

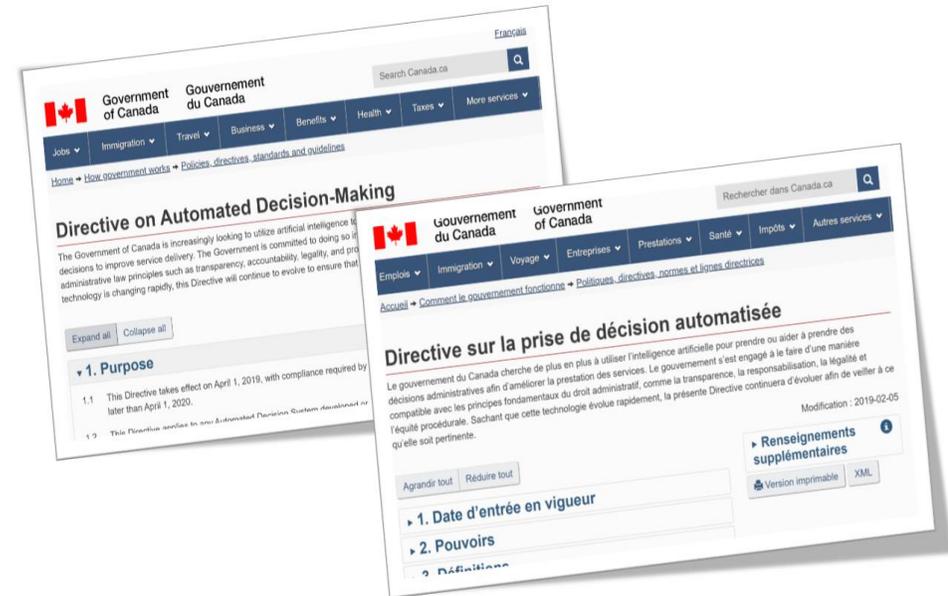


Troisième examen de la Directive sur la prise de décisions automatisée – Résumé des principales questions, des recommandations stratégiques et des modifications proposées

**Dossier de consultation
Printemps 2022**

Objectif

- **Donner un aperçu du 3^e examen de la Directive sur la prise de décisions automatisée (DPDA).**
- **Obtenir des commentaires sur les recommandations stratégiques et les modifications provisoires.**



Le Secrétariat du Conseil du Trésor du Canada (SCT) est tenu d'examiner régulièrement la DPDA afin de veiller à ce qu'elle demeure pertinente et adaptée à l'évolution du paysage de l'automatisation dans le gouvernement fédéral.

Renseignements

- Ces dernières années, les gouvernements et les organisations internationales ont poursuivi divers efforts de réglementation pour gérer les risques des systèmes d'intelligence artificielle (IA) et favoriser la transparence et la responsabilité algorithmiques.
- L'approche du gouvernement du Canada (GC) à l'égard de l'IA responsable favorise l'équité et l'inclusion dans la prise de décisions automatisée en veillant à ce que les résultats des systèmes décisionnels automatisés soient explicables et impartiaux.
- En 2019-2020, plus de 300 projets d'IA ont été consignés dans 80 % des institutions fédérales. Certains cas d'utilisation automatisent les décisions ayant une incidence sur les bénéficiaires de services à l'intérieur et à l'extérieur du gouvernement.

Europe Is in Danger of Using the Wrong Definition of AI

Some intelligent systems are at risk of being excluded from oversight in the EU's proposed legislation. This is bad for both businesses and citizens.

**Federal rules on AI too narrow and risk 'damaging public trust':
Internal review**

White House science advisers call for an "AI Bill of Rights"

Canada's opportunity to ensure AI remains a force for good

Research shows AI is often biased. Here's how to make algorithms work for all of us

LCO report comparing European Union, Canadian AI regulation stresses development of 'trustworthy AI'

Directive sur la prise de décisions automatisée

- Le gouvernement fédéral utilise parfois des systèmes informatiques pour automatiser la prise de décisions. Dans d'autres cas, les systèmes informatiques complètent une partie de l'analyse menant à une décision.
- Lorsqu'il s'agit de décisions de services qui ont une incidence sur les droits juridiques, les intérêts ou les privilèges des personnes, les exigences de la Directive sur la prise de décisions automatisée s'appliquent.
- Les exigences de la Directive garantissent que :
 - ✓ **les personnes sont informées du moment et de la façon dont l'automatisation est utilisée,**
 - ✓ **les personnes reçoivent des explications significatives sur les décisions qui les concernent,**
 - ✓ **les décisions sont justes et exactes,**
 - ✓ **les incidences négatives possibles de l'automatisation sont déterminées et réduites au minimum.**

Une directive du Conseil du Trésor établit des exigences obligatoires pour la façon dont les organisations fédérales doivent fonctionner.

Approche du 3^e examen

- L'examen en cours fait le point de l'état actuel de la directive et cerne les risques et les défis à l'engagement du gouvernement à l'égard d'une IA responsable dans le secteur public fédéral.
- L'examen a étudié la gamme croissante de services en cours d'automatisation, mettant en lumière les lacunes et les « angles morts » critiques qui limitent la pertinence et l'efficacité de la directive à appuyer la transparence, la responsabilisation et l'équité dans la prise de décisions automatisée.
- Les questions liées à la terminologie, à la faisabilité et à la cohérence avec d'autres instruments de politique du Conseil du Trésor ont également éclairé l'objectif et l'orientation de l'examen.
- Les examens périodiques ne sont pas exhaustifs. Ils cherchent à adapter la directive aux tendances pertinentes du paysage canadien et mondial de l'IA, tout en affinant progressivement le texte de l'instrument pour appuyer l'interprétation et faciliter la conformité.

