

**RAPPORT DE LA COMMISSION THÉMATIQUE DES SYSTÈMES D'INFORMATION
chargée d'examiner l'objet suivant :**

**Postulat Carine Carvalho et consorts au nom du groupe socialiste - L'état doit prévenir la
discrimination algorithmique dans son utilisation de l'intelligence artificielle (IA)**

1. PRÉAMBULE

La Commission thématique des systèmes d'information (CTSI) s'est réunie le mardi 20 mai 2025 à la salle du Bicentenaire, Place du Château 6 à Lausanne, pour traiter de cet objet.

Elle était composée de Mmes et MM. Maurice Gay (président et rapporteur), Céline Baux, Carine Carvalho, Michael Demont, Yann Glayre, Vincent Jaques, Didier Lohri, Laurent Miéville, Charles Monod, Bernard Nicod, Alexandre Rydlo, Théophile Schenker, Michael Wyssa, Valérie Zonca.

Excusées : Carole Dubois (remplacée par Marion Wahlen), Olivier Gfeller (remplacé par C. Carvalho), Regula Zellweger (remplacée par M. Bernard Nicod).

Mme Nuria Gorrite, cheffe du Département de la culture, des infrastructures et des ressources humaines (DCIRH), a participé à la séance ; elle était accompagnée de M. Patrick Amaru, directeur général de la Direction générale du numérique et des systèmes d'information (DGNSI) et Catherine Pugin, déléguée au numérique au sein de la DGNSI.

M. Yvan Cornu, secrétaire de la commission, a tenu les notes de séance et contribué à la rédaction de ce rapport de commission, ce dont nous le remercions.

2. POSITION DE LA POSTULANTE

Le postulat porte sur les systèmes algorithmiques, et plus particulièrement sur ceux qui permettent des décisions automatisées. En Suisse comme en Europe, l'usage par les autorités publiques de ces algorithmes décisionnels automatisés s'intensifie dans de nombreux domaines. On peut notamment citer l'imposition et la fiscalité, les demandes d'aides sociales, la détection des fraudes, la sélection des candidats aux postes ouverts dans la fonction publique, le soutien aux prises de décisions policières et à la surveillance, ainsi que les réponses automatiques aux questions des citoyennes et citoyens. L'intelligence artificielle (IA) est largement mobilisée dans ces secteurs, sans s'y limiter. Un postulat récent de Michael Wyssa évoque également d'autres domaines où l'IA pourrait se développer, comme la santé, la justice, l'enseignement et la formation.

L'utilisation de l'IA dans l'automatisation des décisions présente une promesse importante : plus d'efficacité et une prise de décision plus rapide, ce qui peut significativement améliorer le fonctionnement de l'État. Cependant, il est essentiel de mettre en place des garde-fous pour éviter ce que l'on appelle la discrimination algorithmique.

En effet, les systèmes algorithmiques peuvent reproduire, voire amplifier, les discriminations déjà présentes dans la société. Cela peut provenir de la conception même des systèmes, notamment lorsque ceux-ci sont entraînés sur des données reflétant des inégalités sociales existantes, ou encore à cause de biais introduits durant le développement des modèles d'IA. Il est aussi nécessaire de s'interroger sur les acteurs impliqués dans la création de ces systèmes.

À partir d'une série de sources, telles que l'association AlgorithmWatch ou le dernier rapport de la Commission fédérale des questions féminines sur les discriminations algorithmiques, la postulante propose au Conseil d'État de structurer la réflexion autour de trois axes principaux :

1. La création d'un registre public des systèmes de prise de décision automatisée.
2. Une analyse systématique des risques liés à l'utilisation de ces systèmes.
3. Un contrôle individuel et démocratique de leur usage.

Cette proposition s'inscrit dans la continuité d'interventions parlementaires récentes déjà déposées au Grand Conseil, notamment le postulat Wyssa déjà mentionné, ainsi que la réponse à l'interpellation Raedler, qui offrent des indications intéressantes et rassurantes sur l'utilisation actuelle et la réflexion en cours par l'État concernant ces technologies.

