Livre blanc sur l'impact de la loi IA de l'UE relative à l'utilisation de l'IA conversationnelle à des fins professionnelles

www.spitch.ai

Foire aux questions sur l'impact de la loi européenne sur l'intelligence artificielle (IA) relative à l'utilisation de l'IA conversationnelle à des fins professionnelles

mars 2024

Pertinence et impact potentiel de la loi européenne sur l'IA pour les entreprises

Les premiers adeptes figurant parmi certaines des plus grandes entreprises capitalisent déjà sur les vastes possibilités offertes par GenAI[1] et les LLM[2]. Ils tiennent toutefois à trouver un équilibre entre les impératifs de confidentialité et la sécurité des données, qui sont également perçus comme fondamentaux par les clients finaux. Les solutions d'IA conversationnelle pour le service à la clientèle, telles que les assistants virtuels (chatbots textuels et vocaux), la biométrie vocale et l'analyse de la parole, entre autres fournies par Spitch, font réellement la différence pour les performances de l'entreprise et la sécurité des données. L'équilibre entre les avantages commerciaux des techniques innovants, une réglementation efficace pour contrôler les risques et la responsabilité sociale semble essentielle pour une adoption durable de l'IA.

L'introduction de la loi européenne sur l'IA fournit le premier cadre réglementaire complet pour la gouvernance de l'IA. Elle aura forcément un impact sur les fournisseurs et les utilisateurs de certaines formes de GenAI et de biométrie de reconnaissance faciale qui peuvent potentiellement présenter des risques si elles sont utilisées de manière non réglementée et incontrôlée. Nous pensons que l'impact de la nouvelle réglementation européenne sera largement positif pour les cas d'utilisation dans différentes industries et dans le secteur public, avec une influence relativement plus faible sur les fournisseurs d'IA conversationnelle spécifique à un domaine d'activité. Néanmoins, de nouvelles exigences devront être respectées par tous les fournisseurs et déployeurs de systèmes d'IA.

Ces exigences devraient être encore plus développées par les organismes de normalisation européens. Par la suite, les autorités nationales devront veiller à ce que les entreprises se conforment aux nouvelles exigences et normes en matière de gouvernance et de gestion des risques liés à l'IA, tout en évaluant dans quelle mesure des orientations sectorielles plus détaillées pourraient être nécessaires.

Résumé de la loi européenne sur l'IA*

La loi sur l'IA classe l'IA en fonction des risques qu'elle présente :

- Les risques inacceptables sont interdits (par exemple, les systèmes de notation sociaux et l'IA manipulatrice).
- La majeure partie du texte porte sur les systèmes d'IA à haut risque, qui sont réglementés.
- Une section plus restreinte traite des systèmes d'IA à risque limité, soumis à des obligations de transparence plus légères : les développeurs et les déployeurs doivent s'assurer que les utilisateurs finaux savent qu'ils interagissent avec une IA (chatbots et deepfakes).
- Le risque minimal n'est pas réglementé (y compris la majorité des applications d'IA actuellement disponibles sur le marché unique de l'UE, telles que les jeux vidéo et les filtres anti-spam activés par l'IA, au moins en 2021 ; cette situation est en train de changer avec l'IA générative).

La majorité des obligations incombent aux fournisseurs (développeurs) de systèmes d'IA à haut risque.

- Ceux qui ont l'intention de mettre sur le marché ou de mettre en service des systèmes d'IA à haut risquent dans l'UE, qu'ils soient basés dans l'UE ou dans un pays tiers.
- Et également les fournisseurs de pays tiers où les résultats du système d'IA à haut risque sont utilisés dans l'UE.

Les utilisateurs sont des personnes physiques ou morales qui déploient un système d'IA à titre professionnel, et non des utilisateurs finaux concernés.

- **Les utilisateurs (déployeurs) de systèmes d'IA à haut risque** ont certaines obligations, mais moins que les fournisseurs (développeurs).
- Cela s'applique aux utilisateurs situés dans l'UE et aux utilisateurs de pays tiers lorsque les résultats du système d'IA sont utilisés dans l'UE.

* Résumé de haut niveau de la loi sur l'IA

Spitch estime qu'il peut être utile pour les clients et partenaires existants et potentiels de consulter les réponses à certaines des questions fréquemment posées sur les nouvelles exigences, résumées ci-dessous.

