Artificial Intelligence (High-Risk Systems) Bill

A

BILL

TO

prohibit high-risk AI practices and introduce regulations for greater AI transparency and market fairness, and for connected purposes.

BE IT ENACTED by the King's Most Excellent Majesty, by and with the advice and consent of the Lords Temporal, and Commons, in this present Parliament, assembled, and by the authority of the same, as follows —

Chapter 1: General Provisions

Section 1: Definitions

For the purposes of this Act, the following terms apply —

(1) artificial intelligence system' (AI system) refers to software that is developed with one or more of the techniques and approaches listed in Schedule 3 and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;

(2) 'provider' refers to a natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark, whether for payment or free of charge;

(3) 'small-scale provider' refers to a provider ~~that is a micro or small enterprise~~ With fewer than 50 employees

(4) Any distributor, importer, user or other party shall be considered a provider and shall be subject to the obligations of the provider under Section 15, in any of the following circumstances —

> (a) they place on the market or put into service a high-risk AI system under their name or trademark;
>
> (b) they modify the intended purpose of a high-risk AI system already placed on the market or put into service;
>
> (c) they make a substantial modification to the high-risk AI system.

(5) ‘user’ means any natural or legal person, public authority, agency or other body using an AI system under its authority, except where the AI system is used in the course of a personal non-professional activity;

(6) ‘authorised representative’ means any natural or legal person established in the UK who has received a written mandate from a provider of an AI system to, respectively, perform and carry out on its behalf the obligations and procedures established by this Act;

(7) ‘importer’ means any natural or legal person established in the UK that places on the market or puts into service an AI system that bears the name or trademark of a natural or legal person established outside the UK;

(8) ‘distributor’ means any natural or legal person in the supply chain, other than the provider or the importer, that makes an AI system available on the UK market without affecting its properties;

(9) ‘operator’ means the provider, the user, the authorised representative, the importer and the distributor;

Section 2: Purpose

(1) This Act provides the provisions for the the establishment of —

> (a) prohibitions of certain artificial intelligence practices;
>
> (b) specific requirements for high-risk AI systems and obligations for operators of such systems;
>
> (c) transparency rules for AI systems intended to interact with natural persons, emotion recognition systems and biometric categorisation systems, and AI systems used to generate or manipulate image, audio or video content;
>
> (d) rules on market monitoring and surveillance.

Section 3: Scope

(1) The scope of this Act shall apply to the following —
>
> (a) providers placing on the market or putting into service AI systems in the United Kingdom, irrespective of whether those providers are established within the UK or in a foreign country;
>
> (b) users of AI systems located within the United kingdom;
>
> (c) providers and users of AI systems that are located in a foreign country, where the output produced by the system is used in the United Kingdom;

~~(2) This Act shall not apply to AI systems developed or used exclusively for military purposes.~~

(2) This Act shall not apply to public authorities in a foreign country nor to international organisations falling within the scope of this Act pursuant to Subsection (1), where those authorities or organisations use AI systems in the framework of international agreements for law enforcement and judicial cooperation with the United Kingdom.

Chapter 2: Prohibited Practices

Section 4: Prohibited Artificial Intelligence Practices

(1) The following artificial intelligence practices shall hereby be prohibited —

> (a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm;
>
> (b) the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour 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;
>
~~> (c) the placing on the market, putting into service or use of AI systems by public authorities or on their behalf for the evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known or predicted personal or personality characteristics, with the social score leading to either or both of the following~~

(c) the placing on the market, putting into service or use of AI systems for the evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known or predicted personal or personality characteristics, with the social score leading to either or both of the following—
>
>> (i) detrimental or unfavourable treatment of certain natural persons or whole groups thereof in social contexts which are unrelated to the contexts in which the data was originally generated or collected;
>>
>> (ii) detrimental or unfavourable treatment of certain natural persons or whole groups thereof that is unjustified or disproportionate to their social behaviour or its gravity;
>
> (d) the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement, ~~unless and in as far as such use is strictly necessary for one of the following objectives —~~
~~>~~

>> (i) the targeted search for specific potential victims of crime, including missing children;
>>
>> (ii) the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or of a terrorist attack.

> (e) the placing on the market, putting into service or use of an AI system for the purpose of influencing political processes, including elections or referenda, in a manner that manipulates or distorts democratic discourse or electoral outcomes.

(2) The use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement for any of the objectives referred to in Subsection (1)(d) shall take into account the following elements —

> (a) the nature of the situation giving rise to the possible use, in particular the seriousness, probability and scale of the harm caused in the absence of the use of the system;
>
> (b) the consequences of the use of the system for the rights and freedoms of all persons concerned, in particular the seriousness, probability and scale of those consequences.

In addition, the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement for any of the objectives referred to in Subsection (1)(d) shall comply with necessary and proportionate safeguards and conditions in relation to the use, in particular as regards the temporal, geographic and personal limitations.