Cependant, compte tenu de la rapide évolution de ces systèmes, ainsi que des pressions croissantes en faveur de leur déploiement, il est essentiel que l'État s'engage dans une réflexion approfondie et définisse des principes solides pour encadrer leur usage, tant aujourd'hui que dans le futur.

3. POSITION DU CONSEIL D'ÉTAT

Le Conseil d'État est en phase avec les éléments développés par la postulante, dans la mesure où il a déjà été fait mention, dans le cadre de la CTSI, de tout le travail accompli par un groupe interdépartemental spécifiquement dédié à l'IA. Le rapport, qui synthétise non seulement les enjeux mais aussi la posture de l'administration cantonale vaudoise (ACV) vis-à-vis de l'IA, est accessible sur le site internet du canton :

https://www.vd.ch/fileadmin/user_upload/organisation/dinf/dsi/fichiers_pdf/IA_Rapport_interdepartemental.pdf.

Au sein de l'ACV, l'IA n'est pas développée pour elle-même, mais dans une démarche d'analyse des enjeux que l'État doit considérer, entre risques et opportunités. Plusieurs enjeux ont déjà été identifiés et méritent d'être rappelés.

Tout d'abord, l'État porte une **responsabilité** particulière. Contrairement aux acteurs privés qui peuvent générer des contenus via l'IA sans garantie de fiabilité, l'État doit s'assurer de la fiabilité de ce qu'il produit, car il engage sa crédibilité auprès des citoyennes et citoyens. Cette responsabilité est incontournable.

Ensuite, la **souveraineté** constitue un enjeu majeur. Il n'est pas envisageable de confier des données sensibles à des intelligences génératives s'appuyant sur des bases de données internationales, sans une position souveraine solide. Cette précaution est indispensable avant d'adopter une IA.

Par ailleurs, les questions de **transparence** s'imposent, notamment la nécessité d'indiquer ce qui est compris et généré par une IA. La **sécurité** demeure un enjeu fondamental pour l'État de Vaud, tout comme la **participation** citoyenne, qui implique d'encourager une meilleure compréhension des mécanismes algorithmiques, thématique directement abordée par le postulat.

En ce qui concerne l'état actuel au sein de l'ACV, le Conseil d'État a clairement exprimé qu'il encourage les services à utiliser l'IA dans un objectif d'efficacité. L'utilité de l'IA pour une administration est évidente, mais son usage doit impérativement s'inscrire dans un cadre légal. Par ailleurs, l'État doit anticiper les risques liés à l'IA, notamment ceux liés aux biais - à la sécurité ou à la souveraineté - comme cela est mis en lumière dans le rapport interdépartemental sur l'IA.

Le Conseil d'État a également décidé de mettre en place une gouvernance spécifique liée à l'IA. À cette fin, l'Office de la transformation numérique et d'appui à l'administration (ONA) a été créé, avec la nomination de Sébastien Dufaut en tant que directeur. Par la suite, l'ONA, en collaboration avec la Direction générale des systèmes d'information (DGNSI), accompagnera les services de l'ACV dans leurs processus internes afin d'intégrer l'IA dans les tâches quotidiennes, dans le cadre défini par l'État de Vaud. Une réorganisation a également eu lieu à la DGNSI, avec la désignation d'un responsable des missions IA chargé de faire le lien entre le support technique, la DGNSI, les métiers et l'ONA.

Un programme de formation est en cours de déploiement, afin que les utilisatrices et utilisateurs connaissent et appliquent les règles encadrant l'usage de l'IA. Un guide des bonnes pratiques sera en principe diffusé avant l'été 2025 et mis à la disposition de toutes et tous.

