

Décisions automatisées dans l'administration publique

La prise de décision automatisée, c'est-à-dire la prise de décisions par l'ordinateur sans intervention humaine, sera au centre des débats politiques dans les mois et les années à venir.

CH++ se réjouit de ce débat démocratique et souhaite y contribuer avec le présent papier. Notre objectif est d'éviter les malentendus techniques et de présenter et justifier notre point de vue sur la prise de décisions automatisées.

Les décisions automatisées sont principalement utilisées dans les administrations publiques pour accroître l'efficacité. Cependant, le débat public est marqué par des préoccupations concernant la transparence et la discrimination. CH++ est d'avis que non seulement l'on puisse répondre à ces préoccupations de manière satisfaisante *en posant les bonnes conditions*, mais que *plus* de transparence et d'équité dans les décisions administratives sont possibles.

Pour y parvenir, il est nécessaire d'instaurer les conditions-cadres légales adéquates et CH++ travaillera dans ce sens. **Les conditions de base sont exposées dans ce document.**

Point de départ

Dans les discussions sur les décisions automatisées, il est souvent question d'algorithmes et d'intelligence artificielle (IA). Il est donc important de comprendre et de différencier ces termes.

Un **algorithme** est une séquence d'instructions uniques qui, semblable à une recette de cuisine, conduisent à un résultat spécifique en fonction de certaines entrées. Les algorithmes peuvent être très simples, par exemple sous la forme de règles simples "Si-Alors". Mais ils peuvent aussi être très complexes. Plus un algorithme est complexe, plus il est difficile à expliquer. Tant que l'algorithme peut être divisé en étapes claires - même s'il y en a beaucoup - il est explicable.

L'intelligence artificielle moderne est un modèle informatique qui fonctionne de manière similaire à un algorithme : leur but est de transformer des entrées en un résultat. La différence avec les algorithmes classiques est qu'il s'agit principalement de modèles mathématiques (surtout sous la forme de réseaux de neurones artificiels), qui peuvent contenir des millions, voire des milliards de paramètres. Par exemple, le

modèle GPT-3, l'un des modèles informatiques les plus avancés en matière de reconnaissance linguistique, contient environ 175 milliards de paramètres.

L'avantage des modèles d'IA est qu'ils peuvent traiter une quantité beaucoup plus importante de données par rapport aux algorithmes classiques, en particulier les ensembles de données non structurées. Ces données peuvent être des images, des textes ou des données audio. Un autre avantage est que les modèles d'IA peuvent prendre en charge des tâches pratiquement impossibles à réaliser avec des algorithmes classiques. Par exemple, la reconnaissance d'images et de la parole est aujourd'hui presque exclusivement gérée par l'IA. Dans presque tous les domaines, l'IA permet de prendre en charge des tâches que l'on pensait jusqu'à récemment impossibles à être effectuées par des ordinateurs. Cette évolution continuera à s'accélérer.

La complexité des modèles d'IA pose cependant des problèmes. D'une part, il est impossible d'expliquer directement de tels modèles complexes contenant des milliards de paramètres. D'autre part, les décisions de nombreux modèles d'IA dépendent fortement des données d'entraînement utilisées pour ajuster les paramètres. Les données d'entraînement sont des données pour lesquelles un résultat accepté est déjà disponible, et avec lesquelles le modèle d'IA tente de reconnaître des structures au cours de sa phase d'entraînement (qui sont ensuite reflétées par les paramètres). Il est donc important de comprendre les données d'entraînement utilisées pour créer les modèles d'IA.

Ces systèmes ont paradoxalement un potentiel incroyable pour résoudre les problèmes causés par leur utilisation. Un exemple est la reconnaissance de modèles de décision : les préjugés des humains peuvent être difficilement détectés, alors que ceux des systèmes peuvent être facilement détectés en jouant les processus de décision des milliers, voire des millions de fois. Il est essentiel que des organisations tierces puissent vérifier ces systèmes à tout moment et sans restriction.

Prise de décision

Le recours aux systèmes informatiques d'aide à la décision n'est pas nouveau: on les retrouve dans de nombreux domaines, du diagnostic médical à la détection de fraude en passant par l'évaluation de la solvabilité. Le modèle est toujours le même : le système reçoit des données nécessaires à la prise de décision, puis fournit un résultat en fonction d'algorithmes classiques ou d'un modèle d'IA. Il s'agit souvent de situations dans lesquelles le résultat agit en soutien à une décision rendue par un humain. Dans ce cas,

on parle de **prise de décision assistée par ordinateur**. Si le résultat de l'ordinateur représente la décision finale, c'est-à-dire qu'aucun être humain ne prend la décision dans le cas individuel concret, alors on parle de **prise de décision automatisée**.

La distinction entre la prise de décision assistée par ordinateur et la prise de décision automatisée est importante. La première est aujourd'hui largement répandue dans de nombreux domaines, tels que l'industrie ou la médecine, et est généralement largement acceptée. L'acceptation généralisée est principalement due au fait que l'être humain est finalement l'instance décisive. Mais pour toutes les formes de décision - qu'elles soient humaines, assistées par ordinateur ou automatisées - la décision doit être compréhensible, vérifiable, contrôlable et non discriminatoire.

Dans le contexte de la nouvelle Loi fédérale sur la protection des données (art. 21 : « décision individuelle automatisée ») au niveau fédéral ou de l'avant-projet de règlement du canton d'Argovie sur les procédures entièrement automatisées, il s'agit respectivement de la prise de décision automatisée, c'est-à-dire l'émission entièrement automatisée de décisions par les autorités. Les lois ou projets de loi mentionnés fixent des exigences minimales du point de vue de la protection des données ou du droit de la procédure administrative qui doivent être respectées si une loi permet à l'autorité d'utiliser la prise de décision automatisée dans un domaine spécifique. Plusieurs lois prévoient cette possibilité depuis la révision de la loi sur la protection des données, sans aucun doute de nombreuses autres suivront.

Les conditions cadres

CH++ est favorable à la prise de décision automatisée si les conditions suivantes sont remplies :

1. Droit à l'évaluation de la décision par l'humain

Toute décision automatisée doit être examinée par une personne physique sur demande. Pour ce faire, chaque décision automatisée doit être clairement déclarée en tant que telle.

2. Systèmes vérifiables de manière systématique

Les systèmes doivent être mis à disposition pour une vérification publique et systématique. C'est la seule façon de rendre visible et de corriger toute discrimination causée par les systèmes.

3. Visibilité des systèmes

La loi prévoit la tenue d'un registre public sur l'utilisation de systèmes algorithmiques décisionnels (registre ADM) et les contenus minimums du registre sont ancrés.

Janvier 2023

CH++

hallo@chplusplus.org

chplusplus.org