(3) As regards Subsection (1)(d) and Subsection (2), each individual use for the purpose of law enforcement of a 'real-time' remote biometric identification system in publicly accessible spaces shall be subject to a prior authorisation granted by a judicial authority or by an independent administrative authority in which the use is to take place, issued upon a reasoned request and in accordance with the detailed rules of relevant law referred to in Subsection (4) where —

> (a) the competent judicial or administrative authority shall only grant the authorisation where it is satisfied, based on objective evidence or clear indications presented to it, that the use of the 'real-time' remote biometric identification system at issue is necessary for and proportionate to achieving one of the objectives specified in Subsection (1)(d), as identified in the request; and

> (b) in deciding on the request, the competent judicial or administrative authority shall take into account the elements referred to in Subsection (2).

Chapter 3: Classification Systems

Section 5: Classification Rules of High-risk AI systems

(1) Irrespective of whether an AI system is placed on the market or put into service independently from the products referred to in points (a) and (b), that AI system shall be considered high-risk where both of the following conditions are fulfilled —

> (a) the AI system is intended to be used as a safety component of a product, or is itself a product; and
>
> (b) the product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a conformity assessment with a view to the placing on the market or putting into service of that product.

(2) In addition to the high-risk AI systems referred to in Subsection (1), AI systems referred to in Schedule 1 shall also be considered high-risk.

Section 6: Powers of the Secretary of State

(1) The Secretary of State shall be empowered to set regulations, via secondary legislation, to update the list in Schedule 1 by adding or removing high-risk AI systems where both of the following conditions are fulfilled —

> (a) the AI systems are intended to be used in any of the areas listed in points 1 to 8 of Schedule 1;
>
> (b) the AI systems pose a risk of harm to the health and safety, or a risk of adverse impact on fundamental rights, that is, in respect of its severity and probability of occurrence, equivalent to or greater than the risk of harm or of adverse impact posed by the high-risk AI systems already referred to in Schedule 1.

(2) When assessing for the purposes of Subsection (1) whether an AI system poses a risk of harm to the health and safety or a risk of adverse impact on fundamental rights that is equivalent to or greater than the risk of harm posed by the high-risk AI systems already referred to in Schedule 1, the Secretary of State shall take into account the following criteria —

> (a) the intended purpose of the AI system;
>
> (b) the extent to which an AI system has been used or is likely to be used;
>
> (c) the extent to which the use of an AI system has already caused harm to the health and safety or adverse impact on the fundamental rights or has given rise to significant concerns in relation to the materialisation of such harm or adverse impact, as demonstrated by reports or documented allegations submitted to the competent authorities;
>
> (d) the potential extent of such harm or such adverse impact, in particular in terms of its intensity and its ability to affect a plurality of persons;
>

> (e) the extent to which potentially harmed or adversely impacted persons are dependent on the outcome produced with an AI system, in particular because for practical or legal reasons it is not reasonably possible to opt-out from that outcome;
>
> (f) the extent to which potentially harmed or adversely impacted persons are in a vulnerable position in relation to the user of an AI system, in particular due to an imbalance of power, knowledge, economic or social circumstances, or age;
>
> (g) the extent to which the outcome produced with an AI system is easily reversible, whereby outcomes having an impact on the health or safety of persons shall not be considered as easily reversible;
>
> (h) the extent to which existing legislation provides for —
>
>> (i) effective measures of redress in relation to the risks posed by an AI system, with the exclusion of claims for damages;
>>
>> (ii) effective measures to prevent or substantially minimise those risks.

Chapter 4: High-Risk Systems Requirements

Section 7: Requirement Compliance

(1) High-risk AI systems shall comply with the requirements established in this Chapter.

(2) The intended purpose of the high-risk AI system and the risk management system referred to in Section 8 shall be taken into account when ensuring compliance with those requirements.

Section 8: Risk management system

(1) A risk management system shall be established, implemented, documented and maintained in relation to high-risk AI systems.

(2) The risk management system shall consist of a continuous iterative process run throughout the entire lifecycle of a high-risk AI system, requiring regular systematic updating. It shall comprise the following steps —

> (a) identification and analysis of the known and foreseeable risks associated with each high-risk AI system;
>
> (b) estimation and evaluation of the risks that may emerge when the high-risk AI system is used in accordance with its intended purpose and under conditions of reasonably foreseeable misuse;
>

> (c) evaluation of other possibly arising risks based on the analysis of data gathered from the post-market monitoring system referred to in Section 23;
>
> (d) adoption of suitable risk management measures in accordance with the provisions of the following paragraphs.

(3) The risk management measures referred to in Subsection (2)(d) shall give due consideration to the effects and possible interactions resulting from the combined application of the requirements set out in this Chapter, where —

> (a) they shall take into account the generally acknowledged state of the art, including as reflected in relevant standards or specifications.

(4) The risk management measures referred to in Subsection (2)(d) shall be such that any residual risk associated with each hazard as well as the overall residual risk of the high-risk AI systems is judged acceptable, provided that the high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, and where those residual risks shall be communicated to the user. In identifying the most appropriate risk management measures, the following shall be ensured —

> (a) elimination or reduction of risks as far as possible through adequate design and development;
>
> (b) where appropriate, implementation of adequate mitigation and control measures in relation to risks that cannot be eliminated;
>
> (c) provision of adequate information pursuant to Section 12, in particular as regards the risks referred to in Subsection (2)(b) of this Section, and, where appropriate, training to users.

