On the Use of Large Language Models for Schema Mapping Generation in Ontology-Based Data Integration

Yannis Marketakis · Milio Lintanff-Castel · Yannis Tzitzikas









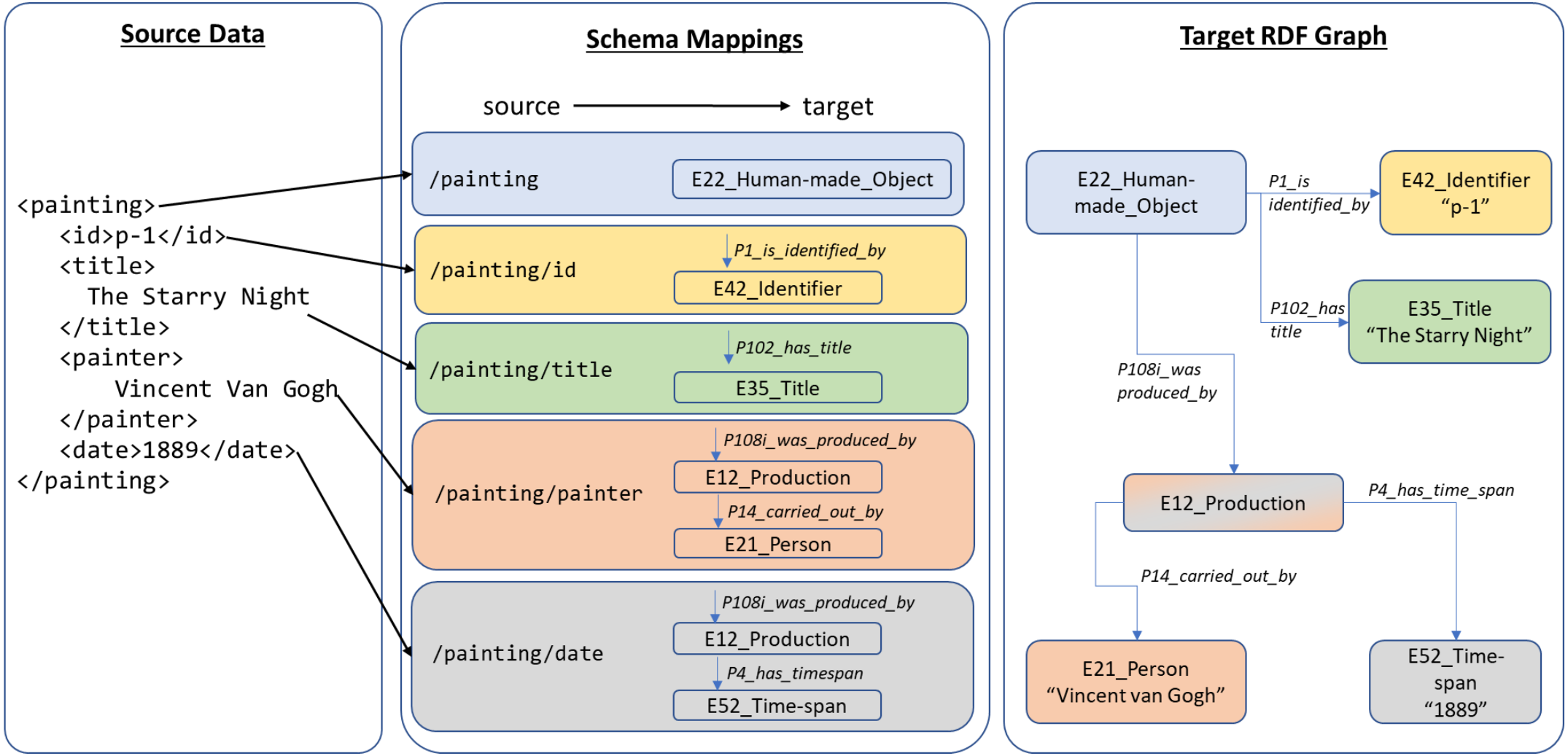
Outline





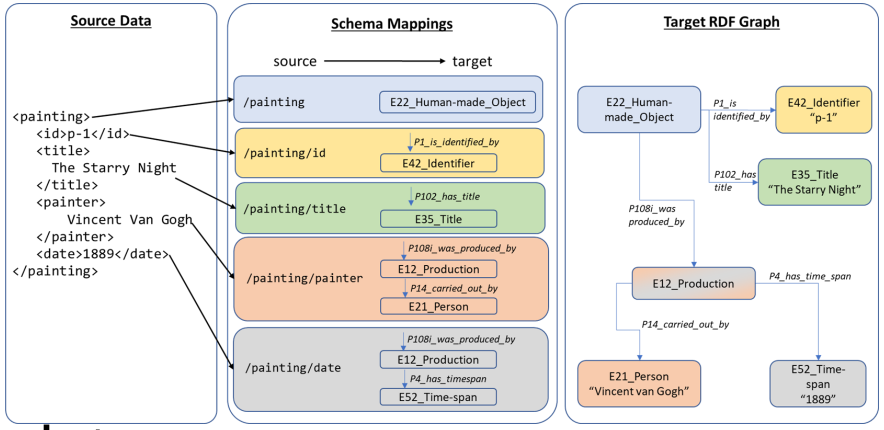
Motivation and Positioning	3-6	4’
LLM-assisted Schema Mapping	7-9	2’
X3MLBench and Evaluation Setup	10-13	4’
Results and Lessons Learned	14-18	4’
Next steps & Future work	19	1’

What is a schema mapping?



Why schema mappings matter?

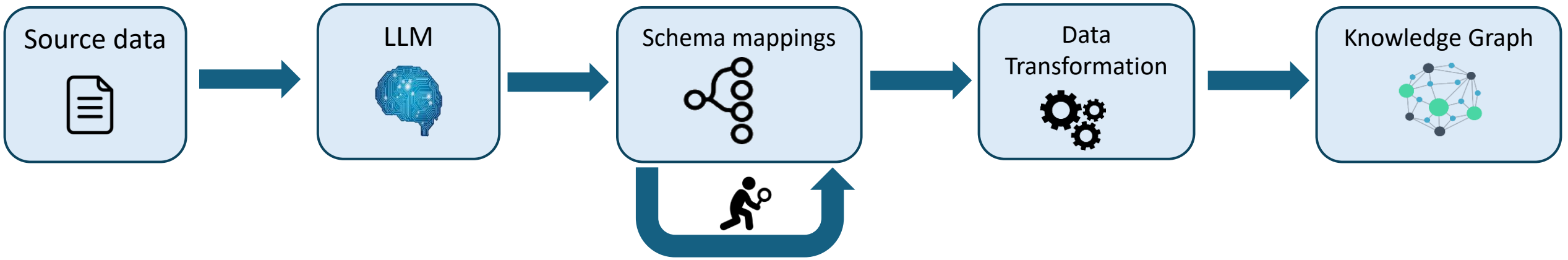
- Explicit transformation logic
  - How source elements correspond to classes and properties
- Reusability
  - Same mappings executed repeatedly on similar data
- Validation and debugging
  - Inspected, corrected and validated before RDF transformation
- Ontological correctness
  - Ability to express complex target patterns
- Human control
  - Experts can review and refine mappings



Schema mappings: powerful but costly

- Schema mappings **can**:
  - **Connect** data elements to ontology classes and properties
  - **Define the transformation logic** from source data to RDF Knowledge Graphs
  - Support **complex ontology** patterns
- …but their definition is cumbersome
  - **Manual** and **time-consuming**
  - Requires good **knowledge** of both source and target schemata
  - Requires **experience** with formal mapping languages and technology
- Schema mapping definition becomes a **bottleneck** as data volume and data heterogeneity increase
  - We need methods than can **accelerate** schema mapping definition while preserving correctness and expert control

LLMs as mapping assistants vs LLM-to-RDF



Aspect	Direct LLM-to-RDF	LLM-assisted Schema Mappings
Input	Full Data	**Sample data** ✓
Transparency	Implicit	**Explicit** ✓
Reproducibility	Limited	**High** ✓
Debugging	Limited	**High** ✓
Variability	High	**Limited** ✓
Scalability	Limited	**High** ✓
Human Control	Minimal	**High** ✓

Main Contributions

Formalize the task

- Schema mapping generation as a constrained output task
- X3ML as the target mapping language
- Requirements: syntactic validity, semantic alignment, structural consistency

Build the benchmark

- Real-world, expert validated X3ML schema mappings
- Gold-standard collection for comparison
- Automated evaluation framework and metrics

Evaluate and learn

- Multiple LLMs and prompting strategies
- Lexical baseline for comparison
- Quantitative results and qualitative error analysis
- Insights for semi-automated mapping workflows

