

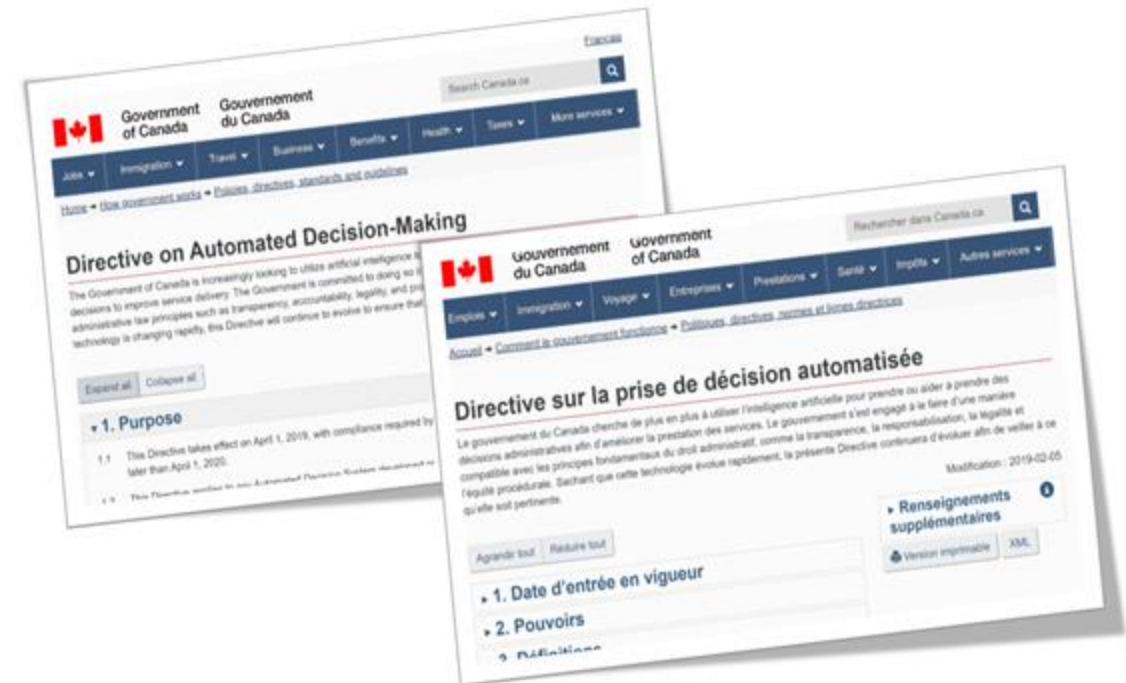


Troisième examen de la Directive sur la prise de décisions automatisée – Résumé des principales questions, des recommandations stratégiques et des modifications proposées

**Dossier de consultation
Automne 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 directive 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 2019-2010, 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*

Portée		Examen périodique	Clients touchés	Gouvernance des données	Biais du modèle
<ul style="list-style-type: none"> ▪ L'orientation externe exclut les décisions automatisées qui ont une incidence sur les employés du gouvernement fédéral. 	<ul style="list-style-type: none"> ▪ Le langage utilisé pour définir la portée doit être clarifié. 	<ul style="list-style-type: none"> ▪ Le délai actuel de six mois pour l'examen crée des défis stratégiques et opérationnels. 	<ul style="list-style-type: none"> ▪ Le renvoi aux Canadiens dans certaines parties de la directive ne reconnaît pas d'autres clients possibles. 	<ul style="list-style-type: none"> ▪ Les mesures d'assurance qualité ne répondent pas à la nécessité de tracer, de protéger, de conserver et d'éliminer les données utilisées et générées par un système. 	<ul style="list-style-type: none"> ▪ Les mesures de test de biais sont limitées aux données et ne tiennent pas compte d'autres sources possibles de biais telles que le modèle.
Inclusion	Explication	Raisons de l'automatisation	Examen par les pairs	Planification d'urgence	Date de publication de l'EIA
<ul style="list-style-type: none"> ▪ Les mesures soutenant les approches intersectionnelles de la conception et de la mise en œuvre des systèmes sont manquantes. 	<ul style="list-style-type: none"> ▪ Il n'y a pas de critères pour ce qui constitue une explication valable. ▪ L'approche de la publication des explications n'est pas claire. 	<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. 	<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'achèvement de l'examen par les pairs n'est pas clair. 	<ul style="list-style-type: none"> ▪ La terminologie n'est pas harmonisée avec la Politique sur la sécurité du Conseil du Trésor. 	<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.

* Une description plus détaillée de ces questions se trouve en annexe.

