## Table of contents

| Text of model law | Remarks |
|---|---|
| *1. Scope of this law and related definitions*<br>This law applies to the development, the operation and the use of software that constitutes artificial intelligence or of items that use artificial intelligence (both called hereafter: AI systems), where the development, the operation and the use:<br>● take place on the territory of … (the jurisdiction in question), or<br>● have effects on that territory[, regardless of where they have been developed and from where they are operated]. | Better to include items using software so as to avoid that physical items using software escape the law.<br><br>Referring to the effects on the territory of the jurisdiction in question prevents the possibility of operating AI systems in other jurisdictions to circumvent legal requirements. It also creates a level playing field. The addition "regardless where …" is not necessary, but further clarifies this idea.<br><br>Keeping the first dot point ensures that the law can be used to address AI systems intended for other jurisdictions which is important for the mutual support of jurisdictions: many jurisdictions request reciprocity when being asked to enforce in the interest of another jurisdiction. |
| It applies / does not apply to administrations and public institutions controlled by the state [with the exception of the military]. | Only needed if the respective Section 35 below is kept. |
| It also establishes some obligations for traders of AI systems. | We recommend covering all AI systems and not to exempt any of them in order to ensure that it is verified whether there are risks linked to the AI system. E.g.:<br>● Many presumably harmless games are built in such a way that may cause dependencies, psychological damage and possibly an inability to live in the real world, particularly for minors (young adults up to age … to be specified in the respective law). Hence, we recommend a low risk class with a few (see below) rather than exempting entirely a certain category of AI systems. However, legislators might prefer to explicitly exempt some AI systems, e.g. to limit the administrative burden or to better focus scarce enforcement resources.<br>● Even AI systems developed for research purposes should be subject to some rules, e.g. the prohibition of dissemination to prevent abuse by terrorists. |
| The following AI systems are exempted:<br>● … ;<br>● … . | |

| | |
|---|---|
| *2. Definitions*<br>Artificial intelligence:<br>● Var. 1: software or software and hardware combined that learns, self-amends or solves problems without human interference.<br>● Var. 2: software or software and hardware combined that, without human interference after its initial training phase:<br>○ learns or self-amends,<br>○ solves problems in a way that is not predetermined by humans, or<br>○ takes or recommends decisions based on more than two independent parameters<br>● Var. 3: software or software and hardware combined that consumes data and does one or more of the following:<br>○ takes actions based on that data,<br>○ predicts values or labels from that data,<br>○ transforms data in a non-trivial way, thus by adding substantial information to the data set other than by simple aggregations,<br>○ where the mechanism of action, prediction or transformation is either:<br>■ learned from a body of data ("machine learning"), or<br>■ set directly ("expert system"). | The first definition is very narrow and close to the scientific understanding of AI. But it does not include many cases which might be regarded as ethically problematic, meaning automated decision-making systems (ADS). Therefore, we recommend the second alternative.<br>The same result could be reached by using a narrow definition like the one in Variant 1 and then include this in the scope of the law regulating automated decision-making systems (ADS).<br>Not to include ADS in one or the other way creates borderline issues and leaves uncontrolled too many important IT systems that merit control.<br><br>See as an alternative the technology-driven "listing" approach suggested by the European Commission in Annex I to its proposal for an AI Regulation. It is likely that this "listing" approach will need frequent and cumbersome updating, and still create loopholes. The risk-based approach recommended here is faster, more flexible and more complete, but it is slightly less precise than the proposal of the European Commission. |
| Human-independent elements of definitions above do not exclude further human contribution to and supervision of an operating AI system (e.g. updating, corrections, maintenance etc.). | The best flexible / future-proof solution, which is still very precise, is to complement our approach with a so-called positive list (of software and items covered under an abstract definition) so that the abstract definition and list go hand in hand. The positive list approach ensures almost the same degree of legal certainty as the pure "listing" approach.<br>The proposal of the European Commission has listed, in its Annex I, a few technological features or characteristics justifying the qualification of certain software as AI. However, we recommend assessing the situation precisely at the time of law-making as new features or characteristics might have emerged by then. |
| The following always qualifies as artificial intelligence:<br>● …<br>● … | |

| | |
|---|---|
| AI systems:<br>Software that constitutes artificial intelligence or items that use artificial intelligence.<br><br>Developers:<br>Natural or legal persons creating the AI system, namely by conceiving the software.<br><br>Operators:<br>Natural or legal persons running an AI system (Var. 1), regardless of whether providing access to it to users or not (Var. 2) where they provide access to it or where they use the AI system in relation to clients.<br><br>Users:<br>Natural or legal persons using the access to an AI system offered for use by an operator.<br><br>Clients:<br>Natural or legal persons (directly / not just remotely) impacted by the AI system, regardless of their contractual relationship with the operator or user or not.<br><br>Users with clients:<br>Users who have clients in the meaning of the above.<br><br>Traders:<br>Natural or legal persons other than developers, operators and users with clients who broker, publicise, distribute, import, export or otherwise support the dissemination and the use of AI systems, regardless whether in return for payment or for free.<br><br>Discrimination:<br>Unequal treatment based on sexual orientation, biological, ethnic, language, or religious criteria.<br><br>Social scoring:<br>Evaluation of the credit-worthiness or of the generic legal, social or contractual behaviour or trustworthiness of natural [and legal] persons [trespassing a concrete context such as behaviour on a trading platform] [where this leads to negative consequences]. | Below, we suggest a fine-tuned system distinguishing between the different degrees of responsibility of different natural or legal persons. To that end, we need these definitions.<br><br>If you wish to exempt operators who use AI systems only for themselves without any immediate external effect, choose Variant 2.<br><br>In particular, we recommend distinguishing between users who use an AI system exclusively for themselves and those whose AI system affects other persons (“clients”). The latter are not “operators” in so far as they do not provide access to the AI system, but still should be incumbent to more obligations in so far as the “clients” are impacted by the AI system, some legislators will certainly find. Therefore, we offer this opportunity for differentiation.<br><br>Further criteria might need to be added, in particular where such criteria (e.g. place of residence) mirrors one of the listed criteria. Age and disabilities might be further candidate criteria.<br><br>Check whether legal persons shall be included.<br>An evaluation system limited to one specific field, like the behaviour on a trading platform, might be justified.<br>The limitation to cases with negative |

| | |
|---|---|
| | consequences permits still to provide privileges linked to positive social scoring, which is not much less debatable. |
| Large natural habitats:<br>Alt. 1: Natural habitats of a size of … (e.g. the jurisdiction in question, ¼ of the jurisdiction in question or performing an important planetary, ecological function e.g. Amazon rainforest).<br>Alt. 2: Large natural habitats [other group(s) of natural habitats] according to national legislation. | Any criterion (criteria) for defining large natural habitats for the benefits of this law can be chosen by the jurisdiction in question, including using existing legal classifications of natural habitats. |
| Property:<br>All goods and rights with monetary value. | |
| Deep fake:<br>Assembly of elements aimed at presenting an item or information as authentic even though it is not. | |
| Conformity assessment body:<br>Legal person independent from control by another legal person or state [with place of business in … (the jurisdiction in question)] that has been accredited by … (ministry or national authority) for the verification of conditions and obligations set out in this law [or that has obtained a similar status in the following jurisdictions … and has committed in writing to fulfil the information obligations set out in this law].<br>… | This is only needed where the involvement of third party bodies for verification of compliance is deemed necessary, see below.<br><br>Further potentially useful definitions can be found in the Canadian Directive on Automated Decision-Making. |
| *3. Risk Classes*<br>The development, operation and use of AI systems is grouped into risk classes. | Avoiding Risk Classes would lead to not so desirable effects like wasting scarce authority's resources for unproblematic cases and unnecessarily burdening operators or, if the opposite solution is chosen, missing problematic cases. |
| Risk Class 3:<br>An AI system belongs to Risk class 3 where its use potentially has an impact on the following values:<br>● Mankind;<br>● The earth as habitat ("geo-engineering");<br>● Large natural habitats;<br>● Biological species (to an extent of extinction or severe risk of extinction or severe | The Risk Classes suggested here are generic, not technology-linked and thus open to future developments. They do not rely on permanent updating. For instance: Risk Class 3 is consciously much broader than European Commission approach, referring to certain software technologies and covering "safety relevant software components for products subject to a third |