Aperçu des questions clés identifiées dans le 3^e examen*

<p>Portée</p> <ul style="list-style-type: none"> ▪ L'orientation externe exclut les décisions automatisées qui ont une incidence sur les employés du gouvernement fédéral. ▪ Cadre linguistique : il faut clarifier la portée. 	<p>Examen périodique</p> <ul style="list-style-type: none"> ▪ Le délai actuel de six mois pour l'examen mène à des défis stratégiques et opérationnels. 	<p>Clients touchés</p> <ul style="list-style-type: none"> ▪ Le renvoi aux Canadiens dans certaines parties de la DPDA ne reconnaît pas d'autres clients possibles. 	<p>Gouvernance des données</p> <ul style="list-style-type: none"> ▪ Il faut des mesures visant à assurer la traçabilité, la protection, la conservation et la destruction des données utilisées et générées par un système. 	<p>Biais du modèle</p> <ul style="list-style-type: none"> ▪ Des mesures d'assurance de la qualité sont nécessaires pour corriger les biais découlant du modèle sous-jacent à un système.
<p>Explication</p> <ul style="list-style-type: none"> ▪ Il n'y a pas de critères pour ce qui constitue une explication valable. 	<p>Raisons de l'automatisation</p> <ul style="list-style-type: none"> ▪ Une justification de l'adoption de l'IA par rapport aux besoins et aux objectifs d'un programme n'est pas actuellement nécessaire. 	<p>Examen par les pairs</p> <ul style="list-style-type: none"> ▪ L'obligation de publier des renseignements sur les examens par les pairs n'est pas incluse. ▪ Le moment de l'examen par les pairs n'est pas clair. 	<p>Planification d'urgence</p> <ul style="list-style-type: none"> ▪ La terminologie n'est pas harmonisée avec la Politique sur la sécurité du Conseil du Trésor. 	<p>Date de publication de l'EIA</p> <ul style="list-style-type: none"> ▪ Les exigences de l'évaluation de l'incidence algorithmique (EIA) ne précisent pas la date de publication de l'EIA.

Recommandations stratégiques de haut niveau

1. Élargir la portée pour couvrir les services internes.
2. Préciser que la portée comprend des systèmes à l'appui des décisions administratives.*
3. Remplacer l'intervalle d'examen de six mois par un examen biennal et désigner à la DPI du Canada le rôle de demander un examen en fonction des besoins.
4. Remplacer les renvois aux Canadiens par un langage plus englobant, comme les clients.
5. Mettre en place des mesures favorisant le suivi, la protection, la conservation et la destruction appropriés des données utilisées et générées par un système décisionnel automatisé.
6. Élargir l'exigence relative aux essais de préproduction pour couvrir les essais du biais du modèle.
7. Établir des critères d'explication à l'appui de l'exigence d'explication.
8. Élargir l'EIA pour y inclure des questions sur les raisons poussant le ministère à adopter l'automatisation.
9. Donner le mandat de publier un résumé des examens par les pairs et d'exiger l'achèvement avant la production du système.
10. Harmoniser l'exigence d'urgence avec la terminologie pertinente établie dans la Politique sur la sécurité du Conseil du Trésor.
11. Donner le mandat de diffuser les EIA avant la production d'un système.

* Voir l'annexe pour des exemples de fonctions qu'un système automatisé dans la portée peut effectuer dans un processus décisionnel.

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, section 5 (Portée)	N° 1 – Modifier le paragraphe 5.1 : « La présente Directive s'applique uniquement aux systèmes qui fournissent des services externes ou internes, conformément à la tels que définis dans la Politique sur les services et le numérique. »*	L'adoption de systèmes décisionnels automatisés pour améliorer les services internes (p. ex., l'embauche) dans plusieurs institutions fédérales souligne la nécessité d'élargir la portée de la DPDA afin de réduire au minimum les risques pour les droits, les intérêts et les privilèges des employés fédéraux. La modification proposée permet d'y parvenir en incluant les systèmes fournissant des services internes qui relèvent de la Politique sur les services et le numérique. La suggestion de remplacer « tels que défini dans » par « conformément à » découle du fait que la Politique (annexe A) ne définit que les « services internes intégrés », un sous-ensemble de services internes. (Veuillez consulter l'annexe pour un aperçu des services internes du GC.)
DPDA, section 5 (Portée)	N° 2 – Modifier le paragraphe 5.2 : « La présente Directive s'applique à tout système, outil ou modèle statistique utilisé pour appuyer ou prendre une décision administrative au sujet d'un client. »	La distinction entre recommander et prendre une décision n'est pas claire. Le cadre actuel a causé la confusion quant aux conditions dans lesquelles un système automatisé qui participe à un processus décisionnel administratif serait assujéti à la DPDA. Le terme « recommander » occulte l'intention d'inclure tout système automatisé influant sur le processus décisionnel dans la portée de l'instrument (cela se voit dans la définition du système décisionnel automatisé à l'annexe A de la DPDA). On peut l'interpréter à tort comme établissant un seuil d'applicabilité élevé pour les systèmes qui ne font pas partie du processus décisionnel. La modification proposée vise à remédier à cette situation en optant pour un terme plus générique qui clarifie la vaste gamme de cas d'utilisation que la portée de la DPDA doit couvrir. Il convient de prendre note que la modification ne change pas la portée actuelle; elle ne fait que l'exprimer en termes plus appropriés.
DPDA, section 1 (Date d'entrée en vigueur)	N° 3 – Modifier le paragraphe 1.2 : « La Directive sera assujéti à un processus d'examen automatique; on prévoit que celui-ci se déroulera tous les six mois à partir de sa date d'entrée en vigueur examinée tous les deux ans, comme déterminé par la dirigeante principale de l'information du gouvernement du Canada. »	Un mécanisme d'examen plus souple aiderait à relever les défis stratégiques et opérationnels de l'exigence actuelle. Une période de deux ans rendrait mieux compte du long processus de modification d'une directive, tout en allégeant les charges de capacité imposées par un intervalle d'examen de six mois. L'approche proposée donnerait également plus de temps aux décideurs pour recueillir des données sur l'adoption et la conformité de l'IA dans l'ensemble du GC, tout en apportant plus de stabilité et de prévisibilité aux institutions fédérales assujétiées à la DPDA. En permettant à la DPI du Canada de demander des examens à tout moment, la modification garantit que la DPDA peut rester réceptive aux besoins au fur et à mesure qu'ils se présentent.
DPDA, section 4 (Objectifs et résultats escomptés)	N° 4 – Modifier le paragraphe 4.1 : « La présente Directive a pour objet de veiller à ce que les systèmes décisionnels automatisés soient déployés d'une manière qui permet de réduire les risques pour les clients Canadiens et les institutions fédérales, et qui donne lieu à une prise de décisions plus efficace, exacte et conforme, qui peut être interprétée en vertu du droit canadien. »	Les décisions administratives dans la portée actuelle de la DPDA ne concernent pas seulement les citoyens canadiens. Elles peuvent également toucher les résidents permanents, les demandeurs d'asile, les demandeurs de visa, ou d'autres personnes (ou entreprises) qui reçoivent un service du gouvernement fédéral. Il serait donc plus approprié d'utiliser un terme plus générique comme « clients » dans l'énoncé de l'objectif et dans l'ensemble de l'instrument. Ce changement appuierait également l'élargissement proposé de la portée de la DPDA afin d'inclure les employés fédéraux, qui ne sont pas exclusivement composés de citoyens canadiens. Il appuierait également l'harmonisation avec la terminologie utilisée dans la Politique sur les services et le numérique.