In eliminating or reducing risks related to the use of the high-risk AI system, due consideration shall be given to the technical knowledge, experience, education, training to be expected by the user and the environment in which the system is intended to be used.

(5) High-risk AI systems shall be tested for the purposes of identifying the most appropriate risk management measures, where testing shall ensure —

> (a) that high-risk AI systems perform consistently for their intended purpose; and
>
> (b) they are in compliance with the requirements set out in this Chapter.

(6) Testing procedures shall be suitable to achieve the intended purpose of the AI system and do not need to go beyond what is necessary to achieve that purpose.

(7) The testing of the high-risk AI systems shall be performed, as appropriate, at any point in time throughout the development process, and, in any event, prior to the placing on the market or the putting into service, where testing shall be made against preliminarily defined

metrics and probabilistic thresholds that are appropriate to the intended purpose of the high-risk AI system.

(8) When implementing the risk management system described in Subsections (1) to (7), specific consideration shall be given to whether the high-risk AI system is likely to be accessed by or have an impact on children.

Section 9: Data and Data Governance

(1) High-risk AI systems which make use of techniques involving the training of models with data shall be developed on the basis of training, validation and testing data sets that meet the quality criteria referred to in Subsections (2) to (5).

(2) Training, validation and testing data sets shall be subject to appropriate data governance and management practices. Those practices shall concern in particular —

> (a) the relevant design choices;
>
> (b) data collection;
>
> (c) relevant data preparation processing operations, such as annotation, labelling, cleaning, enrichment and aggregation;
>
> (d) the formulation of relevant assumptions, notably with respect to the information that the data are supposed to measure and represent;
>
> (e) a prior assessment of the availability, quantity and suitability of the data sets that are needed;
>
> (f) examination in view of possible biases; and
>
> (g) the identification of any possible data gaps or shortcomings, and how those gaps and shortcomings can be addressed.

(3) Training, validation and testing data sets shall be relevant, representative, free of errors and complete where they shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons on which the high-risk AI system is intended to be used, in which these characteristics of the data sets may be met at the level of individual data sets or a combination thereof.

(4) Training, validation and testing data sets shall take into account, to the extent required by the intended purpose, the characteristics or elements that are particular to the specific geographical, behavioural or functional setting within which the high-risk AI system is intended to be used.

(5) Appropriate data governance and management practices shall apply for the development of high-risk AI systems other than those which make use of techniques involving the training of models in order to ensure that those high-risk AI systems comply with Subsection (2).

Section 10: Technical Documentation

(1) The technical documentation of a high-risk AI system shall be drawn up —

> (a) before that system is placed on the market or put into service and shall be kept up-to-date; and
>
> (b) in such a way to demonstrate that the high-risk AI system complies with the requirements set out in this Chapter and provide competent authorities and notified bodies with all the necessary information to assess the compliance of the AI system with those requirements. It shall contain, at a minimum, the elements set out in Schedule 2.

(2) Where a high-risk AI system related to a product is placed on the market or put into service one single technical documentation shall be drawn up containing all the information set out in Schedule 2 as well as the information required under those legal acts.

(3) The Secretary of State shall set regulations, via secondary legislation, to amend Schedule 2 where necessary to ensure that, in the light of technical progress, the technical documentation provides all the necessary information to assess the compliance of the system with the requirements set out in this Chapter.

(4) Regulations set under this Section shall be subject to negative procedures.

Section 11: Record-Keeping

(1) High-risk AI systems shall be designed and developed with capabilities enabling the automatic recording of events (‘logs’) while the high-risk AI systems are operating, in which those logging capabilities shall conform to recognised standards or specifications.

(2) The logging capabilities shall ensure a level of traceability of the AI system’s functioning throughout its lifecycle that is appropriate to the intended purpose of the system.

(3) In particular, logging capabilities shall enable the monitoring of the operation of the high-risk AI system with respect to the occurrence of situations that may result in the AI system presenting a risk or lead to a substantial modification, and facilitate the post-market monitoring referred to in Section 23.

(4) For high-risk AI systems referred to in Subsection (1)(a) of Schedule 1, the logging capabilities shall provide, at a minimum —

> (a) recording of the period of each use of the system (start date and time and end date and time of each use);

>
> (b) the reference database against which input data has been checked by the system;
>
> (c) the input data for which the search has led to a match;
>
> (d) the identification of the natural persons involved in the verification of the results, as referred to in Section 13(5).

(5) the retention period for logs shall be no longer than is necessary to serve the purposes for which they are processed, with clear guidelines on the conditions for their deletion or anonymization once the retention period expires.

Section 12: Transparency and provision of information to users

(1) High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately where an appropriate type and degree of transparency shall be ensured, with a view to achieving compliance with the relevant obligations of the user and of the provider.

(2) High-risk AI systems shall be accompanied by instructions for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to users.

