

ÉTUDE « L'ÉVALUATION D'IMPACT ALGORITHMIQUE : UN OUTIL QUI DOIT ENCORE FAIRE SES PREUVES »

La présente étude a été réalisée par **Elisabeth Lehagre** à la demande d'**Etalab** (direction interministérielle du numérique – etalab.gouv.fr) dans le cadre des travaux sur les algorithmes publics.

RÉSUMÉ

Alors que le recours à des systèmes algorithmiques (notamment dans le cadre de la prise de décisions) s'est considérablement développé ces dernières années, les risques et enjeux associés à leurs effets sur les personnes et la société dans son ensemble n'ont cessé de se préciser. C'est dans ce cadre qu'ont émergé plusieurs propositions d'**« évaluation d'impact algorithmique »** (ou « **EIA** »), développés au niveau gouvernemental ou par des organisations de la société civile.

À quoi correspondent ces EIA, que contiennent-elles, comment sont-elles mises en œuvre et peut-on d'ores et déjà juger de leur efficacité ? Cette étude s'est attachée à examiner ces questions sur la base de quelques exemples récents d'EIA (voir ci-contre).

Outils d'évaluation examinés	
Société civile	Gouvernement
<ul style="list-style-type: none"> Liste d'évaluation pour une IA de confiance par le groupe d'experts de haut-niveau en IA ou « GEHN IA » (constitué par la Commission européenne)⁴ [Juillet 2020] Rapport « <i>Algorithmic Impact Assessments</i> » de l'AI Now Institute (États-Unis)⁶ [Avril 2018] « <i>Ethics and Algorithms Toolkit</i> » proposé par un collectif de chercheurs et d'experts (États-Unis)⁸ [2018] 	<ul style="list-style-type: none"> « Évaluation de l'incidence algorithmique » - Secrétariat du Conseil du Trésor, Canada⁵ [2018 - 2020] « <i>Stakeholder Impact Assessment</i> » ou « <i>SIA</i> » - Government Digital Services, Office for Artificial Intelligence, Royaume-Uni⁷ [2019] « <i>Algorithm Charter for Aotearoa New-Zealand</i> » - Stats NZ, Nouvelle-Zélande⁹ [Juillet 2020]

Un premier constat s'est rapidement imposé au regard des outils examinés : il n'existe pas de modèle type d'EIA. Malgré un objectif commun d'évaluation des effets d'un système algorithmique sur les personnes et/ou la société, tout au long de son cycle de vie, les EIA présentent des formes et contenus variés. L'évaluation des systèmes sera ainsi plus ou moins guidée, entre recommandations pratiques d'ordre général et questions ciblant de manière directe ou indirecte les impacts à évaluer. L'évaluation pourra également être plus ou moins orientée vers l'identification de niveaux de risques associés aux impacts. Néanmoins, quels que soient les formes et contenus retenus, **l'EIA apparaît être, en premier lieu, un support privilégié de réflexion et de discussion** invitant les personnes à s'interroger sur les effets de leur projet de déploiement d'un système algorithmique et à envisager les mesures pouvant éventuellement limiter les risques qui pourraient y être associés.

La pertinence de l'exercice suppose toutefois une utilisabilité de ces outils, notamment au regard de la compréhension même de l'objet et des effets à évaluer. Or, l'accessibilité de l'EIA ne relève pas de l'évidence notamment du fait