* Cette modification serait accompagnée de modifications mineures à l'EIA (p. ex., aux questions dans les sections « À propos de la décision » et « Consultations »).

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, section 6 (Exigences)	N° 5 – Ajouter un nouveau paragraphe 6.3 intitulé « Gouvernance des données » : « Établir des mesures afin de veiller à ce que les données utilisées et générées par le système décisionnel automatisé soient traçables, protégées, conservées et éliminées de façon appropriée conformément à la Directive sur les services et le numérique, à la Directive sur les pratiques relatives à la protection de la vie privée et à la Directive sur la gestion de la sécurité. »	Même si elle établit des exigences à l'appui de la gestion de la qualité et de l'atténuation des biais pour les données utilisées par les systèmes décisionnels automatisés, la DPDA ne répond pas au besoin plus vaste de régir les intrants et les extrants des systèmes tout au long de leur cycle de vie. En particulier, il est nécessaire de veiller à ce que les données utilisées et générées par les systèmes soient traçables, protégées, conservées et éliminées de manière appropriée. Lorsque de telles données contiennent des renseignements personnels ou peuvent mener à de tels renseignements en combinaison avec d'autres données, la DPDA doit veiller à ce qu'elles ne soient pas partagées, réutilisées ou autrement traitées sans les autorisations requises. Le paragraphe proposé intégrerait les mesures de protection pour ces données uniques, en veillant à ce qu'elles soient traçables (p. ex., pour les audits, les examens, les litiges, les explications), protégées (p. ex., contre le partage ou l'utilisation non autorisés), conservées et éliminées de façon appropriée. Ces mesures favoriseraient une approche plus holistique de la gouvernance des données de l'IA.
DPDA, section 6 (Exigences)	N° 6 – Modifier le paragraphe 6.3.1 : « Avant d'amorcer la production, élaborer les processus d'évaluation des données et de l'information utilisés dans les systèmes de prise de décisions automatisée, ainsi que les modèles sous-jacents des systèmes , visant à s'assurer de l'absence de biais imprévus dans les données et d'autres facteurs qui pourraient influencer injustement les résultats. »	Les biais dans l'IA peuvent provenir de sources multiples, y compris les données utilisées pour former un système et le modèle utilisé pour les traiter. La DPDA exige des essais de préproduction des données d'entrée pour détecter le biais, mais il ne tient pas compte de la possibilité que le biais puisse également résulter des hypothèses et des paramètres intégrés dans un modèle. Pour atteindre les résultats escomptés de la DPDA, il faut surveiller non seulement les données d'entrée (et de sortie), mais aussi le modèle utilisé pour obtenir les extrants à l'appui ou constituant des décisions. La modification proposée garantirait que les problèmes liés aux modèles sont abordés au début du cycle de vie, avant le déploiement du système.
DPDA, annexe C (Exigences par niveau d'incidence)	N° 7 – Modifier les mesures d'explication pour les quatre niveaux d'incidence afin d'introduire des critères d'explication : « Il s'agit de fournir des renseignements décrivant : <ul style="list-style-type: none"> • le rôle du système dans le processus décisionnel; • les données d'apprentissage et des clients, leur source et leur méthode de collecte, le cas échéant; • les critères utilisés pour évaluer les données des clients et les opérations effectuées pour les traiter; • l'extrait produit par le système et tous renseignements pertinents nécessaires pour l'interpréter dans le contexte de la décision administrative. » 	L'exigence d'explication actuelle ne précise pas ce qui constitue une explication valable. Elle est susceptible de faire l'objet de nombreuses interprétations, ce qui crée plusieurs problèmes pour les organisations fédérales, les responsables des politiques du SCT et les clients. Le manque de clarté quant aux renseignements nécessaires pour répondre à l'exigence pourrait entraîner des pratiques incohérentes, ce qui pourrait entraîner des explications incomplètes et des disparités dans le traitement des clients. Cela crée également une approche ponctuelle d'explicabilité au gouvernement, les organisations fédérales cherchant à obtenir une orientation interprétative des responsables des politiques du SCT, au cas par cas. Cette tâche est lourde non seulement pour le SCT, mais aussi pour les programmes visant à assurer la conformité efficace à l'exigence. Inspirée par la <i>Loi pour une République numérique</i> de la France, la modification proposée formule des critères d'explication visant à garantir le droit d'un client à un décideur juste et impartial, ainsi qu'aux motifs des décisions qui les concernent. Les critères permettraient à la DPDA de mieux tenir compte du caractère numérique des décisions automatisées, ce qui exige des mesures uniques afin de veiller à ce que les institutions qui cherchent à augmenter ou à remplacer les décideurs humains puissent continuer de respecter les normes du droit administratif.

