

## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

"Duolingo-for-Schools est un tableau de bord intégré au compte Duolingo de l'enseignant qui lui permet de créer des classes et des devoirs, et de suivre l'activité des élèves. Nous sommes ravis de travailler avec les éducateurs pour apporter l'application d'apprentissage des langues numéro un au monde dans les salles de classe, avec des fonctionnalités conçues pour maximiser l'efficacité des enseignants et l'apprentissage des élèves".

Voyons comment cela fonctionne !

🚩 Pour Duolingo-for-Schools, nous utilisons les deux éléments donnés sur Duolingo et Duolingo-for-Schools car le second consiste en tableau de bord pour l'enseignant ajouté sur les fonctionnalités classiques de Duolingo.

#### Note

Il faut tenir compte du fait que les éléments donnés dans ce modèle sont pertinents pour un scénario d'utilisation spécifique de Duolingo-for-Schools où :

1. Les enseignants créent des comptes pour les élèves
2. sans éléments d'identification personnelle (nom, prénom, adresse électronique).

Dans d'autres scénarios possibles (par exemple, un enseignant utilise des comptes d'élèves déjà existants (ID/email), certaines réponses données dans ce modèle ne seront pas pertinentes.

## Groupe 1

*Ce groupe a pour but de clarifier l'objectif et les usages des outils d'IA en éducation et d'aider les utilisateurs à comprendre leur rôle dans l'utilisation de ces ressources Éducatives en Intelligence Artificielle (AIER).*

## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

- **Couche 1 : Scénario d'utilisation**

- À qui s'adresse la ressource ?



- L'IA orientée vers l'apprentissage,
- L'IA orientée vers l'instructeur
- L'IA axée sur les systèmes institutionnels.

- Quels sont les objectifs de la ressource ?



**Duolingo-for-Schools** "Avec Duolingo-for-Schools, les enseignants peuvent créer une (ou plusieurs !) classe(s) et ajouter des étudiants ou permettre à leurs étudiants de s'inscrire eux-mêmes. Les enseignants peuvent gérer l'accès des élèves au contenu des cours et suivre leurs progrès tout au long de leur parcours d'apprentissage en suivant l'activité des élèves, y compris les XP gagnés et le temps passé à travailler sur un cours, ainsi que les activités individuelles des élèves, telles que les leçons terminées et les scores obtenus".

<https://duolingoschools.zendesk.com/hc/en-us/articles/6830454446093-What-is-Duolingo-for-Schools->

<https://blog.duolingo.com/duolingo-for-Schools/>

**Duolingo** : Duolingo propose des exercices et des parcours d'apprentissage adaptés pour améliorer la maîtrise de nombreuses langues. Il peut être utilisé pour apprendre le français.

La couche "**scénario d'utilisation**" décrit les utilisations et applications de l'AIER. Ce qui est souvent la première information que les enseignants souhaitent connaître. Les AIER peuvent être classés en 3 types :

- ◆ L'IA orientée vers l'apprentissage,
- ◆ Instructions d'IA orientées vers l'enseignement et
- ◆ L'IA orientée vers les systèmes institutionnels.

Il existe de nombreux scénarios d'utilisation différents en fonction du type d'AIER. Par exemple, une AIER orienté vers l'apprenant pourrait être un système informatique intelligent conçu pour aider les étudiants à apprendre un cours, tandis qu'une AIER orientée vers l'enseignant pourrait être un


## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

système qui génère automatiquement des questions et des tests.

- **Couche 2 : Niveau de décision**

- Quelle est le niveau décision de l'IA ?

 **Duolingo-for-Schools** : Adapter le parcours d'apprentissage des élèves ou d'un groupe d'élèves lorsque les enseignants leur assignent des tâches via un tableau de bord.

**Duolingo** : Les algorithmes d'IA aident à fournir, par exemple, le bon exercice au bon moment tout en prenant en compte de nombreux concepts éducatifs tels que :


- ◆ La **courbe d'oubli** de Hermann Ebbinghaus

(<https://blog.duolingo.com/how-we-learn-how-you-learn/>).

- ◆ La **frontière de l'apprentissage** (pousser votre apprentissage au niveau suivant, mais seulement lorsque vous êtes prêt à le faire - développement proximal),

<https://blog.duolingo.com/keeping-you-at-the-frontier-of-learning-with-adaptive-lessons/>

- Quel est le degré de sensibilité de la décision proposée par l'IA ??