Le rapport décrit aussi l'ambition du Conseil d'État de renforcer la gouvernance numérique par la création d'un budget dédié aux partenariats et la désignation d'un interlocuteur cantonal fort. Ces mesures sont en cours de réalisation, tout comme les formations.

Concernant spécifiquement le postulat, la dimension de la confiance est centrale. La mise en place d'un registre des IA figure parmi les préconisations, afin que les services documentent les finalités des algorithmes, les données utilisées, ainsi que les modalités d'anticipation des biais et des risques liés à la sécurité ou à la souveraineté. Ce registre, accompagné d'une campagne de sensibilisation destinée au grand public, vise à diffuser les bonnes pratiques et à informer la population sur les opportunités et les risques liés à l'IA. L'objectif est d'éviter la reproduction de biais présents dans les données existantes, car l'IA n'est pas « intelligente » en soi, mais agrège des éléments existants, eux-mêmes imparfaits.

Une consultation plus large est également prévue, impliquant notamment le Parlement des Jeunes, la Commission des Jeunes, et d'autres groupes cibles, comme les seniors. Enfin, la mise en place d'un Conseil d'éthique est envisagée pour accompagner le déploiement de cette technologie.

Voilà ce que contient le rapport du groupe de travail interdépartemental, désormais disponible en ligne. Il apparaît clairement que le Conseil d'État est en phase avec la démarche proposée dans ce postulat qui permettra de rendre plus public ce que l'administration a déjà anticipé.

La DGNSI a dès le début privilégié une démarche interdépartementale. Il aurait été possible de produire un rapport avec uniquement les ingénieur·es en informatique, mais l'objectif était d'impliquer toutes et tous afin de faire ressortir la diversité des enjeux. Ce travail collectif, mené en 2024, est reflété dans le contenu du rapport.

Aujourd'hui, la DGNSI échange beaucoup avec ses homologues des autres cantons et avec la Confédération. Plusieurs groupes de travail, notamment pilotés par l'Administration numérique suisse (ANS), auxquels la DGNSI et l'État de Vaud participent, sont fortement impliqués sur ces thématiques. La DGNSI collabore aussi avec le milieu académique, notamment l'EPFL et l'UNIL, ainsi qu'avec « l'Innovation Park » via Innovaud. De nombreuses start-up actives dans le canton permettent des échanges enrichissants.

Le responsable de mission IA à la DGNSI permet d'avoir une vision globale des demandes des services et d'y répondre progressivement.

Enfin, un aspect spécifique dans l'approche du Canton de Vaud concerne la durabilité. Le label numérique responsable obtenu en février 2025 témoigne de cette attention portée à l'impact environnemental et énergétique de l'IA, sujet qui fait partie des axes de travail actuels et figure également dans le rapport.

4. DISCUSSION GÉNÉRALE

Groupe interdépartemental

En 2023, un commissaire avait déjà déposé une interpellation portant sur l'utilisation de ChatGPT. Dans sa réponse, le Conseil d'État mentionnait l'existence d'un groupe interdépartemental et soulignait l'importance de clarifier, informer et expliquer lorsqu'une décision est prise par un algorithme, notamment pour gagner en efficacité.

Décisions automatisées

Conformément à la loi fédérale sur la protection des données, les citoyennes et les citoyens sont protégés à l'égard des décisions automatisées, autrement dit, produites par l'IA. Elles et ils doivent par ailleurs être informé·es lorsque la décision a été produite par de l'IA. Il existe néanmoins des décisions dites partielles, comme une recommandation, avec une intervention de l'administration. La question qui se pose c'est de savoir si ces décisions partielles sont utilisées par l'ACV.

La déléguée au numérique indique que la notion de la souveraineté a été largement étudiée par le Canton de Vaud de concert avec d'autres cantons latins. Ensemble, ils se sont entendus sur une définition de la souveraineté numérique qui se comprend comme la capacité de prendre une décision, mais à ne pas se trouver dans une forme de dépendance.