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
EIA (Section Facteur opérationnel/ Incidence positive)	<p>N° 8 – Ajouter une nouvelle série de questions sur les raisons de l'automatisation :</p> <ul style="list-style-type: none"> • « À quel besoin d'utilisateur le système doit-il répondre? [texte libre] • Comment le système sera-t-il utilisé pour répondre aux besoins des utilisateurs? [texte libre] • Dans quelle mesure le système répondra-t-il efficacement aux besoins des utilisateurs? [Légèrement efficace; Modérément efficace; Très efficace] • Veillez expliquer pourquoi vous vous attendez à ce que le système atteigne le niveau d'efficacité indiqué ci-dessus. [texte libre] • Veillez décrire comment vous vous assurerez que le système se limite à répondre au besoin de l'utilisateur identifié ci-dessus? [texte libre] • D'autres processus manuels ont-ils été envisagés? [Oui/Non] • Si l'on a envisagé des processus manuels, pourquoi l'automatisation a-t-elle été considérée comme l'option préférée? [texte libre] • Quelles sont les conséquences du non-déploiement du système? (Sélectionnez tous les éléments qui s'appliquent) [Le service ne peut pas être fourni du tout; Le service ne peut être fourni en temps opportun ou de manière efficace; Les frais de service sont trop élevés; La qualité du service n'est pas aussi élevée; La prestation du service ne peut atteindre les objectifs de rendement; Autres [texte libre]] <p>Modifier le titre de la section : « Facteur opérationnel – Incidence positive Raisons de l'automatisation »</p>	<p>L'un des principes directeurs de l'IA responsable au GC souligne la valeur de « [se fonder] sur un besoin clair des utilisateurs et l'intérêt du public ». Toutefois, la DPDA ne demande pas aux ministères d'expliquer pourquoi ils ont décidé d'utiliser l'automatisation dans le cadre d'un processus décisionnel. Même si l'EIA comporte des questions demandant aux ministères de décrire leur projet d'automatisation et d'identifier les facteurs opérationnels pertinents, on ne s'attend pas à ce que les ministères fournissent des raisons justifiant la nécessité d'avoir recours à l'automatisation pour répondre aux besoins spécifiques des utilisateurs. De même, la DPDA et l'EIA ne tiennent pas compte de la manière dont l'utilisation du système soit appropriée pour répondre aux besoins des utilisateurs et aux objectifs du programme. La série de questions proposée positionne l'EIA comme un espace où les ministères peuvent fournir une justification de leur projet d'automatisation, décrivant non seulement pourquoi il est nécessaire, mais aussi son degré de compatibilité avec les besoins des utilisateurs et les objectifs du programme. Ces renseignements seraient accessibles ouvertement aux intervenants fédéraux et publics, créant de nouvelles possibilités de dialogue intersectoriel sur les mérites de l'automatisation et les limites appropriées de l'utilisation de l'IA dans la prise de décisions administratives.</p>
DPDA, section 6 (Exigences)	<p>N° 9 – Modifier le paragraphe 6.3.4 : « Consulter l'expert qualifié approprié pour demander une évaluation du système décisionnel automatisé et publier un résumé en langage clair des constatations avant la production du système, conformément à ce qui est prévu à l'annexe C. »</p>	<p>L'absence d'un mécanisme imposant la publication d'examens par les pairs (ou de renseignements connexes) crée une occasion manquée de renforcer la confiance du public dans l'utilisation des systèmes automatisés au moyen d'une évaluation par des experts externes. La publication d'au moins un résumé des examens par les pairs effectués (compte tenu des défis liés à l'exposition de données sensibles sur les programmes, de secrets commerciaux ou de renseignements sur les systèmes propriétaires) peut renforcer la transparence et la responsabilité en permettant aux intervenants de valider les renseignements dans l'EIA. L'exigence actuelle est également muette quant au calendrier des examens par les pairs, ce qui crée une incertitude tant pour les ministères que pour les examinateurs quant à savoir s'il faut effectuer un examen avant ou pendant le déploiement du système. Contrairement aux audits, les examens sont plus efficaces lorsqu'ils sont disponibles avec une EIA, avant la production d'un système, de sorte qu'ils puissent servir de couche d'assurance supplémentaire. Les modifications proposées abordent ces questions en élargissant l'exigence de publication obligatoire et en précisant le calendrier des examens. Les résumés d'examen par les pairs publiés complèteraient la documentation sur les résultats des audits ou d'autres examens que la DPDA exige que les chefs de projet communiquent dans le cadre de l'exigence d'avis (voir l'annexe C de la DPDA).</p>

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, annexe C (Exigences par niveau d'incidence)	<p>N° 9 (suite) – Modifier les mesures d'examen par les pairs pour les niveaux d'incidence II et III : « Veillez consulter au moins l'un des experts suivants et publiez un résumé des constatations en langage clair : » « OU Publier les spécifications du système décisionnel automatisé dans une revue à comité de lecture. Lorsque l'accès à l'examen publié est restreint, assurez-vous qu'un résumé en langage simple des conclusions est librement disponible. » (Cette dernière entrée serait placée à la fin de la liste des options.)</p> <p>Modifier les mesures d'examen par les pairs pour le niveau d'incidence IV : « Veillez consulter au moins deux des experts suivants et publiez un résumé des constatations en langage clair : » « Publier les spécifications du système décisionnel automatisé dans une revue à comité de lecture. Lorsque l'accès à l'examen publié est restreint, assurez-vous qu'un résumé en langage simple des conclusions est librement disponible. »</p>	<p>Les modifications proposées visent à harmoniser les mesures d'examen par les pairs de l'annexe C avec l'exigence d'examen par les pairs mise à jour (paragraphe 6.3.4). Elles soulignent la nécessité de publier un résumé des constatations de l'examen par les pairs. Pour les niveaux d'incidence II et III, l'option de publier les spécifications du système décisionnel automatisé dans une revue à comité de lecture a été présentée comme une solution de rechange à la consultation d'un ou plusieurs groupes d'experts. Conformément à l'approche adoptée dans le cadre du niveau d'incidence IV, il s'agit d'une revue à comité de lecture avec l'examen effectué par l'un des experts énumérés.</p>
DPDA, section 6 (Exigences)	<p>N° 10 – Modifier le paragraphe 6.3.6 : « Mettre en place des stratégies, des plans et/ou des mesures pour appuyer la TI et la gestion de continuité des activités conformément à ce qui est prévu à l'annexe C, en vertu de la Directive sur la gestion de la sécurité ».</p> <p>Modifier le titre du paragraphe 6.3.6 en remplaçant « Éventualités » par « TI et gestion de la continuité des activités ».</p>	<p>Les mesures requises en vertu de l'exigence d'urgence sont bien établies dans la Politique sur la sécurité du gouvernement (PSG) et la Directive sur la gestion de la sécurité (DGS). Toutefois, le terme « urgence » n'est ni défini ni décrit dans ces instruments. La DPDA ne donne pas non plus de définition. La formulation de l'exigence en matière de TI et de gestion de la continuité des activités, et l'établissement de liens clairs avec la PSG et les instruments de politique à l'appui pourraient faciliter l'interprétation, améliorer la coordination avec les responsables de la sécurité du Ministère et réduire au minimum le chevauchement des efforts de conformité. En s'éloignant du fait que la planification d'urgence est une exigence unique plutôt qu'une exigence qui a des ancrages clairs dans d'autres instruments de politique, cela peut également contribuer à la cohérence des politiques. Les modifications proposées à l'annexe C visent à s'harmoniser avec le langage proposé pour la section sur les exigences. Toutefois, elles fournissent également des détails supplémentaires, en s'appuyant sur les procédures obligatoires des mesures de sécurité dans la DGS, en particulier dans le contexte de la TI et de la gestion de la continuité des activités.</p>
DPDA, annexe C (Exigences par niveau d'incidence)	<p>N° 10 (suite) – Modifier les mesures de planification d'urgence pour les niveaux d'incidence III et IV : « Veiller à ce que des stratégies de récupération du système, des plans et d'autres systèmes des mesures de sécurité pertinents soient établis en coordination avec les responsables désignés, dans l'éventualité où le système décisionnel automatisé ne soit pas disponible. »</p> <p>Modifier le titre de la présente section en remplaçant « Planification des mesures d'urgence » par « TI et gestion de la continuité des activités ».</p>	<p>Les modifications proposées visent à harmoniser les mesures d'examen par les pairs de l'annexe C avec l'exigence d'examen par les pairs mise à jour (paragraphe 6.3.4). Elles soulignent la nécessité de publier un résumé des constatations de l'examen par les pairs. Pour les niveaux d'incidence II et III, l'option de publier les spécifications du système décisionnel automatisé dans une revue à comité de lecture a été présentée comme une solution de rechange à la consultation d'un ou plusieurs groupes d'experts. Conformément à l'approche adoptée dans le cadre du niveau d'incidence IV, il s'agit d'une revue à comité de lecture avec l'examen effectué par l'un des experts énumérés.</p>

