

Modèle de caractérisation de l'IA

Étude de cas : Duolingo pour les écoles

Cette étude de cas est réalisée en utilisant le modèle de format prêt à l'emploi qui a été conçu par LORIA¹ et Inria dans le cadre du projet AI4T. Elle est basée sur le rapport LORIA sur le modèle d'analyse des caractéristiques liées à l'IA dans les ressources d'apprentissage².

"Duolingo for Schools est 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 élèves. Nous sommes ravis de collaborer avec les éducateurs afin d'introduire l'application d'apprentissage des langues numéro un au monde dans les salles de classe, avec des fonctionnalités conçues pour optimiser l'efficacité des enseignants et l'apprentissage des élèves".

Voyons comment cela fonctionne !

► Pour Duolingo pour les écoles, nous utilisons les deux éléments donnés sur Duolingo et Duolingo pour l'école, car le second est une couche (tableau de bord pour l'enseignant) au-dessus de Duolingo.

Note

Veuillez considérer que les éléments donnés dans ce modèle sont pertinents pour un scénario d'utilisation spécifique de Duolingo pour l'école 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.

¹ Le Loria (Laboratoire Lorrain de Recherche en Informatique et ses Applications) fait partie d'une unité de recherche (UMR 7503), commune au [CNRS](#), à l'[Université de Lorraine](#) et à l'INRIA. Il est membre du consortium AI4T et apporte son expertise en IA dans l'éducation (et sur l'analyse de l'apprentissage) au projet Erasmus+ AI4T.

² [Report on template for analyzing AI-related features in learning resources](#) - Jiajun PAN, Azim ROUSSANALY, Anne BOYER - AI4T Erasmus+ project, 2022.

Modèle de caractérisation de l'IA

Étude de cas : Duolingo pour les écoles

Groupe 1 : Scénario d'utilisation et niveau de décision

Ce cours a pour but de clarifier l'objectif et les utilisateurs d'une AIER et d'aider les utilisateurs à comprendre quel est leur rôle dans l'utilisation des Ressources Educatives en Intelligence Artificielle - AIER.

- Couche 1 : Scénario d'utilisation

- À qui s'adresse cette ressource ?



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

- Quels sont les objectifs de la ressource ?



Duolingo pour les écoles "Avec Duolingo pour les écoles, les enseignants peuvent créer une (ou plusieurs !) classe et soit ajouter des étudiants, soit autoriser leurs étudiants à se joindre eux-mêmes. Les enseignants peuvent gérer l'accès des élèves au contenu des cours et suivre les progrès des élèves tout au long de leur parcours d'apprentissage en suivant l'activité des élèves, notamment les XP obtenus 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-pour-les-écoles>

<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.

🗣️ Duolingo pour les écoles propose des formations dans 4 des 5 langues du Consortium AI4T

: DE, EN, FR, IT, DE

La couche des scénarios d'utilisation décrit les utilisateurs et les applications de l'AIER, ce qui est souvent la première information que les enseignants veulent connaître. Les AIER peuvent être classés en 3 types :

- ◆ L'IA orientée vers l'apprenant,
- ◆ IA orientée vers l'instructeur et
- ◆ L'IA axée sur les systèmes institutionnels.


Il existe de nombreux scénarios d'utilisation différents en fonction du type d'AIER. Par exemple, une IAER orientée vers l'apprenant pourrait être un système informatique intelligent conçu pour aider les étudiants à apprendre un cours, tandis qu'une IAER orientée vers l'enseignant pourrait être un système qui génère automatiquement des questions et des tests pour des didacticiels préparés par l'enseignant pour un cours particulier.

Modèle de caractérisation de l'IA

Étude de cas : Duolingo pour les écoles

- Couche 2 : Niveau de décision

■ Quelle est la décision pour l'IA ?

 **Duolingo pour les écoles** : Adaptez le parcours d'apprentissage des élèves ou des groupes d'élèves lorsque les enseignants leur assignent des tâches via un tableau de bord.