| | |
|---|---|
| population decline);<br>● Lives;<br>● Health; or<br>● Functioning of the society and governance at international or national level, including for elections. | party conformity assessment procedure", as very few jurisdictions' legislation cover all technologies requiring a "safety component", to complete a third party conformity assessment procedure – technical progress is faster than legislative progress. E.g. the following items are rarely or poorly regulated across the globe, despite their high risk potential: satellites, geo-engineering tools and software, navigation tools and software, (water-) drones, health-relevant software. |
| Risk Class 2:<br>An AI system not belonging to Risk Class 3 belongs to Risk Class 2 where its use potentially has an impact on the following values:<br>● Animals;<br>● Natural habitats other than those covered under Risk class 3; | Jurisdictions may determine for themselves what size and type of natural habitat facing endangerment should trigger the application of requirements for the highest Risk Class. |
| ● Personal liberties; or<br>● Property, | The "personal liberties" merit a definition fitting to the jurisdiction in question, e.g. by reference to a section in the constitution. |
| or where it is likely to impact humans with regard to the following:<br>● Management of public infrastructure with effect on any of the values listed for Risk Classes 3 and 2, including namely: satellites, air, train, ship and road traffic; storage facilities for essential goods; hospitals, ambulances, fire-brigades, and civil protection; the supply of water, gas, heating, electricity and of the internet;<br>● Functioning of the society or governance at sub-national level; | Here again we recommend the use of a positive list. |
| ● State [or private] sanction systems;<br>● Forensic evaluations;<br>● Prediction of criminal and other offences by natural [and legal] persons and profiling of these persons; | Private sanction systems might be deemed less problematic.<br>The rights of legal persons might be deemed less in need of protection. |
| ● Management of the penal OR judicial system;<br>● Assignment to institutions limiting the freedom of persons;<br>● Access to the national territory and residence right;<br>● Access to identification or travel documents or | The penal system is the most intrusive part of the judicial system and some judicial systems might already deem the use of AI problematic enough to justify Risk Class 2. |

| | |
|---|---|
| other means of identification and authenticity verification of such documents or means;<br>● Access to employment, self-employment and registration of businesses;<br>● Access to education and vocational training and respective institutions;<br>● Access to social benefits and private payments ensuring social protection;<br>● Access to essential services;<br>● Health-relevant life-style and consumption;<br>● Biometric identification and categorisation of natural persons;<br>● Evaluation relating to any of the above;<br>● Evaluation of the credit-worthiness or of the generic legal, social or contractual behaviour or trustworthiness of natural [and legal] persons; and<br>● Detection of emotional states.<br><br>Risk Class 1:<br>All other AI systems.<br><br>The term “access to” in the meaning of this section includes the maintenance of a given legal or factual status. | . |
| *4. Ethical rules*<br>AI systems shall be developed, operated and used in such a way that the following ethical principles and rules are respected to the extent possible:<br><br>● Where several lives stand against each other, the solution saving the maximum number of lives shall be sought for;<br>● The lives of all persons have the same value, in particular regardless of origin and wealth or any of the criteria listed in the definition of “discrimination”;<br>● The different life expectancy of persons may / may not be taken into account / may only be taken into account where one life stands against another life; | The ethical principles and rules will certainly not fit for all jurisdictions alike and they are not intended to do so. They are solely meant as a basis for discussion of legislators worldwide, like the rest of the model law. Hence, the list should be rather comprehensive and readers should not be surprised by one or the other principle or rule not fitting their particular tradition.<br><br>See also the last bullet which might be seen as encompassing this one.<br><br>This is a particularly problematic balancing question. If the life expectancy is fully taken into account, the life of a child could be worth more than the life of 5 persons 80 or more years old, which might be deemed very strange. On the other hand it might |

| | |
|---|---|
| | also seem strange not to favour a child against an elderly person where just one life stands against another.<br><br>Alternatively, one might attribute a lower value. |
| ● As from … months after conception, unborn lives have the same value as born lives;<br>● Where either lives or health can be protected, preference is to be given to protect lives;<br>● Where either casualties or injuries are caused, preference is to be given to the avoidance of casualties[, unless the number of severe injuries is expected to be at least 10 times higher]; | It might be useful to establish such an exception as severe injuries might also trigger casualties. |
| ● Where there is a choice between protecting lives or health and the protection of property and financial interests, preference is to be given to the former, lives or health;<br>● Where there is a choice between protecting animals and protecting property and financial interests, preference is to be given to the former; | |
| ● Where there is a choice between protecting nature and protecting property and financial interests, preference is to be given to the former / latter; | We find this principle particularly debatable. A trade-off curve would be most appropriate, but complicated to be laid down in law. |
| ● Where the interests of … (jurisdiction X) as a whole are in conflict with the interests of individuals, preference is to be given to the former;<br>● Where the interests of … (jurisdiction X) are in conflict with the interests in another jurisdiction, preference is to be given to the former[, unless the repercussions to be expected outweigh the advantage]; | |
| ● Short term and long term interests are to be valued in the same way (no discounting of long term interests); | Some people see future interests as equally valuable, some not. Hence, the legislator has to reflect on whether future interests should be slightly discounted. Still the main problem, in particular in democracies, is that politicians discard opposing future interests, e.g. to be re-elected or to avoid other immediate negative reactions like protests; future interests are mostly not appropriately represented and defended. |
| ● The likelihood of positive or negative effects is to be taken into account as discounting factor; | |

| | |
|---|---|
| ● Where the self-interests of the AI system stand against any of the interests listed so far, preference is to be given to the latter; OR<br>● The self-interests of the AI system may not influence decisions;<br>● No distinction shall be made on the basis of physical characteristics of natural persons unless this is justified for medical reasons;<br>● No discriminatory distinction between natural persons shall be made. However, biological criteria may be used in medical AI systems where the health benefit outweighs the damage caused by the discriminatory effect of the AI system. | Some thinkers consider that AI systems might, in the future, be construed as having a kind of self-interest.<br>This is the more radical solution.<br><br>More such rules can and should be developed for a given jurisdiction in order to optimise the adaptation to societal values. |
| *5. Prevention of accidents*<br>AI systems shall be developed and operated in such a way that all risks, including the following, are, where possible, eliminated and otherwise reduced and mitigated to the extent possible:<br>● Unintended and uncontrollable results;<br>● Erroneous input data;<br>● Erroneous processing;<br>● Erroneous results;<br>● Misleading presentation of correct results;<br>● Loss of connection with the human controller;<br>● Uncontrolled self-proliferation;<br>● Disabling other IT systems;<br><br>● Absorbing disproportionate internet data transmission capacities;<br>● Transgression of legal rules;<br><br>● Hacking;<br><br>● Being hacked;<br><br>● Tampering of data;<br>● Manipulation of data in view of inviting the AI system to draw erroneous conclusions (“data poisoning”). | Preference is to be given to the elimination, followed by the reduction and finally the mitigation of aforementioned risks. Risk mitigation comes into play where a certain risk cannot be further reduced, but where protective measures still can reduce the consequences. Some of the aforementioned risks may be inappropriate for mitigation e.g. transgression of legal rules.<br><br>Each of the risks could be further defined or exemplified.<br><br>E.g., some software might “find” the creative solution to copy itself into other hardware to increase computational capacities.<br><br>In particular in jurisdictions with limited internet transmission capacities, high volume internet consumption can cause important damage.<br><br>Here we refer to the AI system entering other IT systems.<br>Here we refer to the risk of the AI system itself being hacked.<br><br>The obligation to create “fall-back” solutions |

| | |
|---|---|
| The prevention of accidents shall include back-up and other subsidiary solutions. | should apply to all risks. |
| *6. Conflicts between the previous principles*<br>Where a conflict arises amongst the ethical principles, amongst the rules for the prevention of accidents and between the two, | In this section, we invite legislators to reflect further on how AI systems should be construed and operated. We offer different approaches to resolve questions of conflict. |
| Var. 1<br>… risk management principles laid down in … (ISO, regional or national standard or another national or third party document) shall be applied. | This approach sounds plausible and rational at first sight, but when diving into the content of such standards, one finds little guidance of the type needed for our purposes. |
| Var. 2<br>… preference is to be given to the ethical principles set out in Section 10 on "Risk management" in the order of their enumeration. | This approach gives full control to the legislator, who can determine the rule by consciously establishing an order of priority of rules. But it is a simplistic solution. |
| Var. 3<br>… the following principles shall apply:<br>● Where or more differing interests in question can be fairly well protected, either by limiting the harm, limiting the probability of harm or a combination thereof, this fairly good protection shall be sought for. If thereafter there is still margin of discretion, the interest(s) with higher value shall be protected as a priority. Where two or more interests have the same value, the degree of protection of the interests shall be optimised so that, if the values were on the same scale, the overall degree of protection would be highest.<br>● Where the interests in question cannot be fairly well protected, the interest(s) with higher value shall be protected as a priority, unless the probability of harm is negligible. Where two or more interests have the same value, the degree of protection of the interests shall be optimised so that, if the values were on the same scale, the overall degree of protection would be highest. | This approach offers more differentiation than Variant 2 whilst still being relatively easy to be applied. |
| Var. 4<br>… the following shall apply: The value of two or more interests shall be multiplied with the probability of harm thereto. A solution shall be sought where the sum of products of two or more interests multiplied with their respective probability of harm is minimal. | This model is suitable in jurisdictions where quantification of values, interests and harms are relatively commonplace, like the Anglo-American ones. It provides ethical optimisation, but reduced control by the legislator. |