(3) The information referred to in Subsection (2) shall specify —

> (a) the identity and the contact details of the provider and, where applicable, of its authorised representative;
>
> (b) the characteristics, capabilities and limitations of performance of the high-risk AI system, including —
>
>> (i) its intended purpose;
>>
>> (ii) the level of accuracy, robustness and cybersecurity referred to in Section 14 against which the high-risk AI system has been tested and validated and which can be expected, and any known and foreseeable circumstances that may have an impact on that expected level of accuracy, robustness and cybersecurity;
>>
>> (iii) any known or foreseeable circumstance, related to the use of the high-risk AI system in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, which may lead to risks to the health and safety or fundamental rights;
>>
>> (iv) its performance as regards the persons or groups of persons on which the system is intended to be used;
>>

>> (v) when appropriate, specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets used, taking into account the intended purpose of the AI system.
>
> (c) the changes to the high-risk AI system and its performance which have been pre-determined by the provider at the moment of the initial conformity assessment, if any;
>
> (d) the human oversight measures referred to in Section 13, including the technical measures put in place to facilitate the interpretation of the outputs of AI systems by the users;
>
> (e) the expected lifetime of the high-risk AI system and any necessary maintenance and care measures to ensure the proper functioning of that AI system, including as regards software updates.

Section 13: Human oversight

(1) High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use.

(2) Human oversight shall aim at preventing or minimising the risks to health, safety or fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, in particular when such risks persist notwithstanding the application of other requirements set out in this Chapter.

(3) Human oversight shall be ensured through either one or all of the following measures —

> (a) identified and built, when technically feasible, into the high-risk AI system by the provider before it is placed on the market or put into service;
>
> (b) identified by the provider before placing the high-risk AI system on the market or putting it into service and that is appropriate to be implemented by the user.

(4) The measures referred to in Subsection (3) shall enable the individuals to whom human oversight is assigned to do the following, as appropriate to the circumstances —

> (a) fully understand the capacities and limitations of the high-risk AI system and be able to duly monitor its operation, so that signs of anomalies, dysfunctions and unexpected performance can be detected and addressed as soon as possible;
>
> (b) remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system (‘automation bias’), in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons;
>

> (c) be able to correctly interpret the high-risk AI system's output, taking into account in particular the characteristics of the system and the interpretation tools and methods available;
>
> (d) be able to decide, in any particular situation, not to use the high-risk AI system or otherwise disregard, override or reverse the output of the high-risk AI system;
>
> (e) be able to intervene in the operation of the high-risk AI system or interrupt the system through a "stop" button or a similar procedure.

(5) For high-risk AI systems referred to in point 1(a) of Schedule 1, the measures referred to in Subsection (3) of this Section shall be such as to ensure that, in addition, no action or decision is taken by the user on the basis of the identification resulting from the system unless this has been verified and confirmed by at least two natural persons.

Section 14: Accuracy, Robustness and Cybersecurity

(1) High-risk AI systems shall be designed and developed in such a way that they achieve, in the light of their intended purpose, an appropriate level of accuracy, robustness and cybersecurity, and perform consistently in those respects throughout their lifecycle.

(2) The levels of accuracy and the relevant accuracy metrics of high-risk AI systems shall be declared in the accompanying instructions of use.

(3) High-risk AI systems shall be resilient as regards errors, faults or inconsistencies that may occur within the system or the environment in which the system operates, in particular due to their interaction with natural persons or other systems, whereby —

> (a) the robustness of high-risk AI systems may be achieved through technical redundancy solutions, which may include backup or fail-safe plans; and
>
> (b) high-risk AI systems that continue to learn after being placed on the market or put into service shall be developed in such a way to ensure that possibly biased outputs due to outputs used as an input for future operations ('feedback loops') are duly addressed with appropriate mitigation measures.

(4) High-risk AI systems shall be resilient as regards attempts by unauthorised third parties to alter their use or performance by exploiting the system vulnerabilities, where the technical solutions —

> (a) aimed at ensuring the cybersecurity of high-risk AI systems shall be appropriate to the relevant circumstances and the risks; and
>
> (b) to address AI specific vulnerabilities shall include, where appropriate, measures to prevent and control for attacks trying to manipulate the training dataset ('data poisoning'), inputs designed to cause the model to make a mistake ('adversarial examples'), or model flaws.

Chapter 5: Provider and User Obligations

Section 15: Obligations of Providers of High-Risk Systems

(1) Providers of high-risk AI systems shall —

> (a) ensure that their high-risk AI systems are compliant with the requirements set out in Chapter 4 of this Act;
>
> (b) have a quality management system in place which complies with Section 16;
>
> (c) draw-up the technical documentation of the high-risk AI system;
>
> (d) when under their control, keep the logs automatically generated by their high-risk AI systems;
>
> (e) ensure that the high-risk AI system undergoes the relevant conformity assessment procedure, prior to its placing on the market or putting into service;
>
> (f) comply with the necessary registration obligations; and
>
> (g) take the necessary corrective actions, if the high-risk AI system is not in conformity with the requirements set out in Chapter 4 of this Act;

Section 16: Quality management system

(1) Providers of high-risk AI systems shall put a quality management system in place that ensures compliance with this Regulation, in which that system shall be documented in a systematic and orderly manner in the form of written policies, procedures and instructions, and shall include at least the following aspects —

> (a) a strategy for regulatory compliance, including compliance with conformity assessment procedures and procedures for the management of modifications to the high-risk AI system;
>
> (b) techniques, procedures and systematic actions to be used for the design, design control and design verification of the high-risk AI system;
>
> (c) techniques, procedures and systematic actions to be used for the development, quality control and quality assurance of the high-risk AI system;
>
> (d) examination, test and validation procedures to be carried out before, during and after the development of the high-risk AI system, and the frequency with which they have to be carried out;
>

