

BigScience Open RAIL-M License を読んでみよう

弁護士 渡邊 明彦

はじめに

前回、私は、「Carlos Munoz Ferrandis さんの「Towards open and responsible AI licensing frameworks」(Published August 31, 2022) https://huggingface.co/blog/open_rail の紹介」という記事を投稿した。Carlos Munoz Ferrandis さんこの論文（以下、「Ferrandis 論文」と言う）の中の一節に「OpenRAIL ライセンス契約書の 3 つの例は、最近リリースされた BigScience OpenRAIL-M、Stable Diffusion の CreativeML OpenRAIL-M、そして前記の 2 つの原型である BigScience BLOOMRAIL v1.0（このサイトにあるポスティングと FAQ、参照）です。」とあることも紹介した。今回は、その最初に掲げられた BigScience OpenRAIL-M License を検討してみる。（参照の便宜のために、BigScience OpenRAIL-M License を末尾に添付してあります。Stable Diffusion の CreativeML OpenRAIL-M License の内容を紹介した後に、順次「(日本語の) 翻訳」を行っていく予定です。）

I. 前提としての注意事項

BigScience OpenRAIL-M License は、「枠組み (フレームワーク)」としては普遍的な価値がありうるものの、その適用範囲は、具体的にはかなり限られていることに、先ず注意しておかなければならない。この License は、そのまま他の事例に流用できるようなものではない。

BigScience is not a consortium nor an officially incorporated entity. It's an open collaboration boot-strapped by [HuggingFace](#), [GENCI](#) and [IDRIS](#), and [organised as a research workshop](#). This research workshop gathers academic, industrial and independent researchers from many affiliations and whose research interests span many fields of research across AI, NLP, social sciences, legal, ethics and public policy.

まず、第一に、このライセンスを公表している BigScience は、「コンソーシアムでも、正式な法人でもない。」多数の研究者が集まって作業している「ワークショップ」である。つまり、当事者は、このプロジェクトに参加する個人・企業間の、ネットワークのような契約関係になる (nexus of contracts)。そして、

This Open RAIL-M License was created by BigScience, a collaborative open innovation project aimed at the responsible development and use of large multilingual datasets and Large Language Models (“LLMs”).

第二に、この OpenRAIL-M License は、BigScience が作成しているもので、「大規模な多言語データベース及び大規模言語モデル (LLM) の責任のある開発及び使用を目指すものである。」前提としている対象は、大規模多言語データベース及び大規模言語モデル (LLM) である。画像処理分野等については、想定されていない。

また、The Linux Foundation の「オープンソース コミュニティへの参加」¹ に見て取れるように、「オープンソース ソフトウェア」² の開発は「オープンソース プロジェクト」で行われており、それは「オープンソースコミュニティ」によって行われる。「オープンソースコミュニティ」のガバナンスは、「リーダー」、「メンテナー」、「コミッター」、「コントリビューター」、「ユーザー」という階層で構成される。

- **コントリビューター**：多くの人々が、コード、ドキュメント、およびその他のコントリビューションでオープンソース プロジェクトに貢献しています。これらのコントリビューションは、コントリビューションが採用される前に、経験豊富なコミッターやメンテナーからレビューを受けます。」

さらに、日本でもおなじみの「オープンイノベーション」という言葉も登場する。

オープンイノベーションはハーバード大学経営大学院の教授だった[ヘンリー・チェスブロウ](#) (英語版) によって提唱された概念で、組織内部の[イノベーション](#)を促進するため、企業の内部と外部との技術やアイデアの流動性を高め、組織内で創出されたイ

¹ <https://www.linuxfoundation.jp/resources/open-source-guides/participating-open-source-communities/>

² 「オープンソース」という概念については、Open Source Initiative の The Open Source Definition, <https://opensource.org/osd/> 参照。「オープンソース」という言葉が示唆するように「ソースコードの開示」が義務付けられる。

1. Free Redistribution

The license shall not restrict any party from selling or giving away the software as a component of an aggregate software distribution containing programs from several different sources. The license shall not require a royalty or other fee for such sale.