| | |
|---|---|
| The hierarchy of values presented in Section 10 on “Risk management” shall be used. | This sentence is not needed for Variant 2. |
| *7. Traceability and correctness of data*<br>Developers, operators and users shall ensure that:<br>● the origin of input data can be retrieved;<br>● all input data is checked for plausibility and against errors; and<br>● the persons responsible for the correctness of data can be identified. | This obligation must be incumbent on developers to the extent that they insert data as reference or for training. It must be incumbent on users as they, too, might insert data which influence the outcome.<br><br>For a deeper insight into this topic, see the “Artificial Intelligence Governance Framework Model” of Singapore. |
| *8. Final human control, surveillance and responsibility*<br>AI systems [of Risk Class 2 and 3] shall be designed, manufactured and operated in a way that ensures human control of ethical principles as well as parameters and mechanisms of decision-making, whilst the individual decisions do not need to be controlled by humans. All autonomous sub-systems shall be controlled by at least one human, in addition to the human responsible for the entire AI system. The accountability of humans shall cover all aspects of the AI system and shall at all times be clear and traceable.<br><br>However, final human control may be discarded where it is not possible to have human control in addition to the control by the AI system and where the AI system outperforms humans in terms of protection of the values referred to in Section 10 on Risk management. | To stay proportionate, we recommend limiting this requirement to Risk Classes 2 and 3.<br><br>Full control by humans would take away the advantage of AI. On the other hand, societies fear to be the subject of decisions by AI systems. The provisions suggested here try to strike a balance.<br><br>An alternative, more detailed approach for human control / oversight can be found in Article 14 of the European Commission proposal for an AI Regulation.<br>Human control is not a goal as such, but a means. Where the underlying goals (e.g. safety or security) are better pursued without human control, there should not be an obligation for keeping human control. |
| *9. Societal, environmental and global effects*<br>Developers and operators shall assess potential negative societal, environmental and global effects, if any, of their AI system [of Risk Class 2 and 3] and the decisions taken by it. They shall describe each of the effects with a most likely scenario and two extreme scenarios. They shall, within the framework of their risk management, minimise negative effects, unless these are unavoidable and outweighed by positive effects in view of the values listed in Section 10 on Risk management. | We recommend limiting this obligation to Risk Classes 2 and 3. |

| | |
|---|---|
| *10. Risk management*<br>Developers and operators of AI systems [of Risk Class 2 and 3] shall undertake a comprehensive and continuous risk management that encompasses all potential risks for humans, animals, nature, society, and property, and at least the risks explicitly referred to in this law. The risk management is based on a comprehensive risk identification, quantification of the risks both in terms of harms and probability under all reasonably foreseeable scenarios. Developers and operators shall establish risk indicators and control mechanisms. Concrete corrective measures, which shall include shut down, shall be linked to precisely defined risk thresholds. | We recommend limiting this obligation to Risk Classes 2 and 3.<br><br>As stated above, we do not find that risk management standards necessarily provide for an appropriate methodology for the decisions to be taken in the context of AI systems. Therefore, we provide here for the big lines of what we deem to be an appropriate risk management approach for AI systems. |
| The risk management shall take account and seek to eliminate, and if this is not possible, to minimise, and if this is possible, to mitigate all known and foreseeable risks, including the risk of foreseeable misuse, meaning use for purposes that were not intended by the developer or operator. | |
| Developers or operators may refrain from risk elimination or (further) risk reduction where the risk reduction would disproportionately increase other risks or disproportionately reduce the benefit sought for by the AI system. | We recommend a double weighing of effects:<br>● One risk versus the other; and<br>● The risk versus the benefit.<br>Alternatively, one could refer to the risk-benefit balance which should not become worse. |
| The risk management shall take account of the following overall value hierarchy:<br>● Mankind;<br>● The earth as habitat;<br>● Large natural habitats;<br>● Lives;<br>● Health;<br>● Functioning of the society;<br>● Animals;<br>● Natural habitats;<br>● Personal liberties; and<br>● Property. | If the legislator does not set up a hierarchy of values, each developer or operator will make their own, with very much varying results. Hence, as difficult and arbitrary it might seem, it is still better that the legislator sets the hierarchy of values and thereby steers the core of risky AI systems. |
| The risk management shall also take account of the degree to which these values are at stake, which is equivalent to the probability of harm multiplied with the size of the harm. | Taking into account the degree improves the decisions, but renders decision-making more complicated. |
| The risk management shall give absolute | If we deem lives to be valuable, it is logical |

| | |
|---|---|
| preference to the avoidance of risks for mankind ("existential risks") and to the protection of the earth as habitat for humans and animals.<br><br>When undertaking calculations, including on how to sort out conflicts between the various principles and rules set out in this law, the following scores shall be used:<br>● Mankind: 100;<br>● The earth as habitat: 80;<br>● Large natural habitats: 15;<br>● Lives: 10;<br>● Health: 8;<br>● Functioning of the society: 7;<br>● Animals: 6;<br>● Natural habitats: 5;<br>● Personal liberties: 4; and<br>● Property: 2. | that we need to protect mankind and its habitat.<br><br>We recommended (see above) setting up a hierarchy of values. If the legislator wishes to further steer the developers and operators, the introduction of weighting factors might serve the purpose. |
| *11. Transparency*<br>AI systems shall be developed, operated and used in such a way that:<br>● Decision-making can be probed, understood and reviewed by authorities, supervisory bodies, common interest third parties, operators, users and their clients;<br>● Decisions are explainable [both in technical and non-technical terms], which implies in particular that the processes that extract model parameters from training data and generate labels from testing data can be described and motivated;<br>● Inputs and outputs can be verified;<br>● Records of design processes, decision-making and other events with external effects or system relevant events are established and kept;<br>● The persons steering the processes, decision-making or other operations can be identified, together with the decisions they have been taken during installation or operation of the AI system;<br>● Training, validation and testing datasets are accessible; and<br>● IT interfaces for full remote authority control (e.g. application programming interfaces) are available and can be operated with commonly available **OR** freely available software. | AI systems are often perceived as black holes by the outside. Transparency rules can remedy this to some extent. Evidently, transparency rules can come in conflict with intellectual property rights.<br><br>More detailed provisions on such records can be found in Article 12 of the proposal of the European Commission for an AI regulation.<br><br>The downside is that such interfaces create security risks. |
| *12. Human rights*<br>The development, operation and use of AI | |

| | |
|---|---|
| systems shall not constitute or contribute to a violation of human rights in the meaning of … (Regional legal instrument or UN Human Rights Charter). | |
| *13. Illicit purposes*<br>The development, operation and use of AI systems for the following purposes is banned:<br>● [Full] societal control;<br>● Social scoring of individuals [trespassing a concrete context such as behaviour on a trading platform];<br>● Political profiling and repression;<br>● Manipulation of democratic elections and political processes;<br>● Interrupting public services;<br>● Causing damage to third parties;<br>● Exploitation of psychological or physical weaknesses or vulnerabilities;<br>● Manipulation of opinions and preferences using erroneous information;<br>● Creating psychological dependencies;<br>● Steering and dissemination of internationally banned arms; and<br>● Generating “deep fake”. | The geographic scope of this law is such that the domestic development of illicit AI systems for export is covered too. |
| *14. Non-proliferation*<br>Developers, operators and traders of AI systems shall not make available AI systems for illicit purposes or to terrorists in the meaning of … (national or international definition), to criminals in the meaning of … (national definition) or to the following states … (list of rogue states). In case of doubt, they shall seek clearance from the supervising authority. | As AI systems can also be used to harm, non-proliferation is a useful means.<br>In certain jurisdictions, the inclusion of obligations for traders would need to be reflected in the scope, at the beginning of this law. |
| *15. Items with AI systems*<br>Items with integrated AI systems shall be labelled as such, shall have a kill switch and be shockproof. Requirements for AI systems apply to them. | We recommend establishing a few horizontal physical safety or security requirements for physical properties that might impact the AI system. |
| *16. AI systems for tuning*<br>AI systems intended to alter the behaviour of other software or items may only be made available where the software or items are precisely indicated with high visibility and where the compatibility and the overall fulfilment of the obligations of this law and of the law applicable to the item has been assessed. | Software for tuning vehicles or other mechanical items is common and might already use AI.<br>This provision aims at closing a security loophole by obliging to assess the tuning AI system together with the item to be tuned. |

