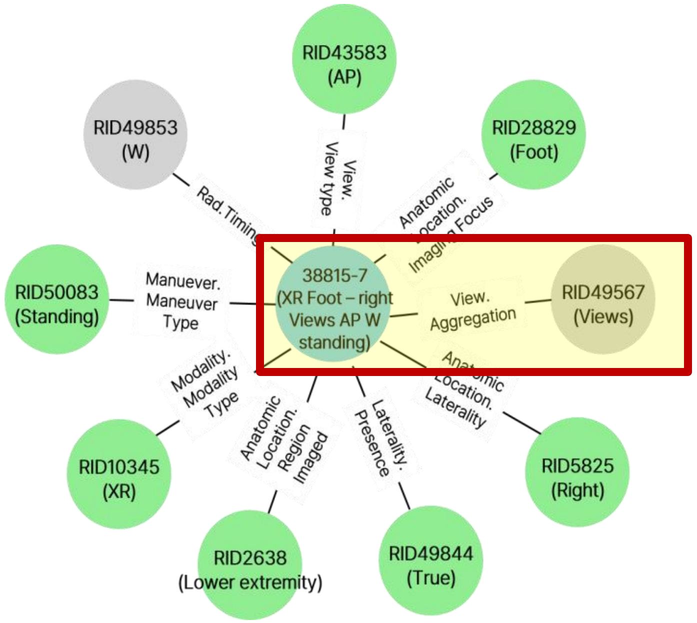

- Name key:
  Translated local terms with special character removed
- Parts key:
  Structured local terms using 18 RadLex-PartTypes

Framework - Step1: Knowledge Graph

- Structure the standard LOINC/RSNA Radiology Playbook into a KG

LOINC Radiology		LOINC/RSNA Radiology Playbook			RadLex	
LoincNumber	LongCommonName	PartNumber	PartTypeName	PartName	RID	PreferredName
38815-7	XR Foot - right Views AP W standing	LP212285-3	Rad.Modality.Modality Type	XR	RID10345	projection radiography
		LP199985-5	Rad.Anatomic Location.Region Imaged	Lower extremity	RID2638	Lower extremity
		LP221475-9	Rad.Anatomic Location.Imaging Focus	Foot	RID28829	foot
		LP199954-1	Rad.Anatomic Location.Laterality.Presence	TRUE	RID49844	Laterality present
		LP199953-3	Rad.Anatomic Location.Laterality	Right	RID5825	RIGHT
		LP220552-6	Rad.View.View Type	AP	RID43583	AP view
		LP221404-9	Rad.View.Aggregation	Views	RID49567	multiple views
		LP200088-5	Rad.Timing	W	RID49853	with
		LP221186-2	Rad.Maneuver.Maneuver Type	Standing	RID50083	Standing maneuver

Central entity node | Semantic Edges | Attribute node



Knowledge Graph

Framework - Step2: Rule-based Scoring

- Attribute-**Weighted Matching**
  - key PartTypes: Modality, Anatomic location, View, Timing, Pharmaceutical

Input Text Parsed into RadLex PartTypes


```
{
  "Name": "Foot Standing AP Right",
  "Parts": {
    "Rad.Modality.Modality Type": "XR",
    "Rad.Modality.Modality Subtype": "",
    "Rad.Anatomic Location.Region Imaged": "Lower extremity",
    "Rad.Anatomic Location.Imaging Focus": "Foot",
    "Rad.Anatomic Location.Laterality.Presence": "True",
    "Rad.Anatomic Location.Laterality": "Right",
    "Rad.View.Aggregation": "",
    "Rad.View.View Type": "AP",
    "Rad.Timing": "",
    "Rad.Maneuver.Maneuver Type": "Standing",
    "Rad.Pharmaceutical.Substance Given": "",
    "Rad.Pharmaceutical.Route": "",
    "Rad.Reason for Exam": "",
    "Rad.Guidance for.Presence": "",
    "Rad.Guidance for.Approach": "",
    "Rad.Guidance for.Action": "",
    "Rad.Guidance for.Object": "",
    "Rad.Subject": ""
  }
}
```