2. Source Code

The program must include source code, and must allow distribution in source code as well as compiled form. Where some form of a product is not distributed with source code, there must be a well-publicized means of obtaining the source code for no more than a reasonable reproduction cost, preferably downloading via the Internet without charge. The source code must be the preferred form in which a programmer would modify the program. Deliberately obfuscated source code is not allowed. Intermediate forms such as the output of a preprocessor or translator are not allowed.

3. Derived Works

The license must allow modifications and derived works, and must allow them to be distributed under the same terms as the license of the original software.

...

ノバージョンをさらに組織外に展開するイノベーションモデルをいう。³

「オープン」という言葉が、多義的に用いられることに注意したい。

以上が、BigScience OpenRAIL-M License は、BigScience の行っているプロジェクトの参加者に適用されること、適用対象となるプロジェクトは「大規模な多言語データベース及び大規模言語モデル (LLM)」であり、プロジェクトは「オープンソースプロジェクト」であること、以上は、このライセンス契約が人工知能 (AI) に関係することとは、いったんは切り離して考えられること。これらの前提条件の認識は、BigScience OpenRAIL-M License と人工知能 (AI) 関連契約の特徴を把握する上で重要であり、またその「責任のある開発及び使用」を理解する上で必要となる。

II. AI システムライセンスを特徴付けるもの

—Model、Derivatives of the Model、Complementary Material—

A. ソフトウェア (コード) ライセンスとは異なる

Ferrandis 論文で強調されているとおり、AI システムのライセンスは、ソフトウェアプログラム/コードのライセンスとは異なるのではないかという疑問を、誰しもが抱くかもしれない。しかしながら、AI システムのライセンス契約書は、ソフトウェアライセンス契約書を引き写したような内容、体裁のものが多く、「異なる点」を明確に示した文献・契約書書式は多くなかった。

Ferrandis 論文では、明示的に本件プロジェクトが機械学習 (ML) であることが強調されている。そして、BigScience OpenRAIL-M License では、

1. Model
2. Derivatives of the Model
3. Complementary Material

という用語が契約書の組立てに使用されている。

B. 「モデル」とは何か？

まず、この「モデル」という用語について、情報を収集しておこう。

³ <https://ja.wikipedia.org/wiki/オープンイノベーション>

「機械学習におけるモデルとは、その機械学習のプロセスそのもののアルゴリズムを指しています。」⁴

「An AI model is a program or algorithm that relies on training data to recognize patterns and make predictions or decisions.」⁵

また、AI Model、ML Model、DL Model と分けるものもある。⁶

C. 「モデル」の定義

BigScience OpenRAIL-M License では、第 1 条定義(d)号に「モデル」という用語が定義されて登場するので、「モデル」の定義を見てみよう。(これらの定義の「翻訳」は、全体の翻訳において行う予定である。)

- (d) “**Model**” means any accompanying machine-learning based assemblies (including checkpoints), consisting of learnt weights, parameters (including optimizer states), corresponding to the model architecture as embodied in the Complementary Material, that have been trained or tuned, in whole or in part on the Data, using the Complementary Material.

ここでの「モデル」の意味を把握するために、2つの方法を試してみたい。まず、BigScience OpenRAIL-M License が適用される対象の The BLOOM project である。

The BLOOM project was started by a co-founder of [Hugging Face](#). Six main groups of people were involved, including HuggingFace's BigScience team, the [Microsoft](#) DeepSpeed team, the [NVIDIA](#) Megatron-LM team, the [IDRIS/GENCI](#) team, the [PyTorch](#) team, and the volunteers in the BigScience Engineering workgroup.^[2] BLOOM was trained using data of 46 natural languages and 13 programming languages. In total, 1.6 TeraByte pre-processed text was converted into 350 billion unique tokens as BLOOM's training datasets.⁷

使い方についてであるが、Generate Content Using Open Source BLOOM AI model: Complete guide を見てみよう。⁸ また、別の文献では、Google Colab を経由しての AI