| | |
|---|---|
| *17. Instructions for use*<br>Developers, operators and users with clients shall establish and give access to instructions for use. Instructions for use shall:<br>● Name its precise version and the characteristics distinguishing that version from previous versions;<br>● Name and contact details, including the physical addresses, of the developer, the operator and of the user with clients;<br>● Present the purpose, including in particular the targeted situations, facts or persons;<br>● List the purposes and forms of use for which the AI system is not made, whilst it could be deemed to be made for it;<br>● Present the features and characteristics of the AI system, including its level of accuracy and robustness and its means to prevent data tampering or manipulation;<br>● Inform on data and data training, testing or validation requirements to be fulfilled for accurate use;<br>● List all risks and the foreseeable situations or circumstances under which they might occur, including those of foreseeable misuse, together with the likelihood of their occurrence and the most appropriate downstream means to reduce or mitigate them;<br>● List in particular the means necessary for ensuring cybersecurity and the limitations of the AI systems in that regard;<br>● Instruct the user on how to operate or use the AI system in a way that is most respectful towards the ethical values set out or referred to in this law;<br>● Highlight warnings for the most problematic aspects of the previous two dot points;<br>● Summarise the legal obligations of this law applicable to the target audience on the instructions for use; and<br>● Contain a link to [a short version of] the technical documentation.<br>Instructions for use may be provided via a permalink in a commonly used electronic format. | Instructions for use can play an important role in terms of risk reduction and mitigation and can increase the overall compliance. |
| *18. Technical documentation*<br>Developers and operators shall establish, keep up to date and keep accessible for ten years after the last making available of the AI system a | A comprehensive technical documentation is key for permitting authorities to verify compliance. However, there is also a |

| | |
|---|---|
| comprehensive technical documentation which shall include at least the following:<br>● A description of the purpose and the features of the AI system;<br>● The precise version and its latest update and the characteristics distinguishing that version from previous ones;<br>● Links to previous versions of the technical documentation for that precise AI system and to the technical documentations of previous versions of the AI systems;<br>● Known but deliberately not eliminated technical issues;<br>● Requirements and necessary conditions for the AI systems use;<br>● The installation instructions;<br>● The instructions for use;<br>● The necessary and possible interactions of the AI system with other systems, software or hardware;<br>● An explanation of how the ethical principles and other rules set out in Sections 4 and 6 to 14 have been fine-tuned and applied in the concrete context;<br>● A detailed description of the Risk management, including an assessment and quantification of risks both in terms of gravity and probability of harm;<br>● A description of the machine learning approaches, including supervised, unsupervised and reinforcement learning and “deep learning”;<br>● A description of the logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;<br>● A description of the statistical approaches, Bayesian estimation, search and optimization methods;<br>● A description of the methods used, at various stages of development, for testing the robustness of the AI;<br>● A comprehensive description of societal, environmental and global risks, if any;<br>● A strategy on how to avoid the abuse for human rights violations (Section 12) or illicit purposes in the meaning of Section 13, if any such risk there is;<br>● A strategy on how to respect the | legitimate interest in protecting crucial information both in terms of intellectual property and security. Legislators hence have to strike a balance. Namely, for the items at the end of the list, a cautious approach is preferable, e.g. by limiting the depth of the necessary disclosure. |

| | |
|---|---|
| non-proliferation obligation set out in Section 14;<br>● A plan on how information on incidents and malfunction shall be collected, evaluated and fed into the improvements of the AI system.<br><br>The technical documentation of operators may build on the technical documentation of developers to the extent that the latter is still available. | Operators often do not have the same technical insight as developers and so they should be authorised to build on the technical documentation of the developers. |
| *19. Requirements for and training of staff*<br>Developers, operators and users with clients shall train their staff (both employees and freelancers) with regard to this law and supplementing decrees, ethics in general and their own ethical code in particular. They shall raise awareness of risks and impacts of the AI systems in question. They shall support their staffs' and freelancers' adherence to professional organisations aiming at the identification and tackling of issues of professional ethics and AI system ethics. | This is a side measure to increase compliance. |
| *20. Protection of whistle-blowers and staff rights*<br>Developers, operators and users with clients shall protect their staff (both employees and freelancers) against any discrimination or sanction where they have disclosed potentially unlawful or ethically problematic practices. They shall also protect their staff in case of disclosure to the outside where the internal disclosure was unsuccessful or would most likely have been unsuccessful, in particular where a general policy was adopted or conscious decisions taken that accept the negative effects or risks. The disclosure to the supervising authority is not subject to these conditions and may not be reprimanded in any form. | Another such side-measure. For more details on whistle-blowing, see this article. |
| *21. Developers', operators' and users' with clients internal surveillance*<br>Developers, operators and users with clients shall establish an internal control unit responsible for proactively tracking and investigating all malfunctions and incidents. They shall publish and label on their products or service-related media the contact data thereof. The responsible unit shall also permit anonymous bilateral communication for whistle-blowers who prefer to stay anonymous. It shall communicate to the | Where nobody looks for and analyses malfunctions and incidents, AI systems may stay unnecessarily problematic. This section provides a clear processing line for information on malfunctions and incidents. |

| | |
|---|---|
| supervising authority all incidents and malfunctions that merit a public response, that are systematic and thus not linked to a specific AI system or the root causes of which cannot be immediately deactivated.<br>Developers and operators shall accredit and, to the extent that this can be done without triggering safety or security risks, provide access to trustworthy persons who are able to detect safety or security risks, possibilities for data tampering or manipulation and ethical flaws of the AI system.<br>Developers, operators and users with clients shall pay out proportionate rewards to persons who detect deficiencies in accordance with the previous paragraph. | Some safety or security risks can be avoided by opening up a parallel, play-ground AI system. |
| *22. Registration of developers, operators and users with clients and of their AI systems*<br>Developers, operators and users with clients shall register themselves and their AI systems with the supervising authority in a procedure set out by the latter. | Registration is a burden, but facilitates the activities of the supervising authority. |
| *23. Additional obligations of developers*<br>Developers shall:<br>● Inform operators, also in their commercial contracts, of their respective obligations and the conditions set out in this law;<br>● Inform operators, also in their commercial contracts, of ethical problematic aspects mentioned in this law, namely by referring to their own ethics code and respective reports;<br>● Keep records of their commercial contacts with operators and inform authorities upon their request; and<br>● Inform the supervising authority of infringements they become aware of, regardless whether these are made by competitors, operators, users or conformity assessment bodies. | The obligations of developers and other actors are cast in such a way that a system of mutual control arises. Such a mutual control system leads to a higher degree of compliance, not only where supervising authorities are weak. |
| *24. Additional obligations of operators*<br>Operators shall:<br>● Inform users, also in their commercial contracts, of their respective obligations and the conditions set out in this law;<br>● Inform their clients, also in their commercial contracts, of ethical problematic aspects mentioned in this law, namely by referring to | Idem. |