Retrieval Top-10 Candidates based on Weighted-Matching Score Method

R_{text}: RadLex PartTypes parsed from the input text

C^k: k–th LOINC candidate ($k = 1, \dots, n$)

R^k_{cand}: RadLex PartTypes corresponding to C^k ($k = 1, \dots, n$)

Step 1 — Candidate–wise Scoring

$$Score^k = \sum_{r \in R_{text} \cap R^k_{cand}} w(r) - \sum_{r \in R^k_{cand} - R_{text}} w(r) - \sum_{r \in R_{text} - R^k_{cand}} w(r)$$

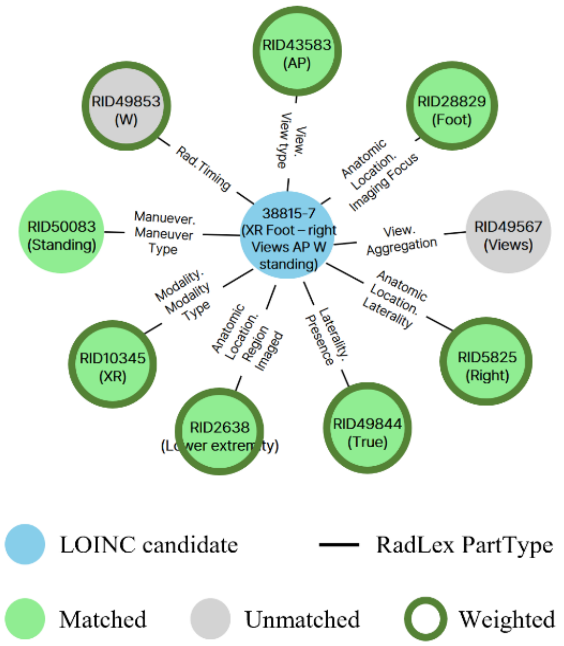
$$w(r) = \begin{cases} 1.5, & \text{if } r \text{ is a high–priority PartType (e.g., Laterality, View)} \\ 1, & \text{otherwise} \end{cases}$$

Step 2 — Ranking and Top–10 Selection

$$\{i_1, i_2, \dots, i_n\} = \mathrm{argsort}_{i=1,\dots,n}\left(Score^i\right)_{desc}$$

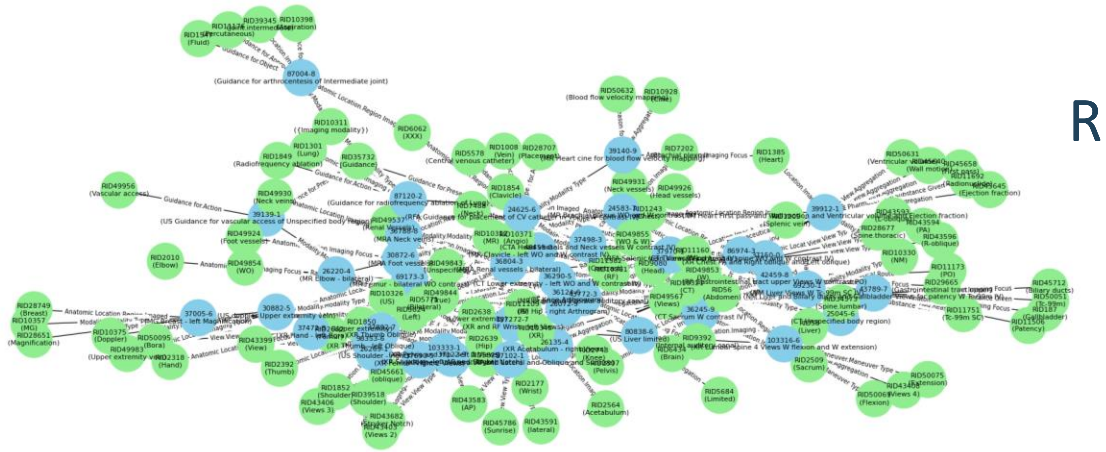
$$Top\text{-}j\ Candidate = C^{i_j}, j = 1, \dots, 10$$

