Metadata Template Documentation

This documentation provides an overview of the metadata structure, relationships.

1. General Model Information

- `Title` : The name of the model; should be descriptive and concise.
- `Editor Note` : Optional note provided by the editor, describing context or decisions.
- `Created by` : The name of the person or organization who created the model (usually the first author of a paper).
- `Creation date` : The date of publishing the paper or model in the repository.
- `Contact email` : Email address for inquiries about the model.
- `References to papers` : Insert one or more links to academic papers describing the model.
- `References to code` : Insert one or more links to code repositories (e.g., GitHub) relevant to the model.
- `Software License` : Type of license governing use of the model and its components. The preferred label should be selected from the Software Ontology (SWO).
- `FAIRmodels image name` : Name of the Docker or container image used for evaluation.
- `Docker image details (exposed port)` : Technical information such as which port is used in the Docker setup.

2. Model Information

- `Input data` : A list describing the data input structure for the model.
  - `Input feature` : The name of the input variable or column. The name can be selected from several ontologies: NCIT, ROO, LOINC, ICD10CM, SNOMEDCT.
  - `Description` : A textual description of what the feature represents, e.g., based on the reference paper where the training dataset was described.
  - `Input label` : The label or identifier used in the model for this feature. Input CSV files should include columns with these specified input labels.
  - `Type of input` : The data type: should be selected between categorical and numerical.
  - `Minimum – for numerical` : The lowest expected value for numerical input features.
  - `Maximum – for numerical` : The highest expected value for numerical input features.
  - `Categories` : If categorical, a list of category labels with:
    - `Category Label` : Name of the category. The name can be chosen from: NCIT, ROO, LOINC, ICD10CM, SNOMEDCT.
    - `Identification for category used in model` : Internal representation of the category. Input data should align with these identifiers.
- `Outcome` : The predicted or generated result(s) from the model, ideally using standardized ontology terms (e.g., NCIT, ROO, LOINC, ICD10CM, SNOMEDCT).
- `Outcome label` : Human-readable label or name of the outcome used in the model. Input CSV files should include a column with this specified outcome label.
- `Applicability criteria` : A list of situations or conditions where the model can be applied.
- `Foundational model or algorithm used` : An underlying model or method. The name should be selected from AIO, AI, or STATO.

Intended use:

- `Primary intended use` : A list of recommended use cases for the model.
- `Primary intended users` : The groups or roles expected to use the model.
- `Out-of-scope use cases` : Cases where the model should *not* be used.

Ethical Considerations:

- `Data` : Does the model use any sensitive data (e.g., protected classes)?
- `Human life` : Is the model intended to inform decisions about matters central to human life or flourishing – e.g., health or safety? Or could it be used in such a way?
- `Mitigations` : What risk mitigation strategies were used during model development?
- `Risks and harms` : What risks may be present in model usage? Identify potential recipients, likelihood, and magnitude of harms. If unknown, note that they were considered but remain undetermined.
- `Use cases` : Are there any known model use cases that are especially fraught? This may connect directly to the intended use section.

Caveats and recommendations:

- `Additional concerns` : A list of model limitations, uncertainties, or warnings.

3. Previous model tests

- `Performance metrics` : List of previous test results. May be added several previous tests if there are several previous tests. However, each test should have the proper decsription.
  - `Metric Label` : Name of the performance metric (e.g., accuracy). The label should be chosen from STATO or SWO.
  - `Measured metric (mean value)` : Average score.
  - `Measured metric (low 95% confidence interval)` : Lower bound of the 95% confidence interval.
  - `Measured metric (up 95% confidence interval)` : Upper bound of the 95% confidence interval.
  - `Acceptance level` : Benchmark or threshold to determine success, if mentioned in references.
  - `Additional information (if needed)` : Notes on how the metric was calculated or interpreted.
- `Link to dataset` : URL or path to the dataset used.
- `Link to reference paper` : Paper DOI documenting the test results or experimental setup.
- `Notes` : Free-text notes relevant to the model's past performance.

4. Evaluation results using FAIVOR tool:

- `Evaluation results` : Evaluation metrics generated using the FAIVOR tool. May be presented as a list of grades if there is more than one test.
  - `sha256 of Docker image` : Hash of the Docker image used.
  - `user/hospital` : Identifier for the evaluation context (e.g., organization).
  - `Performance metric` :
    - `Metric Label` : Name of the performance metric (e.g., accuracy). The label should be chosen from STATO or SWO.

- Measured metric (mean value) : Average score.
  - Measured metric (low 95% confidence interval) : Lower bound of the 95% confidence interval.
  - Measured metric (up 95% confidence interval) : Upper bound of the 95% confidence interval.
  - Acceptance level : Benchmark or threshold to determine success, if mentioned in references.
  - Additional information (if needed) : Notes on how the metric was calculated or interpreted.
- Dataset characteristics :
  - Input feature : Feature name (can be selected from NCIT, ROO, LOINC, ICD10CM, SNOMEDCT).
  - The general number of subjects in the dataset : Sample size.
  - Number of missing values : Count of missing entries.
  - The characteristics of dataset : Specify whether the dataset includes the whole population (ex, without additional filtering) or only complete cases.
  - The number of subjects for evaluation : Number of evaluated subjects.
  - The mean value — for numerical feature : Mean of the numerical feature.
  - The lower bound of 95% confidence interval — for numerical feature : Lower bound of the 95% confidence interval.
  - The higher bound of 95% confidence interval — for numerical feature : Upper bound of the 95% confidence interval.
  - Categories distribution :
    - Category Label : Category name (from ontologies like NCIT, ROO, LOINC, ICD10CM, SNOMEDCT).
    - Distribution for category : Number of cases in the category.
- User Note : Optional note from the evaluator.