⁴ <https://ainow.ai/2021/04/13/254075/>

⁵ <https://www.koombea.com/blog/ai-model/>

⁶ <https://viso.ai/deep-learning/ml-ai-models/>

⁷ [https://en.wikipedia.org/wiki/BLOOM_\(language_model\)](https://en.wikipedia.org/wiki/BLOOM_(language_model))

⁸ <https://harishgarg.com/writing/generate-content-using-bloom-open-source-ai-model/>

Content generation の初歩的なやり方が紹介されている。⁹ 一定の設定のもとでの、生成された文章 3 例の紹介もある。

では、どうやってそのような文章が生成されるのかという点であるが、A one-year long research workshop on large multilingual models and datasets¹⁰ からリンクをたどっていくと¹¹、

You can find the model [here](#) and learn more by reading our [blog post](#).

Technical Specifications

Model Architecture and Objective

- Modified from Megatron-LM GPT2 (see [paper](#), [BLOOM Megatron code](#)):
- Decoder-only architecture
- Layer normalization applied to word embeddings layer (StableEmbedding; see [code](#), [paper](#))
- ALiBI positional encodings (see [paper](#)), with GeLU activation functions
 - 176,247,271,424 parameters:
 - 3,596,615,680 embedding parameters
 - 70 layers, 112 attention heads
 - Hidden layers are 14336-dimensional
 - Sequence length of 2048 tokens used (see [BLOOM tokenizer](#), [tokenizer description](#))

Objective Function: Cross Entropy with mean reduction (see [API documentation](#)).¹²

という「仕様」を確認できる。¹³

「チェックポイント」については、

「TensorFlow のモデルを保存する」という言いまわしは通常、次の 2 つのいずれかを意味します。

1. チェックポイント、または
2. 保存されたモデル (SavedModel)

チェックポイントは、モデルで使用されるすべてのパラメータ ([tf.Variable](#) オブジェ

⁹ <https://towardsdatascience.com/getting-started-with-bloom-9e3295459b65>

¹⁰ <https://bigscience.huggingface.co/>

¹¹ <https://huggingface.co/bigscience/bloom>

¹² <https://huggingface.co/bigscience/bloom>

¹³ <https://github.com/CompVis/stable-diffusion> も参照。

クト)の正確な値をキャプチャします。チェックポイントにはモデルで定義された計算のいかなる記述も含まれていないため、通常は、保存されたパラメータ値を使用するソースコードが利用可能な場合に限り有用です。

一方、SavedModel 形式には、パラメータ値(チェックポイント)に加え、モデルで定義された計算のシリアル化された記述が含まれています。この形式のモデルは、モデルを作成したソースコードから独立しています。したがって、TensorFlow Serving、TensorFlow Lite、TensorFlow.js、または他のプログラミング言語のプログラム(C、C++、Java、Go、Rust、C#などのTensorFlow API)を介したデプロイに適しています。

が分かりやすい。¹⁴

「モデルカード」¹⁵については、

機械学習モデルを適切に選定・運用する際には、そのモデルがどんなデータで学習され、何ができて何ができない(かもしれない)のかを後から確認できるようにすることが重要です。これに対してGoogleは2018年に「[Model Card for Model Reporting](#)」という論文を発表しました。

- モデルカードは機械学習モデルのドキュメント化のためのフレームワークである
- モデルカードにはタスクに関連する様々な条件でのベンチマークスコアを記載する
- 他にモデルの想定用途、詳しい評価方法なども記載する

という解説がある。¹⁶

AIシステムのライセンス契約書を取り扱う法律家としては、一応のイメージが把握できるのではなかろうか。¹⁷

D. 「コンプリメンタリマテリアル」

¹⁴ <https://www.tensorflow.org/guide/checkpoint?hl=ja>

¹⁵ This License governs the use of the model (and its derivatives) and is informed by the model card associated with the model.