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, section 6 (Exigences).	N° 11 – Modifier le paragraphe 6.1.1 : « Effectuer et publier une évaluation de l'incidence algorithmique avant la production de tout système décisionnel automatisé. »	La DPDA exige des institutions fédérales qu'elles effectuent et publient une EIA dans le Portail du gouvernement ouvert. Toutefois, les paragraphes 6.1.1 et 6.1.4 ne précisent pas quand les EIA doivent être publiées. Même si le SCT a encouragé les institutions fédérales à publier leurs EIA avant la production d'un système, la politique n'établit pas explicitement le moment de la publication. Cela crée de l'incertitude quant au moment approprié de la publication et risque d'affaiblir les mesures de transparence de la DPDA en permettant aux institutions de retarder la publication de l'EIA bien dans le cycle de vie d'un système. Le déploiement et l'utilisation d'un système décisionnel automatisé en l'absence d'une EIA accessible au public peuvent avoir des conséquences négatives sur la confiance du public dans l'utilisation de l'IA dans le secteur public fédéral. Tous les clients soumis à une prise de décisions automatisée doivent avoir accès à une EIA achevée sans délai. Plus tôt l'EIA est publiée dans le cycle de vie d'un système, mieux c'est pour la transparence et la responsabilité. Certaines EIA publiées sur le Portail du gouvernement ouvert n'ont pas été publiées avant la production du système. La modification proposée répond à cette question en précisant clairement la nécessité de publier une EIA avant la production d'un système.

Résultats escomptés

La mise en œuvre des modifications proposées effectuerait ce qui suit :

- Garantir que les systèmes décisionnels automatisés touchant les fonctionnaires sont équitables et inclusifs.
- Renforcer la transparence et la responsabilité afin de favoriser la confiance du public.
- Renforcer les mesures de protection contre la discrimination et le préjudice.
- Préciser les exigences et appuyer les besoins opérationnels.

Questions de discussion

- Y a-t-il des questions essentielles ou urgentes dont l'examen actuel ne tient pas compte?
- Les modifications proposées à la DPDA et à l'EIA sont-elles claires et justifiées?
- Prévoyez-vous des problèmes liés à la modification de la DPDA et de l'EIA comme proposée?
- Y a-t-il des intervenants fédéraux ou externes clés que le SCT doit engager dans le cadre de cette consultation?
- Quelles questions le SCT doit-il envisager de prioriser dans le prochain examen de la DPDA?

Prochaines étapes

- Travailler ouvertement, collaborer avec les intervenants nationaux et internationaux afin de cerner et de régler les problèmes avant le processus de modification des politiques.

Mars	Printemps et été	Été et automne
Étape 1 : Consultation préliminaire du BDPI	Étape 2 : Discussion avec les intervenants	Étape 3 : Modifications de politique
<ul style="list-style-type: none"> Assurer l'harmonisation avec les politiques sur la protection des renseignements personnels, la sécurité, le gouvernement ouvert, et le numérique. Sensibiliser le public au 3^e examen. 	<ul style="list-style-type: none"> Affiner les problèmes et les solutions. Consulter les ministères, le Commissariat à la protection de la vie privée du Canada (CPVP), les agents des services et les agents négociateurs. Collaborer avec les instituts de recherche et d'autres gouvernements. Travailler ouvertement. 	<ul style="list-style-type: none"> Démarrer le processus de gérance intégrée des politiques du BDPI. Demander l'approbation des comités supérieurs et sensibiliser (SMA SPI, agents des services, et autres). Demander l'approbation de la DPI du GC et du secrétaire. Publier la directive à jour. Appuyer les ministères (permanent).

