

Virtual-Lab: Challenges-Feature

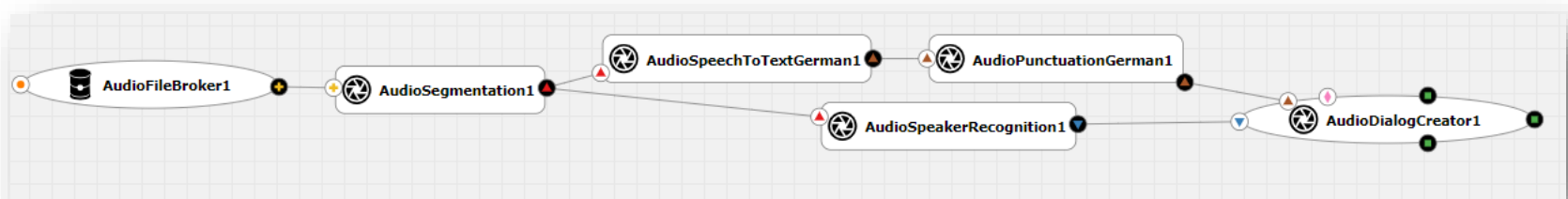
Martin Welss, Fraunhofer IAIS

martin.welss@iais.fraunhofer.de

Precondition: execution-run

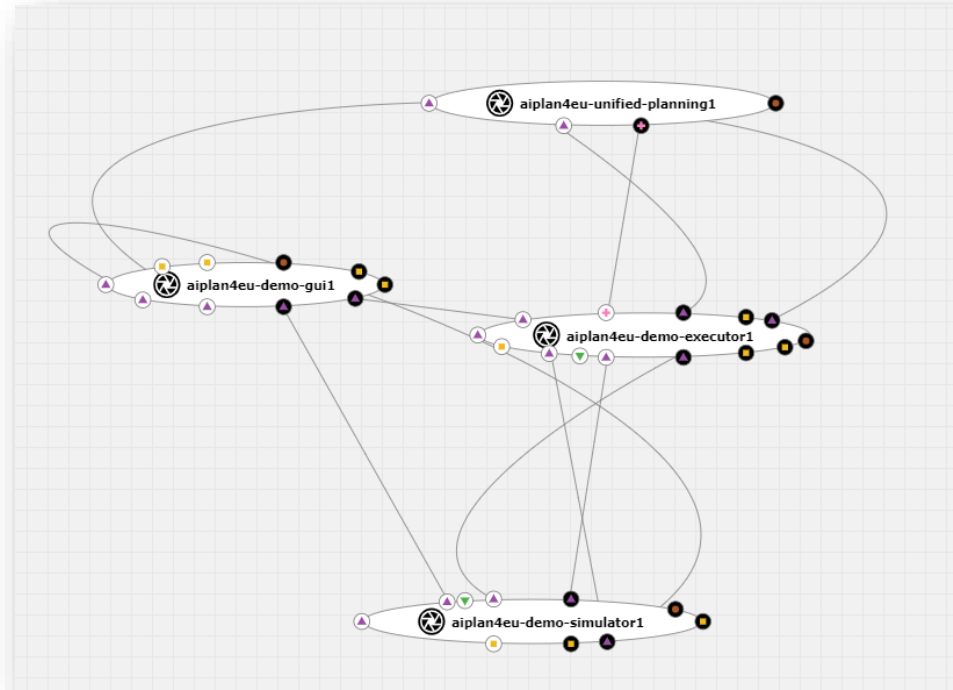
- Built on the reproducibility extension developed in AI4Europe
- Execution-run.json contains
 - Dataset metadata
 - Container images checksums
 - All information to re-run the pipeline
 - Pipeline topology (textual representation)

```
"dataset_features": {  
  "type": "aiod-dataset/v1",  
  "datasetname": "The Reuters Dataset, TensorFlow Dataset(tfds)",  
  "description": "http://kdd.ics.uci.edu/databases/reuters21578/README.txt",  
  "size": "4MB",  
  "DOI_ID": "Not available"  
}
```



Cyclic Topologies

- Cognitive architectures and pipelines contain loops and cycles
- For training as well as inference use cases