¹⁶ <https://zenn.dev/dhirooka/articles/c2b3d52e31cd10>

¹⁷ youwht (著)『コピペで簡単実行! キテレッツおもしろ自然言語処理 Python と Colaboratory で身につく基礎の基礎』、斎藤 康毅 (著)『ゼロから作る Deep Learning — Python で学ぶディープラーニングの理論と実装』、『ゼロから作る Deep Learning ② — 自然言語処理編』が役に立つ。

- (f) “**Complementary Material**” means the accompanying source code and scripts used to define, run, load, benchmark or evaluate the Model, and used to prepare data for training or evaluation, if any. This includes any accompanying documentation, tutorials, examples, etc, if any.

「コンプリメンタリマテリアル」の意味は、斎藤 康毅 (著)『ゼロから作る Deep Learning —Python で学ぶディープラーニングの理論と実装』、『ゼロから作る Deep Learning ② —自然言語処理編』の中で学習する Python script のようなものであると理解してよからう。

III. Section II: INTELLECTUAL PROPERTY RIGHTS の意味と Derivatives of the Model

A. 著作権

BigScience OpenRAIL-M License の Section II は INTELLECTUAL PROPERTY RIGHTS となっている。内容は、著作権に基づく使用許諾と特許権に基づく実施許諾であって、従来のソフトウェア（コード）のライセンス契約では、セクションの標題そのものが「ライセンスの許諾 (Grant of License)」となる箇所である。

3. Grant of Copyright License. Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable copyright license to reproduce, prepare, publicly display, publicly perform, sublicense, and distribute the Complementary Material, the Model, and Derivatives of the Model.

通常のライセンス契約では `grant a license to use,, under the copyright` となるところが、`grant` とされる権利が `copyright license` となっている点で、いわゆる英米系の発想とはことなっている。

使用等できる対象は、the Complementary Material, the Model, and Derivatives of the Model である。Derivatives は、通常、著作権に関しては「二次的著作物」に該当する

十一 二次的著作物 著作物を翻訳し、編曲し、若しくは変形し、又は脚色し、映画化し、その他翻案することにより創作した著作物をいう。(著作権法第2条(定義))

A “derivative work” is a work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted. A work consisting of

editorial revisions, annotations, elaborations, or other modifications, which, as a whole, represent an original work of authorship, is a “derivative work”. (米国著作権法)

ところで、Derivatives of the Model の定義は、

- (e) **“Derivatives of the Model”** means all modifications to the Model, works based on the Model, or any other model which is created or initialized by transfer of patterns of the weights, parameters, activations or output of the Model, to the other model, in order to cause the other model to perform similarly to the Model, including - but not limited to - distillation methods entailing the use of intermediate data representations or methods based on the generation of synthetic data by the Model for training the other model.

となっているが、これは、各国の「二次的著作物」、derivative work と整合させて解釈する必要があろう。

B. 特許権

4. Grant of Patent License. Subject to the terms and conditions of this License and where and as applicable, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable (except as stated in this paragraph) patent license to make, have made, use, offer to sell, sell, import, and otherwise transfer the Model and the Complementary Material, where such license applies only to those patent claims licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Model to which such Contribution(s) was submitted. If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Model and/or Complementary Material or a Contribution incorporated within the Model and/or Complementary Material constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for the Model and/or Work shall terminate as of the date such litigation is asserted or filed.

実施許諾は、to make, have made, use, offer to sell, sell, import, and otherwise transfer the Model and the Complementary Material と一般的な特許実施許諾契約書の体裁になっている。また、実施許諾の対象は、著作権の場合とは異なり、the Model and the Complementary

Material である (Derivatives of the Model は、入っていない)。「モデル」、「コンプリメンタリマテリアル」が特許権の保護対象となっている場合に於てはまろう (わが国 (日本) でも、この分野の特許出願は盛んであり、実務的には適用の機会があるものと思われる。

the Model and the Complementary Material, where such license applies only to those patent claims licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Model to which such Contribution(s) was submitted.