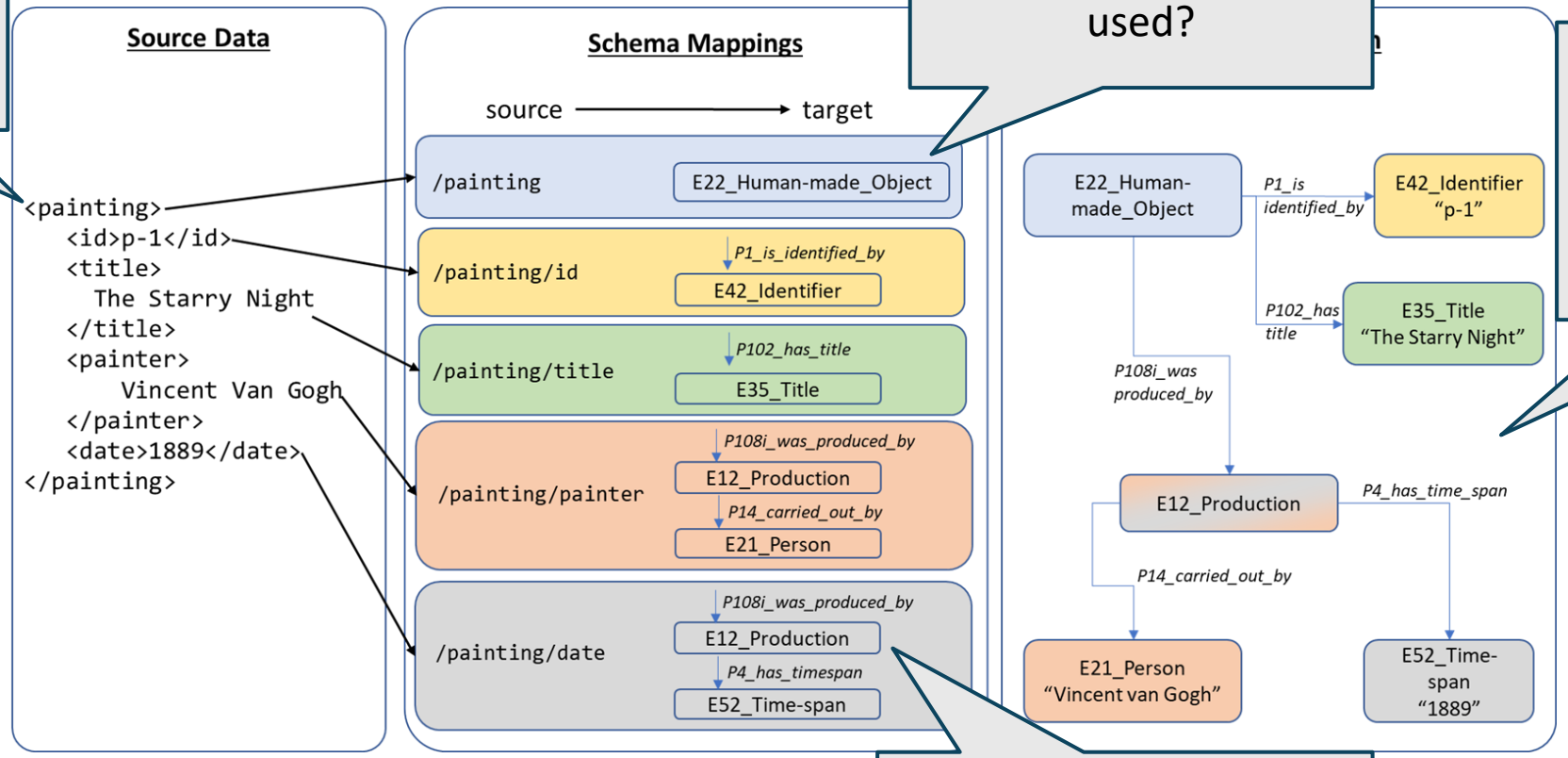
Schema mapping challenges

Interpret source elements and values

Which classes and properties should be used?

Will the transformed data be properly connected?