[1] *GenAI - Intelligence artificielle générative*

[2] *LLMs - Grands Modèles de Langage*

SPITCH

Foire aux questions sur l'impact de la loi européenne sur l'intelligence artificielle (IA) relative à l'utilisation de l'IA conversationnelle à des fins professionnelles

FAQ

1. Quels sont les systèmes d'IA interdits en vertu de la loi européenne sur l'IA ? Les systèmes d'IA conversationnelle pour le service à la clientèle entrent-ils dans la catégorie des solutions « interdites "» ?

Les types de systèmes d'IA suivants sont "« interdits "» en vertu de la loi sur l'IA (Titre II, Art. 5)[3] :

- Déployer **des techniques subliminales, manipulatrices ou trompeuses** pour fausser le comportement et entraver la prise de décision éclairée, causant ainsi un préjudice important.
- **L'exploitation des vulnérabilités** liées à l'âge, au handicap ou à la situation socio-économique pour fausser le comportement et causer un préjudice important.
- **Systèmes de catégorisation biométriques** déduisant des attributs sensibles (race, opinions politiques, appartenance à un syndicat, croyances religieuses ou philosophiques, vie sexuelle ou orientation sexuelle), à l'exception de l'étiquetage ou du filtrage d'ensemble de données biométriques acquis légalement ou lorsque les forces de l'ordre catégorisent des données biométriques.
- **La notation sociale**, c'est-à-dire l'évaluation ou la classification d'individus ou de groupes sur la base de leur comportement social ou de leurs traits personnels, ce qui entraîne un traitement préjudiciable ou défavorable de ces personnes.
- **Évaluer le risque qu'une personne commette des infractions pénales** uniquement sur la base d'un profilage ou de traits de personnalité, sauf lorsque cela sert pour compléter des évaluations humaines fondées sur des faits objectifs et vérifiables directement liés à l'activité criminelle.
- **Compilation de bases de données de reconnaissance faciale** par l'extraction non ciblée d'images faciales d'Internet ou d'images de vidéosurveillance.
- **Inférer des émotions sur les lieux de travail ou dans les établissements d'enseignement**, sauf pour des raisons médicales ou de sécurité.

Résumé de la loi européenne sur l'IA*

IA à usage général (GPAI) :

- Tous les fournisseurs de modèles GPAI doivent fournir une documentation technique, des modes d'emploi, se conformer à la directive sur les droits d'auteur et publier un résumé du contenu utilisé pour la formation.
- Les fournisseurs de modèles GPAI sous licence libre et gratuite doivent uniquement respecter les droits d'auteur et publier le résumé des données de formation, à moins qu'ils ne présentent un risque systémique.
- Tous les fournisseurs de modèles GPAI qui présentent un risque systémique, qu'ils soient ouverts ou fermés, doivent également procéder à des évaluations de modèles, à des tests contradictoires, suivre et signaler les incidents graves et garantir des protections en matière de cybersécurité.

* Résumé de haut niveau de la loi sur l'IA

* Résumé de haut niveau de la loi sur l'IA

- **Identification biométrique à distance (RBI) « en temps réel "» dans les espaces accessibles au public pour les forces de l'ordre**, sauf dans les cas suivants :
  - Recherche de personnes disparues, de victimes d'enlèvement et de personnes victimes de la traite des êtres humains ou de l'exploitation sexuelle ;
  - Prévention d'une menace substantielle et imminente pour la vie ou une attaque terroriste prévisible ; ou
  - Identification de suspects d'infractions graves (meurtre, viol, vol à main armée, trafic de stupéfiants et d'armes illégales, crime organisé, crimes contre l'environnement, etc.).

Aucun des systèmes d'IA conversationnelle généralement acceptés et couramment utilisés à des fins professionnelles dans le cadre du service à la clientèle n'entre dans ces catégories.

[3] *Source : https://artificialintelligenceact.eu/high-level-summary/*

De nombreuses entreprises, en particulier les banques, ont commencé à déployer la forme en temps réel de la vérification d'identité biométrique vocale pour leurs clients. Dans la plupart des cas, elle sert à renforcer un processus de vérification de l'identité contrôlé par l'homme au cours d'une conversation en direct avec un agent du centre de contact. Pendant l'appel, la voix du client est comparée à l'empreinte vocale créée précédemment (une représentation mathématique des caractéristiques vocales de l'utilisateur) afin de s'assurer que la voix de l'appelant correspond.