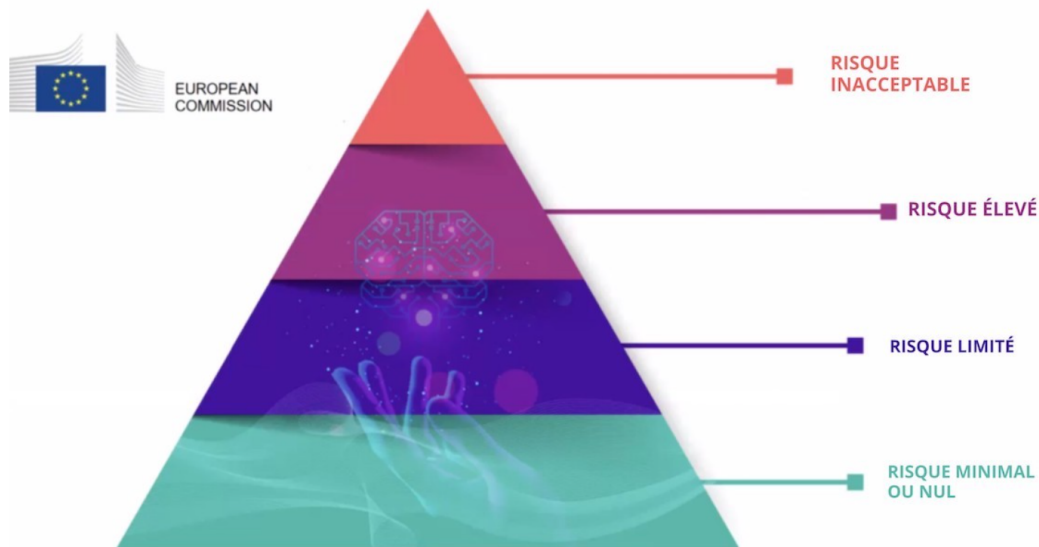
 **Duolingo-for-Schools** : Dans la mesure où les éléments du tableau de bord (compétences/niveaux/leçons) ne sont pas directement transformés en notes ou que duolingo n'est pas utilisé pour l'évaluation sommative, le risque peut être considéré comme faible.

*Rappel :*

Cette considération doit être envisagée dans le contexte du **cadre réglementaire sur l'IA** proposé par l'UE, dans lequel tous les systèmes d'enseignement de l'IA sont classés par défaut comme présentant un risque élevé.

## Modèle de caractérisation de l'IA


### Étude de cas : Duolingo-for-Schools



*Les 4 niveaux risques liés à l'intelligence artificielle utilisés dans la proposition de cadre réglementaire l'IA de la Commission européenne.*


La **proposition de cadre réglementaire sur l'intelligence artificielle présentée par l'UE** (<sup>1</sup>) classe les systèmes d'intelligence artificielle utilisés dans les domaines suivants comme présentant un risque élevé : "la formation initiale ou professionnelle, qui peut déterminer l'accès à l'éducation et le parcours professionnel d'une personne (par exemple, la notation d'examens)".

#### ■ Qui porte la responsabilité de la décision ?

 **Duolingo-for-Schools** : La décision de confier une tâche ou une évaluation à un apprenant est prise par l'enseignant.

**Duolingo** : Les exercices et leur disposition sont proposés par l'AIER.

#### ■ A quel point l'enseignant est-il impliqué dans la boucle de décision ?

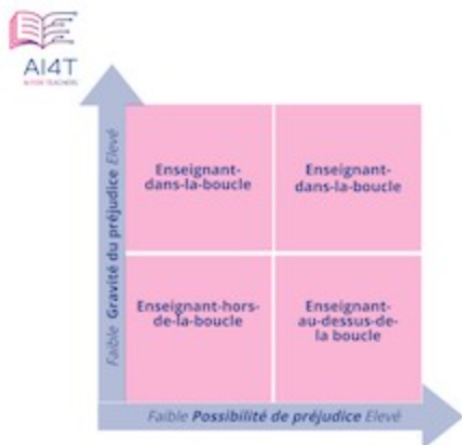
 **Duolingo-for-Schools** : Les enseignants peuvent proposer deux types de missions aux apprenants, la première consiste à collecter des points d'expérience (ou XP) et la seconde à améliorer une compétence.

<sup>1</sup> <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> (consulté le 10/11/2022)

## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

**Duolingo :** Une fois les décisions prises par les enseignants, c'est duolingo for schools qui met en place les tâches détaillées pour les apprenants. Il s'agit d'une **automatisation des décisions**<sup>2</sup> et les enseignants, dans ce cas précis, sont hors de la boucle<sup>3</sup>.