Precondition: Metrics and Metadata

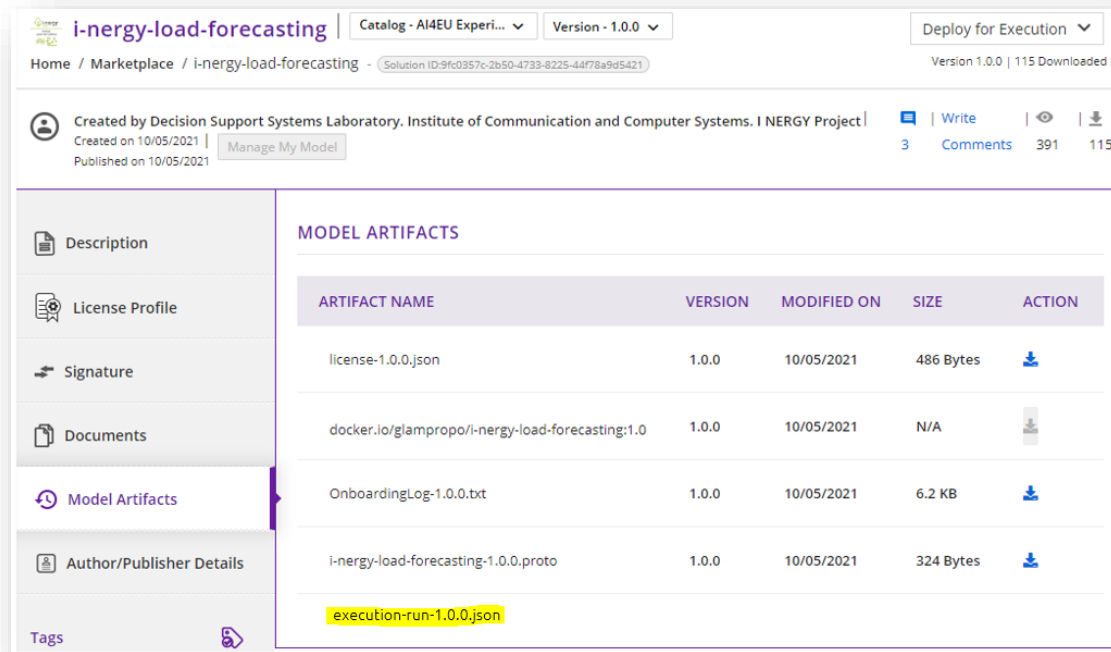
- Metrics data format developed in AI4Europe
 - Supports different types (regression, classification,...)
 - Support for automatic comparison
 - Important for ranking
 - Integration with central Metadata-Service for of other services

```
"metrics": {  
  "type": "regression-metrics/v1",  
  "date_time": "2023-11-22 15:36:44",  
  "status_text": "success"  
  "more_is_better": {  
    "r_squared": 0.831178119365146,  
    "adjusted_r_squared": 0.8270170166734419,  
  },  
  "less_is_better": {  
    "mse": 0.0025680503998073007,  
    "rmse": 0.05067593511527242,  
  }  
}
```





Tutorials: Metrics and Metadata

- Metrics and other execution metadata already implemented in some tutorials:
 - News-Trainer Pipeline
 - Houseprices-Prediction Pipeline
 - Sentiment-Analysis Pipeline
 - Updated Container Specification
 - <https://gitlab.eclipse.org/eclipse/graphene/tutorials>
- The AI Playground supports creation of the execution-run.json datastructure