Personnes-ressources

Secrétariat du Conseil du Trésor du Canada

Bureau de la dirigeante principale de l'information

Équipe de la Politique sur les données et de l'IA

ai-ia@tbs-sct.gc.ca

Annexe

Questions clés identifiées dans le 3^e examen

- **Portée** : La portée de la DPDA exclut les décisions administratives automatisées qui ont une incidence sur les employés du gouvernement fédéral. Cela crée des vulnérabilités pour les employés assujettis à l'automatisation dans l'embauche, l'évaluation du rendement ou d'autres décisions appuyant la prestation de services internes. De plus, les termes utilisés pour définir la portée de la DPDA ont causé la confusion quant aux conditions qui la déclenchent.
- **Examen périodique** : L'intervalle d'examen de six mois présente des défis stratégiques et opérationnels pour le SCT. Cela est dû à la durée des processus de consultation et d'approbation, l'incidence des examens réguliers sur la capacité des équipes, le rythme relativement lent de l'adoption de l'automatisation au GC, et les incertitudes découlant de changements fréquents de la politique administrative.
- **Clients touchés par les systèmes décisionnels automatisés** : Lorsqu'elle s'adresse particulièrement aux Canadiens, la DPDA ne reconnaît pas son applicabilité possible aux cas qui touchent d'autres clients au Canada ou à l'étranger (p. ex., les résidents permanents, les réfugiés, les citoyens d'autres pays).

Questions clés identifiées dans le 3^e examen

- **Gouvernance des données** : Même si la DPDA comprend des dispositions appuyant la gestion des données recueillies pour un système et utilisées par celui-ci (p. ex., pour réduire au minimum le biais, assurer la qualité), elle n'établit pas de mesures appuyant la traçabilité, la protection, la conservation et l'élimination appropriées de ces données. Cela est également nécessaire pour les extrants du système (p. ex., recommandations, notes) qui ne sont pas abordés dans la DPDA. Les deux types de données peuvent présenter des risques pour la vie privée ou la sécurité si elles sont communiquées, réutilisées, conservées ou éliminées de manière inappropriée.
- **Biais du modèle** : Les exigences d'assurance de la qualité de la DPDA ne tiennent pas compte des biais découlant du modèle sous-jacent à un système (plutôt que des données utilisées pour l'élaborer). Cela pourrait amener les utilisateurs à négliger les essais de modèles de préproduction.
- **Explication** : L'exigence d'explication actuelle ne précise pas ce qui constitue une « explication valable ». L'absence de critères d'explication pourrait conduire à une interprétation et une application incohérentes.

Questions clés identifiées dans le 3^e examen

- **Raisons de l'automatisation** : La DPDA ne tient pas compte du but et de la portée des projets d'automatisation. Cette lacune laisse les clients et les intervenants publics sans justification claire de la décision d'adopter l'IA et sans description de la façon dont le système sera déployé pour répondre aux besoins des utilisateurs et aux objectifs du programme.
- **Examen par les pairs** : L'absence d'obligation de publier des renseignements sur les examens par les pairs constitue une occasion manquée pour le GC, qui pourrait tirer parti de ce mécanisme pour renforcer la confiance du public dans les systèmes décisionnels automatisés utilisés au gouvernement. Le calendrier approprié des examens par les pairs est également flou.
- **Planification d'urgence** : La terminologie utilisée dans cette exigence n'est pas conforme à ce qui est bien établi dans la Politique de sécurité du Conseil du Trésor, qui établit des mesures de sécurité pour la TI, la gestion de la continuité des activités, et d'autres domaines. Cette inadéquation pourrait entraîner un double emploi des efforts de conformité et avoir une incidence négative sur la cohérence des politiques.

Questions clés identifiées dans le 3^e examen

- **Calendrier de publication de l'EIA :** La DPDA ne précise pas le moment de la publication des EIA. L'exigence relative à la publication d'EIA ne fixe que le format et l'emplacement de la publication. Cela crée de l'incertitude quant au moment approprié de la publication et risque d'affaiblir les mesures de transparence de la DPDA en permettant aux institutions de retarder la publication de l'EIA bien dans le cycle de vie d'un système. Cela a des conséquences négatives sur la confiance du public dans l'utilisation de l'IA dans le secteur public fédéral.

Principes directeurs du GC pour une intelligence artificielle responsable

1. **Comprendre et mesurer** l'incidence de l'utilisation de l'IA.
2. **Transparence** sur la façon et le moment d'utiliser l'IA.
3. **Explications claires sur l'IA** dans la prise de décisions.
4. **Être le plus ouvert possible** en communiquant le code source, les données d'apprentissage et d'autres renseignements pertinents.
5. **Offrir une formation suffisante** qui permet aux fonctionnaires d'élaborer et d'utiliser des solutions d'IA qui ont une conception, une fonction et une mise en œuvre responsables.

Aperçu de la Directive sur la prise de décisions automatisée

Exigences de la Directive sur la prise de décisions automatisée

Évaluation de l'incidence algorithmique

Comprendre

- EIA avant la production
- EIA lorsque la portée change
- Publication des résultats

Transparence

Communiquer

- Avis avant décision
- Explication après décision
- Accès aux composants
- Distribution du code source
- Documenter les décisions

Assurance de la qualité

Prévenir

- Essai et suivi des résultats
- Qualité des données
- Examen par les pairs
- Formation des employés
- Réserve
- Sécurité
- Consultation de services juridiques
- Intervention humaine

Recours

Corriger

- Options de recours pour contester les décisions

Aperçu des exigences de la DPDA

- Publiée en 2019, la DPDA cherche à assurer la transparence, la responsabilisation et l'équité procédurale dans l'utilisation des systèmes décisionnels automatisés dans le gouvernement fédéral.
- La portée de la DPDA couvre les systèmes utilisés pour prendre ou appuyer des décisions administratives ayant une incidence sur les clients externes (p. ex., les citoyens, les entreprises). Elle s'applique aux systèmes mis au point ou achetés à compter du 1^{er} avril 2020.
- La DPDA officialise la responsabilisation algorithmique en tenant des sous-ministres adjoints (SMA) qui supervisent les projets d'automatisation pertinents responsables de se conformer aux exigences de la Politique.
- Les institutions fédérales assujetties à la DPDA sont tenues de remplir et de publier une Évaluation de l'incidence algorithmique (EIA) au Portail du gouvernement ouvert. L'outil d'EIA est un questionnaire qui détermine le niveau d'incidence d'un système décisionnel automatisé.