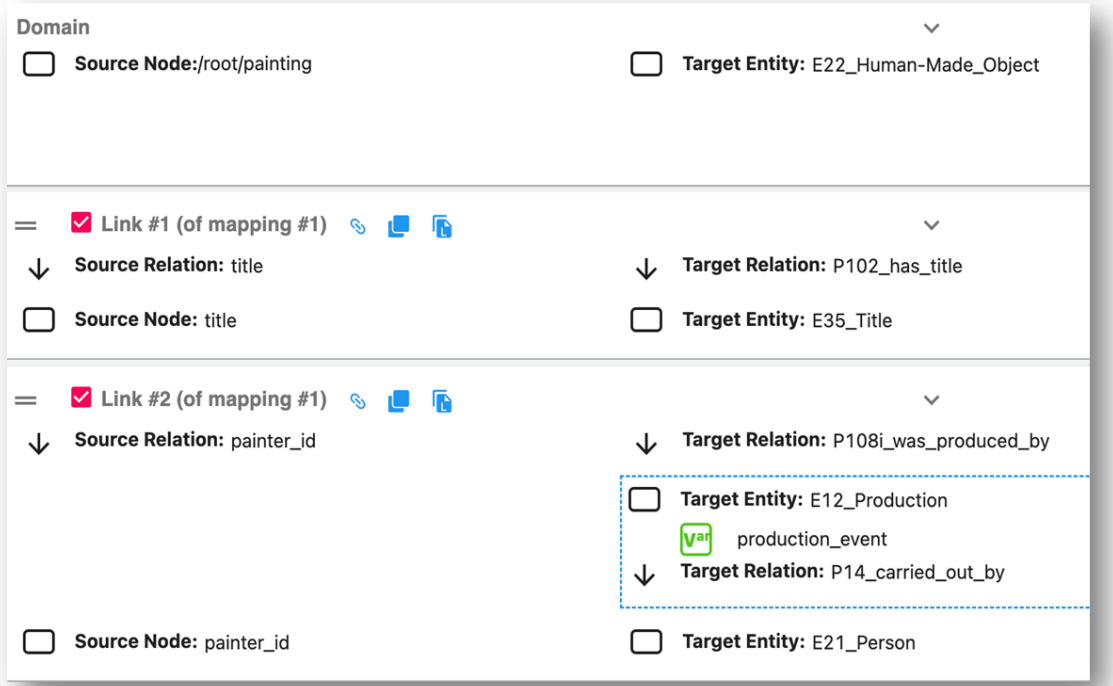
When are intermediate nodes needed?



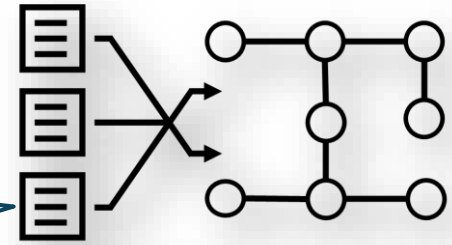
X3ML-specific challenges

```
<mapping>
    <domain>
        <source_node>/painting</source_node>
        <target_node>
            <entity>
                <type>crm:E22_Human-Made_Object</type>
                <instance_generator name="LocalTermURI">
                    <arg name="hierarchy" type="constant">painting</arg>
                    <arg name="term" type="xpath">id/text()</arg>
                </instance_generator>
                <label_generator name="CompositeLabel">
                    <arg name="term1" type="constant">Painting with title:</arg>
                    <arg name="term2" type="xpath">title/text()</arg>
                </label_generator>
            </entity>
        </target_node>
    </domain>
    <link>
        <path>
            <source_relation>
                <relation>title</relation>
            </source_relation>
            <target_relation>
                <relationship>crm:P102_has_title</relationship>
            </target_relation>
        </path>
        <range>
            <source_node>title</source_node>
            <target_node>
                <entity>
                    <type>crm:E35_Title</type>
                    <instance_generator name="LocalTermURI">
                        <arg name="hierarchy" type="constant">title</arg>
                        <arg name="term" type="xpath">text()</arg>
                    </instance_generator>
```