Duolingo : Les algorithmes d'IA permettent de proposer, par exemple, le bon exercice au bon moment.

tout en considérant 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 supérieur, mais seulement lorsque vous êtes prêt à le faire) <https://blog.duolingo.com/keeping-you-at-the-frontier-of-learning-with-adaptive-leçons/>

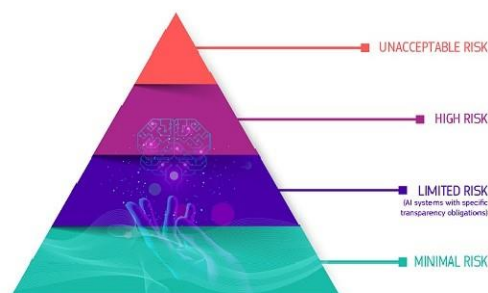
■ A quel point cette décision est sensible ?

 **Duolingo pour les écoles** : 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 une évaluation sommative, le risque peut être considéré comme faible.

Rappel :

Cette considération doit être considérée dans le contexte du **cadre réglementaire sur l'IA**, proposé par l'UE, où tous les systèmes d'enseignement de l'IA sont classés à haut risque par défaut.

A risk-based approach



The Regulatory Framework defines 4 levels of risk in AI:

- Unacceptable risk
- High risk
- Limited risk
- Minimal or no risk


La **proposition de cadre réglementaire sur l'intelligence artificielle mise en place par l'UE**

classe comme à haut risque les systèmes d'IA utilisés dans : " 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) ".

Modèle de caractérisation de l'IA


Étude de cas : Duolingo pour les écoles

■ Qui porte la responsabilité de la décision ?

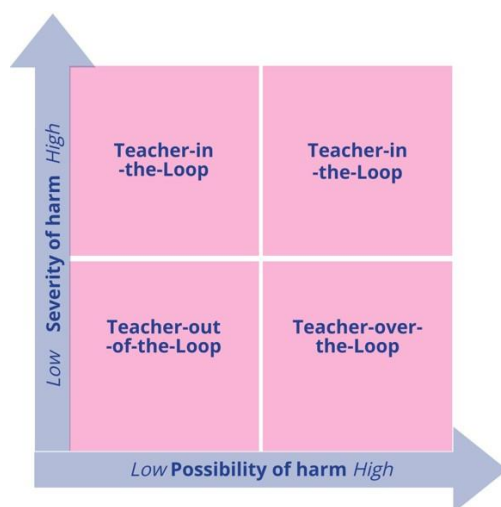
 **Duolingo pour l'école** : La décision de donner une tâche ou une évaluation à un apprenant est prise par l'enseignant.

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

■ A quel point l'enseignant se situe-t-il dans la boucle de décision ?

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

Duolingo : Une fois que les décisions sont prises par les enseignants, c'est Duolingo pour les écoles qui met en place les tâches détaillées réelles pour les apprenants. C'est l'**automatisation des décisions** et les enseignants, dans ce cas précis, sont hors circuit.



- **Enseignant hors de la boucle** : Dans une situation où la probabilité et la gravité des dommages causés par, par exemple, une application éducative utilisée en dehors de l'école sont faibles, la surveillance de l'éducateur n'est pas nécessaire (figure 1, en bas à gauche).

La couche du niveau de décision décrit la proportion du résultat final qui est attribuable à la faculté lorsque l'enseignant utilise l'AIER. Normalement, le niveau de décision dépend de la sortie de l'AIER. Par exemple, pour une IA qui utilise des stratégies de notation pour évaluer automatiquement les résultats d'apprentissage des étudiants, le niveau de décision dépend de la question de savoir si les notes sont envoyées directement à l'étudiant ou à l'enseignant et qui fait les résultats d'évaluation ultérieurs. Le niveau de décision aide les utilisateurs à connaître leur droit et leur rôle dans l'utilisation d'une IAER.

Modèle de caractérisation de l'IA


Étude de cas : Duolingo pour les écoles

Groupe 2 : Technologie et algorithme

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 abordée par l'IA ?

