# Use Case 1: Machine Learning Dataset Search

## Description

Searching for machine learning datasets can be a cumbersome process due to the scattered nature of data repositories and the lack of standardized formats for representing datasets and their attributes. Users often encounter challenges in discovering relevant datasets for their specific tasks, comparing dataset characteristics, and understanding how datasets have been used in past machine learning experiments. Faced with this problem, this ontology aims to provide a comprehensive solution by representing machine learning datasets in a standardized and interpretable format, as well as their characteristics, features and relationships with machine learning experiments, software and models.

## Actors

Researcher/practitioner, ontology-based application

## Flow

A user is interested in finding large datasets for object detection tasks. Using a search application based on this ontology, they can specify the field of machine learning they are interested in for discovering datasets and they can retrieve relevant datasets from diverse sources, view their characteristics, locations, features. They can also view their data loader locations, machine learning implementations and software that leverages them.

# Use Case 2: Machine Learning Experiment Data Search

## Description

Finding and interpreting machine learning data is a difficult task because of the small number of repositories that provide them and the absence of a universal format to represent machine learning problems. Most of the time that users want to find information about machine learning tasks they have taken on, they manually search on the web for related datasets, examples of past implementations, pipelines that achieve the wanted results and their performance characteristics. This is a time consuming process that often involves the search in many different repositories and scientific papers. Also, it is difficult to associate the findings between each other since repositories and scientific works often represent similar concepts using different formats and words. Faced with this problem, this ontology aims to represent machine learning experiment data from different sources in a universal and easier to interpret format, from their higher level concepts down to details of their implementation details, parameters used, datasets used and evaluations.

## Actors

Researcher/practitioner, ontology-based application

## Flow

A user is interested in the field of natural language processing and wants to find relative tasks. Using a search application based on this ontology, they can specify the field of machine learning they are looking to find more about. The application will provide related datasets, algorithms, tasks and workflows. The user can browse these data, view tasks and projects that interest him along with details about their implementation, executions and evaluations.

# 

# Use Case 3: Machine Learning Software Search

## Description

Locating software implementations in the field of machine learning can be a challenging endeavor due to the multitude of repositories, varied coding styles, and the absence of a standardized representation for software characteristics. Researchers and practitioners often face difficulties in discovering machine learning code, understanding its functionalities, assessing its compatibility with their specific needs. Faced with this problem, this ontology aims to represent machine learning software, their characteristics and dependencies. Also, it aims to match machine learning software with relevant datasets, machine learning implementations, experiment and models around the web.

## Actors

Researcher/practitioner, ontology-based application

## Flow

A user is interested in discovering machine learning software solutions that leverage a specific dataset. Using a search application based on this ontology, they can specify the dataset they are interested in and retrieve machine learning software notebooks or code repositories that use this leverage to conduct machine learning experiments. The user can view the location, characteristics and dependencies of the software, as well as related machine learning evaluation results, published models and scientific publications that the software is referenced in.

# Use Case 4: Machine Learning Models Search

## Description

The quest for discovering machine learning models is often complicated by the fragmented nature of repositories and a lack of standardized representations for their evaluations. Researchers and practitioners looking for specific models often have to search over different platforms and are often unable to discover related scientific publications, code repositories and evaluations for the models found. Faced with this problem, this ontology aims to represent machine learning models, their characteristics and evaluations as well as link them with relevant scientific publications and notebooks and code repositories that leverage them.

## Actors

Researcher/practitioner, ontology-based application

## Flow

A user is interested in discovering pre-trained machine learning models related with object tracking. Using a search application based on this ontology, they can specify the task they are interested in solving, or a name of a specific model they are searching for and discover platforms that make this model available, evaluation results on different datasets, code notebooks that configure and use the model, the code repository of the model as well as publications that reference it.

# Use Case 5: Machine Learning Publications Search

## Description

The exploration of machine learning literature is a critical aspect of staying abreast of the latest research trends, methodologies, and breakthroughs. However, the dispersed nature of scientific publications across various platforms, along with the difficulty in discovering publications’ related sources such as software and datasets used and experiments conducted, makes it challenging for researchers to efficiently search, compare, and comprehend the vast landscape of machine learning literature. Faced with this problem, this ontology aims to represent machine learning publications and link them with relevant datasets, software, models and evaluations.

## Actors

Researcher/practitioner, ontology-based application

## Flow

A user is interested in discovering machine learning publications related to localization algorithms in robotics. Using a search application based on this ontology, they can specify the task they are interested in solving, the name of a relevant algorithm or a publication and retrieve related publications, their code repositories, datasets and algorithms they leverage, their evaluation results and models.

# Use Case 6: Machine Learning Recommendation System

## Description

Building effective machine learning recommendation systems relies on a nuanced understanding of diverse datasets, experiment configurations, software implementations, and model evaluations. However, there is an absence of interconnected representations of the machine learning landscape, making it challenging to develop a sophisticated machine learning recommendation system able to reason over diverse data. Faced with this problem, this ontology aims to represent the machine learning lifecycle, its configurations and evaluation results, in order to potentially support machine learning recommendation systems in recommending machine learning solutions based on user input.

## Actors

Researcher/practitioner, ontology-based application

## Flow

A user is interested in discovering the best pre-trained object recognition model for a specific dataset of their choice. Using a recommendation application based on this ontology, the user can share their dataset and the task they are trying to solve and the system can recommend high-performance models for datasets with semantic, textual or statistical similarities to the one possessed by the user as well as relevant code notebooks and configuration settings.

# Use Case 7: Machine Learning Analytics

## Description

Unveiling meaningful insights from the expansive field of machine learning data requires a sophisticated approach to data analytics. The current fragmented state of machine learning repositories and the absence of standardized structured representations for machine learning information hinder the process of conducting data analysis on machine learning information and potentially discovering useful insights and trends. Faced with this problem, this ontology aims to provide a structured, and interconnected representation that bridges the gaps in machine learning data, fostering a more seamless and effective analytics experience.

## Actors

Researcher, ontology-based application

## Flow

A researcher is interested in creating scatterplots of the most popular pre-trained machine learning models. Using a database modeled with this ontology, the researcher can leverage techniques such as graph embeddings to cluster and visualize machine learning models in an embedding space, based on their discrete attributes as well as their intricate relationships with other entities such as as evaluations, datasets and tasks the models are deployed for.