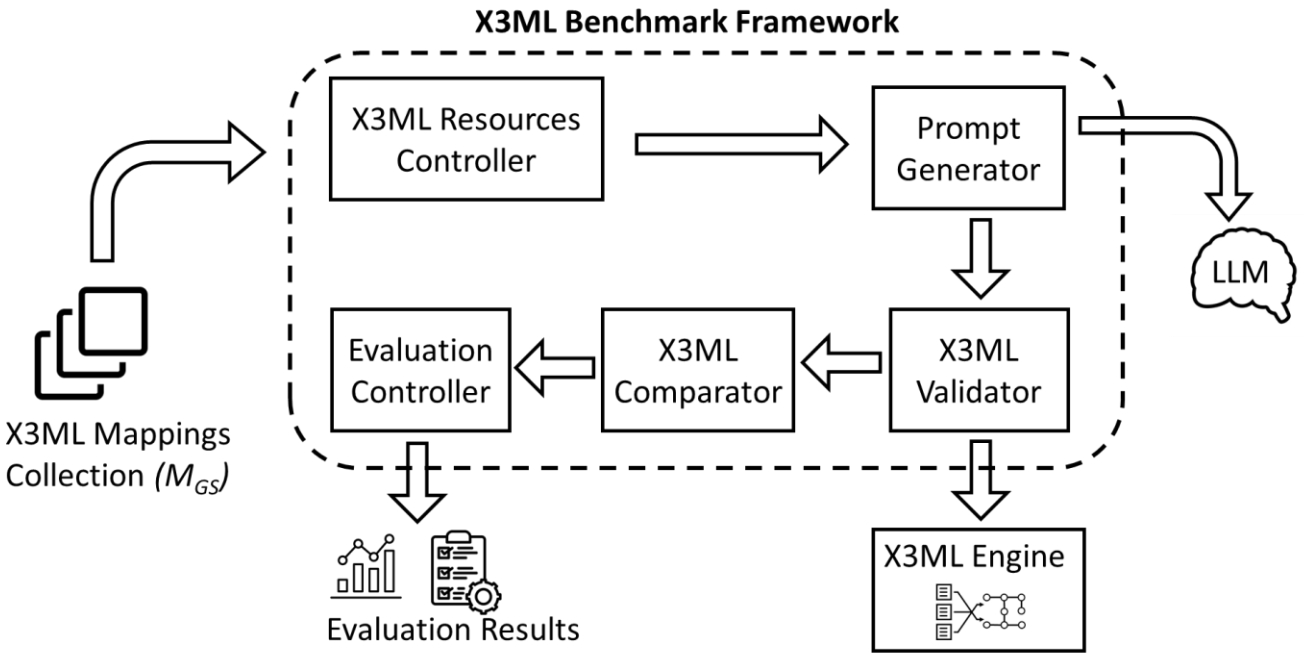
Does it conform to the X3ML ?



Can they be used with X3ML-related services and tools?



X3MLBench: design and workflow



- Input:
  - Source data
  - Prompt
  - X3ML-related resources
- Output:
  - generated X3ML mappings
  - Evaluation metrics

X3MLBench compares LLM-generated X3ML mappings against expert-created schema mappings

X3MLBench – The gold standard collection

- Consists of X3ML mapping projects, validated by experts
  - Target Ontology: **ISO 21127:2023 CIDOC CRM**
  - Reference entities: **Persons, Organizations, Locations, Events**
  - Derived from real-world use cases and projects: **RICONTRANS**, **ARIADNEplus**, **China Biographical Database Project (CBDB)**

13

mapping projects

53

mapping domains

308

mapping links

X3MLBench: Evaluation focus

- Generated X3ML mappings vs gold-standard X3ML mappings
- Validity Check
  - **Syntactic Validity (SV)**: can the mapping be parsed/executed?
- Main reported metric
  - **Accuracy (A)**: what proportion of the generated mapping elements are correct?

$$Accuracy = \frac{|gen.mapping\ elements\ \cap gold\ standard\ mapping\ elements|}{|gen.mapping\ elements|}$$

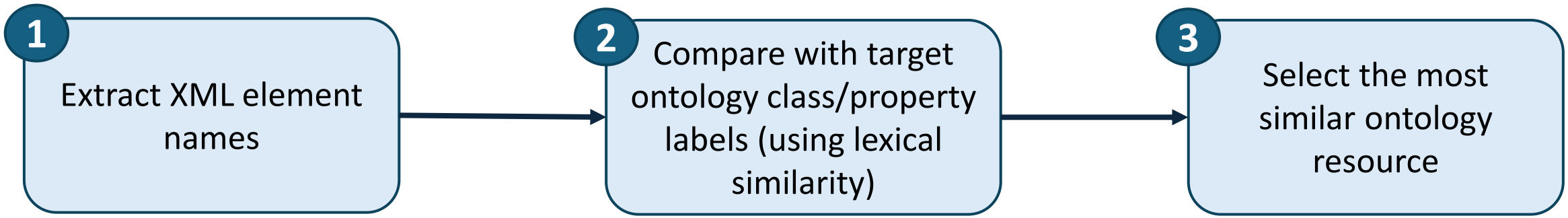
- We evaluate mapping components, not RDF triples
  - source-target correspondences
  - target classes & properties
  - links and target paths