#### **Enseignant-hors-de-la-boucle :**

Dans une situation où la probabilité et la gravité des dommages causés par une application éducative utilisée en dehors de l'école sont faibles, la surveillance de l'enseignant n'est pas nécessaire (fig. 1, en bas à gauche).

Le niveau de décision décrit la proportion du résultat final qui est attribuable à l'outil lorsqu'un enseignant utilise une AIER. Normalement, le niveau de décision dépend de ce que "produit" l'AIER. Prenons l'exemple d'une IA qui utilise des stratégies de notation pour évaluer automatiquement les résultats d'apprentissage des étudiants.

Le niveau de décision délégué à l'IA va dépendre de ce qui est fait de ces résultats d'évaluation automatique. Si les notes sont envoyées directement à l'étudiant ou si elles sont transmises à l'enseignant et que c'est ce dernier qui fige les résultats définitifs de l'évaluation, le niveau de décision délégué à l'IA n'est pas le même.

Le niveau de décision aide donc les utilisateurs à connaître leur droits, et leur responsabilité lorsqu'ils utilisent une AIER.

<sup>2</sup> cf. Mooc 4.1.1. Prendre des décisions avec l'IA

<sup>3</sup> cf. Mooc 4.1.3 Enseignant-dans-la- boucle

## Modèle de caractérisation de l'IA


### Étude de cas : Duolingo-for-Schools

#### Groupe 2

*Ce cours a pour but de clarifier les approches utilisées par l'AIER pour atteindre son objectif et d'aider les enseignants à comprendre son fonctionnement.*

- **Couche 3 : Technologie**

- Quelle est la question traitée par l'IA ?

 **Duolingo-for-Schools** : Il s'agit d'un **tableau de bord** intégré au compte Duolingo d'un enseignant qui lui permet de créer des classes et des devoirs, et de suivre l'activité des étudiants.

<https://blog.duolingo.com/duolingo-for-Schools/#:~:text=Duelingo%20pour%20les%20écoles%20est%20a,et%20pour%20suivre%20l'activité%20des%20étudiants>).

**Duolingo** : C'est une **application d'apprentissage des** langues qui revendique 500 millions d'utilisateurs dans le monde.

<https://www.duolingo.com/learn>

La "couche technologique" décrit les technologies d'IA utilisées dans un AIER. Étant donné qu'une AIER peut être : une plateforme éducative, une application autonome ou en ligne, etc., plusieurs technologies différentes peuvent être utilisées dans une seule ressource. L'objectif de cette couche est de préciser si des technologies basées sur l'IA **sont effectivement utilisées** et de **présenter aux utilisateurs le type de technologies d'IA utilisées**.

La technologie peut faire partie de la liste suivante :

- ◆ Génération automatique de contenu éducatif (cours, textes, etc.)
- ◆ Amélioration du contenu éducatif
- ◆ Aide de l'IA aux enseignants pour fournir un enseignement personnalisé à chaque élève
- ◆ Retour d'information rapide aux élèves
- ◆ Aide au suivi des élèves.
- ◆ L'évaluation automatique du comportement et des parcours d'apprentissage des élèves
- ◆ L'enregistrement automatique du processus d'apprentissage de l'élève.
- ◆ Amélioration ciblée basée sur l'analyse de l'apprentissage
- ◆ Chatbot entre les enseignants, les élèves, les parents ou les groupes concernés
- ◆ Autre technologie dédiée aux tâches éducatives.

## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

#### • Couche 4 : Algorithmes

##### ■ À quelle famille d'approches les algorithmes se rattachent-ils ?



- Systèmes basés sur la connaissance :
- Machine Learning : pour les deux exemples trouvés
- Apprentissage profond :

##### ■ Quels sont les algorithmes ?



#### **Duolingo :**

Ex1. Sleeping recovering Bandit = pour les notifications sur duolingo  
<https://research.duolingo.com/papers/yancey.kdd20.pdf>

Ex2. Formation bayésienne des connaissances = pour la traçabilité des connaissances (estimation de l'état actuel des connaissances d'un étudiant et prédiction de ses performances lors d'interactions futures, les étudiants peuvent recevoir des supports d'apprentissage personnalisés  
<https://sharedtask.duolingo.com/papers/chen.slam18.pdf>

La "couche algorithmes" décrit les algorithmes utilisés dans l'AIER.