| | |
|---|---|
| the developer's and their own ethics code and respective reports;<br>● Delete data on and provided by users and their clients on their request;<br>● Trace their commercial contacts with developers and users and inform authorities upon their request;<br>● Inform the responsible authority of infringements they become aware of, regardless whether these are made by competitors, developers, users or conformity assessment bodies;<br>● Inform the developer of any incidents or malfunctions or non-fulfilment of requirements;<br>● Repair in these cases or in cases of non-compliance the AI system as soon as possible; and<br>● Stop providing access to the AI system where important harm or continued infringement of this law cannot be prevented otherwise. | |
| *25. Additional obligations of users*<br>Users [with clients] shall:<br>● Inform their clients, also in their commercial contracts, of ethical problematic aspects mentioned in this law, namely by referring to the operator's and their own ethics code and respective reports;<br>● Delete data on and provided by clients on their request;<br>● Trace their commercial contacts with operators and inform authorities upon their request;<br>● Inform the responsible authority of infringements they become aware of, regardless whether these are made by developers, operators, other users or conformity assessment bodies;<br>● Inform the operator and the developer of any incidents, malfunctions or non-fulfilment of requirements and request them to repair the AI system as soon as possible; and<br>● Stop using the AI system where important harm or continued infringement of this law cannot be prevented otherwise. | Idem.<br>Legislators might wish to reflect on whether to cover all users or only users with clients. Users who do not have clients mostly have a limited effect on the outside world. This would justify exempting them. However, some of the listed obligations also make sense for users without clients. |
| *26. Additional obligations for Risk Class 2 systems*<br>Developers and operators of Risk Class 2 AI systems and users with clients of such systems | Here again, we offer legislators possibilities to increase the likelihood of compliance by other means than simple authority |

| | |
|---|---|
| shall in addition:<br>● Establish [and publish] an ethics code;<br>● Establish a centralised, high-responsibility unit in charge of verifying the fulfilment of obligations of this law and compliance with the ethics code;<br>● Permit the anonymous deposit of documents and information indicating an infringement of this law or the ethics code in an accessible way e.g. via a website and announce this publicly;<br>● Protect against any discrimination or sanction staff (employees or freelancers) or contractual partners pointing at such infringements;<br>● Establish and manage a risk management system;<br>● Publish annually a report on the most recent risk estimations for the risks referred to in this law;<br>● Study best practices in view of optimising compliance with this law and addressing ethical concerns; and<br>● Register the AI system or its use with the supervising authority. | intervention, bearing in mind that authorities are rather weak in quite some jurisdictions.<br><br>Evidently, the obligations increase with each step up the ladder of Risk Classes, whilst there can be debate on whether to attribute a certain obligation to Risk Class 2 or to Risk Class 3. |
| *27. Additional obligations for Risk Class 3 systems*<br>In addition to the obligations for Risk Class 2 systems, developers and operators and users with clients shall:<br>● Establish a quality management system fulfilling the requirements set out in Standard … (e.g. ISO 9001-2015);<br>● Publish annually a report on the compliance with this law that lists all malfunctions, incidents and facts that might be regarded as problematic in terms of compliance with this law;<br>● Publish the latest auditing report of their conformity assessment body; and<br>● Offer stakeholders [and the general public] an annual dialogue forum. | Idem.<br><br>Quality management standards can be applied in our context too, see this page for the ISO 9000 family and in particular the software specific guidance at the end.<br>A good high-level structure of a quality management system can also be found in Article 17 of the European Commission proposal for an AI Regulation.<br>A combination of elements of both these sources would be ideal. |
| *28. Procedure for Risk Class 2 activities*<br>Operators of Risk Class 2 AI systems [and users with clients] shall in the first three months of their activity and thereafter every three years undergo an audit by an independent, accredited third party conformity assessment body verifying the fulfilment of the conditions and obligations set out | We recommend a nuanced set of procedural obligations, subject to the effective risks. The development of AI systems is less risky than their operation and for Risk Class 2, only the operation, not the development, shall be subject to a |

| | |
|---|---|
| in this law.<br>AI systems of Risk Class 2 shall undergo the procedure foreseen for AI systems of Risk Class 3 where they affect … (one / ten million inhabitants **OR** one tenth of the population). | procedure.<br><br>On one hand, it would be useful to also cover users with clients. On the other hand, the examination capacities might be so limited that it is preferable to exempt them.<br><br>To keep the burden of the supervising authority limited, the respective procedure can be outsourced to a (public or private) third party conformity assessment body. However, we only recommend that step to those jurisdictions which already have some experience with these bodies as it is very cumbersome to set up a respective designation and supervision system from scratch. For jurisdictions that wish nonetheless to establish such a system from scratch, we outline below some basic requirements for such bodies. Please contact the Regulatory Institute if you need further advice on this specific topic for which we have not yet put in writing our knowledge. |
| *29. Procedure for Risk Class 3 activities*<br>Developers and operators of Risk Class 3 AI systems [and users with clients] shall request an authorisation by … [ministry / authority] before starting their activities. [Except for users with clients and] [Except where the AI system has already been subject to an authorisation procedure] The authorisation procedure shall be preceded by the initial audit foreseen for Risk Class 2 activities.<br>The procedure set out in this Section is to be renewed every … (2, 3 or 4) years.<br>The authorisation may be subject to conditions or be limited in time or in scope. It may impose additional procedural steps to stay valid in time. | For Risk Class 3, we deem it proportionate to include developers into the procedure so that the supervising authority is informed and can intervene early on. |
| *30. Procedure for Risk Class 1 activities*<br>Operators [and users with clients] of AI systems of Risk Class 1 shall undergo the procedure foreseen for AI systems of Risk Class 2 where they affect … (one / ten million persons / inhabitants **OR** one tenth of the population).<br><br>Alt. 1:<br>Otherwise, no procedural obligation applies. | Choose “persons” where you wish to refer both to your own and foreign population and “inhabitants” where you wish to refer to your own population alone. |

| | |
|---|---|
| Alt. 2:<br>Otherwise, the AI system is subject to a self-certification procedure. To that end, the natural person representing the legal person **OR** the highest ranking employee in charge of ensuring compliance shall sign, publish and send to the supervising authority a declaration listing the items s/he has verified.<br><br>Developers, operators and users with clients of Risk Class 1 AI systems may voluntarily undergo the procedure for Risk Class 2. | The classic self-certification does not foresee any notification to an authority, whilst such notification provides the supervising authority with the possibility to make a quick (plausibility) check and to give feedback.<br><br>Capacity constraints make this option less attractive, but in general terms, it ensures a higher likelihood of compliance. |
| *31. Requirements for conformity assessment bodies*<br>The … [ministry / authority] shall verify the following conditions before accrediting a conformity assessment body under this law:<br>● Independence from control by another legal body [in another jurisdiction] or by a foreign state when executing tasks under this law;<br>● Availability of qualified permanent employees proportionate to the activities and to the number of clients;<br>● Technical equipment necessary to undertake verification tasks under this law;<br>● A verification scheme encompassing all the verifications to be undertaken under this law;<br>● A quality management system, ensuring the fulfilment of the body's own obligations under this law. | As stated above, we recommend using conformity assessment bodies only where there is already such practice in the respective jurisdiction. Nonetheless, we list here a few useful requirements … |
| *32. Obligations of conformity assessment bodies*<br>Conformity assessment bodies shall:<br>● Examine compliance with the conditions and fulfilment of obligations set out in this law as well as the structural ability of their clients to fulfil these conditions and obligations [on the basis of clear, predetermined pass / fail criteria];<br>● In that context also inspect the internal working methods and structures of their clients that might have an influence on the fulfilment of conditions and obligations, such as quality management systems, ethical codes, mechanisms permitting and protecting whistle-blowers;<br>● Refuse certification in case of grave infringements of this law or in case of structural inability to fulfil the conditions and | … and obligations. |

| | |
|---|---|
| requirements of this law;<br>● Withdraw certificates where the conditions for issuing the certificate were from the beginning not fulfilled or are not fulfilled anymore;<br>● Inform their peers and … [ministry / authority in charge of these bodies] and the supervising authority of any withdraw certificates;<br>● Inform the supervising authority of particular grave infringements and issues or questions that might be relevant for other operators or users;<br>● Follow the instructions of the … [ministry / authority in charge of these bodies] and of the supervising authority;<br>● Seek alignment of practices with their peers;<br>● Inform the … [ministry / authority in charge of these bodies] and the supervising authority of cases where peers have doubtful practices or deviate from the common practice of conformity assessment bodies; and<br>● Publish a register of certificates issued. | |
| *33. Recognition of foreign approvals and certificates*<br>The … [ministry / supervising or other authority] may recognise foreign authorities' decisions [and foreign conformity assessment bodies' certificates] as equivalent to domestic approvals [and domestic conformity assessment bodies' certificates].<br>The recognition may be subject to additional verifications and conditions. | It is an elegant way for reaching a high level of safety and security whilst economising own administrative resources to recognise foreign approvals and certificates where they emanate from a – in that regard – trustworthy foreign jurisdiction. |
| *34. Urgency admission procedure*<br>The supervising authority may, in the public interest and namely to protect the values listed in Section 10 on Risk management, authorise AI systems to be developed and to be used without the conditions and obligations of this law being fulfilled. It shall specify which conditions and obligations are to be fulfilled and not fulfilled. | In some cases, it might be in the public health or another public interest to quickly apply an innovative AI system. |
| *35. Obligations of traders*<br>Traders shall not trade with, broker or otherwise support the dissemination or the use of AI systems that do not fulfil the conditions set out in this law and whose developers, operators and users with clients do not fulfil their respective obligations. To that end, traders shall in particular verify: | By obliging traders to verify certain items, compliance of AI systems can be increased.<br><br>Traders are also very important to avoid the dissemination of AI systems with illicit purposes or to avoid the dissemination to criminals. |