X3MLBench: Prompting strategies & LLMs

- **Zero-shot** with input:
  - source data
- **Syntax-aware** with input:
  - source data
  - X3ML structure
- **In-context mapping** with input:
  - source data
  - relevant X3ML mapping

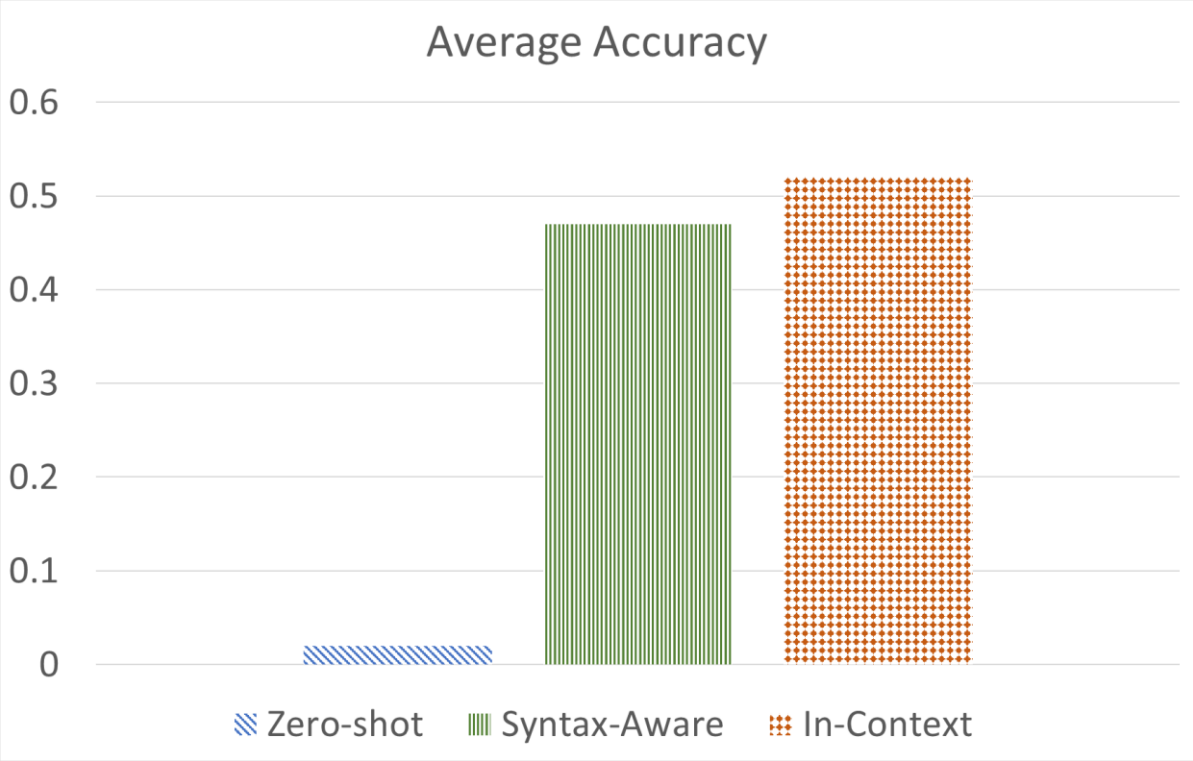
- Evaluated across 6 LLMs
  - OpenAI GPT 5.3 Instant
  - Gemini 3 Flash
  - DeepSeek V3-0324
  - Mistral Medium-2505
  - Grok-3
  - Llama 4-Scout-17B-16E-Instruct

Lexical Baseline approach



- A simple, lower-bound reference for the mapping selection task
  - Does not guarantee complete X3ML mapping file and validity with respect to X3ML schema
  - Evaluated only with respect to mapping accuracy
- For example:
  - **<person> → E21_Person**
  - **<title> → P102_has_title → E35_Title**