Par rapport à la couche précédente, cette couche se concentre davantage sur le niveau scientifique que sur le niveau technique. Nous définissons la famille des algorithmes à partir de trois groupes communs : les systèmes basés sur la connaissance, l'apprentissage automatique et l'apprentissage profond. Nous sommes conscients du débat sur la question de savoir si l'apprentissage profond est, ou non, de l'apprentissage automatique. Cependant, nous les avons séparés ici car cela est plus facile à comprendre pour les enseignants. Pour chaque groupe, l'algorithme peut être choisi dans la liste suivante :

- ◆ **Systèmes basés sur la connaissance** : Systèmes basés sur des règles (ou systèmes experts), ontologie, réseaux sémantiques, etc.
- ◆ **Apprentissage automatique** : Regroupement, possibilité approximative, régression

## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

Analyse, représentation et réduction de la dimensionnalité, apprentissage actif, arbres de décision, etc.

- ◆ **Apprentissage profond** : Réseaux convolutifs (CNN), réseaux à mémoire longue à court terme (LSTM), réseaux neuronaux récurrents (RNN), réseaux antagonistes génératifs (GAN).


Attention : ici nous nous contentons d'identifier les algorithmes utilisés dans cette couche et de les présenter brièvement, sans les comparer ni les évaluer.

### Groupe 3

Cette classe a pour but de clarifier les données d'entrée utilisées par l'AIER et le traitement sécurisé de ces données. On s'intéresse à la manière dont une AIER démontre la transparence de ses mécanismes internes et d'aider les utilisateurs à avoir confiance dans la sécurité de son utilisation..

- **Couche 5 : Données personnelles**

- Quel type de données personnelles est utilisé ?

 **Duolingo** : Afin de proposer des exercices et des parcours d'apprentissage personnalisés aux apprenants, il est nécessaire que l'AIER suive les apprenants. Dans le cadre de Duolingo pour les écoles, il est possible de ne pas utiliser les **données personnelles** des élèves.

**Duolingo-for-Schools** : permet de créer des profils sans nom, prénom ou adresse e-mail :

- Les enseignants peuvent également créer des comptes pour leurs élèves.
- Il n'est pas nécessaire d'ajouter un nom, un prénom ou une adresse électronique.
- Exemple d'inscription d'un enseignant à duolingo pour les écoles :  
<https://drive.google.com/file/d/1a71TUmkdXjSr8fldU4RAe5BYdRE9UJpC/view?usp=sharing>

Note : Les enseignants doivent avoir un compte duolingo (création : nom d'utilisateur + adresse email)

<https://support.duolingo.com/hc/en-us/articles/204829870-How-do-I-create-an-account-and-or-login-#:~:text=To%20create%20a%20new%20account,your%20email>



## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

[l%2Fusername%20and%20password](#)

Dans ce cas, la politique générale en matière de données de duolingo s'applique : <https://www.duolingo.com/privacy>

👉 Tout utilisateur peut accéder à ses données personnelles et demander à ce qu'elles soient supprimées.

La couche des données personnelles énumère les données personnelles utilisées dans l'AIER. De nos jours, la sécurité de l'information est de plus en plus importante. Outre les données explicitement proposées pour la saisie dans l'AIER, des sources externes, telles que les comptes sociaux partagés, peuvent collecter des informations personnelles supplémentaires. La création de cette couche est nécessaire pour permettre aux utilisateurs de clarifier ces informations.

#### • Couche 6 : Sécurité des données

- Les données personnelles sont-elles anonymes ? Si oui, comment cela fonctionne-t-il ?

🖋️ **Duolingo-for-Schools** : Dans le scénario d'utilisation choisi pour cet exemple, les élèves sont inscrits par leur enseignant sans utiliser de nom, de prénom ou d'adresse électronique. Les données sont donc anonymes.

- Quelles sont les possibilités d'audit de la ressource par des personnes extérieures ?

🖋️ **Duolingo-for-Schools & Duolingo** : Au vu des informations que nous avons pu trouver au cours de nos recherches sur le sujet, il n'y a pas de possibilité pour une personne extérieure d'auditer les ressources.

Le [programme de protection de la vie privée Common Sense](#) a attribué à Duolingo une note globale de 63 % en 2021.