Il existe également des processus de vérification biométrique vocale « hybrides "» dans lesquels l'appelant répète des chiffres ou des mots générés de manière aléatoire et où la voix est comparée à l'empreinte vocale. Cette méthode permet d'accélérer la vérification.

À moins qu'elles ne soient utilisées dans des espaces publics par les forces de l'ordre, ces méthodes d'authentification biométriques vocale à distance n'entrent pas dans la catégorie des systèmes d'IA interdits.

Toutefois, si les systèmes d'IA conversationnelle utilisent la détection et l'inférence des émotions pour prendre des décisions automatisées dans un environnement professionnel ou pour classer les appelants en fonction des émotions détectées, ces systèmes peuvent entrer dans la catégorie des systèmes d'IA interdits en vertu de la loi européenne sur l'IA.

Toutefois, si l'analyse des sentiments est utilisée dans le cadre d'un système d'IA plus large pour le service à la clientèle afin de comprendre les réactions des clients et d'améliorer la qualité du service, sans prendre de décisions automatisées ayant un impact direct sur les personnes, elle ne serait pas considérée comme un cas d'utilisation interdit. Il est important de noter que la loi européenne sur l'IA n'interdit pas l'utilisation de l'analyse des sentiments en général, mais plutôt des applications spécifiques jugées à haut risque ou contraire à l'éthique. Les entreprises doivent évaluer avec soin la manière dont elles prévoient d'utiliser l'analyse des sentiments et s'assurer qu'elle ne tombe pas sous le coup des cas d'utilisation interdits décrits dans la loi.

2. Quels sont les systèmes d'IA classés comme "« à haut risque "» en vertu de la loi européenne sur l'IA ? Est-il possible que l'utilisation de systèmes d'IA conversationnelle crée des risques pour les entreprises en raison de cette classification et de la nécessité de satisfaire à des exigences supplémentaires ?

Certains systèmes d'IA sont considérés comme « à haut risque "» en vertu de la loi sur l'IA (**Titre III**)[4]. Les fournisseurs de ces systèmes seront soumis à des exigences supplémentaires. Systèmes d'IA à haut risque (**Art. 6**) Sont considérés comme ceux qui sont utilisés comme composant de sécurité ou comme produit couvert par les lois de l'UE dans **l'Annexe II** **ET** devant faire l'objet d'une évaluation de la conformité par un tiers en vertu de ces lois de l'annexe II; **OU** ceux dans les cas d'utilisation de **l'Annexe III** (voir l'encadré ci-dessous), sauf si :

- Le système d'IA exécute une tâche procédurale restreinte ;
- Cela améliore le résultat d'une activité humaine déjà réalisée ;
- Cela détecte des modèles de prise de décision ou des écarts par rapport à des modèles de prise de décision antérieure et n'est pas destiné à remplacer ou à influencer l'évaluation humaine réalisée précédemment sans un examen humain approprié ; ou
- Cela effectue une tâche préparatoire à une évaluation pertinente aux fins des cas d'utilisation énumérés à l'Annexe III.

Il est important de noter que le système d'IA est toujours considéré à haut risque s'il sert à **profiler des individus**, c'est-à-dire s'il effectue un traitement automatisé de données à caractère personnel pour évaluer divers aspects de la vie d'une personne, tels que ses performances professionnelles, sa situation économique, sa santé, ses préférences, ses intérêts, sa fiabilité, son comportement, sa localisation ou ses déplacements.

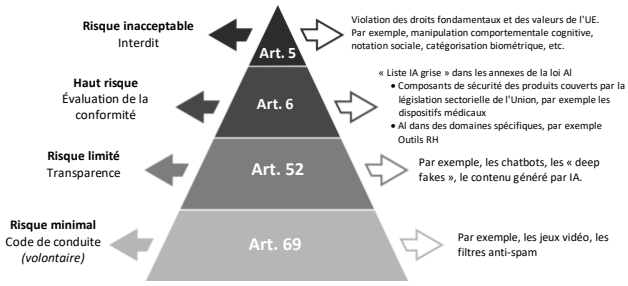
Les fournisseurs qui estiment que leur système d'IA, qui relève de **l'Annexe III**, n'est pas à haut risque, **doivent documenter** cette évaluation avant de le mettre sur le marché ou de le mettre en service.