 **Duolingo pour les écoles** : 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 élèves. <https://blog.duolingo.com/duolingo-for-schools/>
Duolingo : Il s'agit d'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 le rapport. Étant donné qu'une IAER 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 clarifier si des technologies basées sur l'IA sont réellement utilisées et de présenter aux utilisateurs le type de technologies IA utilisées.

La technologie pourrait être l'une des listes suivantes :

- ◆ Génération automatique de contenus pédagogiques (Cours, textes, etc.)
- ◆ Amélioration des contenus pédagogiques
- ◆ Assistance de l'IA aux enseignants pour fournir un enseignement personnalisé à chaque élève
- ◆ Retour rapide aux élèves
- ◆ Aide au suivi des élèves.
- ◆ L'automatique évalue le comportement et les parcours d'apprentissage des élèves
- ◆ L'automatique enregistre le processus d'apprentissage de l'élève.
- ◆ Amélioration ciblée basée sur l'analyse de l'apprentissage
- ◆ Chatbot entre les enseignants, les élèves, les parents et les groupes concernés
- ◆ Autre technologie pour les tâches éducatives.

Modèle de caractérisation de l'IA

Étude de cas : Duolingo pour les écoles

- Couche 4 : Algorithmes

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



- Systèmes basés sur la connaissance :
- **Machine Learning : pour les 2 exemples trouvés**
- Apprentissage profond :

■ Quels sont les algorithmes ?



Duolingo :

Ex1. Bandit récupérateur de sommeil pour les notifications sur Duolingo

<https://research.duolingo.com/papers/yancey.kdd20.pdf>

Ex2. Formation bayésienne des connaissances pour le traçage des connaissances (estimation de l'état actuel des connaissances d'un étudiant et prédiction de ses performances lors de futures interactions, les étudiants peuvent recevoir du matériel d'apprentissage personnalisé) <https://sharedtask.duolingo.com/papers/chen.slam18.pdf>

La couche Algorithme 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 déclarerons la famille des algorithmes à partir de trois groupes communs : les systèmes basés sur la connaissance, l'apprentissage automatique et l'apprentissage profond. Notez que nous sommes conscients du débat sur la question de savoir si l'apprentissage profond est de l'apprentissage automatique ou non. Cependant, nous les avons séparés ici car cela est plus facile à comprendre pour les enseignants. Pour chaque groupe, l'algorithme pourrait être choisi dans la liste suivante :

- ◆ Systèmes basés sur la connaissance : Systèmes à base de règles (ou systèmes experts), ontologie, réseaux sémantiques, etc.
- ◆ Apprentissage automatique : Regroupement, possibilité d'approximation, régression
Analyse, représentation et réduction de la dimensionnalité, apprentissage actif, arbres de décision, etc.
- ◆ Apprentissage profond : Réseaux convolutifs (CNN), réseaux à mémoire à long terme (LSTM), réseaux neuronaux récurrents (RNN), réseaux adversariens génératifs (GAN).

Veuillez noter que nous nous contenterons d'identifier les algorithmes utilisés dans cette couche et de les présenter brièvement, sans comparer ni évaluer les algorithmes utilisés.

Modèle de caractérisation de l'IA


Étude de cas : Duolingo pour les écoles

Groupe 3 : Données personnelles - Sécurité des données - Transparence

Cette classe a pour but de clarifier les données d'entrée utilisées par l'AIER et le traitement sécurisé de ces données, de faire comprendre que l'AIER démontre la transparence des mécanismes internes et d'aider les utilisateurs à avoir confiance dans la sécurité de l'AIER.

- Couche 5 : Données personnelles

■ Quel type de données personnelles est utilisé ?

 **Duolingo** : Afin de proposer aux apprenants des exercices et des parcours d'apprentissage personnalisés, il est nécessaire que l'AIER suive les apprenants.

Dans le contexte de Duolingo pour les écoles, c'est possible sans **données personnelles** des élèves.