> (e) technical specifications, including standards, to be applied and, where the relevant harmonised standards are not applied in full, the means to be used to ensure that the high-risk AI system complies with the requirements set out in Chapter 4;
>
> (f) systems and procedures for data management, including data collection, data analysis, data labelling, data storage, data filtration, data mining, data aggregation, data retention and any other operation regarding the data that is performed before and for the purposes of the placing on the market or putting into service of high-risk AI systems;
>
(g) the risk management system referred to in Section 8 of this Act;
>
> (h) the setting-up, implementation and maintenance of a post-market monitoring system, in accordance with Section 23;
>
> (i) procedures related to the reporting of serious incidents and of malfunctioning in accordance with Section 24;
>
> (j) the handling of communication with national competent authorities, competent authorities, including sectoral ones, providing or supporting the access to data, notified bodies, other operators, customers or other interested parties;
>
> (k) systems and procedures for record keeping of all relevant documentation and information;
>
> (l) resource management, including security of supply related measures;
>
> (m) an accountability framework setting out the responsibilities of the management and other staff with regard to all aspects listed in this paragraph.

(2) The implementation of aspects referred to in Subsection (1) shall be proportionate to the size of the provider’s organisation.

Section 17: Obligation to draw up technical documentation

(1) Providers of high-risk AI systems shall draw up the technical documentation referred to in Section 10 in accordance with Schedule 2.

Section 18: Automatically generated logs

(1) Providers of high-risk AI systems shall keep the logs automatically generated by their high-risk AI systems, to the extent such logs are under their control by virtue of a contractual arrangement with the user or otherwise by law. In which the logs shall be kept for a period that is appropriate in the light of the intended purpose of the high-risk AI system and applicable legal obligations under the relevant law.

Section 19: Corrective actions

(1) Providers of high-risk AI systems which consider or have reason to consider that a high-risk AI system which they have placed on the market or put into service is not in conformity with this Act shall immediately take the necessary corrective actions to bring that system into conformity, to withdraw it or to recall it, as appropriate where they shall inform the distributors of the high-risk AI system in question and, where applicable, the authorised representative and importers accordingly.

Section 20: Obligations of importers

(1) Before placing a high-risk AI system on the market, importers of such system shall ensure that —

> (a) the appropriate conformity assessment procedure has been carried out by the provider of that AI system
>
> (b) the provider has drawn up the technical documentation in accordance with Schedule 2;
>
> (c) the system bears required conformity marking and is accompanied by the required documentation and instructions of use.

(2) Where an importer considers or has reason to consider that a high-risk AI system is not in conformity with this Act, it shall not place that system on the market until that AI system has been brought into conformity, and where the high-risk AI system presents a risk, the importer shall inform the provider of the AI system and the competent market surveillance authorities to that effect.

(3) Importers shall indicate their name, registered trade name or registered trademark, and the address at which they can be contacted on the high-risk AI system or, where that is not possible, on its packaging or its accompanying documentation, as applicable.

(4) Importers shall ensure that, while a high-risk AI system is under their responsibility, where applicable, storage or transport conditions do not jeopardise its compliance with the requirements set out in Chapter 4 of this Act.

(5) Importers shall provide the competent authorities, upon a reasoned request, with all necessary information and documentation to demonstrate the conformity of a high-risk AI system with the requirements set out in Chapter 4 of this Act in the English Language, including access to the logs automatically generated by the high-risk AI system to the extent such logs are under the control of the provider by virtue of a contractual arrangement with the user or otherwise by law.

Section 21: Obligations of distributors

(1) Before making a high-risk AI system available on the market, distributors shall verify that the high-risk AI system bears the required conformity marking, that it is accompanied by the required documentation and instruction of use, and that the provider and the importer of the system, as applicable, have complied with the obligations set out in this Act.

(2) Where a distributor considers or has reason to consider that a high-risk AI system is not in conformity with the requirements set out in Chapter 4 of this Act, it shall not make the high-risk AI system available on the market until that system has been brought into conformity with those requirements. Furthermore, where the system presents a risk, the distributor shall inform the provider or the importer of the system, as applicable, to that effect.

(3) Distributors shall ensure that, while a high-risk AI system is under their responsibility, where applicable, storage or transport conditions do not jeopardise the compliance of the system with the requirements set out in Chapter 4.

(4) A distributor that considers or has reason to consider that a high-risk AI system which it has made available on the market is not in conformity with the requirements set out in Chapter 4 shall take the corrective actions necessary to bring that system into conformity with those requirements, to withdraw it or recall it or shall ensure that the provider, the importer or any relevant operator, as appropriate, takes those corrective actions.

(5) Upon a reasoned request from a national competent authority, distributors of high-risk AI systems shall provide that authority with all the information and documentation necessary to demonstrate the conformity of a high-risk system with the requirements set out in Chapter 4.

Section 22: Obligations of users of High-Risk AI Systems

(1) Users of high-risk AI systems shall use such systems in accordance with the instructions of use accompanying the systems, pursuant to Subsections (2) and (5).