L'article 8-25 de la loi impose un certain nombre d'exigences aux fournisseurs de systèmes d'IA à haut risque. Il s'agit notamment de l'obligation de mettre en place un **système de gestion des risques** tout au long du cycle de vie du système IA à haut risque ; mettre en œuvre une gouvernance des données appropriées, d'établir une documentation technique pour démontrer la conformité et fournir aux autorités les informations nécessaires pour évaluer cette conformité ; de veiller à ce que le système IA à haut risque prévoie **une tenue de registres** enregistrant automatiquement les événements pertinents pour l'identification des risques au niveau national et les modifications substantielles tout au long du cycle de vie du système.

Parmi les autres exigences figure l'élaboration de mode d'emploi par les déployeurs en aval pour permettre à ces derniers de se conformer et de mettre en œuvre une surveillance humaine ; la garantie des niveaux appropriés de **précision, de robustesse et de cybersécurité**, ainsi que la mise en place d'un **système de gestion de la qualité** pour assurer la conformité.

[4] *Source : https://artificialintelligenceact.eu/high-level-summary/*

Certains systèmes biométriques non interdits, tels que les systèmes de biométrie vocale fournis par Spitch, peuvent relever des systèmes énumérés à l'annexe III. D'autres systèmes d'IA conversationnelle, tels que les chatbots, l'analyse de la parole, les bases de connaissances, etc. peuvent relever des catégories "« risque limité » ou « risque minimal », avec des exigences supplémentaires minimes ou inexistantes, comme l'illustre le graphique[5] ci-dessous :



3. Quel est le calendrier de mise en conformité avec la loi européenne sur l'IA ?

Spitch propose des services de conseil à ses clients pour les aider à assurer une conformité réglementaire complète et opportune, conformément à la loi européenne sur l'IA et à d'autres cadres réglementaires dans un pays ou une région spécifique, y compris les réglementations relatives à la protection des données.

Nos juristes aident les prestataires à formuler des contrats clients basés sur les exigences légales existant dans chaque pays qui réglemente le déploiement des systèmes d'IA. Ils doivent être approuvés par les juristes du prestataire avant le lancement du projet.

Vous trouverez ci-dessous le graphique[5] illustrant le calendrier de mise en œuvre, selon le journal officiel de l'UE. Pour la plupart des solutions d'IA conversationnelle, par exemple les biométries non interdites et les chatbots, les exigences supplémentaires devront être satisfaites dans un délai de 12 à 36 mois à compter de mai 2024, date d'entrée en vigueur de la loi européenne sur l'IA.

Cas d'utilisation de l'Annexe III

Systèmes de biométrie non interdits : Les systèmes d'identification biométrique à distance, à l'exclusion des vérifications biométriques qui confirment qu'une personne est bien celle qu'elle prétend être. Systèmes de catégorisation biométriques déduisant des attributs ou des caractéristiques sensibles ou protégés. Systèmes de reconnaissance des émotions.

Infrastructures critiques : Composants de sécurité dans la gestion et l'exploitation des infrastructures numériques critiques, du trafic routier et de la fourniture d'eau, de gaz, de chauffage et d'électricité.

Enseignement et formation professionnelle : Les systèmes d'IA déterminant l'accès, l'admission ou l'affectation dans les établissements d'enseignement et de formation professionnelle à tous les niveaux. Évaluer les résultats de la formation, y compris ceux utilisés pour orienter le processus d'apprentissage de l'étudiant. Évaluer le niveau d'éducation approprié pour un individu. Contrôler et détecter les comportements interdits des étudiants pendant les tests.