この箇所は、例えば Fire Wire Patent Pool の例のように、

1394 必須特許 — 1394 規格 (「本規格」) の特別な構成を前提とすると、製品を本規格に完全に準拠するものとしようとすれば、本規格の要求するとおりに実装しなければならない場合に、本規格のいずれかの部分に準拠し及び/又は実装しようとすると、そのことにより、必然的及び不可避に、また文字どおりにその権利が侵害される、特許権 (複数の場合がある) の一切のクレームを意味するものとするが、但しかかるクレームのみに限るものとする。

の必須特許に相当するものと思われるが、Contribution という用語が、著作権法との関連でしか定義されていないため、理解が困難である。

If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Model and/or Complementary Material or a Contribution incorporated within the Model and/or Complementary Material constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for the Model and/or Work shall terminate as of the date such litigation is asserted or filed.

の箇所も、同じく 1394 特許の License に、

2.4 何れかのサブライセンシーが、本件ライセンサーに対して、或る 1394 必須特許 (複数の場合がある) 及び/又は 1394 規格に準拠し又はこれを実装すれば権利を侵害する可能性のある 1 項又は複数項のクレームをもつ何れかその他の「特許権」 (複数の場合がある) の権利侵害の中止を求める訴訟その他の手続を提起している場合において、かかるサブライセンシーが、訴訟その他の手続の基礎とする 1394 必須特許 (複数の場合がある) 及び/又はその他の「特許権」 (複数の場合がある) に基づくライセンスを、公平かつ合理的な条件で、かかる本件当事者に許諾することを拒絶しているときは、各

本件当事者は、かかるサブライセンシーに対しては、本件ライセンサーの 1394 特許ポートフォリオ内の「特許権」の何れかに基づく特定のサブライセンスを終了させるよう、本件ライセンス管理者に対して指示する権利を有するものとする。

と同様の趣旨の規定と思われる。

IV. 「責任のある使用」

Ferrandis 論文では、「Open & Responsible AI ライセンス（「OpenRAIL」）は、AI 固有のライセンスで、AI 関連制作物（AI artifact）の責任のある使用を義務付ける一方で」、との記述のあるとおり「責任のある使用」が、「オープンであること」と並んでもう一つのテーマとなっている。

この趣旨は、「使用（方法）に基づく制限」として第 9 条に規定され、その内容は、Attachment A に具体化されている。

- 9. Use-based restrictions.** The restrictions set forth in Attachment A are considered Use-based restrictions. Therefore You cannot use the Model and the Derivatives of the Model for the specified restricted uses. You may use the Model subject to this License, including only for lawful purposes and in accordance with the License. Use may include creating any content with, finetuning, updating, running, training, evaluating and/or reparametrizing the Model. You shall require all of Your users who use the Model or a Derivative of the Model to comply with the terms of this paragraph (paragraph 5).

ところで、Attachment A の内容であるが、例えば、サブスクリプション契約には、これと同様な規定があり、

(d) お客様は、本プラットフォーム及び SFDC サービスを、お客様の内部事業目的のためにのみ利用するものとし、以下のことを行わないものとします。(i) 本プラットフォーム又は SFDC サービスを、ライセンス、サブライセンス、販売、再販、賃貸、リース、移転、譲渡、頒布、タイムシェアリング若しくはその他商業上の利用、又は本ユーザ以外の第三者に対して、若しくはその他この SFDC サービス規約で規定された以外の方法で提供すること (ii) 適用ある法令に違反してスパム又はその他の反復メッセージ若しくは迷惑メールを送信すること (iii) 児童に対して有害なもの及び第三者のプライバシーの権利を侵害するものを含め、第三者の権利を侵害するもの、猥褻なもの、脅迫的なもの、第三者の名誉を毀損するもの、その他違法若しくは不法ものを送付し又は保存すること、(iv) ウィルス、ワーム、時限爆弾、トロイの木馬、又はその他の有害若しくは悪意のあるコード、ファイル、スクリプト、エージェント若し

くはプログラムを送信 又は保存すること (v) 本プラットフォーム又はSFDCサービス 又はそれらに含まれるデータの 完全性又は性能を妨害し、又は混乱させること、又は (vi) 本プラットフォーム又はSFDCサービス、又はそれらに関連するシステム若しくはネットワークに対する不正なアクセスを試みるこ と。