| | |
|---|---|
| ● The availability of instructions for use;<br>● The accessibility of the technical documentation;<br>● That the name, electronic addresses and the physical addresses of developers, operators and users with clients are clearly and correctly marked;<br>● That the procedure for Risk Class 2 or 3 has been undertaken, where applicable; and<br>● That the respective procedure has been concluded with a valid approval [or certificate].<br><br>Where traders have any doubt on the compliance of an AI system or the fulfilment of obligations of developers, operators and users with clients, they shall inform the supervising authority.<br><br>Traders shall publish their contact details and indicate their readiness to collect and forward information on incidents or malfunctions to the responsible developer, operator or user with clients and to the responsible supervising authority. | The certificate is only to be referred to where conformity assessment bodies play a role. |
| Traders of AI systems of Risk Class 3 **OR** 2 and 3 shall register with the supervising authority in a procedure set out by that authority. | The registration of traders facilitates the control of the non-proliferation obligation of traders and the prevention of trade with illicit AI systems. |
| *36. Research and its funding*<br>Research and its funding is limited to those applications which are permitted under this law. However, illicit applications may be developed [by public research institutions] with the view to detect possibilities to counter them, provided that the security rules applicable to … (e.g. military undertakings) are applied. | In particular as research also takes place in private or semi-private institutions, we recommend keeping some boundaries and not exempting research from the scope of this law. |
| Researchers or research institutions making available their AI systems shall fulfil the obligations of developers or operators, subject to the way of dissemination. However, they may escape these obligations by handing over their rights to a legal person organisationally and financially capable of assuming all responsibilities of a developer or operator towards third parties and the authorities. Such capability shall be presumed to be given where the legal person has an annual turn-over of … and exists for at least … years. In other cases, the agreement of the | On one hand, it should be avoided that, under the label of “research” problematic AI systems are disseminated and made available. On the other hand, the crucial function of (fundamental) research, which is to trigger technological innovation, should not be hampered. This subsection tries to strike a balance by permitting the handing-over of knowledge provided that the recipient is capable of fulfilling all the obligations of a developer or operator. We suggest defining cases in which this is to be |

| | |
|---|---|
| supervisory authority is required prior to any handing-over of knowledge.<br><br>Researchers and research institutions cooperating in international partnerships or consortia shall still (Alt. 1) **OR** (Alt. 2) shall not be bound by the obligations of this law **OR** (Alt. 3) shall only be bound by the following obligations of this law:<br>● … ;<br>● … .<br><br>They shall not cooperate where a part of the AI system is illicit in the meaning of Sections 12 or 13 or where an illicit application is likely. They shall not cooperate where the non-proliferation obligation of Section 14 is likely to be infringed. | presumed and subjecting the other cases to an authority approval.<br><br>Fundamental research is often so expensive that one jurisdiction alone cannot finance it. International cooperation is needed. At the same time, one cannot expect the international partners to follow the domestic legal requirements. Hence, a balance needs to be struck in light of the respective jurisdiction.<br><br>For further aspects of research regulation, see this prototype regulation and its preparing articles. |
| *37. Copyright violation by AI*<br>The internal processing of copyright protected content by the AI system does not constitute a violation of copyright.<br>For the display of copyright protected content to operators, users or clients, the ordinary copyright rules apply. | This aspect should preferably be regulated in the respective copyright law. However, if the next amendment of the respective copyright law is not expected in the short term, this Section might be useful.<br>It would be preferable to refer to a precise act on copyright instead of “ordinary copyright rules”. |
| *38. Copyright protection of AI*<br>AI systems and their source codes are copyright protected according to ordinary copyright rules. However, the state may impose mandatory (free) copyright licences or may impose (free) services to be provided with the help of AI systems in the following cases:<br>● existential threat for humankind;<br>● epidemics with more than 1.000 casualties to be expected;<br>● disruption of the infrastructure and the public order [with a potential to indirectly cause 1.000 casualties or with a potential to topple the constitutional system];<br>● attacks by foreign powers;<br>● …<br>The state shall compensate the copyright holder or service provider in a proportionate manner. The compensation shall cover at least the additional costs caused by the provision of the licence or service [and 1/3 of the commercial tariff | Idem.<br><br>AI systems can help tackle certain catastrophes. Possibly, copyright should not be a hindrance for addressing these. |

| | |
|---|---|
| for such provision]. | |
| *39. Open source AI systems*<br>Developers of open source AI systems shall deposit the source codes in a way that requires downloaders to register and to commit to ensuring traceability along the distribution chain.<br>Operators shall take over the obligations incumbent on developers.<br>Operators using open source Risk Class 3 AI systems shall fulfil the obligations of Risk Class 3 developers, including the respective procedure set out in Section 29 [with the exception of … ], unless another operator has successfully passed that procedure. | Open source developers often work without payment and contribute to technical progress voluntarily. In view of the absence of any (direct) financial reward for service it would be unfair to impose heavy obligations on open source developers. On the other hand, open source AI systems are not necessarily less risky than others. To strike a balance between these two aspects, we suggest imposing an obligation to ensure traceability of the distribution and obliging the (mostly commercial) operators of open source AI systems to take over the obligations that normally are covered by developers. However, some of the obligations of developers would be too cumbersome to be fulfilled and might need to be exempted. |
| *40. Ownership of data*<br>Users and their clients keep the ownership of data provided by them. However, operators and users have the ownership of aggregated data in which the individual data of users and clients cannot be identified anymore. | Another side aspect possibly to be regulated is the ownership of data. We suggest differentiating between the raw data and aggregated data, produced by the AI system. |
| *41. Right to access and correct data*<br>For all AI systems except the following [natural] persons concerned have the right to access their personal data and the right to request corrections thereof:<br>● Customs, police and prosecutors investigative AI systems;<br>● … ;<br>● … . | This topic might already be governed by another law of the jurisdiction in question. But if not, it is useful to provide access and a right to correct data. Correct data is also in the interest of the operators or users with clients. |
| *42. Right to consent*<br>Alt. 1:<br>For the use of the following AI systems the explicit consent of the [natural] person concerned is required:<br>● AI systems that operate personal data to an extent a data subject's consent is needed under … (respective domestic legislation);<br>● AI systems integrating five or more personal parameters in addition to the name and contact data;<br>● AI systems integrating political or religious | Alternative 1 enumerates the applications for which consent is needed, whereas Alternative 2 requires consent by default and just exempts certain applications. Alternative 2 provides a better protection of citizens, but might stifle innovation or create an unnecessary burden. The opposite is true for Alternative 1. |

| | |
|---|---|
| beliefs or preferences, sexual orientation and / or gender identity;<br>● AI systems tracking preferences for certain cultural products or traditions;<br>● AI systems tracking social and political engagement, activities and positions;<br>● AI systems processing and deciding on natural persons OR consumers;<br>● AI systems presenting 'deep fake';<br>● AI systems operating emotion recognition systems; and<br>● AI systems operating biometric categorisation.<br>The consent must be linked to one of the categories listed above.<br><br>Alt. 2:<br>For the use of any AI systems except the following the explicit consent to the [natural] person concerned is required:<br>● AI system processing elections;<br>● AI systems operated by public authorities;<br>● … | We propose in this Model Law that deep fakes be prohibited (see above Section 13), however, should this option not be pursued, requiring consent for presenting deep fakes will facilitate authority surveillance of its use. |
| Explicit consent is given where it is provided in writing or by ticking a box followed by an acceptance click, by oral agreement or by a similarly explicit and unequivocal expression of acceptance. | If Alt. 2 is selected then the following Sections 43 right to know and 44 right to refuse should not be selected for reasons of consistency. |
| *43. Right to know*<br>Anybody put in contact with an AI system that acts similarly to a person to an extent where it may be mistaken for a person needs<br>(Var. 1) to be explicitly informed thereof in a clear way.<br>(Var. 2) to provide explicit consent before the contact.<br>Explicit consent is given where it is given in writing or by ticking a box followed by an acceptance click, by oral agreement or by a similarly explicit and unequivocal expression of acceptance. | Subject to the variants chosen, this section can be merged with the previous section. |
| *44. Right to refuse the processing by AI systems*<br>For the following applications / uses / cases clients / consumers may refuse to be processed or subject to decision-making by AI systems:<br>● processing of medical data;<br>● processing of data on religious, philosophical or political preferences; | Subject to the society in question, various types of data or their processing might be regarded as problematic or worth being protected by the right to refuse the processing. |

