Universität Zürich[UZH]

CENTER FOR INFORMATION TECHNOLOGY SOCIETY AND LAW — ITSL

État des lieux juridique du développement et de l'utilisation de l'IA dans l'espace de formation suisse

EXECUTIVE SUMMARY

L'intelligence artificielle (IA) a un grand potentiel pour transformer le secteur de la formation. L'utilisation de cette technologie est toutefois liée à différents défis. La notion d'IA est aujourd'hui définie de manière très large. Le présent rapport se concentre sur les systèmes d'apprentissage automatique (machine learning), dont la caractéristique principale est qu'ils ne suivent pas les règles imposées par les êtres humains. Au lieu de cela, ils développent des règles de manière autonome en identifiant des modèles statistiques dans les données afin de passer de l'input à l'output. Ils sont ainsi en mesure d'identifier d'autres relations et d'effectuer d'autres analyses que les êtres humains ou les algorithmes classiques. Le chemin concret d'un système d'IA de l'input à l'output n'est toutefois pas explicable et compréhensible pour les êtres humains.

Une question fondamentale lors de l'utilisation de l'IA dans le domaine de la formation est celle de la **responsabilité** en matière de protection des données. Lors de l'utilisation d'outils d'IA, l'école doit généralement être qualifiée de responsable, le fournisseur de l'outil d'IA et un éventuel fournisseur de services cloud qualifiés de sous-traitants.

La question de la **base légale** est également centrale. L'action de l'État nécessite toujours une base légale. Cela vaut également pour le droit de la protection des données. Une base dans une loi au sens formel est nécessaire pour le traitement de données personnelles sensibles et parfois pour un profilage. La base légale détermine également la **finalité du traitement des données**, car les données ne peuvent être traitées que dans le but prévu par la loi (**principe de finalité**).

Les **lois scolaires** contiennent souvent des autorisations de traitement de données personnelles (sensibles) qui servent à l'accomplissement d'une tâche publique de l'école, ces autorisations étant le plus souvent de nature assez générale. Ces bases légales devraient toutefois suffire dans la mesure où le traitement des données par un outil d'IA sert à l'accomplissement d'une tâche publique. Cela peut être le cas, par exemple, lorsque l'école utilise un outil d'IA pour un apprentissage personnalisé, car la mission éducative de l'école comprend généralement aussi le soutien individuel des élèves.

Il est plus difficile de répondre à la question de savoir si les données personnelles peuvent être utilisées comme données d'entraînement pour le développement (ultérieur) d'outils d'IA. En ce qui concerne l'amélioration des outils d'IA par les écoles, une utilisation peut être considérée comme admissible. En revanche, l'utilisation de données personnelles pour l'entraînement et le développement par le fournisseur de l'outil d'IA à des fins (commerciales) propres ne devrait pas être couverte par la base légale. Dans la mesure où le fournisseur de l'outil d'IA souhaite utiliser les données à ses propres fins, il y a **changement de finalité**, ou au moins une **communication à un tiers**.

Une telle communication n'est autorisée que s'il existe une base légale, un consentement au cas par cas, si la communication a lieu à des fins non personnelles ou si les données sont anonymisées. Une base légale devrait régulièrement faire défaut, et le motif justificatif du consentement dans un cas particulier ne devrait pas non plus entrer en ligne de compte, car il n'y a pas de **cas particulier** lorsque des données concernant un grand

nombre de personnes sont transmises de manière standard et sur une longue période à un fournisseur d'IA. En outre, l'obtention d'un consentement n'est pas praticable, car un consentement valable est lié à une série de conditions préalables, par exemple la possibilité de révoquer le consentement à tout moment. Dans les situations où il existe une relation de subordination, comme c'est le cas entre l'école et les élèves, il est en outre douteux que le consentement puisse être volontaire. La communication des données au développeur d'un outil d'IA est toutefois autorisée dans la mesure où le droit cantonal prévoit la possibilité de transmettre des données personnelles à des tiers à des fins non personnelles, par exemple pour la recherche, la planification ou les statistiques. Cela présuppose toutefois que le traitement de données personnelles pour l'entraînement et le développement de l'IA soit classé comme traitement non personnel. Bien que ce point de vue semble correct, il existe actuellement une grande insécurité juridique sur cette question. Une clarification de la situation juridique par le législateur ou par les autorités de surveillance fédérales ou cantonales serait souhaitable.

L'utilisation de données pour le développement (ultérieur) d'un outil d'IA peut, dans certaines circonstances, constituer une **réutilisation des données,** c'est-à-dire une utilisation de données à des fins autres que celles prévues par la loi. Une telle utilisation n'est autorisée que dans des cas exceptionnels sans base légale pertinente. Dans le canton de Zurich, par exemple, il est possible d'obtenir un consentement au cas par cas, ce qui ne devrait toutefois pas être praticable pour les raisons mentionnées. Tout comme la communication de données à des tiers, le traitement de données à des fins non personnelles est autorisé au niveau fédéral et souvent aussi au niveau cantonal. Cela peut inclure l'utilisation de données issues d'outils d'IA pour le développement par le fournisseur de l'outil d'IA à des fins propres. Il serait également envisageable d'anonymiser les données, ce qui rendrait la loi sur la protection des données inapplicable.

Lors de l'utilisation d'outils d'IA, il convient de veiller tout particulièrement à ce que les exigences du **principe de proportionnalité** soient respectées, notamment l'aspect de la **minimisation des données**. Celle-ci se trouve dans un rapport de tension fondamental par rapport au fonctionnement des outils d'IA qui, en règle générale, fonctionnent d'autant mieux qu'ils traitent un grand nombre de données. Il convient donc d'examiner, dans le cadre d'une pesée des intérêts, si la finalité du traitement des données peut justifier la quantité de données traitées.

Selon le canton, certaines dispositions prises par les écoles peuvent être qualifiées de décisions contestables. Il s'agit par exemple de l'évaluation des performances des élèves (pour la promotion) ou de la répartition des écoles et des classes. Les **décisions contestables** doivent être motivées, du moins sur demande. Comme les systèmes d'IA se basent sur des corrélations et non sur des causalités, les décisions de ces systèmes ne sont pas compréhensibles pour les êtres humains et ne peuvent pas être justifiées juridiquement de manière suffisante. Tant qu'il n'est pas possible d'identifier les critères pertinents pour les dispositions prises par les outils d'IA, ces outils ne sont pas adaptés à la prise de décisions sous forme de résolution.

Certaines décisions prises par des outils d'IA doivent être qualifiées de **décisions individuelles automatisées** lorsqu'elles affectent considérablement les personnes concernées. C'est notamment le cas des décisions de promotion et, en partie, des décisions d'affectation à une école ou à une classe. Des dispositions légales particulières s'appliquent (en partie) aux décisions individuelles non automatisées. Les personnes concernées doivent notamment être informées qu'une décision a été prise de manière automatisée. En vertu du droit d'accès à la protection des données, les personnes concernées peuvent en outre exiger d'être informées de la logique qui sous-tend la décision automatisée. Les personnes concernées doivent donc être informées de l'utilisation d'un outil d'IA et (sur demande) du fonctionnement de l'algorithme et des objectifs assignés à l'outil d'IA, ainsi que des catégories de données utilisées comme données d'entraînement. Cette information doit être adaptée au destinataire.