文面上は、特に目新しいものではないが、人工知能 (AI)、特に「生成 AI」との関係では、文言が同じ、または類似していても、まったく新しい意味を持ちうる可能性がある。

Ferrandis 論文では、次のように述べられている。

責任のあること : OpenRAIL ライセンスは、特定された緊要なシナリオの中での使用許諾 AI 制作物の使用に関する一連の具体的な規制事項が埋め込まれています。機械学習 (ML) 開発に対するエビデンスに基づくアプローチから、ユーザーベースの規制事項が導かれており、また機械学習への幅広いアクセスと使用を促進することと、人工知能 (AI) 制作物のオープンに使用許諾の有害な使用から生ずる潜在的な社会的費用との間に一線を画することを強いる、使用法の規制が生まれます。したがって、機械学習 (ML) モデルにオープンなアクセスを認めることから生ずる便益を享受する一方で、ユーザーは、指定された規制に該当するシナリオではそのモデルを使用できなくなります。

(完)

BigScience Open RAIL-M License

dated August 18, 2022

Section I: PREAMBLE

This Open RAIL-M License was created by BigScience, a collaborative open innovation project aimed at the responsible development and use of large multilingual datasets and Large Language Models (“LLMs”). While a similar license was originally designed for the BLOOM model, we decided to adapt it and create this license in order to propose a general open and responsible license applicable to other machine learning based AI models (e.g. multimodal generative models).

In short, this license strives for both the open and responsible downstream use of the accompanying model. When it comes to the open character, we took inspiration from open source permissive licenses regarding the grant of IP rights. Referring to the downstream responsible use, we added use-based restrictions not permitting the use of the Model in very specific scenarios, in order for the licensor to be able to enforce the license in case potential misuses of the Model may occur. Even though downstream derivative versions of the model could be released under different licensing terms, the latter will always have to include - at minimum - the same use-based restrictions as the ones in the original license (this license).

The development and use of artificial intelligence (“AI”), does not come without concerns. The world has witnessed how AI techniques may, in some instances, become risky for the public in general. These risks come in many forms, from racial discrimination to the misuse of sensitive information.

BigScience believes in the intersection between open and responsible AI development; thus, this License aims to strike a balance between both in order to enable responsible open-science in the field of AI.

This License governs the use of the model (and its derivatives) and is informed by the model card associated with the model.

NOW THEREFORE, You and Licensor agree as follows:

5. Definitions

- (e) **"License"** means the terms and conditions for use, reproduction, and Distribution as defined in this document.
- (f) **"Data"** means a collection of information and/or content extracted from the dataset used with the Model, including to train, pretrain, or otherwise evaluate the Model. The Data is not licensed under this License.
- (g) **"Output"** means the results of operating a Model as embodied in informational content resulting therefrom.

- (h) **“Model”** means any accompanying machine-learning based assemblies (including checkpoints), consisting of learnt weights, parameters (including optimizer states), corresponding to the model architecture as embodied in the Complementary Material, that have been trained or tuned, in whole or in part on the Data, using the Complementary Material.
- (i) **“Derivatives of the Model”** means all modifications to the Model, works based on the Model, or any other model which is created or initialized by transfer of patterns of the weights, parameters, activations or output of the Model, to the other model, in order to cause the other model to perform similarly to the Model, including - but not limited to - distillation methods entailing the use of intermediate data representations or methods based on the generation of synthetic data by the Model for training the other model.
- (j) **“Complementary Material”** means the accompanying source code and scripts used to define, run, load, benchmark or evaluate the Model, and used to prepare data for training or evaluation, if any. This includes any accompanying documentation, tutorials, examples, etc, if any.
- (k) **“Distribution”** means any transmission, reproduction, publication or other sharing of the Model or Derivatives of the Model to a third party, including providing the Model as a hosted service made available by electronic or other remote means - e.g. API-based or web-access.
- (l) **“Licensor”** means the copyright owner or entity authorized by the copyright owner that is granting the License, including the persons or entities that may have rights in the Model and/or distributing the Model.
- (m) **"You"** (or **"Your"**) means an individual or Legal Entity exercising permissions granted by this License and/or making use of the Model for whichever purpose and in any field of use, including usage of the Model in an end-use application - e.g. chatbot, translator, image generator.
- (n) **“Third Parties”** means individuals or legal entities that are not under common control with Licensor or You.
- (o) **"Contribution"** means any work of authorship, including the original version of the Model and any modifications or additions to that Model or Derivatives of the Model thereof, that is intentionally submitted to Licensor for inclusion in the Model by the copyright owner or by an individual or Legal Entity authorized to submit on behalf of the copyright owner. For the purposes of this definition, **“submitted”** means any form of electronic, verbal, or written communication sent to the Licensor or its representatives, including but not limited to communication on electronic mailing lists, source code control systems, and issue tracking systems that are managed by, or on behalf of, the Licensor for the purpose of discussing and improving the Model, but excluding communication that is conspicuously marked or otherwise designated in writing by the copyright owner as **"Not a Contribution."**
- (p) **"Contributor"** means Licensor and any individual or Legal Entity on behalf of whom a Contribution has been received by Licensor and subsequently incorporated within the Model.