Aperçu des exigences de la DPDA

- Les incidences de l'automatisation d'une décision administrative sont classées en quatre niveaux, allant du niveau I (faible incidence) au niveau IV (incidence très élevée). L'EIA permet de cerner les risques et évaluer les incidences dans un large éventail de domaines liés aux droits et aux intérêts des personnes et des collectivités.
- La directive établit des mesures d'assurance de la qualité pour aider à assurer la légalité d'un projet d'automatisation, la qualité des données d'entrée, la sécurité du système, la surveillance humaine, l'examen par les pairs et la connaissance des employés.
- Les systèmes en production doivent être surveillés afin de veiller à ce qu'ils ne produisent pas de résultats non intentionnels et qu'ils respectent les politiques et les lois applicables.
- La directive exige des institutions fédérales qu'elles fournissent aux clients soumis à une prise de décisions automatisée un mécanisme de recours approprié leur permettant de contester une décision.
- Le SCT utilise plusieurs mécanismes de gouvernance pour assurer le respect de la directive, y compris le Cadre pour la gestion de la conformité, les cas conceptuels ministériels, les propositions d'architecture intégrée, et les présentations au Conseil du Trésor.

Exemples de fonctions système dans un processus décisionnel

Recommandations



- Présenter des renseignements pertinents au décideur.
- Avertir le décideur des conditions inhabituelles.
- Présenter des renseignements provenant d'autres sources (« correspondance des données »).
- Fournir des évaluations, p. ex., en générant des cotes, des prédictions ou des classifications.
- Recommander une ou plusieurs options au décideur.
- Prendre des décisions partielles ou intermédiaires dans le cadre d'un processus décisionnel.
- Prendre la décision définitive.

Décisions

Évaluation de l'incidence algorithmique (EIA)

<https://ouvert.canada.ca/eia>

Description, instructions et méthode de notation expliquées à l'adresse :

<https://www.canada.ca/fr/gouvernement/systeme/gouvernement-numerique/innovations-gouvernementales-numeriques/utilisation-responsible-ai/evaluation-incidence-algorithmique.html>.

ALGORITHMIC IMPACT ASSESSMENT



Government of Canada / Gouvernement du Canada

Algorithmic Impact Assessment

Home > Open Government

Algorithmic Impact Assessment

Information in the AIA is only stored locally on your computer, and the Government of Canada does not have access to the information you place into the tool. If you wish to keep your work, please save the data locally for future use by using the 'Save' button. You can reload a previously saved AIA form using the 'Upload JSON File' button.

Save Upload JSON File Start Again

Navigate to a Specific Page (Out of 13)

Section 8: Impact Assessment

Page 8 of 13

Impact Assessment

Will the system only be used to assist a decision-maker?

Yes
 No

Will the system be replacing a decision that would otherwise be made by a human?

Yes
 No

Will the system be replacing human decisions that require judgement or discretion?

Yes
 No

Is the system used by a different part of the organization than the ones who developed it?

Yes
 No

Are the impacts resulting from the decision reversible?

Likely reversible

How long will impacts from the decision last?

Impacts are most likely to be brief

Please describe why the impacts resulting from the decision are as per selected option above.

The ultimate consequence of the decision is a letter being sent to the company advising them that corresponding actions should be taken. These actions all require less than 10 minutes to implement.

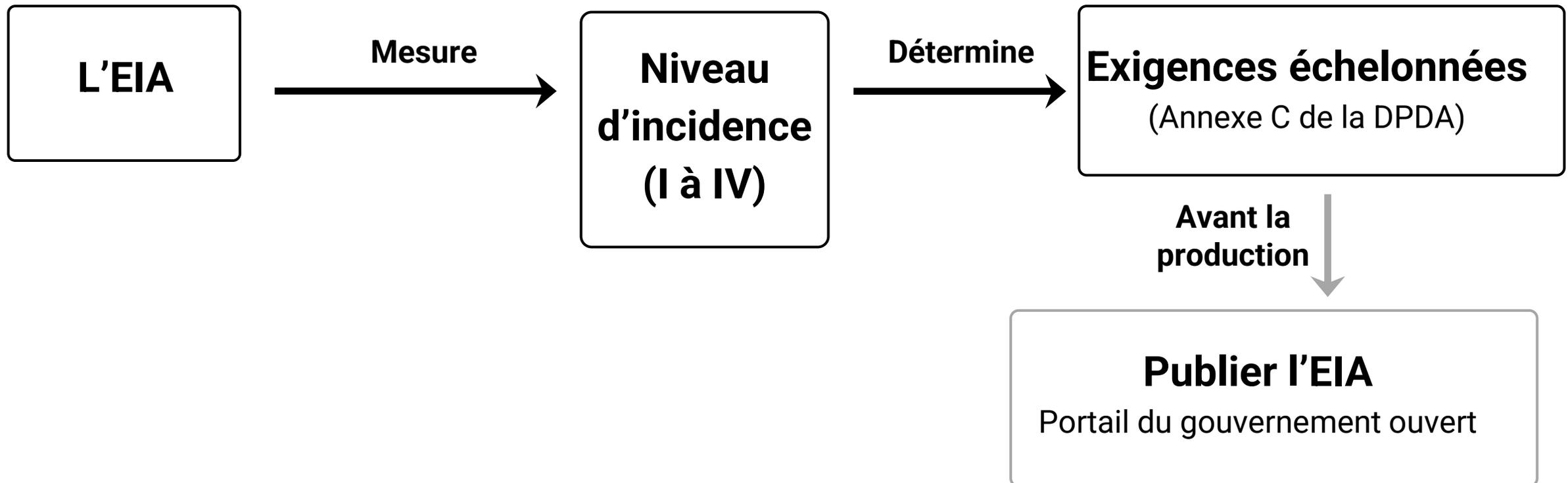
The impacts that the decision will have on the rights or freedoms of individuals will likely be:

Little to no impact

Please describe why the impacts resulting from the decision are (as per selected option above).