(2) The obligations in Subsection (1) are without prejudice to other user obligations under the competent law and to the user’s discretion in organising its own resources and activities for the purpose of implementing the human oversight measures indicated by the provider.

(3) Without prejudice to Subsection (1), to the extent the user exercises control over the input data, that user shall ensure that input data is relevant in view of the intended purpose of the high-risk AI system.

(4) Users shall monitor the operation of the high-risk AI system on the basis of the instructions of use. When they have reasons to consider that the use in accordance with the instructions of use may result in the AI system presenting a risk they shall —

> (a) inform the provider or distributor and suspend the use of the system;
>
> (b) inform the provider or distributor when they have identified any serious incident or any malfunctioning within the meaning of Section 24 and interrupt the use of the AI system.

(5) Users of high-risk AI systems shall keep the logs automatically generated by that high-risk AI system, to the extent such logs are under their control, in which the logs shall be kept for a period that is appropriate in the light of the intended purpose of the high-risk AI system and applicable legal obligations in national law.

Chapter 6: Post-Market Monitoring and Information Sharing

Section 23: Post-Market Monitoring

(1) Providers shall establish and document a post-market monitoring system in a manner that is proportional to the nature of the artificial intelligence technologies and the risks of the high-risk AI system.

(2) The post-market monitoring system shall —

> (a) actively and systematically collect, document and analyse relevant data provided by users or collected through other sources on the performance of high-risk AI systems throughout their lifetime, and
>
> (b) allow the provider to evaluate the continuous compliance of AI systems with the requirements set out in Chapter 4.

(3) The post-market monitoring system shall be based on a post-market monitoring plan.

(4) The Secretary of State may set regulations, via secondary legislation, laying down detailed provisions establishing a template for the post-market monitoring plan and the list of elements to be included in the plan.

(5) Regulations set under this Section shall be subject to negative procedure.

Section 24: Incident and Malfunction Information Sharing

(1) Providers of high-risk AI systems placed on the market shall be required to report any serious incident or any malfunctioning of those systems which constitutes a breach of obligations under the relevant law intended to protect fundamental rights to the competent market surveillance authorities where that incident or breach occurred.

(2) Pursuant to subsection (1) such notification shall be made immediately after the provider has established a causal link between the AI system and the incident or malfunctioning or the reasonable likelihood of such a link, and, in any event, not later than 15 days after the providers becomes aware of the serious incident or of the malfunctioning.

(3) Upon receiving a notification related to a breach of obligations under law intended to protect fundamental rights, the competent market surveillance authority shall inform the competent national public authorities or bodies.

Chapter 7: Governance and Administration

Section 25: National Authorities

(1) The Secretary of State shall establish or designate a national competent authority for the purpose of ensuring the application and implementation of this Act, in which national authorities are to be organised in a way that safeguards the objectivity and impartiality of their activities and tasks.

(2) The national competent authority shall have the following duties include but not be limited to —

> (a) a notifying authority and market surveillance authority under the provisions of this Act;
>
> (b) the provision of guidance and advice on the implementation of this Act, including towards small-scale providers; and
>
>(c) monitoring and reporting duties of its activities and market developments under the provisions of this Act.

Section 26: Funding

(1) The Secretary of State shall appropriate the necessary funding for the establishment and running of the national competent authority and its operations under the relevant Department.

Section 27: Confidentiality

(1) National competent authorities and notified bodies involved in the application of this Act shall respect the confidentiality of information and data obtained in carrying out their tasks and activities in such a manner as to protect, in particular —

> (a) intellectual property rights, and confidential business information or trade secrets of a natural or legal person, including source code against their unlawful acquisition, use and disclosure apply.
>
> (b) the effective implementation of this Act, in particular for the purpose of inspections, investigations or audits; public and national security interests; and
>
> (c) integrity of criminal or administrative proceedings.

Chapter 8: Enforcement

Section 28: Liability

(1) Violation of the provisions of this Act in use and distribution of high-risk systems and compliance may result in penalties, among other criminal charges under applicable law, specified in Section 28(2) as determined by the regulatory authority or the Secretary of State.

(2) Regulations set the Secretary of State, via secondary legislation, may make provisions for —

> (a) a regulatory body to issue the following —
>
>> (i) a compliance notice, and
>>
>> (ii) a stop notice, or
>
> (b) where the Secretary of State or an regulatory body are to issue a monetary penalty notice.

(3) Regulations may provide for a requirement imposed by a stop notice to be enforceable, on the application of the Secretary of State, by injunction.

(4) Regulations under this Section must secure necessary review and appealment procedures are included.

(5) Regulations under this Section are subject to affirmative procedure.

Section 29: Compliance Notices

(1) Regulations which provide for the issue of a compliance notice must secure that —

> (a) a compliance notice may only be issued where the issuing inspector of the notice is satisfied that person to whom it is issued has committed or is committing a relevant breach,
>
> (b) the steps specified in relation to the notice are steps that the inspector considers will ensure that the relevant breach does not continue or reoccur, and
>
> (c) the period specified in relation to the notice is not less than 14 days beginning on the day on which the notice is received.