Section II: INTELLECTUAL PROPERTY RIGHTS

Both copyright and patent grants apply to the Model, Derivatives of the Model and Complementary Material. The Model and Derivatives of the Model are subject to additional terms as described in Section III.

6. Grant of Copyright License. Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable

copyright license to reproduce, prepare, publicly display, publicly perform, sublicense, and distribute the Complementary Material, the Model, and Derivatives of the Model.

7. Grant of Patent License. Subject to the terms and conditions of this License and where and as applicable, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable (except as stated in this paragraph) patent license to make, have made, use, offer to sell, sell, import, and otherwise transfer the Model and the Complementary Material, where such license applies only to those patent claims licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Model to which such Contribution(s) was submitted. If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Model and/or Complementary Material or a Contribution incorporated within the Model and/or Complementary Material constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for the Model and/or Work shall terminate as of the date such litigation is asserted or filed.

Section III: CONDITIONS OF USAGE, DISTRIBUTION AND REDISTRIBUTION

8. Distribution and Redistribution. You may host for Third Party remote access purposes (e.g. software-as-a-service), reproduce and distribute copies of the Model or Derivatives of the Model thereof in any medium, with or without modifications, provided that You meet the following conditions:

- a. Use-based restrictions as referenced in paragraph 5 **MUST** be included as an enforceable provision by You in any type of legal agreement (e.g. a license) governing the use and/or distribution of the Model or Derivatives of the Model, and You shall give notice to subsequent users You Distribute to, that the Model or Derivatives of the Model are subject to paragraph 5. **This provision does not apply to the use of Complementary Material.**
- b. You must give any Third Party recipients of the Model or Derivatives of the Model a copy of this License;
- c. You must cause any modified files to carry prominent notices stating that You changed the files;
- d. You must retain all copyright, patent, trademark, and attribution notices excluding those notices that do not pertain to any part of the Model, Derivatives of the Model.

You may add Your own copyright statement to Your modifications and may provide additional or different license terms and conditions - **respecting** paragraph 4.a. - for use, reproduction, or Distribution of Your modifications, or for any such Derivatives of the Model as a whole, provided Your use, reproduction, and Distribution of the Model otherwise complies with the conditions stated in this License.

9. Use-based restrictions. The restrictions set forth in Attachment A are considered Use-based restrictions. Therefore You cannot use the Model and the Derivatives of the Model for the specified restricted uses. You may use the Model subject to this License, including only for lawful purposes and in accordance with the License. Use may include creating any content with, finetuning, updating, running, training, evaluating and/or reparametrizing the Model. You shall require all of Your users who use the Model or a Derivative of the Model to comply with the terms of this paragraph (paragraph 5).

10. The Output You Generate. Except as set forth herein, Licensor claims no rights in the Output You generate using the Model. You are accountable for the Output you generate and its subsequent uses. No use of the output can contravene any provision as stated in the License.