<https://privacy.commonsense.org/privacy-report/duolingo>

Le programme Common Sense Privacy évalue les **politiques de protection de la vie privée afin que les** parents et les enseignants puissent faire des choix

## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

judicieux concernant les outils d'apprentissage qu'ils utilisent avec leurs enfants et leurs élèves, et que les écoles et les districts puissent participer à l'amélioration de la technologie utilisée dans les salles de classe de la maternelle à la 12e année).

La couche "sécurité des données" précise si les **visiteurs non-utilisateurs** et les **visiteurs externes** de l'AIER peuvent accéder aux données et comment les données doivent être traitées en toute sécurité. Cette couche est liée à la précédente. Différentes données personnelles doivent être traitées de différentes manières, telles que l'anonymat, le cryptage et le refus d'accès.

◆ **L'anonymisation** signifie que le destinataire de l'information n'est absolument pas en mesure d'identifier directement ou brièvement la personne concernée. Il existe cinq opérations courantes d'anonymisation des données : la généralisation, la suppression, l'anonymisation, la permutation et la perturbation. ◆ **Le cryptage**, également connu sous le nom de **pseudonymisation** signifie que le récepteur de l'information ne peut pas identifier directement la personne concernée, mais que l'information peut être désanonymisée par le gestionnaire des données après identification croisée avec d'autres informations. ◆ **Le refus d'accès** est un moyen simple et direct de conserver l'information uniquement dans le gestionnaire de l'information, sans en donner l'accès à d'autres personnes.

Dans cette couche, il est également précisé si les données seront utilisées par des acteurs externes. De nos jours, de nombreux logiciels partagent des données avec leurs propres partenaires, et si c'est le cas, cela peut également entraîner différents problèmes concernant la sécurité des données.

#### ● **Couche 7 : Transparence**

- Existe-t-il un mécanisme d'explication des décisions ? Si oui, comment fonctionne-t-il ?

## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

#### **Duolingo-for-Schools / Duolingo**

Il n'existe aucun mécanisme permettant d'expliquer les nombreuses microdécisions prises par l'AIER pour fournir des tâches adaptées à chaque étape du parcours d'apprentissage dans chaque langue proposée dans Duolingo.

D'ailleurs, de nombreux articles scientifiques sont publiés par les équipes de Duolingo dans des publications scientifiques pour expliquer le fonctionnement de certains éléments du modèle : <https://research.duolingo.com/>

Pour compléter cette couche sur la transparence, même s'il ne s'agit pas de la transparence de l'IA, Duolingo offre aux apprenants la possibilité de passer le "Duolingo English Test".

Comme l'apprenant peut utiliser le résultat de ce test pour faire valoir ses compétences en anglais afin de s'inscrire dans les établissements d'enseignement supérieur, de nombreuses informations sont publiées à ce sujet :

- **Le manuel technique du test d'anglais de Duolingo** : il donne un aperçu de la conception, du développement, de l'administration et de la notation du test d'anglais de Duolingo. En outre, le manuel technique fait état de la validité, de la fiabilité et de l'équité, ainsi que des caractéristiques démographiques des participants et des caractéristiques statistiques du test.  
<https://duolingo-papers.s3.amazonaws.com/other/det-technical-manual-current.pdf>
- Une évaluation externe de la fiabilité du test, notamment en termes de prévention de la fraude, est effectuée.

La "couche transparence" clarifie le niveau de compréhension du mécanisme interne de l'AIER. Par exemple, si les tâches éducatives consistent à évaluer les étudiants en fonction de leurs comportements d'apprentissage, le mécanisme interne de l'AIER concerne les comportements d'apprentissage qui ont été sélectionnés et les paramètres qui ont été utilisés pour générer le rapport final. Si le rapport d'évaluation final est remis à l'enseignant avec ces paramètres ou si le site web d'introduction de l'AIER explique le mécanisme interne de cet AIER, cela aidera les enseignants à comprendre les possibilités,

## Modèle de caractérisation de l'IA

### Étude de cas : Duolingo-for-Schools

les limites et les risques de cet AIER dans le domaine de l'éducation.

En outre, la transparence pour les enseignants des résultats intermédiaires générés par les AIER dans les tâches éducatives mérite d'être déclarée à ce niveau. Par exemple, un AIER fournissant un enseignement personnalisé peut avoir à générer d'abord un profil de chaque étudiant qu'il utilise pour recommander différents contenus éducatifs dans le processus d'accomplissement de cette tâche. Si le profil est disponible pour les enseignants, cela aidera les utilisateurs à comprendre l'AIER et à lui faire confiance.