Evaluation Results: Prompting matters



Zero-shot: ~ 0

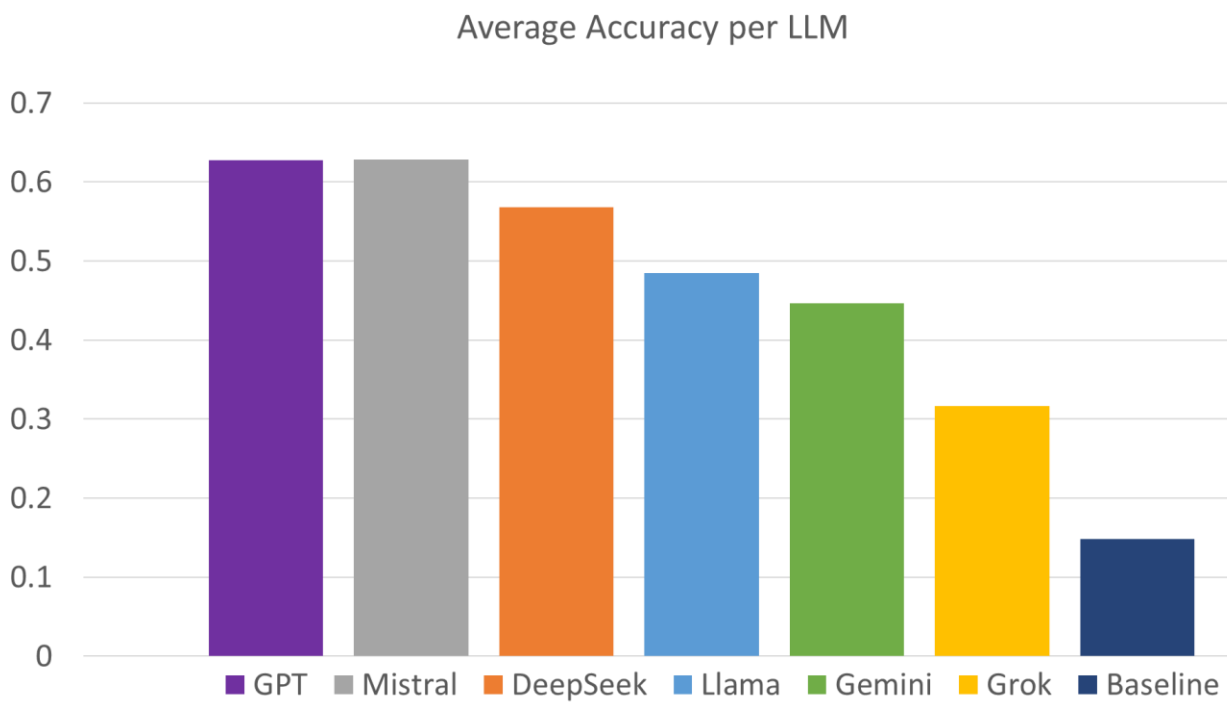
Syntax-aware: ~ 0.47

In-context: ~ 0.53

- **Zero-shot** performs very poorly
- **Syntax-aware** improves both validity and accuracy
- **In-context** gives the best results by providing both syntax and domain guidance