Section IV: OTHER PROVISIONS

11. Updates and Runtime Restrictions. To the maximum extent permitted by law, Licensor reserves the right to restrict (remotely or otherwise) usage of the Model in violation of this License, update the Model through electronic means, or modify the Output of the Model based on updates. You shall undertake reasonable efforts to use the latest version of the Model.

12. Trademarks and related. Nothing in this License permits You to make use of Licensors' trademarks, trade names, logos or to otherwise suggest endorsement or misrepresent the relationship between the parties; and any rights not expressly granted herein are reserved by the Licensors.

13. Disclaimer of Warranty. Unless required by applicable law or agreed to in writing, Licensor provides the Model and the Complementary Material (and each Contributor provides its Contributions) on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied, including, without limitation, any warranties or conditions of TITLE, NON-INFRINGEMENT, MERCHANTABILITY, or FITNESS FOR A PARTICULAR PURPOSE. You are solely responsible for determining the appropriateness of using or redistributing the Model, Derivatives of the Model, and the Complementary Material and assume any risks associated with Your exercise of permissions under this License.

14. Limitation of Liability. In no event and under no legal theory, whether in tort (including negligence), contract, or otherwise, unless required by applicable law (such as deliberate and grossly negligent acts) or agreed to in writing, shall any Contributor be liable to You for damages, including any direct, indirect, special, incidental, or consequential damages of any character arising as a result of this License or out of the use or inability to use the Model and the Complementary Material (including but not limited to damages for loss of goodwill, work stoppage, computer failure or malfunction, or any and all other commercial damages or losses), even if such Contributor has been advised of the possibility of such damages.

15. Accepting Warranty or Additional Liability. While redistributing the Model, Derivatives of the Model and the Complementary Material thereof, You may choose to offer, and charge a fee for, acceptance of support, warranty, indemnity, or other liability obligations and/or rights consistent with this License. However, in accepting such obligations, You may act only on Your own behalf and on Your sole responsibility, not on behalf of any other Contributor, and only if You agree to indemnify, defend, and hold each Contributor harmless for any liability incurred by, or claims asserted against, such Contributor by reason of your accepting any such warranty or additional liability.

16. If any provision of this License is held to be invalid, illegal or unenforceable, the remaining provisions shall be unaffected thereby and remain valid as if such provision had not been set forth herein.

END OF TERMS AND CONDITIONS

Attachment A

Use Restrictions

You agree not to use the Model or Derivatives of the Model:

- (e) In any way that violates any applicable national, federal, state, local or international law or regulation;
- (f) For the purpose of exploiting, harming or attempting to exploit or harm minors in any way;
- (g) To generate or disseminate verifiably false information and/or content with the purpose of harming others;
- (h) To generate or disseminate personal identifiable information that can be used to harm an individual;
- (i) To generate or disseminate information and/or content (e.g. images, code, posts, articles), and place the information and/or content in any context (e.g. bot generating tweets) without expressly and intelligibly disclaiming that the information and/or content is machine generated;
- (j) To defame, disparage or otherwise harass others;
- (k) To impersonate or attempt to impersonate (e.g. deepfakes) others without their consent;
- (l) For fully automated decision making that adversely impacts an individual's legal rights or otherwise creates or modifies a binding, enforceable obligation;
- (m) For any use intended to or which has the effect of discriminating against or harming individuals or groups based on online or offline social behavior or known or predicted personal or personality characteristics;
- (n) To exploit any of the vulnerabilities of a specific group of persons based on their age, social, physical or mental characteristics, in order to materially distort the behavior of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm;
- (o) For any use intended to or which has the effect of discriminating against individuals or groups based on legally protected characteristics or categories;
- (p) To provide medical advice and medical results interpretation;
- (q) To generate or disseminate information for the purpose to be used for administration of justice, law enforcement, immigration or asylum processes, such

as predicting an individual will commit fraud/crime commitment (e.g. by text profiling, drawing causal relationships between assertions made in documents, indiscriminate and arbitrarily-targeted use).