| | |
|---|---|
| ● processing of requests for review of an administrative decision (e.g. by an applicant for social security or a taxpayer);<br>● …<br>[The right to refuse does not cover cases where an AI system processes depersonalised data.] | Where a public authority uses AI in the processing of claims for social security or taxation matters, applicants or taxpayers may request a review of the decision and such a review of an administrative decision should be made by a human. |
| *45. Liability and its insurance*<br>Alt. 1:<br>Operators shall be liable to their users and users shall be liable to their clients in accordance with the common rules for contractual and non-contractual liability. However, no proof of negligence is required **OR** strict liability applies.<br><br>Alt. 2:<br>Operators shall be liable to their users and users shall be liable to their clients in accordance with the following principles:<br>● Strict liability applies;<br>● The minimum damage assumed by law is: … ;<br>● In case of gross negligence, the damage is increased by a factor of 2;<br>● In cases where at least one staff was aware of an infringement, the damage is increased by a factor of 4;<br>● For all other aspects, common rules for contractual and non-contractual liability apply.<br><br>These liability obligations cannot be reduced by contracts or unilateral declarations.<br><br>Joint liability applies in case of multiple persons contributing to the damage.<br><br>Operators and users with clients shall contract liability insurance covering their liability risks. The liability insurance shall cover at least damages up to three times the annual turn-over of the operator or user or … (put a monetary value), whatever is higher. | Liability obligations have a double function: to repair damage and to deter. The second aspect can be reinforced by choosing Alternative 2.<br><br>Both alternatives use strict liability because it is almost never possible for users or clients to prove negligence.<br><br>More aspects might need to be regulated, such as prescription periods.<br><br>The involvement of insurance ensures on one hand solvency. On the other hand, insurance companies sometimes apply a kind of indirect control of compliance to reduce their own risks. |
| *46. Injunction and compensation*<br>Competitors[, public consumer protection bodies and associations recognised for defending the public interest] may sue by injunction, operators or users with clients for infringement, at the … (court / tribunal) after having requested in writing, | Injunction by competitors, consumer protection bodies and public interest associations is a useful tool to increase compliance, in particular where authorities are weak. However, it works only to the |

| | |
|---|---|
| without success, the injunction. Success is only given where the infringing operator or user with clients recognises in writing the right to injunction and effectively stops the infringement.<br>Any successful injunction claim gives right to compensation for the actual damage in accordance with the previous Section or a minimum compensation of … or 1/10 … 1/50 of the annual turn-over. | extent that the judicial system is effective.<br><br>Establishing obligatory pre-trial procedures can be a useful option. It reduces the burden of courts. But it also can be an additional obstacle to those seeking compensation. |
| *47. Supervision*<br>The responsible authority in charge of supervising the application of this law is …, hereafter the "supervising authority". | In some jurisdictions, several authorities should be designated, based on their specialisation or geographic proximity. |
| *48. Empowerments*<br>Subject to the Risk Class, the supervising authority has the following empowerments with regard to developers, operators and users with clients:<br><br>For Risk Class 1:<br>● Requesting information [on the AI system, on contractual partners, on contracts on economic data], and this also where the informant is a third party or located in a third country or both;<br>● Communicating warnings and recommendations to the population;<br>● Ordering the infringing developers, operators and users with clients and their media and internet service providers to communicate warnings and recommendations;<br>● Blocking or removing content from internet websites offering AI systems or access thereto;<br>● Interrupting or fully controlling telephone, media and internet services of continuously infringing developers, operators or users with clients or ordering respective service providers to do so;<br>● Requesting developers, operators and users with clients to take certain steps in order to stop an infringement or to reduce the likelihood of further infringements;<br>● Recovering costs triggered by the investigation and enforcement costs from infringing developers, operators or users with clients;<br>● Sanctioning developers, operators and users | In the following, we differentiate again in accordance with the Risk Classes, namely to ensure proportionality of state measures. However, such differentiation is not necessary or pertinent in certain jurisdictions.<br><br>All the empowerments need to, of course, be scrutinised in the light of the respective constitution or other law of higher order. This applies in particular for confiscations, intrusion into facilities and supervision of communication.<br><br>Readers should feel free to shift empowerments from Risk Class 1 to Risk Class 2 and vice versa. |

| | |
|---|---|
| with clients who do not respect the conditions or do not fulfil the requirements set out in this law with administrative sanctions up to three times the annual turn-over; and<br>● Enforcing financial obligations and financial sanctions or penalties via confiscation of AI systems, rights, money or other items in possession of the infringing person.<br><br>For Risk Classes 2 and 3:<br>● The empowerments listed for Risk Class1;<br>● Obliging contractual partners of infringing developers, operators and users with clients to stop, limit or modify their cooperation;<br>● Obliging developers, operators and users with clients to display information on the conformity assessment of regulated products or services on their website;<br>● Requiring operators to inform users of infringements affecting them and requiring users with clients to inform their clients of infringements affecting them;<br>● Compelling the attendance of witnesses, including third parties, to provide evidence under subpoena, when there are reasons to believe or there is evidence of infringement;<br>● Creating financial or other incentives for persons to provide or confirm information;<br>● Inspecting, without notice, offices, factories, warehouses, wholesaling establishments, retailing establishments, laboratories, research institutions and other premises or vehicles in which AI systems are produced or kept;<br>● Taking samples or copies of AI systems or purchasing them, openly or covertly;<br>● Reverse engineer AI systems;<br>● Supervising the AI system during the course of an investigation of infringement.<br>● Confiscating documents, data and AI systems;<br>● Targeted dissemination of warning information to partners of the infringing developer, operator or user with clients;<br>● Confiscating assets of infringing developers, operators or users with clients;<br>● Sanctioning natural and legal persons who contributed to an infringement;<br>● Requesting securities (as guarantee for the fulfilment of non-financial obligations)<br>● Publishing a blacklist of natural and legal | |

| | |
|---|---|
| persons who committed or contributed to infringements;<br>● Excluding those persons from public tenders and grants;<br>● Extending the measures listed above to agents of the infringing person;<br>● Extending the measures listed above to mother and sister companies of the infringing legal person and their agents;<br>● Extending the measures listed above to commercial partners of the infringing person where these have contributed to the infringement. | |
| *49. Penal sanctions*<br>In case of deliberate infringement of the obligations set out in this law, the following penal sanctions apply to the natural persons responsible for the infringement, regardless of whether they are employees[ or freelancers] of the infringing legal person or contractors or staff of contractors:<br>● For infringements of Sections 13, 14, 28 and 29 from … to … years of imprisonment and/or a fine of up to triple their annual net salary.<br>● For infringements of Sections 4 to 6, 8, 10, 12, 17, 18, 20, 21, 23 to 27, 31, 32, 35, 36, 39 and 41 to 45 up to … years of imprisonment or a fine of up to double their annual net salary.<br>● For infringements of Sections 7, 9, 11, 19 and 22 up to … years of imprisonment or a fine of up to their annual net salary.<br><br>In case of unintentional non-compliance of obligations [or where the deliberate character of the infringement cannot be proven], the following penal sanctions apply to the natural persons responsible for the infringement, regardless of whether they are employees or freelancers of the infringing legal person or contractors or staff of contractors:<br>● For infringements of Sections 13, 14, 28 and 29 up to … years of imprisonment or a fine of up to their annual net salary.<br>● For infringements of Sections 4 to 6, 8, 10, 12, 17, 18, 20, 21, 23 to 27, 31, 32, 35, 36, 39 and 41 to 45 up to … years of imprisonment or a fine of up to 2/3 their annual net salary.<br>● For infringements of Sections 7, 9, 11, 19 and | There are evidently many ways to design a system of sanctions. Hence, the provisions should only be regarded as inspiration for the development of own provisions fitting to the respective domestic penal practice.<br><br>We recommend distinguishing between deliberate and unintentional infringements. The part in square brackets might be necessary or not. We avoid here the term “negligent” which is often understood as implying the violation of a duty of care and therefore renders sanctioning more difficult. However, in some jurisdictions, it would not be appropriate to sideline this duty of care aspect.<br><br>Only for deliberate infringements of the most important obligations, a minimum imprisonment sanction seems appropriate.<br><br>For this and the following bullet, we would not deem imprisonment to be appropriate. |