Le Canton de Vaud travaille avec Microsoft et ce dernier lui a mis à disposition un outil d'IA dénommé Copilot. Ce travail est encadré par des contrats. Il est difficile de développer un modèle de type ChatGPT vaudois voire suisse pour des raisons liées aux infrastructures et aux besoins énergétiques importants nécessaires pour assurer les capacités de calcul requises. Ainsi, tout n'est pas fait au sein de la DGNSI, il y a des systèmes d'IA qui sont externalisés mais dans un cadre contractuel.

Le directeur de la DGNSI souligne qu'à ce jour, l'ACV ne dispose d'aucune solution d'IA autonome capable de prendre des décisions. Même dans les cas où des solutions sont acquises auprès de prestataires externes, celles-ci ne reposent pas sur des systèmes d'IA prenant des décisions à la place de l'administration. À ce titre, la DGNSI avait recensé, avec AlgorithmWatch, les prestations faisant appel à des algorithmes déterministes.

En matière de protection des données, domaine où plusieurs questions émergent, l'ACV applique les principes de la loi fédérale sur la protection des données (LPD), notamment en ce qui concerne l'information des personnes concernées sur l'usage de l'IA. D'autres enjeux apparaissent, notamment en lien avec la finalité des données personnelles, telle que définie par la loi. L'utilisation de l'IA complexifie ces questions, dans la mesure où les données peuvent être utilisées à des fins différentes de celles initialement prévues.

Aujourd'hui, aucun algorithme d'IA utilisé par l'ACV ne fonctionne sans contrôle humain.

De manière générale, une commissaire est sidérée de la manière dont l'IA est utilisée sans prise en compte des enjeux écologiques.

Conseil d'éthique

Le but est d'introduire la notion d'éthique dans le débat de l'IA. Le Conseil d'éthique, dont le but serait d'accompagner le Canton de Vaud dans ces questions liées à l'IA, n'a pas encore été constitué. A fortiori, il serait composé de personnes externes à l'ACV avec notamment des juristes, des éthiciens-informaticiens ou encore avec des spécialistes de solutions en termes d'économie. Dans un précédent projet, le Canton de Vaud avait déjà expérimenté l'aide d'un groupe d'accompagnement afin de cadrer l'action publique en intégrant la notion d'éthique, avec un franc succès.

Publication des outils utilisés par l'Etat et « Biais dans les données liées au genre »

Un commissaire ne soutient pas ce postulat pour plusieurs raisons. Tout d'abord, le Canton de Vaud ne partage pas massivement les données sur ses algorithmes, qui sont justement protégées par la LPD. Bien que des incidents, notamment aux Pays-Bas, aient eu lieu, il considère que c'est loin d'être le cas dans le Canton de Vaud. Par ailleurs, dans le secteur de la santé, le postulat fait état d'un manque de données féminines, où des données masculines sont utilisées comme les normes créant par-là des biais dans le diagnostic et dans le traitement des patientes. Le commissaire souhaite avoir des éclaircissements sur ce point. S'agissant du résultat produit par l'IA, la discrimination proviendrait d'une mauvaise interprétation humaine, sans volonté de nuire de la part de l'algorithme. Enfin, il s'oppose à la publication en ligne et à la mise à disposition publique des outils de l'État, craignant qu'ils ne soient manipulés pour modifier les résultats.

Le directeur général de la DGNSI précise qu'il n'est pas question de publier les outils mais plutôt les algorithmes listés et les sujets pour lesquels des décisions automatisées seraient générées. Il n'y a pas d'information communiquée sur les systèmes, les logiciels ou d'autres informations qui mettraient à risque l'aspect sécuritaire.

Il est fondamental qu'il soit publié quand des algorithmes sont utilisés pour prendre des décisions. Il y a un intérêt public à ce que le destinataire de la décision automatisée dispose de ces informations. Evidemment, il faudrait que la mise en œuvre se fasse en toute sécurité afin d'éviter tout risque de piratage.