*Structure guidance and context is not optional, it is **essential***

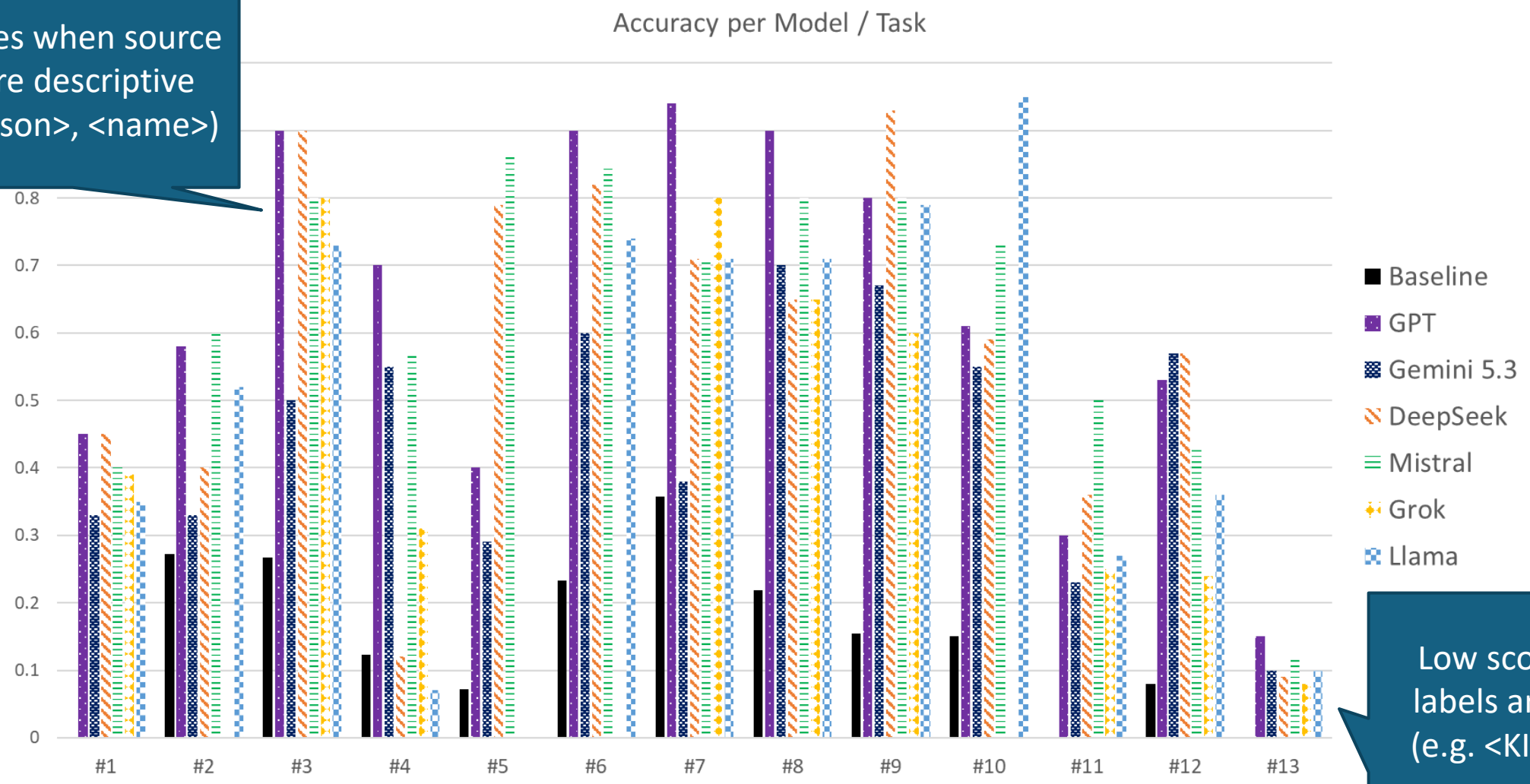
Evaluation Results: LLMs vs lexical baseline



- All LLMs outperform lexical baseline (**0.14**)
- GPT 5.3 and Mistral achieve the best average accuracy (**0.63**)
- DeepSeek follows with **0.57**
- Grok 3 achieved the lowest average accuracy (**0.32**)
- All LLMs captured more than simple lexical similarity

Evaluation Results: Accuracy depends on the dataset

High scores when source labels are descriptive (e.g. <person>, <name>)



Low scores when labels are opaque (e.g. <KIN_DATA>)

Error patterns and lessons learned

Recurring Error Patterns

- Incorrect ontology terms used
- Invalid or incomplete X3ML structures
- Erroneous target-path structure
- Missing or inconsistent references for URI and label generation

Lessons Learned

- Structural guidance and in-context examples are useful
- Source-schema clarity strongly affects mapping quality
- LLMs are useful as assistants, but expert validation remains necessary
- **LLMs can accelerate the schema mapping process**

What's next? Improving guidance

- Can LLMs generate valid and meaningful X3ML mappings?
  - Our work shows that this is possible with proper guidance
- How can we guide LLMs more effectively?

➔"*Exploiting Mapping Corpora and Expert Knowledge for LLM-based Schema Mapping Generation*" to appear in GLOW@ISWC 2026 on October 2026*

- Exploration of alternative guidance strategies: ontology documentation, lexically retrieved examples, semantically retrieved examples (RAG), expert rules
- Investigation of iterative self-verification prompting and fine-tuning

*: Y. Marketakis, A. Buckel, and Y. Tzitzikas, Exploiting Mapping Corpora and Expert Knowledge for LLM-based Schema Mapping Generation, Graph-enhanced LLMs for trustworthy Web data management (GLOW), in ISWC 2026 Workshops Joint Proceedings, October 25–26, 2026, Bari, Italy

Thank you for your attention!

“a good schema mapping is like a good joke:
***every relation must connect**”*

Questions ?

marketak@ics.forth.gr
https://github.com/ymark/x3ml_comparator