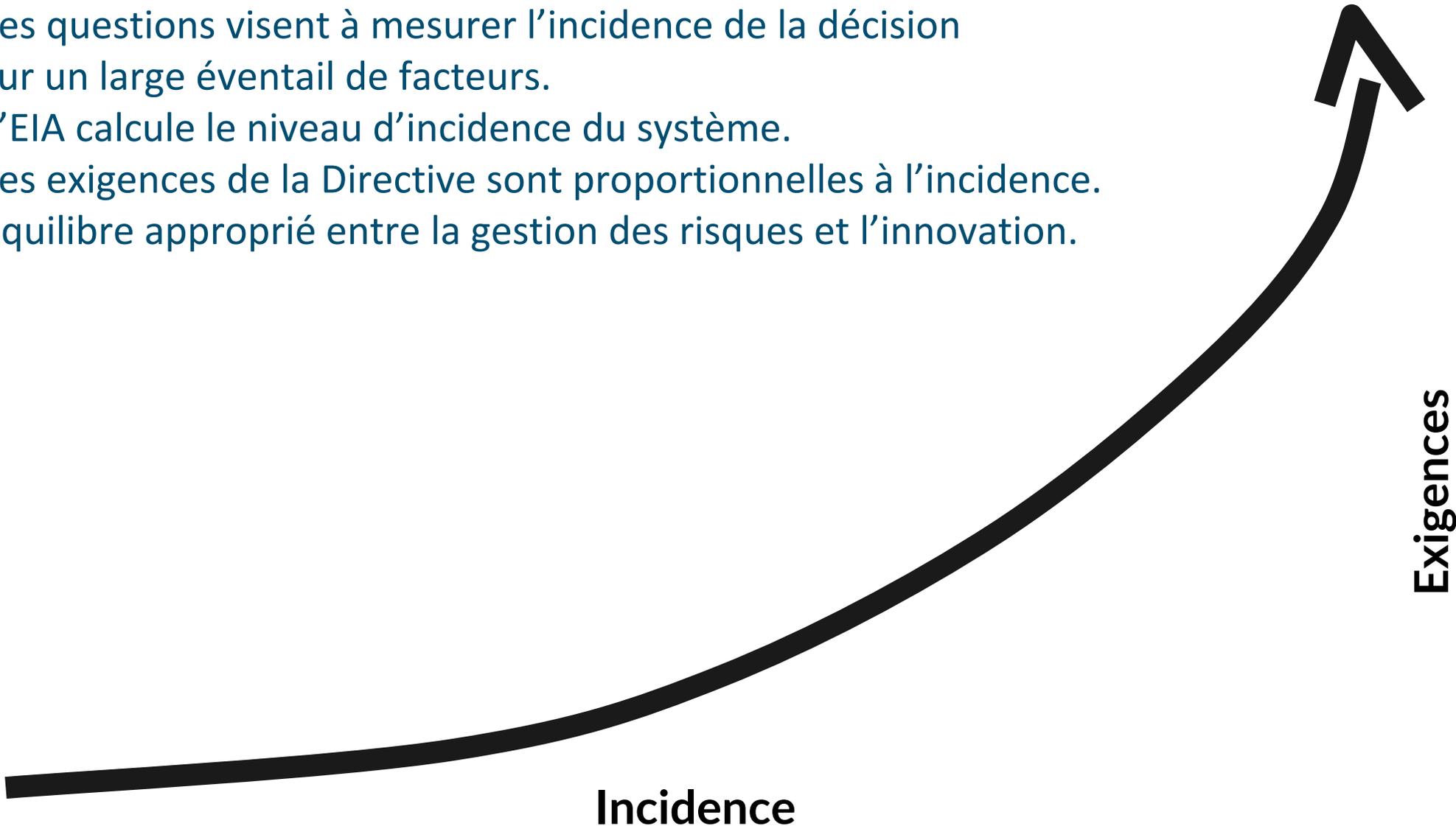
Impact Level: 1 Current Score: 20 Raw Impact Score: 20 Mitigation Score: 0

L'EIA : Aperçu du processus



L'EIA : Exigences progressives

- Les questions visent à mesurer l'incidence de la décision sur un large éventail de facteurs.
- L'EIA calcule le niveau d'incidence du système.
- Les exigences de la Directive sont proportionnelles à l'incidence.
- Équilibre approprié entre la gestion des risques et l'innovation.



L'EIA : Exigences progressives (exemple de l'annexe C)

Exigence	Niveau I	Niveau II	Niveau III	Niveau IV
Maillon humain	Des décisions peuvent être prises sans participation humaine directe.		Des décisions ne peuvent être prises sans qu'il y ait des points d'intervention humaine précis pendant le processus décisionnel. La décision définitive doit être prise par un humain.	
Avis	Aucune	Avis en langage simple publié par l'entremise de tous les modes de prestation de services utilisés (Internet, en personne, par courrier, par téléphone).	Un avis en langage simple publié par l'entremise de tous les modes de prestation de services utilisés (Internet, en personne, par courrier, par téléphone). De plus, publier de la documentation sur les sites Web pertinents au sujet du système décisionnel automatisé, en langage simple, décrivant : <ul style="list-style-type: none"> • le fonctionnement des composants; • la façon dont il appuie la décision administrative; • les résultats de tout examen ou audit; et • [u]ne description des données d'apprentissage ou un lien vers les données d'apprentissage anonymisées si ces données sont accessibles au public. 	

Services internes au GC

- La portée de la [Politique sur les services et le numérique](#) couvre tous les services, y compris ceux qui sont fournis aux employés fédéraux et à d'autres au gouvernement.
- La [Ligne directrice sur les services et le numérique](#) définit les services internes comme « des groupes d'activités et de ressources connexes que le gouvernement fédéral considère comme des services à l'appui de programmes ou nécessaires pour satisfaire aux obligations générales d'une organisation. »
- Le [Répertoire de services du GC](#) fournit la liste officielle des services du GC assujettis à la Politique sur les services et le numérique. Le SCT a déterminé qu'il était nécessaire d'inclure de façon uniforme les services internes dans le Répertoire.
- Les catégories et les termes utilisés pour classer et définir les services internes sont définis dans le Modèle de référence du gouvernement du Canada (MRGC) et le Modèle de référence stratégique du GC (MRS GC). Le MDGC et le MRS GC indiquent les programmes internes (p. ex., la gestion des RH) et les services internes connexes (p. ex., le recrutement) et les extrants des services (p. ex., les ressources).
- Les modèles de référence et le Répertoire de services du GC peuvent être utilisés pour informer le SCT des services qui pourraient être assujettis à la DPDA, si sa portée devait être élargie pour couvrir les services internes.