Visual Example of Input–Candidate Matching (Step 1)



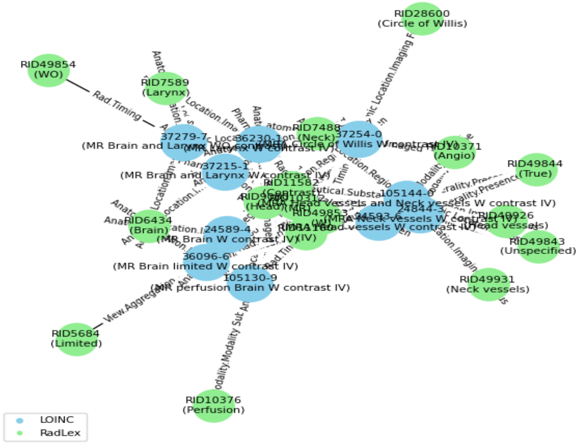
Framework - Step3: Retrieval

- Rank candidates using weighted matching by comparing JSON-structured local terms with LOINC–RadLex attributes in the KG
- Return the **top 10 highest-scoring candidates as match suggestions**



Knowledge Graph

Rule-based scoring



Top-10 Retrieval

Framework - Step4: Final Selection

- LLM with **reasoning** to capture clinical nuances beyond rule-based scoring

Input

Foot Standing AP Rt

Top-10 Candidates

Top-1

39374-4
XR Foot - right AP and Lateral W standing

LLM Prompting

…

(General instruction)

- Select only from the provided candidates.
- Do not add details not present in the input term.
- Preserve all explicitly stated specificity (e.g., views, vessels, contrast).
- Maintain modality/type exactly.
- If similar, choose the most general candidate.

…

Final 1 Selection

{
"**Reasoning**": "The input term 'Foot Standing AP Rt' specifies an X-ray of the right foot with an anteroposterior (AP) view while standing.… Therefore, 38815-7 is the most appropriate choice."

"**Selected_loinc**": "38815-7",

"**Long_common_name**": "XR Foot - right Views AP W standing"
}

Evaluation

- Labelling (Gold Reference)
  - A board-certified radiologist identified the correct match from Top-10 candidates
- Comparison with other methods
  - Candidate retrieval performance, final mapping accuracy
    - **Using Playbook**
      - (ours) weighted-matching, / non-weighted matching
      - Other similarity-based : jaccard, F1
    - **w/o Playbook**
      - Direct matching using cosine similarity
- Performance Metrics
  - Top-N Hit Rate (1/3/5/10), NDCG@10, MRR, Accuracy

Results - Data Source Characteristic

Modality - Type

Variables	YUHS			SNUH
	Training Set N=1,821	Validation Set N=911	Internal Test Set N=304	External Test Set N=290
XR	461 (25.2)	237 (26.0)	73 (23.9)	47 (15.9)
MR	323 (17.7)	161 (17.6)	52 (17.0)	92 (31.2)
US	318 (17.4)	134 (14.7)	59 (19.3)	34 (11.5)
CT	196 (10.7)	102 (11.2)	34 (11.1)	50 (16.9)
RF	178 (9.7)	101 (11.1)	32 (10.5)	27 (9.2)
NM	65 (3.6)	29 (3.2)	9 (3.0)	0 (0.0)
MG	18 (1.0)	11 (1.2)	3 (1.0)	2 (0.7)
PT	11 (0.6)	5 (0.5)	1 (0.3)	5 (1.7)
PT+CT	53 (2.9)	24 (2.6)	9 (3.0)	2 (0.7)
DXA	1 (0.1)	0 (0.0)	0 (0.0)	1 (0.3)
NM.SPECT+CT	7 (0.4)	9 (1.0)	1 (0.3)	0 (0.0)
Image-guided procedures	197 (10.8)	100 (11.0)	32 (10.5)	35 (11.9)

