

# FAQ on the model license

The license is an adapted version of the Apache 2.0 license, with added use-based restrictions. This document aims to give some guidance on the terms that depart from the standard Apache 2.0 license text.

## What licenses apply to the use of the model?

The model is released under the NB-license, and, where warm trained, dual-licensed with the license to the base model. Where the model has been warm trained, both licenses are binding on you as a user of the model.

## What kind of license is the NB-license?

The NB license is a partially reciprocal license.<sup>1</sup> The license is non-reciprocal in the sense that reuse of is free under any license you chose on the condition that you fulfil the conditions of the licenses that applies to the model. For the NB license you must document included copyright notices, the disclaimer and this license, and maintain the use-based restrictions section 4 for any downstream use of the provided NB model, either in its original form or as a derivative.

The National library has included use-based restrictions to ensure that the interests of intellectual property right holders in training material are protected in the downstream use of the model. This predominantly concerns different ways the model might give access to the original training data or might directly compete with the business model of the rights holders.

## What are the demands for use of the NB-license?

You must document the use of the model through including the copyright notice, the disclaimer and this license in the documentation of the model. You must also adhere to the use-based restrictions for the model and make these restrictions binding also on downstream users of the model or derivatives of the model.

## What are the use-based restrictions of the license?

---

<sup>1</sup> <https://opensource.org/blog/permmissive-and-copyleft-are-not-antonyms>

The license requires you to adhere to any applicable national, state, local, or other jurisdictional law, rule, regulation, or sanction. This means that the use of the model shall be lawful. You must adhere to applicable legislation in your use of our model, hereunder on intellectual property rights, data protection, anti-discrimination, and any sectorial legislation from e.g the AI act. This also includes regulation on deceptive business practices etc.

We also require you to refrain from intentionally using the model to recreate data the model has been trained on. It is well known that large language models can reproduce training data in model output, in particular when subject to so-called extraction attacks. This raises the issue of whether representations of training data in protected form remain in the model. The National library follows best practices to avoid that the model retains representations of the training data and reproduce examples of it. To protect the interest of right holders who have right to training data in the corpora we use, we require you to refrain from intentionally trying to gain access to this data.

The license further requires you to refrain from using the names of the National library and creators and publishers of training data in endorsement or promotion of your derivatives of the model without separate permission. As use of the model in adherence with the license is free, the model is supplied for a wide range of lawful uses in original or derivative form. Neither the National library nor right holders in training data grants any permission for use of their names to endorse or promote your use of derivatives of the model, in line with the choices you have made to build on it and use it in derivative form. Such endorsement or promotion requires specific prior written permission.

The license finally requires you to not use the model or its output to provide end-user services with the primary purpose of giving access to licensed press publications in training data. This provision is directed at business models that give alternative access to such content in the form of summaries, restatement or plain copies with the effect that it replaces the need to consult the content through the services of the press publications. A prime example is services that provide users with functional substitutes for access to press content and therefore remove the need to access the originals. Such use must be licensed with the relevant copyright holders. It should be noted that it is end-user directed services who has a substitutive provision of press publications as its primary purpose that are covered by the restriction and therefore requires separate agreements with the right holders. The relevant press publications are the ones licensed to train the model, you can find them listed here: []

## Can I lose the license to use the model in original or derivative form?

Yes, in special cases the license can be terminated. This is most relevant where you fail to adhere to the terms of the license in relation to down-stream uses or where you are in violation of the use-based restrictions. Prominently the intentional use of the model to gain access to training data will be considered such a violation. If such misuse is discovered you will get a warning from the National library, and if you do not remedy your violation(s) of the license terms, the license will be terminated.

Termination of this license does not affect any irrevocable license to parts of the model.