La Confédération maintient un tel registre des projets de l'IA de l'administration fédérale, qui est régulièrement mis à jour (site du centre de compétences en matière d'IA), et qui peut être consulté ci-après : <https://cnaï.swiss/fr/products/projektdatenbank/> Il y a donc une volonté d'indiquer ce qui est fait dans les différents offices à la Confédération.

Concernant ce que l'on appelle le « gender data gap », le manque de données médicales spécifiques aux femmes est bien documenté. À ce titre, le CHUV dispose d'ailleurs d'une Unité santé et genre, dédiée à la recherche sur la santé des femmes et visant à combler cette lacune. Aujourd'hui, de nombreuses applications médicales (par ex. celles liées au suivi du cycle menstruel ou à la prévention de l'hypertension), accompagnent

les utilisatrices et utilisateurs dans la gestion de leur santé, sur des données issues d'échantillons de population limités. En revanche, dans le cas de risques spécifiques, l'IA peut s'avérer inadaptée.

À ce jour, on ne dispose pas d'informations claires sur la manière dont l'État utilise ou encadre le recours à ce type d'outils.

Communication sur l'usage des décisions automatisées

Un commissaire souligne que de nombreuses questions soulevées par le postulat ont déjà trouvé des réponses. Sur le plan décisionnel, on distingue d'une part les décisions automatisées, et d'autre part l'usage individuel de l'IA, notamment par les professions libérales telles que les médecins, mais également dans de nombreux autres métiers.

IA souveraine suisse

Concernant l'IA suisse, un commissaire est convaincu que le pays dispose des capacités nécessaires afin de développer son IA. D'ailleurs, l'EPFL et l'EPFZ ont récemment inauguré le supercalculateur « Alps » hébergé au Tessin, qui est le sixième mondial en puissance de calcul, avec un financement exclusivement suisse. Le besoin en infrastructures et financement se réduit laissant ainsi la place pour la création d'une IA souveraine.

Difficulté de répertorier les algorithmes utilisés au niveau communal

S'agissant du postulat, un commissaire propose une modification afin de supprimer la dimension « communale » du point 1, considérant le travail demandé comme compliqué et déraisonnable : « point 1 : [...] répertorier les systèmes algorithmiques utilisés à des fins décisionnelles aux niveaux cantonal ~~et communal~~. ». Il est proposé de sensibiliser et encourager les communes à créer également un registre public de systèmes de prise de décision algorithmique.

Une formulation trop directive

Une commissaire approuve le fond du postulat tout en relevant un problème de forme liée à des demandes trop directives adressées au Conseil d'État. Elle suggère que le postulat demande plutôt « un rapport sur les possibilités de... ».

5. PRISE EN CONSIDÉRATION PARTIELLE DU POSTULAT / MODIFICATION

Mme Carine Carvalho

La postulante adhère aux propositions de modification de son texte qu'elle reformule de la manière suivante :

« Afin d'assurer que l'utilisation des algorithmes dans le secteur public soit transparente, éthique et bénéfique pour toutes et tous, le présent postulat ~~demande au~~ invite le Conseil d'état à établir un rapport sur les possibilités de : »

Point 1 : [...] répertorier les systèmes algorithmiques utilisés à des fins décisionnelles ~~aux niveaux cantonal et communal~~ au niveau cantonal et sensibiliser les communes à adopter une démarche similaire [...].

6. VOTES SUR LA PRISE EN CONSIDÉRATION PARTIELLE DU POSTULAT (25_POS_13)

Modification du postulat :

Vote sur les propositions de modifications : 12 voix pour, 0 contre et 3 abstentions.

Prise en considération partielle du postulat :

La Commission recommande ainsi au Grand Conseil la prise en considération partielle de ce postulat par 12 voix pour, 1 contre et 2 abstentions.

*Le rapporteur :
(Signé) Maurice Gay*

Nyon, le 5 août 2025