Anatomic Location - Region Imaged

Variables	YUHS			SNUH
	Training Set N=1,821	Validation Set N=911	Internal Test Set N=304	External Test Set N=290
Abdomen	351 (18.8)	191 (20.1)	66 (21.0)	64 (20.9)
Lower extremity	248 (13.3)	137 (14.4)	49 (15.6)	55 (18.0)
Chest	236 (12.6)	114 (12.0)	37 (11.7)	37 (12.1)
Upper extremity	232 (12.4)	94 (9.9)	26 (8.3)	25 (8.2)
Head	178 (9.5)	103 (10.8)	37 (11.7)	46 (15.0)
Neck	114 (6.1)	67 (7.0)	25 (7.9)	18 (5.9)
Pelvis	132 (7.1)	49 (5.2)	27 (8.6)	13 (4.2)
Breast	55 (2.9)	26 (2.7)	10 (3.2)	3 (1.0)
Whole body	32 (1.7)	28 (2.9)	8 (2.5)	5 (1.6)
Extremity	21 (1.1)	9 (0.9)	0 (0.0)	3 (1.0)
Unspecified	254 (13.6)	126 (13.2)	28 (8.9)	37 (12.1)
Others	13 (0.7)	7 (0.8)	2 (0.6)	0 (0.0)

Results - Candidate Retrieval Performance

- ARKE's weighted-matching retrieval outperformed other rule-based and cosine-similarity methods across both test sets, showing consistently higher Top-N hit rates and ranking quality

Internal Test Set (N=237)									
Method		Final selection		Top-1	Top-3	Top-5	Top-10 candidates		
		Accuracy	p-value*	Accuracy	Hit Rate	Hit Rate	Hit Rate	MRR	NDCG@10
LOINC-RadLex mapping	Weighted-matching	**0.696**	Ref	0.490	**0.654**	**0.785**	**0.899**	0.605	**0.675**
	Simple overlap	0.654	0.024	**0.494**	**0.654**	0.764	0.873	**0.607**	0.670
	Jaccard similarity	0.675	0.302	0.473	**0.654**	0.772	0.878	0.595	0.663
	F1-score	0.675	0.267	0.473	**0.654**	0.772	0.878	0.595	0.663
Direct LOINC mapping	Cosine similarity	0.481	<0.001	0.245	0.418	0.511	0.616	0.355	0.418

External Test Set (N=290)									
Method		Final selection		Top-1	Top-3	Top-5	Top-10 candidates		
		Accuracy	p-value*	Accuracy	Hit Rate	Hit Rate	Hit Rate	MRR	NDCG@10
LOINC-RadLex mapping	Weighted-matching	**0.635**	Ref	0.507	0.645	0.731	**0.852**	0.604	0.662
	Simple overlap	0.624	0.663	0.503	0.648	0.741	0.838	0.601	0.657
	Jaccard similarity	0.621	0.556	**0.514**	**0.666**	**0.755**	0.841	**0.611**	**0.666**
	F1-score	0.628	0.838	**0.514**	**0.666**	**0.755**	0.841	**0.611**	**0.666**
Direct LOINC mapping	Cosine similarity	0.376	<0.001	0.090	0.272	0.366	0.493	0.208	0.276

Results - Final Mapping Accuracy

- ARKE's LLM-based final selection achieved the highest final-mapping accuracy across both test sets, with performance differences reaching statistical significance ($p = 0.024$; $p < 0.001$).

Internal Test Set (N=237)

Method		Final selection		Top-1	Top-3	Top-5	Top-10 candidates		
		Accuracy	p-value*	Accuracy	Hit Rate	Hit Rate	Hit Rate	MRR	NDCG@10
LOINC-RadLex mapping	Weighted-matching	**0.696**	Ref	0.490	**0.654**	**0.785**	**0.899**	0.605	**0.675**
	Simple overlap	0.654	0.024	**0.494**	**0.654**	0.764	0.873	**0.607**	0.670
	Jaccard similarity	0.675	0.302	0.473	**0.654**	0.772	0.878	0.595	0.663
	F1-score	0.675	0.267	0.473	**0.654**	0.772	0.878	0.595	0.663
Direct LOINC mapping	Cosine similarity	0.481	<0.001	0.245	0.418	0.511	0.616	0.355	0.418