de son caractère multidimensionnel qui implique souvent d'appréhender des notions techniques liées aux systèmes algorithmiques et des considérations sociologiques, juridiques voire philosophiques. Réaliser une EIA, c'est dans un premier temps pouvoir identifier son « objet algorithmique », qu'il soit système d'IA ou système de décision automatisée. C'est aussi appréhender des notions comme le bien-être, l'équité, l'autonomie ou la vie privée afin de déterminer comment le déploiement d'un système algorithmique peut affecter des individus dans ces domaines, tant au niveau individuel que collectif. Les EIA regorgent en effet de notions larges et complexes qui posent la question de l'expertise nécessaire pour les considérer avec pertinence mais aussi celle de la marge et des divergences d'interprétation possibles. L'élaboration de documents de travail, le développement de formations pour les utilisateurs des EIA ou la mise en place d'équipes pluridisciplinaires pour leur réalisation sont quelques exemples de mesures envisagées pour tenter d'assurer la compréhension des outils d'évaluation d'impact algorithmique. Toutefois, les difficultés pratiques à cet égard sont réelles. Dès lors, **faudrait-il repréciser l'ambition des EIA en tant qu'outil de sensibilisation aux effets potentiels des systèmes algorithmiques plutôt qu'outil destiné à l'identification précise de ces derniers ?** Loin d'être un outil auto-suffisant, l'EIA aurait alors surtout pour objectif de contribuer à un travail d'acculturation destiné à accompagner progressivement l'intégration des questions sur les effets sociaux des systèmes algorithmiques, au niveau opérationnel.

Outre cet objectif de sensibilisation interne, la grande majorité des propositions examinées associent également l'EIA à un cadre plus large invitant à l'envisager comme un outil de transparence et de redevabilité. La publicité de l'EIA est en effet régulièrement recommandée afin de permettre au public d'être informé, de comprendre, de vérifier voire de contester les systèmes algorithmiques et les évaluations qui en sont faites par les administrations. L'EIA prend alors une dimension plus externe, orientée vers une plus grande inclusion du public pouvant se faire *ex ante* (dans le cas d'une publicité assurée avant l'utilisation du système algorithmique) et/ou *ex post* (après son déploiement). La publicité permet ainsi de contrebalancer l'autonomie des utilisateurs dans leur exercice d'auto-arbitrage où les personnes réalisant l'EIA sont à la fois juges et parties. Cette publicité est d'ailleurs perçue par beaucoup d'associations de défense des droits sur Internet comme un élément indispensable à l'opposabilité et à l'efficacité réelle d'une EIA. Cette visée de transparence vis-à-vis du public serait-elle alors la « raison d'être » de l'évaluation d'impact algorithmique, au-delà de sa dimension de sensibilisation interne ? Difficile d'apporter une réponse tranchée au regard des EIA examinées, ces deux objectifs semblant guider le recours à ce type d'outils.

On perçoit néanmoins l'influence du cadre dans lequel s'inscrit l'EIA quant à sa portée. **La question du caractère facultatif ou obligatoire de l'EIA et de ses modalités de mise en œuvre peut alors se poser.** Parmi les EIA examinées seul l'outil proposé par le gouvernement du Canada relève d'un cadre réglementaire spécifique, prévoyant l'obligation pour les ministères et services fédéraux de réaliser une EIA. Les autres EIA relèvent plutôt de démarches internes volontaires. Le caractère obligatoire de l'EIA mériterait-il cependant d'être généralisé et faut-il y voir une nécessité pour assurer la crédibilité et l'opposabilité de l'évaluation réalisée ? Tout semble de nouveau dépendre de la vision dans laquelle peut s'inscrire l'EIA entre outil de sensibilisation et d'acculturation aux effets sociétaux des systèmes algorithmiques et outil de redevabilité et d'arbitrage vis-à-vis du public. En effet, les conséquences de l'EIA, si elles sont peu abordées dans les outils examinés (à l'exception de l'initiative canadienne) mériteraient d'être précisées. Une EIA serait-elle alors susceptible de conditionner la décision de recours même à un système algorithmique via une « note éliminatoire » en raison des risques identifiés ?

Les évaluations d'impact algorithmique en sont à leurs débuts et beaucoup d'interrogations demeurent. Le recul manque encore pour en appréhender l'utilité et l'efficacité et un suivi de leur mise en œuvre paraît indispensable. **Toutefois, à ce jour, les EIA proposées ont le mérite d'exister et si l'évaluation d'impact algorithmique doit encore faire ses preuves, elle contribue à une prise de conscience des enjeux sociétaux des systèmes algorithmiques dont il serait dommage de se priver.**