Collecting Execution Metadata

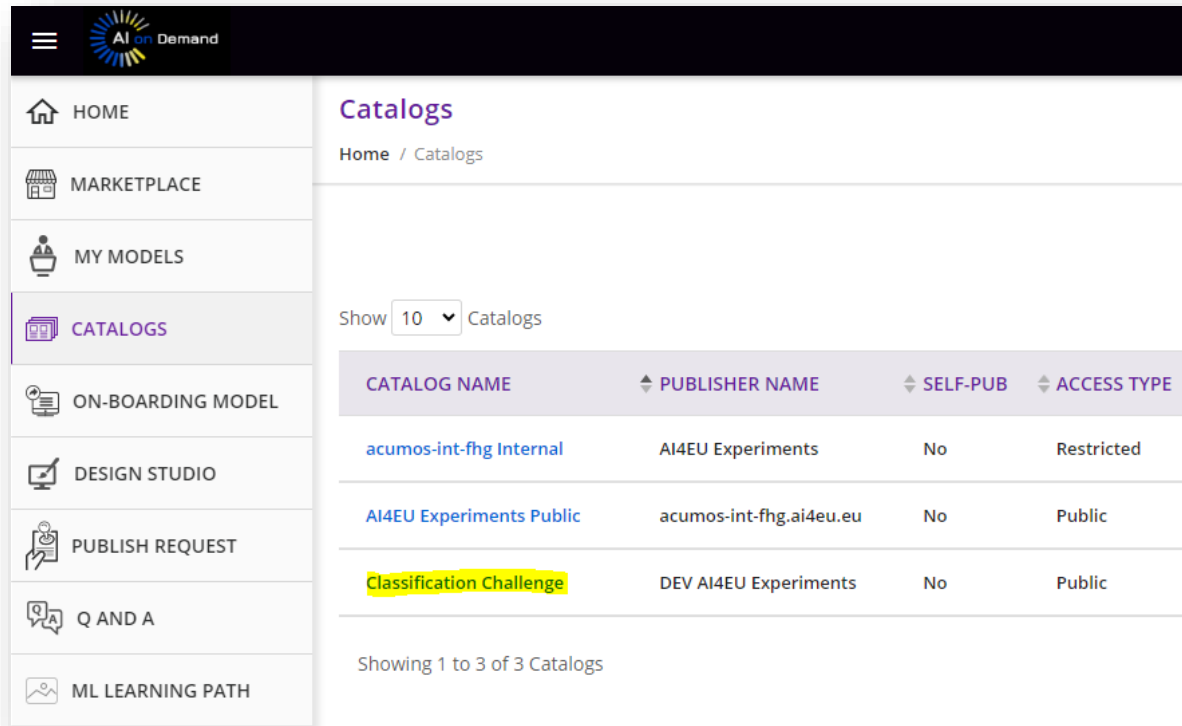


The screenshot displays the 'i-nergy-load-forecasting' model page. The sidebar on the left includes links to Description, License Profile, Signature, Documents, Model Artifacts (selected), Author/Publisher Details, and Tags. The main content area, titled 'MODEL ARTIFACTS', contains a table with the following data:

ARTIFACT NAME	VERSION	MODIFIED ON	SIZE	ACTION
license-1.0.0.json	1.0.0	10/05/2021	486 Bytes	
docker.io/glampropo/i-nergy-load-forecasting:1.0	1.0.0	10/05/2021	N/A	
OnboardingLog-1.0.0.txt	1.0.0	10/05/2021	6.2 KB	
i-nergy-load-forecasting-1.0.0.proto	1.0.0	10/05/2021	324 Bytes	
execution-run-1.0.0.json				

- execution-run.json can be received from different execution environments
- It will be associated with the respective model in the catalog