Section 30: Stop Notices

(1) Regulations which provide for the issue of a stop notice must secure that —

> (a) a stop notice may be issued to a person only where the inspector issuing the notice reasonably believes that the person to whom it is issued has committed or is likely to commit a relevant breach, and
>
> (b) the steps specified in relation to stop notices are steps that the inspector issuing the notice considers will ensure that the specified activity will be carried on in a way that does not involve the person committing a relevant breach.

Section 31: Monetary Penalty Notices

(1) Regulations which provide for the issue of a monetary penalty notice must ensure that the Secretary of State or an inspector may issue a monetary penalty notice only where satisfied that the person to whom it is issued had committed a relevant breach.

(2) Regulations which provide for the issue of a monetary penalty notice must require the notice to state —

> (a) how the payment may be made,
>
> (b) the period within which payment must be made, and
>
> (c) the consequences of late payment or failure to pay.

(3) Regulations which provide for the issue of a monetary penalty notice may make provision —

> (a) for the payment of interest on late payment,
>
> (b) as to how any amounts payable by virtue of the regulations are to be recoverable.

Section 32: Extent, Commencement and Short Title

(1) This Act extends to the United Kingdom.

(2) The provisions of this Act shall come into force three months after this Act is passed and has received Royal Assent.

(3) This Act may be cited as the Artificial Intelligence (High-Risk Systems) Act.

SCHEDULE 1: High-Risk Systems

High-risk AI systems are the AI systems listed in any of the following areas:

(1) Biometric identification and categorisation of natural persons —

> (a) AI systems intended to be used for the ‘real-time’ and ‘post’ remote biometric identification of natural persons;

(2) Management and operation of critical infrastructure —

> (a) AI systems intended to be used as safety components in the management and operation of road traffic and the supply of water, gas, heating and electricity.

(3) Education and vocational training —

> (a) AI systems intended to be used for the purpose of determining access or assigning natural persons to educational and vocational training institutions;
>
> (b) AI systems intended to be used for the purpose of assessing students in educational and vocational training institutions and for assessing participants in tests commonly required for admission to educational institutions.

(4) Employment, workers management and access to self-employment —

> (a) AI systems intended to be used for recruitment or selection of natural persons, notably for advertising vacancies, screening or filtering applications, evaluating candidates in the course of interviews or tests;
>
> (b) AI intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation and for monitoring and evaluating performance and behavior of persons in such relationships.

(5) Access to and enjoyment of essential private services and public services and benefits —

> (a) AI systems intended to be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for public assistance benefits and services, as well as to grant, reduce, revoke, or reclaim such benefits and services;
>
> (b) AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems put into service by small scale providers for their own use;
>
> (c) AI systems intended to be used to dispatch, or to establish priority in the dispatching of emergency first response services, including by firefighters and medical aid.

(6) Law enforcement —

> (a) AI systems intended to be used by law enforcement authorities for making individual risk assessments of natural persons in order to assess the risk of a natural person for offending or reoffending or the risk for potential victims of criminal offences;
>
> (b) AI systems intended to be used by law enforcement authorities as polygraphs and similar tools or to detect the emotional state of a natural person;

>
> (c) AI systems intended to be used by law enforcement authorities to detect deep fakes;
>
> (d) AI systems intended to be used by law enforcement authorities for evaluation of the reliability of evidence in the course of investigation or prosecution of criminal offences;
>
> (e) AI systems intended to be used by law enforcement authorities for predicting the occurrence or reoccurrence of an actual or potential criminal offence based on profiling of natural persons or assessing personality traits and characteristics or past criminal behaviour of natural persons or groups;
>
> (f) AI systems intended to be used by law enforcement authorities for profiling of natural persons in the course of detection, investigation or prosecution of criminal offences;
>
> (g) AI systems intended to be used for crime analytics regarding natural persons, allowing law enforcement authorities to search complex related and unrelated large data sets available in different data sources or in different data formats in order to identify unknown patterns or discover hidden relationships in the data.

(7) Migration, asylum and border control management —

> (a) AI systems intended to be used by competent public authorities as polygraphs and similar tools or to detect the emotional state of a natural person;
>
> (b) AI systems intended to be used by competent public authorities to assess a risk, including a security risk, a risk of irregular immigration, or a health risk, posed by a natural person who intends to enter or has entered into UK territory;
>
> (c) AI systems intended to be used by competent public authorities for the verification of the authenticity of travel documents and supporting documentation of natural persons and detect non-authentic documents by checking their security features;
>
> (d) AI systems intended to assist competent public authorities for the examination of applications for asylum, visa and residence permits and associated complaints with regard to the eligibility of the natural persons applying for a status.

(8) Administration of justice and democratic processes —

> (a) AI systems intended to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts.