| | |
|---|---|
| 22 up to … years of imprisonment or a fine of up to 1/3 their annual net salary.<br><br>In addition to or instead of the sanctions listed in this section, the supervising agency may impose the following collateral sanctions against the infringing legal persons:<br>● Administrative sanctions of up to three times their annual budget or turn-over;<br>● Exclusion from public tenders for up to X years; and<br>● Citation in the public list of law infringing / criminal economic operators for up to X years.<br><br>● The supervising agency may also:<br>● Publish the names and further data permitting the identification of natural or legal persons who have deliberately infringed this law;<br>● Confiscate and destroy, as sanction and thus regardless of their illegal character, a proportionate amount of AI systems, and this in particular where fines are not paid;<br>● Suspend the commercial licenses of the infringing persons; and<br>● In case of particular grave or repetitive deliberate infringements close the facilities of the infringing persons.<br><br>Sanctions and the collateral measures set out in this section may be extended to parent or subsidiary companies or other legal or natural persons and the staff of all these persons if these legal or natural persons controlled the infringing person to such an extent that they were in reality responsible for the infringement.<br><br>The supervising agency may oblige any natural or legal person to cooperate for purposes of enforcement of sanctions and collateral measures set out in this section, including with regard to the disclosure of confidential information, the hand-over of assets of all forms, the temporary closure of websites, the suspension of services supporting the economic activity of the infringing persons. | Some legal persons are in reality controlled by another legal person. Some legal persons create companies as shields or shell companies without assets to pay sanctions. This provision empowers the authorities to counter this situation.<br><br>We have, above in Section 48, listed a full range of empowerments serving the enforcement of obligations of this law. However, these empowerments do not cover the enforcement of sanctions and their collateral measures. Hence, separate empowerments are needed in order to enforce the sanctions with the help of third persons. |
| *50. Legal remedies*<br>Decisions of the supervising authority taken in accordance with this law may be challenged within X months in writing and by [authenticated] | Such provisions are evidently not necessary where generic administrative law contains sufficient provisions. |

| | |
|---|---|
| electronic email at …<br>(higher administration, ministry or court).<br>[Remedies against the decisions of the …<br>(higher administration, ministry) shall be addressed at … (one or several courts). | |
| *51. Incidents and alert portal*<br>The supervising authority shall provide an electronic interface for the, if so desired, anonymous deposit of information on incidents and malfunctions. It shall evaluate this information.<br><br>Any confirmed information on incidents and malfunctions shall, regardless of its origin, be listed in a public incidents and alert portal. That portal may contain a non-public section for confidential information and the authority's own investigation. | Parallel to an interface for the deposit of information on incidents and malfunctions to be established by developers, operators and users with clients, the supervising authority, too, should create such an interface in particular for the cases where developers, operators and users with clients try to hide facts.<br><br>Transparency with regard to confirmed incidents and malfunctions might have a deterring and sanctioning effect, but might also inform competitors on structural risks. |
| *52. Forum for developers, operators and users with clients*<br>The supervising authority shall offer developers, operators and users with clients and their employees a forum in which they can exchange good practices aimed at reducing incidents and malfuncts. It may offer respective training[, attendance of which is mandatory for developers and operators]. | This is another measure to reduce the likelihood of incidents and malfunctions. |
| *53. Rating and labelling*<br>The supervising authority may establish a rating of AI systems in terms of risks, relative frequency of incidents, responsiveness and compliance of operators and users with clients.<br>Operators and users with clients are obliged to publish and label on their products or service related media obtained rating in a well visible way. | Such a rating and labelling system is an incentive for compliance and risk avoidance.<br><br>This system is enhanced by the obligation to publish and label the rating. |
| *54. Monitoring of AI development*<br>The supervising authority and the responsible ministry shall at least every second year inform the Parliament on new technological developments, the issues caused by them and the appropriateness of this law. | A rather short reporting cycle is useful in this quickly evolving technological context. |

| | |
|---|---|
| *55. Intra-state cooperation*<br>State authorities and public bodies shall inform the supervising authority about any plans to develop, operate or use AI systems. They shall seek the advice of the supervising authority for the respective project and shall provide access to the project at any time.<br>State authorities and public bodies shall inform the supervising authority of any potential infringement of obligations or non-respect of conditions set out in this law. They shall respond to requests for support of the supervising authority. | In particular where AI systems of state authorities and public bodies are exposed to reduced requirements, it is important that the supervising authority takes a closer look at them.<br><br>Moreover, state authorities and public bodies can help to pursue infringements of this law. |
| *56. International cooperation*<br>The supervising authority may cooperate with its peers in partner jurisdictions and with international organisations. It may share with peers [and international organisations] information, including sensitive information, on developers, operators and users with clients where this is necessary to pursue the activities foreseen in this law with regard to natural or legal persons governed by it, regardless whether having residence or place of business on the domestic territory or not.<br><br>The supervising authority may also use the empowerments set out in Section 48 to pursue potential infringements of the law of partner jurisdictions where the law of the partner jurisdictions respects the principles of rule of law and human rights. | Internet business is international, therefore any jurisdiction is likely to need sooner or later the support of other jurisdictions to pursue its compliance policy. International cooperation is hence paramount.<br><br>Enforcement on the territory of another jurisdiction is often only possible where there is reciprocity of support, which requires that enforcement empowerments may also be used in favour of a foreign jurisdiction. |
| *57. AI in non-state governed areas and space*<br>The supervising authority shall monitor the operation of AI systems in areas of the earth not governed by any state and in space where the AI operating might now or in the future have effects for the domestic territory, its citizens or the interests of natural or legal persons with residence or place of business therein.<br>The supervising authority shall seek the cooperation of international peers to hold the operation of those AI systems under check, namely by limiting the access thereto and the dissemination of its outputs. The supervising authority may, to that end, use the empowerments set out for AI systems of Risk | Beyond the issue of territories which are not effectively governed by any state, there is the issue of activities launched from the international seas (so far only sporadic events of this kind) or from space. AI systems operating from space could already today be launched by certain private space companies that can outplay any state control. |

| | |
|---|---|
| Class 3. | |
| *58. Human decisions based on AI systems*<br>The following sections of this law apply also to human decisions that are based on the use of AI systems:<br>●<br>●<br>● | Legislators might find that some human decisions based on AI systems should trigger the same consequences as the use of AI systems as such. In this event, the scope (1st section) might need to be enlarged. |
| *59. Adaptation of this law to technical progress and closing regulatory loopholes*<br>The government may adapt this law to technical progress and may close regulatory loopholes, whilst respecting the principles set out in this law. | The fast evolving topic renders such an empowerment useful. Contrary to the empowerment for (executing) decrees or other subordinate legislation, below, the level of intervention is the level of the law itself. |
| *60. Parliamentary control of adaptations*<br>Adaptations of this law to technical progress and closing loopholes may be revoked, suspended or limited by decision of a 2/3 majority of the parliament. | Nonetheless, parliamentary control might need to be ascertained. |
| *61. Government decrees, subordinate legislation*<br>The government may issue decrees setting out details on the execution of the empowerments and on the management of this law. [Where there is no urgency, it shall give the parliament one month notice prior to adoption and shall take account of the reaction of the parliament.]<br><br>The decrees may / may not further restrain data protection law and the protection of confidential information.<br><br>The decrees may / may not further limit other rights of legal and natural persons. | Decrees, subordinate legislation or similar regulatory tools of the government can complement the law appropriately. However, given the very sensitive character of the matter, it might also be deemed appropriate to give the parliament the possibility to informally react to a decree project. This is also helpful where, as suggested below, the parliament can formally revoke a decree. The informal reaction at an earlier stage can help to avoid a later conflict which would let the parliament revoke the decree. |
| *62. Parliamentary control of government decrees*<br>The parliament may revoke, suspend or modify the decrees adopted by the government by ordinary majority decision. | As stated above, the rights-sensitive character of the matter might be regarded as justifying a tight control of the government by the parliament. |