Duolingo pour les écoles : permet de créer des profils sans nom, prénom ou adresse électronique :

- Les enseignants peuvent également créer des comptes pour leurs élèves.
- Il n'est pas nécessaire d'ajouter le nom, le prénom ou l'adresse électronique.
- Exemple d'enregistrement effectué par un enseignant dans Duolingo pour les écoles : https://drive.google.com/file/d/1a71TUmkdXjSr8fldU4RAe5BYdRE9UJpC/view?usp=s_haring

Note : Les enseignants doivent avoir un compte Duolingo (création : nom d'utilisateur + adresse e-mail) <https://support.duolingo.com/hc/en-us/articles/204829870-How-do-I-create-an-account-and-or-login>

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 leur suppression.


La couche de données personnelles énumère les données personnelles utilisées dans l'AIER. De nos jours, la sécurité des informations est de plus en plus valorisée. En plus des 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.

Modèle de caractérisation de l'IA


Étude de cas : Duolingo pour les écoles

- Couche 6 : Sécurité des données

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

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

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

 **Duolingo pour les écoles et Duolingo** : Dans la mesure où des informations ont été trouvées sur le sujet, il n'y a pas de possibilité pour une personne extérieure de vérifier 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 judicieux concernant les outils d'apprentissage qu'ils utilisent avec leurs enfants et leurs élèves, et afin que les écoles et les

La couche de sécurité des données spécifie si les visiteurs non-utilisateurs et les visiteurs externes de l'AIER peuvent accéder aux données et comment les données doivent être traitées de manière sécurisée. Cette couche est liée à la précédente. Des données personnelles différentes 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 est totalement incapable d'identifier directement ou brièvement la personne concernée. Il existe cinq opérations courantes d'anonymisation des données : généralisation, suppression, anatomisation, permutation et 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 les informations uniquement dans le gestionnaire d'informations, sans en donner l'accès aux autres.

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

Modèle de caractérisation de l'IA

Étude de cas : Duolingo pour les écoles

- Couche 7 : Transparence
- Existe-t-il un mécanisme d'explication des décisions ? Si oui, comment fonctionne-t-il ?

Duolingo pour l'école / Duolingo

Il n'existe aucun mécanisme permettant d'expliquer les nombreuses micro-décisions prises par l'AIER pour fournir des tâches adaptées à chaque étape du parcours d'apprentissage dans chaque langue proposée dans Duolingo. Au demeurant, 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 si elle ne concerne pas la transparence de l'IA, Duolingo offre aux apprenants la possibilité de passer le "Duolingo English Test".

Comme les apprenants peuvent utiliser le résultat de ce test pour faire valoir leurs compétences en anglais et s'inscrire dans des établissements d'enseignement supérieur, de nombreuses informations sont publiées à ce sujet :

- **Le manuel technique du test d'anglais Duolingo** : il donne un aperçu de la conception, du développement, de l'administration et de la notation du test d'anglais Duolingo. En outre, le manuel technique fait état des preuves de validité, de fiabilité et d'équité, ainsi que des données démographiques des testeurs et des caractéristiques statistiques du test. <https://duolingo-papers.s3.amazonaws.com/other/det-technical-manual-current.pdf>

La couche de transparence clarifie le niveau de compréhension du mécanisme interne de l'AIER. Par exemple, dans les tâches éducatives visant à évaluer les élèves à travers leurs comportements d'apprentissage, le mécanisme interne de cet AIER concerne les comportements d'apprentissage sélectionnés et les paramètres utilisés pour générer le rapport final. Si le rapport d'évaluation final est remis à l'enseignant en indiquant 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, les limites et les risques de cet AIER dans l'éducation.

En plus de cela, la transparence pour les enseignants des résultats intermédiaires générés par les IAER dans les tâches éducatives mérite d'être déclarée à ce niveau. Par exemple, une IAER fournissant un enseignement personnalisé peut avoir à générer d'abord un profil de chaque élève qu'elle 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 et à faire confiance à cette IAER.