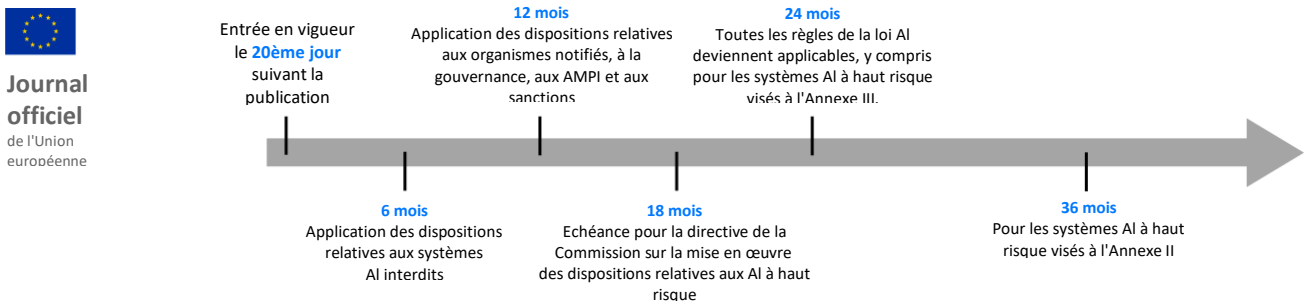
Emploi, gestion des salariés et accès au travail indépendant : Systèmes d'IA utilisés pour le recrutement ou la sélection, en particulier les annonces d'emploi ciblées, l'analyse et le filtrage des candidatures, et l'évaluation des candidats. Promotion et résiliation des contrats, attribution des tâches en fonction des traits de personnalité ou des caractéristiques et du comportement, suivi et évaluation des performances.

Accès à et jouissance des services privés et publics essentiels : Les systèmes d'IA utilisés par les autorités publiques pour évaluer l'éligibilité aux prestations et aux services, y compris leur attribution, leur réduction, leur révocation ou leur recouvrement. Évaluation de la solvabilité, sauf en cas de détection d'une fraude financière.Évaluer et classer les appels d'urgence, y compris la répartition des priorités de la police, des pompiers, de l'aide médicale et des services de tri des patients urgents. Évaluation des risques et tarification de l'assurance-maladie et de l'assurance-vie.

Application de la loi : Systèmes d'IA utilisés pour évaluer le risque qu'une personne soit victime d'un crime. Polygraphes : évaluation de la fiabilité des preuves dans le cadre d'enquêtes ou de poursuites pénales. L'évaluation du risque de délinquance ou de récidive d'un individu ne repose pas uniquement sur le profilage ou l'évaluation des traits de personnalité ou du comportement criminel antérieur. Le profilage lors de détections, d'enquêtes ou de poursuites pénales.

Gestion de l'immigration, de l'asile et des contrôles aux frontières : Polygraphes. Évaluations des migrations irrégulières ou des risques sanitaires. Examen des demandes d'asile, de visa et de permis de séjour, ainsi que des plaintes liées à l'éligibilité. Détecter, reconnaître ou identifier des personnes, à l'exception de la vérification des documents de voyage.

Administration de la justice et des processus démocratiques : Systèmes d'IA utilisés pour la recherche et l'interprétation des faits et l'application de la loi à des faits concrets ou utilisés dans la résolution alternative des litiges. Influencer les résultats des élections et des référendums ou le comportement des électeurs, à l'exclusion des résultats qui n'interagissent pas directement avec les personnes, tels que les outils utilisés pour organiser, optimiser et structurer les campagnes politiques.



[5] *Source : https://www.engage.hoganlovells.com/knowledgeservices/news/the-eu-ai-act-an-impact-analysis-part-1*

4. Existe-t-il des dispositions de la loi européenne sur l'IA qui rendraient impossible ou trop risquée l'utilisation de solutions d'IA conversationnelle fournies par Spitch ?

En bref, il n'y en a pas.

Certaines dispositions imposeraient des exigences supplémentaires au fournisseur des systèmes d'IA, c'est-à-dire Spitch, mais pas aux clients. Spitch veille à ce que toutes les solutions utilisant la biométrie vocale en temps réel, l'analyse des sentiments et la détection des émotions, celles faisant référence au GPT-4 ou à d'autres modèles d'IA à usage général pour la synthèse et la catégorisation, ainsi que les systèmes classés dans la loi sur l'IA de l'UE comme étant « à risque limité "», tels que les assistants virtuels (chatbots d'IA), soient entièrement conformes dans les délais prescrits.

5. Quelles sont les nouvelles obligations des clients déployant les systèmes d'IA fournis par Spitch en vertu de la loi européenne sur l'IA ?

Conformément à l'article 52 de la loi, les fournisseurs déployant un système de reconnaissance des émotions ou d'un système de catégorisation biométrique doivent informer du fonctionnement du système les personnes physiques qui y sont exposées et traité les données à caractère personnel conformément au règlement (UE) 2016/679, au règlement (UE) 2016/1725 et à la directive (UE) 2016/280, selon le cas.