SCHEDULE 2: Technical Documentation

The technical documentation shall contain at least the following information, as applicable to the relevant AI system —

(1) A general description of the AI system including —

> (a) its intended purpose, the person/s developing the system the date and the version of the system;
>
> (b) how the AI system interacts or can be used to interact with hardware or software that is not part of the AI system itself, where applicable;
>
> (c) the versions of relevant software or firmware and any requirement related to version update;
>
> (d) the description of all forms in which the AI system is placed on the market or put into service;
>
> (e) the description of hardware on which the AI system is intended to run;
>
> (f) where the AI system is a component of products, photographs or illustrations showing external features, marking and internal layout of those products;
>
> (g) instructions of use for the user and, where applicable installation instructions;

(2) A detailed description of the elements of the AI system and of the process for its development, including —
>
> (a) the methods and steps performed for the development of the AI system, including, where relevant, recourse to pre-trained systems or tools provided by third parties and how these have been used, integrated or modified by the provider;
>
> (b) the design specifications of the system, namely the general logic of the AI system and of the algorithms; the key design choices including the rationale and assumptions made, also with regard to persons or groups of persons on which the system is intended to be used; the main classification choices; what the system is designed to optimise for and the relevance of the different parameters; the decisions about any possible trade-off made regarding the technical solutions adopted to comply with the requirements set out in Chapter 4;
>
> (c) the description of the system architecture explaining how software components build on or feed into each other and integrate into the overall processing; the computational resources used to develop, train, test and validate the AI system;
>
> (d) where relevant, the data requirements in terms of datasheets describing the training methodologies and techniques and the training data sets used, including information about the provenance of those data sets, their scope and main characteristics; how the data was obtained and selected; labelling procedures (e.g. for supervised learning), data cleaning methodologies (e.g. outliers detection);
>
> (e) assessment of the human oversight measures needed in accordance with Section 13, including an assessment of the technical measures needed to facilitate the interpretation of the outputs of AI systems by the users, in accordance with Section 12;

>
> (f) where applicable, a detailed description of pre-determined changes to the AI system and its performance, together with all the relevant information related to the technical solutions adopted to ensure continuous compliance of the AI system with the relevant requirements set out in Chapter 4;
>
> (g) the validation and testing procedures used, including information about the validation and testing data used and their main characteristics; metrics used to measure accuracy, robustness, cybersecurity and compliance with other relevant requirements set out in Chapter 4 as well as potentially discriminatory impacts; test logs and all test reports dated and signed by the responsible persons, including with regard to pre-determined changes as referred to under point (f).

(3) Detailed information about the monitoring, functioning and control of the AI system, in particular with regard to: its capabilities and limitations in performance, including the degrees of accuracy for specific persons or groups of persons on which the system is intended to be used and the overall expected level of accuracy in relation to its intended purpose; the foreseeable unintended outcomes and sources of risks to health and safety, fundamental rights and discrimination in view of the intended purpose of the AI system; the human oversight measures needed in accordance with Section 13, including the technical measures put in place to facilitate the interpretation of the outputs of AI systems by the users; specifications on input data, as appropriate;

(4) A detailed description of the risk management system in accordance with Section 8;

(5) A description of any change made to the system through its lifecycle;

(6) A detailed description of the system in place to evaluate the AI system performance in the post-market phase in accordance with Section 23, including the post-market monitoring plan referred to in Section 23.

SCHEDULE 3: Artificial Intelligence Approaches and Techniques

(1) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;

(2) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

~~(3) Statistical approaches, Bayesian estimation, search and optimisation methods.~~

(Meta: Relevant and Inspired Documents)

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

This Bill was submitted by The Honourable u/Waffel-lol **LT CMG, Spokesperson for Business, Innovation and Trade, and Energy and Net-Zero, on behalf of the Liberal Democrats**

Opening Speech:

Deputy Speaker,

As we stand on the cusp of a new era defined by technological advancements, it is our responsibility to shape these changes for the benefit of all. The Liberal Democrats stand firmly for a free and fair society and economy, however the great dangers high-risk AI systems bring, very much threaten the integrity of an economy and society that is free and fair. This is not a bill regulating all AI use, no, this targets the malpractice and destruction systems and their practices that can be used in criminal activity and exploitation of society. A fine line must be tiptoed, and we believe the provisions put forward allow for AI development to be done so in a way that upholds the same standards we expect for a free society. This Bill reflects a key element of guarding the freedoms of citizens, consumers and producers from having their fundamental liberties and rights encroached and violated by harmful high-risk AI systems that currently go unregulated and unchecked.

Artificial Intelligence, with its vast potential, has become an integral part of our lives. From shaping our online experiences to influencing financial markets, AI's impact is undeniable. Yet, equally so has its negative consequences. As it stands, the digital age is broadly unregulated and an almost wild west, to put it. Which leaves sensitive systems, privacy and security matters at risk. In addressing this, transparency is the bedrock of a fair and just society. When these high-risk AI systems operate in obscurity, hidden behind complex algorithms and proprietary technologies, it becomes challenging to hold them accountable. We need regulations that demand transparency – regulations that ensure citizens, businesses, and regulators alike can understand how these systems make decisions that impact our lives.

Moreover, market fairness is not just an ideal; it is the cornerstone of a healthy, competitive economy. Unchecked use of AI can lead to unfair advantages, market distortions, and even systemic risks. The regulations we propose for greater safety, transparency and monitoring can level the playing field, fostering an environment where innovation thrives, small businesses can compete, and consumers can trust that markets operate with integrity. We're not talking about stifling innovation; we're talking about responsible innovation. These market monitors and transparency measures will set standards that encourage the development of AI systems that are not only powerful but also ethical, unbiased, and aligned with our societal values. So it is not just a bill that bashes on these high-risk systems, but allows for further monitoring alongside their development under secure and trusted measures.