External Test Set (N=290)

Method		Final selection		Top-1	Top-3	Top-5	Top-10 candidates		
		Accuracy	p-value*	Accuracy	Hit Rate	Hit Rate	Hit Rate	MRR	NDCG@10
LOINC-RadLex mapping	Weighted-matching	**0.635**	Ref	0.507	0.645	0.731	**0.852**	0.604	0.662
	Simple overlap	0.624	0.663	0.503	0.648	0.741	0.838	0.601	0.657
	Jaccard similarity	0.621	0.556	**0.514**	**0.666**	**0.755**	0.841	**0.611**	**0.666**
	F1-score	0.628	0.838	**0.514**	**0.666**	**0.755**	0.841	**0.611**	**0.666**
Direct LOINC mapping	Cosine similarity	0.376	<0.001	0.090	0.272	0.366	0.493	0.208	0.276

Results - Attribute-wise Accuracy

- Comparing the RadLex PartTypes of the parsed input and the final candidates revealed low agreement in anatomic location—particularly for the region imaged and **imaging focus** attributes

Attribute	YUHS Internal Test Set (Accuracy, 95% CI)	SNUH External Test Set (Accuracy, 95% CI)
Modality Type	0.945 (0.908–0.968)	0.852 (0.806–0.888)
Region Imaged	0.781 (0.724–0.829)	0.731 (0.677–0.779)
Imaging Focus	0.650 (0.587–0.708)	0.621 (0.564–0.675)
Laterality	0.844 (0.792–0.885)	0.855 (0.810–0.8910
Pharma + Timing Group	0.899 (0.854–0.931)	0.879 (0.837–0.912)

Results - Analysis of Misclassified Cases

- In many cases, ARKE produced outputs with greater semantic specificity than the reference labels, offering more detailed procedural or anatomic information

Input Text		Gold Reference	Final Mapping Candidate
Not captured by weighted attributes	OB US 1st Trimester Detailed US Twin 11w13w6d	US for pregnancy in first trimester	US for multiple gestation pregnancy in fir st trimester
Clinical relevance	MRI Lt Upper Arm noncontrast	MR Upper extremity - left WO contrast	MR Upper arm-left WO contrast
Specific attributes missed by human reviewers	Tibia Both Oblique Lt	XR Tibia and Fibula Oblique Views	XR Tibia and Fibula –left Oblique Views
Playbook coverage	MRI15T Whole body contrast	MR Whole body	MR Whole body WO and W contrast IV

Significance

- Leveraged underused radiology-specific vocabularies **(RadLex, Radiology Playbook**) to enrich imaging data granularity

- **ARKE Framework** is not just a mapping tool—it automates the transformation of local radiology terms into standardized conceptual knowledge
  – **Knowledge Encoding**: Semantic Meaning of Local Radiology term → Structured attributes (JSON) → Ontology (RadLex) → standardized codes (LOINC–RadLex)
    1) Enables semantic identification of heterogeneous imaging data across institutions
    2) Support stable integration of imaging data based on identified and standardized radiology terms
  – Provides empirical evidence supporting the suitability of LOINC–RadLex for real-world mapping and OMOP Vocabulary discussions

- Improves data quality and downstream analytic usability

Contribution to Community

- Incorporating **LOINC/RSNA Playbook** and **RadLex** into OMOP Vocabulary, enables more accurate imaging-related retrieval & analysis

- The ontology-based mapping approach can be applied to other terminology-mapping tasks

- **Your data contributions are welcome** — offering new opportunities for network studies and helping support diverse evidence generation

Thank You!

- Acknowledgement:
  - Seng Chan You
  - Soonho Yoon
  - Dept. of Biomedical Informatics, Yonsei University College of Medicine
    - Yiju Park, Min Seong Kim, Juhyeon Jin, Changhoon Han
  - OHDSI Medical Imaging Working Group
  - OHDSI Early-Stage Researchers Working Group
- Contact:
  - lsm0801@yuhs.ac, jklee320@yuhs.ac
  - CHANDRYOU@yuhs.ac