En outre, les fournisseurs déployant un système d'IA qui génère ou manipule un texte publié dans le but d'informer le public sur des questions d'intérêt public doivent indiquer que le texte a été généré ou manipulé artificiellement.

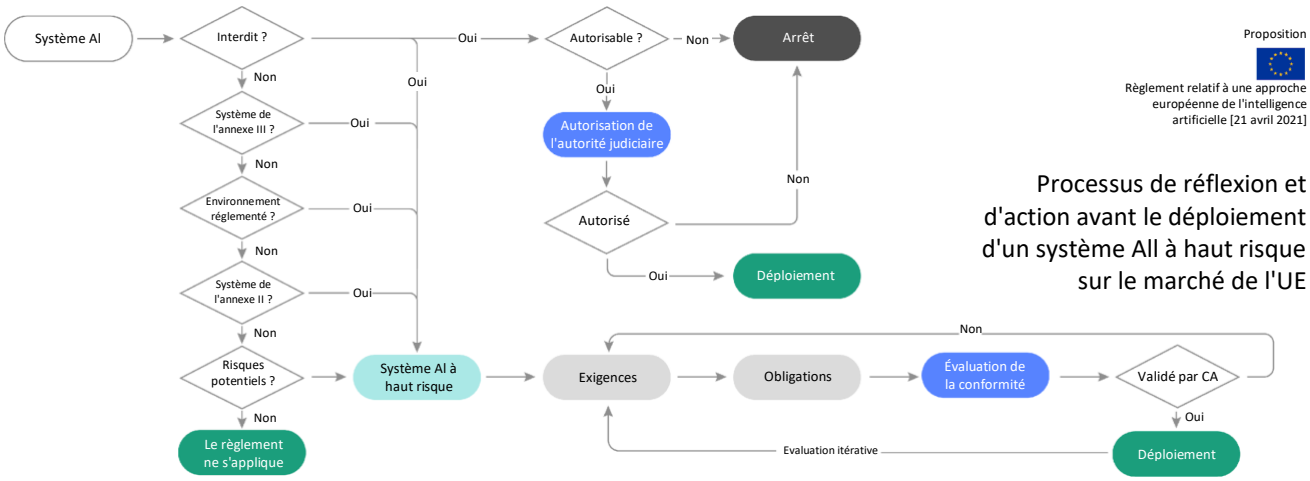
Les informations visées ci-dessus doivent être fournies aux personnes physiques concernées de manière claire et distincte au plus tard au moment de la première interaction ou exposition. Les informations doivent respecter les exigences applicables en matière d'accessibilité.

La manière de mettre en œuvre les exigences peut varier en fonction de la solution Spitch. Notre équipe de services professionnels vous assistera pendant la mise en œuvre afin de vous assurer que vous êtes en conformité avec les dernières réglementations.

6. L'extraction de données de performance sur le lieu de travail à des fins de gestion de la qualité est-elle autorisée par la loi ? Quelles actions garantissent la légalité ?

L'extraction de ces données par des systèmes d'IA est autorisée à moins qu'elle ne soit utilisée pour prendre des décisions automatisées sans la participation d'êtres humains, déduire des émotions sur le lieu de travail, déduire des attributs sensibles (race, opinions politiques, appartenance à un syndicat, croyances religieuses ou philosophiques, vie sexuelle ou orientation sexuelle) ou procéder à une évaluation sociale, c'est-à-dire évaluer ou classer des individus ou des groupes sur la base d'un comportement social ou de traits personnels, ce qui entraîne un traitement préjudiciable ou défavorable de ces personnes.

Afin de déterminer si les dispositions du nouveau règlement s'appliquent à votre cas particulier, ou si l'un des systèmes d'IA est susceptible de relever de la catégorie « à haut risque », conformément à la loi européenne sur l'IA, il est recommandé de suivre la procédure[6] décisionnelle suivante pour garantir la conformité :



Proposition

Règlement relatif à une approche européenne de l'intelligence artificielle [21 avril 2021]

Processus de réflexion et d'action avant le déploiement d'un système AII à haut risque sur le marché de l'UE

6 *Source : https://futurium.ec.europa.eu/en/european-ai-alliance/open-discussion/more-visual-guide-proposed-eu-artificial-intelligence-act?language=da*