

Des notions à distinguer

Systemes d'IA et modes d'IA à usage g n ral

Par Yannick Caballero Cuevas le 4 novembre 2024

Depuis le 1^{er} ao t 2024, le r glement europ en sur l'intelligence artificielle (RIA) est en vigueur (cf. cdbf.ch/1359/). Il s'applique aussi bien aux systemes d'intelligence artificielle (SIA) qu'aux modes d'IA   usage g n ral ([art. 2 RIA](#)). Ce commentaire se concentre sur les d finitions de ces deux notions cl s dans l'application du RIA et essaie de mettre en  vidence leurs caract ristiques et sp cificit s.

A. La notion de SIA

Selon l'[art. 3 ch. 1 RIA](#), un SIA est « un **systeme automatis ** qui est con u pour fonctionner   **diff rents niveaux d'autonomie** et peut faire preuve d'une **capacit  d'adaptation** apr s son d ploiement, et qui, pour des objectifs explicites ou implicites, **d duit,   partir des entr es qu'il re oit**, la mani re de g n rer des sorties telles que des pr dictions, du contenu, des recommandations ou des d cisions qui peuvent influencer les environnements physiques ou virtuels ». Les diff rentes notions mises en  vidence m ritent quelques pr cisions.

Tout d'abord, pour d finir un SIA, la version fran aise utilise la formulation « systeme automatis  » alors que les versions anglaise et allemande utilisent « *machine-based system* » et « *maschinengest tztes System* ». Ces formulations indiquent simplement qu'un SIA est un systeme fonctionnant   l'aide d'une machine, comme un ordinateur, une voiture autonome ou un t l phone portable.

Ensuite, il faut que le SIA d duise   partir des entr es (*input*) la mani re de g n rer des sorties (*output*) qui peuvent prendre la forme d'une pr diction, d'un contenu (par ex. texte, image, vid o, etc.), d'une recommandation ou encore d'une d cision. La liste d'exemples n'est pas exhaustive. La sortie g n r e par le SIA doit influencer l'environnement physique ou virtuel.   notre sens, cette pr cision doit  tre lue en relation aux buts du RIA qui visent notamment   garantir un niveau  lev  de protection de la sant , de la s curit  et des droits fondamentaux (consid. 1 RIA). De nombreux algorithmes d terministes g n rent aussi des sorties   partir d'entr es, de sorte que cette caract ristique ne permet pas de distinguer un SIA d'un systeme de type *if-then*.

Le SIA doit faire preuve d'un certain niveau d'autonomie, ce qui signifie qu'il b n ficie d'une ind pendance dans son action (consid. 12 RIA). Le RIA ne pr cise toutefois pas le degr  d'autonomie qu'un systeme doit avoir pour  tre qualifi  de SIA.   notre avis, ceux con us en se fondant sur une approche de *machine learning* pr sentent un degr  d'autonomie suffisant.

Ce critère d'autonomie d'un SIA permet de distinguer un SIA d'un applicatif informatique ordinaire. Par exemple, les systèmes fondés sur une pure logique déterministe sont exclus de la définition de SIA. Il est utile de relever qu'une supervision humaine d'un SIA ne permet pas de conclure que le système est dépourvu d'autonomie, car le RIA exige explicitement un contrôle humain pour les SIA à haut risque (cf. art. 14 RIA).

Enfin, la définition inclut une dernière caractéristique, précisant que les SIA peuvent s'adapter après leur déploiement. L'emploi du verbe « pouvoir » montre que ce critère n'est pas essentiel pour qualifier un système de SIA, contrairement au critère du degré d'autonomie.

B. La notion de modèle d'IA à usage général

Selon l'[art. 3 ch. 63 RIA](#), un modèle d'IA à usage général est « un modèle d'IA, y compris lorsque ce modèle d'IA est entraîné à l'aide d'un **grand nombre de données utilisant l'auto-supervision à grande échelle**, qui présente une **généralité significative** et est capable d'exécuter de manière compétente un **large éventail de tâches distinctes**, indépendamment de la manière dont le modèle est mis sur le marché, et qui peut être **intégré dans une variété de systèmes ou d'applications en aval**, à l'exception des modèles d'IA utilisés pour des activités de recherche, de développement ou de prototypage avant leur mise sur le marché » (mise en évidence ajoutée). Ce concept a été introduit lors des discussions au Parlement européen et répond à une volonté de réglementer les IA comme ChatGPT ou Gemini. Toutefois, son ajout pose des questions. Un modèle d'IA à usage général est-il forcément un SIA ? Les définitions de SIA et de modèles d'IA à usage général s'excluent-elles ?

Un modèle d'IA à usage général présente quatre caractéristiques principales. Tout d'abord, l'entraînement de ces modèles doit être effectué sur des ensembles de données vastes et diversifiés, en s'appuyant sur l'auto-supervision (par ex. *deep learning*). Ensuite, le modèle ne doit pas être conçu pour une tâche spécifique, mais doit être capable de réaliser une large gamme de tâches variées, telles que la traduction de textes, la génération de contenus textuels, visuels ou audiovisuels, ou encore l'analyse de documents. La généralité d'un modèle peut également être déterminée par le nombre de paramètres que peut compter le modèle (consid. 98 RIA). Finalement, un modèle d'IA doit pouvoir être intégré dans divers systèmes (par ex. smartphones, objets connectés, etc.) ou applications.

La notion de modèles d'IA à usage général doit être distinguée de celle de SIA (consid. 97 RIA). Un modèle d'IA à usage général n'est pas, en soi, un SIA, mais peut constituer un composant essentiel d'un SIA. Dans ce cas, on parle d'un SIA à usage général lorsque ce système est capable de répondre à divers usages (consid. 100 RIA). Prenons l'exemple de ChatGPT : ce dernier utilise un *large language model* (LLM) pour générer du contenu. Le LLM utilisé est un modèle d'IA à usage général au sens du RIA. OpenAI a ensuite développé une interface permettant aux utilisateurs d'interagir avec le LLM. Par conséquent, le LLM constitue un composant essentiel de l'interface, qui, selon nous, peut être qualifié de SIA.

C. Conclusion

L'application matérielle du RIA repose essentiellement sur la notion de SIA et de modèle d'IA à usage général. Tandis qu'un SIA est un système automatisé doté d'un certain degré d'autonomie, capable de générer des sorties influençant des environnements physiques ou virtuels, un modèle d'IA à usage général se concentre sur la capacité d'une IA à exécuter un

large éventail de tâches. Ce qui est important à retenir est qu'un modèle d'IA n'est pas en soi un SIA, mais peut en être un composant essentiel. L'intégration des modèles d'IA se fait généralement à l'aide d'interface utilisateur ou API, à l'image d'OpenAI avec ChatGPT ou son [API](#). Le Bureau d'IA – qui est compétent pour la mise en œuvre du RIA au niveau européen – devra vraisemblablement aider les entreprises à naviguer entre ces deux notions fondamentales.

Reproduction autorisée avec la référence suivante: Yannick Caballero Cuevas, Systèmes d'IA et modèles d'IA à usage général, publié le 4 novembre 2024 par le Centre de droit bancaire et financier, <https://cdbf.ch/1382/>