Ontologies Tutorial 2 Report.

The search for reusable ontologies was not as successful as we expected but we got some useful resources nonetheless.

The main ontologies we found were:

<https://www.w3.org/TR/owl-time/#time:year>

We can use owl-time to ground the release of specific LLMs, mark key dates in the developments of LLM architecture or look at current trends. Being associated with owl gives us confidence that the ontology is well-designed and has been used before.

<https://www.cs.umd.edu/projects/plus/SHOE/onts/univ1.0.html>

It can be useful when dealing with meta concepts related to academic work, for example, the class of ResearchAssistant. We feel confident using it since it comes from the University of Maryland.

For general country, city and language related resources, Dbpedia and Wikimedia provide a vast amount of validated resources we can use. This would come in handy when speaking about the language of the training data of LLMs or the place where certain advancements were developed.

The foundings on LOV were the following :

<https://saidfathalla.github.io/Science-knowledge-graph-ontologies/doc/ModSci_doc/index-en.html#NaturalLanguageProcessing> and

<https://w3id.org/skgo/modsci#NaturalLanguageProcessing>

<https://saidfathalla.github.io/Science-knowledge-graph-ontologies/doc/ModSci_doc/ontology.ttl>

<https://saidfathalla.github.io/Science-knowledge-graph-ontologies/doc/ModSci_doc/index-en.html#ArtificialIntelligence>

They have consistent subclasses that we can find on the list we grasp from the isrt session and that we can reuse for sure each one plays a specific role in the development of intelligent systems. However I find that “pattern recognition” can also work as a property that you found in all subclasses because it underpins various AI techniques across all sub-classes, so it's a very broad category that spans multiple domains of AI.

<https://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core/nif-core.html#d4e59>

This ontology is well designated but it's not very constituent to our LLM’s general ontology, it's specific about the utilization of nlp. Indeed it's well-structured, focusing on key aspects like text span, document metadata, and annotations but it's quite complex and require advanced understanding for a “full utility”.We can reuse that ontology if we want to go into more detail.

<http://ontology.cybershare.utep.edu/ELSEWeb/elseweb-lifemapper-parameters.owl#LearningRate>

**Ontology Repositories on AI & LLM**:

[berkeleybop/artificial-intelligence-ontology: An ontology modeling classes and relationships describing deep learning networks, their component layers and activation functions, machine learning methods, as well as AI/ML potential biases. (github.com)](https://github.com/berkeleybop/artificial-intelligence-ontology)  
**Consistency**: It’s not quite consistent with the ontology we intend to build.  
**Design**: It appears well-designed, focusing on AI concepts such as layers, activation functions, and biases.

**Usage**: While it has limited visibility (14 stars, 2 forks at 16/09/2024), it is listed on BioPortal, a known ontology library, suggesting some external usage.

**Reuse**: The layer and activation function classes, as well as AI bias concepts, are reusable for projects in deep learning and fairness in AI.

[OntoBroker and OntoStudio X | semafora semantic AI (semafora-systems.com)](https://www.semafora-systems.com/ontobroker-and-ontostudio-x)

OntoBroker is a commercial ontology repository and middleware system that stands out for its integration of a high-performance reasoning engine.

**Consistency**: It’s not quite consistent with the ontology we intend to build.

**Design**: OntoBroker is not simply an ontology; it’s an entire semantic platform, including a reasoning engine. It is well-designed in that it adheres to multiple Semantic Web standards (OWL, RDF, RDFS, SPARQL) and F-Logic, which allows for both expressive representation and efficient querying.

**Usage**: OntoBroker has been in industrial use for over 18 years, indicating widespread adoption in various sectors

**Reuse**:

**Reasoning Capabilities**: The core reasoning engine of OntoBroker, which supports F-Logic and OWL-based reasoning, could be reused in any system requiring complex inference over ontologies, particularly for AI systems that need to generate natural language explanations.

**Distributed Query Processing**: The architecture allowing distributed query processing could be reused in environments needing high query throughput or scalability, such as knowledge graphs or semantic data lakes.

**Middleware Features**: Its middleware capabilities—allowing seamless integration with Java, Python, and other systems—could be adapted for use in AI systems that require rapid prototyping and hybrid data processing

More generally, we can also use Dbpedia and wikidata as supportive ontologies for broader terms and classes.