Creating Challenges

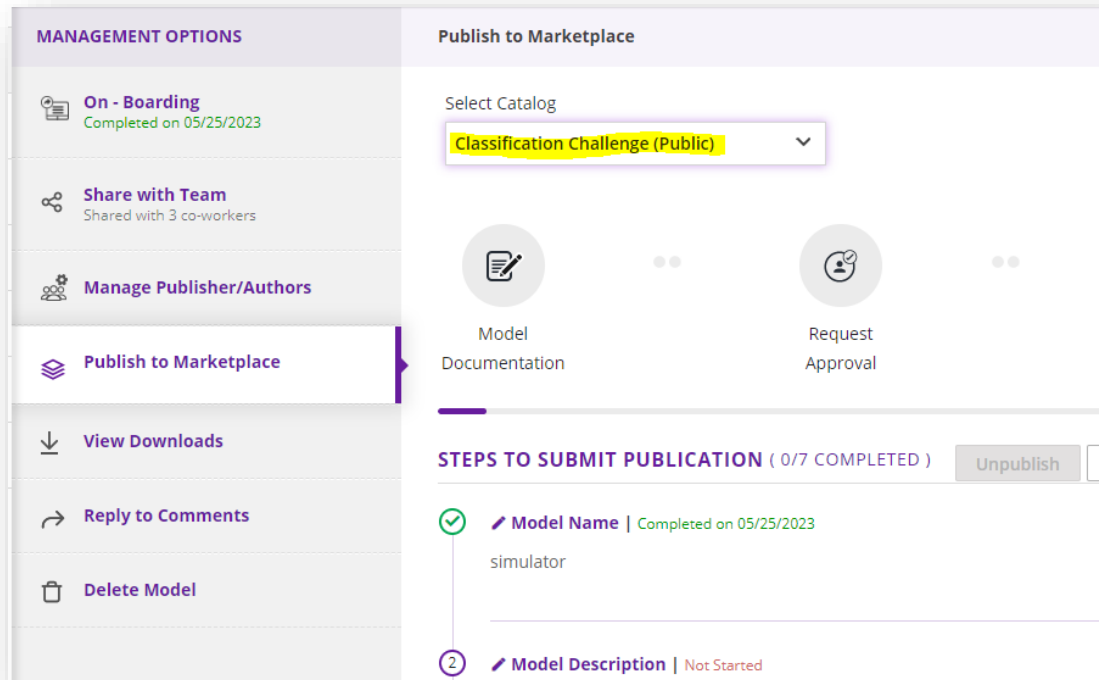


The screenshot shows the 'AI on Demand' marketplace interface. The left sidebar contains navigation links: HOME, MARKETPLACE, MY MODELS, CATALOGS (highlighted), ON-BOARDING MODEL, DESIGN STUDIO, PUBLISH REQUEST, Q AND A, and ML LEARNING PATH. The main content area is titled 'Catalogs' and shows a list of three catalogs. The first two are 'acumos-int-fhg Internal' and 'AI4EU Experiments Public'. The third, 'Classification Challenge', is highlighted in yellow. The table columns are CATALOG NAME, PUBLISHER NAME, SELF-PUB, and ACCESS TYPE. Below the table, it says 'Showing 1 to 3 of 3 Catalogs'.

CATALOG NAME	PUBLISHER NAME	SELF-PUB	ACCESS TYPE
acumos-int-fhg Internal	AI4EU Experiments	No	Restricted
AI4EU Experiments Public	acumos-int-fhg.ai4eu.eu	No	Public
Classification Challenge	DEV AI4EU Experiments	No	Public

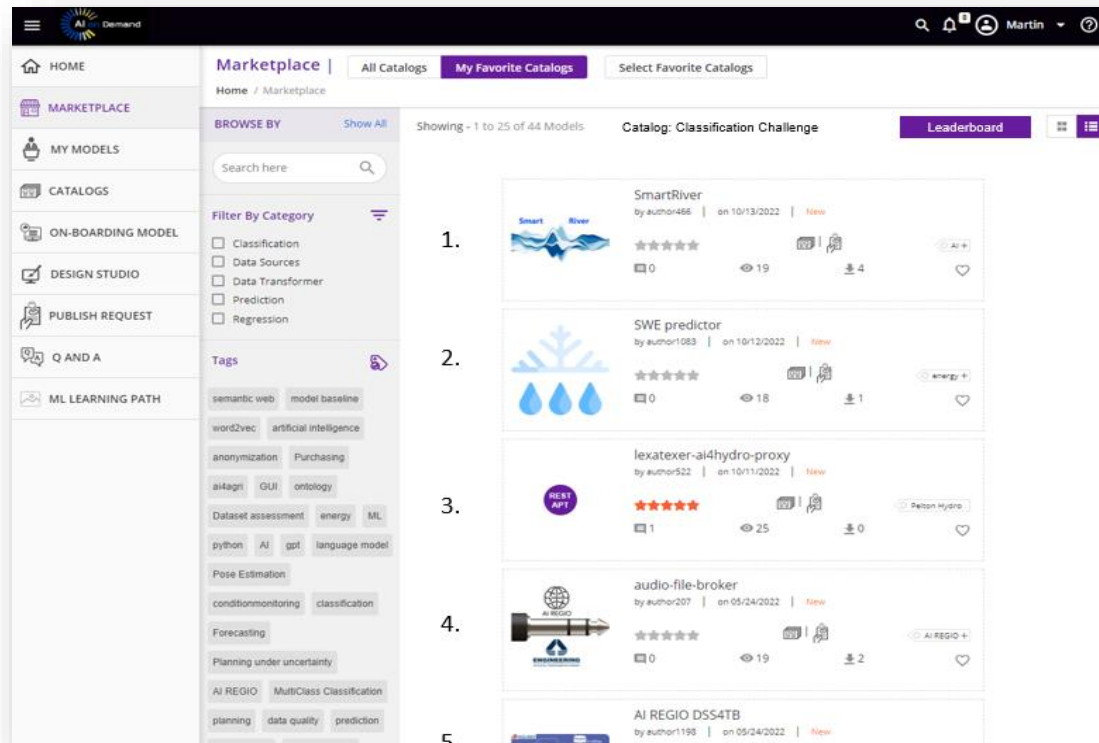
- The marketplace view consists actually of a set of catalogs
- Challenges will be special types of catalogs

Challenge Submissions



- Challenge submissions are handled just as „normal“ onboarding and publication, only the correct target catalog for the challenges must be selected

Leaderboard



- Based on the metrics format with type and comparison information the models can automatically be ranked in the Leaderboard

Timeline

- Metrics and creation of execution-run.json is already available on the production systems
- End of 2023: send execution-run.json from AI Playground to the Catalog
- May 2024: Challenges management completed