# Anonymity form for AI and machine learning models

## Model description

Did you use unsupervised or supervised learning for training the model?

unsupervised

supervised

other, please specify Click or tap here to write.

Is the model a generative, classification or regression model?

generative

classification

regression

other, please specify Click or tap here to write.

What type of model is it (e.g. Is it a neural network? What type of neural network is it and how many hidden layers does it contain? If it is not a neural network, how would you describe the architecture of the model?)

Click or tap here to write.

Did you use a pre-trained model? Which model did you use? How many of the model layers were unfrozen during training? How would you describe the differences between the pre-trained model and the developed model?

Click or tap here to write.

Enter a brief summary of the working principle of the model.

Click or tap here to write.

Enter the number of parameters in the model.

Click or tap here to write.

How has the model’s generalisation ability and the minimisation of overfitting been taken into account in the development of the model (e.g. data augmentation, regularisation methods)?

Click or tap here to write.

What format did you use for saving the model and in which format will you export it? Which files would you like to export in relation to the model?

Click or tap here to write.

## Privacy enhancing technologies

Did you use differential privacy or other privacy-enhancing technologies in the development of the model? Describe how these methods and/or model anonymity were otherwise taken into account during development.

Click or tap here to write.

What parameters did you use for privacy enhancing technologies? Also enter values for hyperparameters.

Click or tap here to write.

## Intended use of the model

Who are the target user group of the model?

Click or tap here to write.

Where is the model stored and who has access to use the model?

Click or tap here to write.

Who has access to view or handle the results of the model?

Click or tap here to write.

What information about the model’s architecture or operating principle are you going to publish?

Click or tap here to write.

Provide an assessment of the risk of the model being used for harmful purposes, particularly concerning the disclosure of data on the individuals in the training data.

Click or tap here to write.

## Description of the training data

What material did you use to train the model? Was the teaching data pre-processed before you used it to train the model? If so, explain in more detail how the training material was pre-processed.

Click or tap here to write.

Did you use individual-level material for training?

yes

no

Does the material used for training contain direct identifiers?

yes

no

How much data did you use to train the model? Indicate the number of instances of the training data and the number of individuals featured in it.

Click or tap here to write.

In what format did you enter the data for the model?

Click or tap here to write.

In the case of a classification model, how many classes did you use for classifying the training data?

Click or tap here to write.

Are there imbalances in the number of training instances belonging to different classes? If there is such class imbalance, describe the imbalance in these classes.

Click or tap here to write.

## Performance of the model and its generalisation ability

How has the model’s performance been assessed?

Click or tap here to write.

If you used separate test data to test the model’s performance (provided that this data has not been used to train the model), how much test data did you use?

Click or tap here to write.

What metrics have you used to assess the performance of the model and its generalisation ability? And what results did the performance metrics produce?

Click or tap here to write.

|  |  |
| --- | --- |
| Performance metric (e.g. accuracy) | Outcome |
| Enter text by clicking here. |  |
| Enter text by clicking here. |  |
| Enter text by clicking here. |  |
| Enter text by clicking here. |  |
| Enter text by clicking here. |  |

## Assessing and verifying the anonymity of the model

Have you tested the model’s resistance against attacks aimed at revealing the information contained in the training data (incl. membership inference attacks, reconstruction attacks)? If so, describe the process and its outcome.

Click or tap here to write.

On what grounds can it be assumed that the model is anonymous?

Click or tap here to write.

How can the anonymity of the model be verified?

Click or tap here to write.