Recommandations stratégiques de haut niveau*

Gouvernance

Élargir la portée pour couvrir les services internes. (1)

Remplacer l'intervalle d'examen de six mois par un examen biennal et permettre à la DPI du Canada de demander des examens hors cycle. (3)

Transparence and responsabilité

Établir des critères d'explication à l'appui de l'exigence d'explication et les intégrer à l'EIA. (8)

Élargir l'EIA pour y inclure des questions sur les raisons poussant le ministère à adopter l'automatisation et les incidences potentielles sur les personnes en situation de handicap. (9)

Donner le mandat de publier des examens par les pairs complets ou résumés et d'exiger l'achèvement avant la production du système. (10)

Donner le mandat de diffuser les EIA avant la production d'un système. (12)

Assurance de la qualité

Mettre en place des mesures favorisant le suivi, la protection, la conservation et la destruction appropriés des données utilisées et générées par un système. (5)

Élargir l'exigence d'évaluer les biais pour couvrir les modèles. (6)

Mandater la réalisation de l'analyse comparative entre les sexes plus (ACS+) lors de l'élaboration d'un système. (7)

Cohérence des politiques

Préciser que la portée comprend des systèmes effectuant des évaluations relatives à des décisions administratives. (2)

Remplacer les renvois aux Canadiens par un langage plus englobant, comme les clients et la société canadienne. (4)

Harmoniser l'exigence d'urgence avec la terminologie pertinente établie dans la Politique sur la sécurité du Conseil du Trésor. (11)

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
Directive sur la prise de décisions automatisée (DPDA), section 1. (Date d'entrée en vigueur)	<p>Modifier le paragraphe 1.1 : « La présente Directive a été modifiée et entre en vigueur le [date de lancement], 2023, et doit être respectée au plus tard le [date; 6 mois après la date de lancement], 2023. »</p> <p>Ajouter une section : « les ministères exploitant des systèmes décisionnels automatisés développés ou achetés avant le [date de lancement] 2023 auront 12 mois pour se conformer aux exigences des paragraphes 6.2.3, 6.3.1, 6.3.X [gouvernance des données], 6.3.X [ACS+], et 6.3.4 de la Directive. »</p>	<p>Le SCT reconnaît le défi de s'adapter aux nouvelles exigences des politiques lors de la planification ou de l'exécution de projets qui y seraient assujettis. En réponse, une « période de grâce » de six mois est proposée pour donner aux ministères le temps de vérifier qu'ils satisfont à la Directive modifiée. Pour les systèmes qui sont déjà en place à la date de publication, le SCT propose d'accorder aux ministères une année complète pour se conformer aux nouvelles exigences de la Directive. L'introduction de cette période permettrait aux ministères de planifier l'intégration de nouvelles mesures dans les systèmes d'automatisation existants, notamment la publication d'examins par les pairs déjà achevés ou la mise en œuvre de nouvelles mesures de gouvernance des données pour les données d'entrée et de sortie. Pendant cette période, ces systèmes resteraient soumis aux exigences actuelles de la Directive.</p>
DPDA, section 5 (Portée)	<p>N° 1 – Modifier le paragraphe 5.1 : « La présente Directive s'applique aux systèmes qui fournissent des services externes ou internes, conformément à la Politique sur les services et le numérique. »</p>	<p>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.)</p>
DPDA, annexe A (Définitions)	<p>N° 1 – Ajouter la définition suivante : « Service interne : Un service où le client visé fait partie du gouvernement du Canada, y compris les employés fédéraux. »</p>	<p>La Directive ne définit pas les services internes, dont beaucoup pourraient être soumis sous une portée élargie. La Politique sur les services et le numérique ne définit que les services externes et les services internes intégrés. La définition proposée, qui s'inspire de la définition des services externes, souligne que les clients du secteur public comprennent les employés fédéraux. (Il existe d'autres types de services internes tels que les services intégrés internes où le client est une institution plutôt qu'un employé individuel.) Clarifier le sens de ce terme répondrait aux préoccupations des intervenants et aiderait à garantir une interprétation et une application cohérentes de la Directive modifiée.</p>
DPDA, annexe B (Niveaux d'évaluation de l'incidence)	<p>N° 1 – Ajouter un nouveau domaine d'incidence sous chacun des quatre niveaux d'incidence : « l'égalité, la dignité et l'autonomie des fonctionnaires fédéraux »</p>	<p>Le mécanisme d'incidence de l'annexe B de la Directive ne reconnaît pas les incidences possibles de l'automatisation sur les fonctionnaires fédéraux en milieu de travail. La mise à jour proposée définit l'égalité, la dignité et l'autonomie comme trois conditions de travail essentielles devant être respectées par les ministères lors de l'introduction de l'automatisation en milieu de travail. Les principes s'inspirent de la publication <i>Good Work Charter</i> (charte du bon travail) de l'Institute for the Future of Work (IFOW), qui fournit un cadre d'organisation pour l'aspiration, l'harmonisation et l'action pour façonner un monde du travail meilleur et plus juste. L'ajout de ce domaine d'incidence à la liste existante élargit la portée de l'approche de la Directive en matière d'évaluation de l'incidence, qui comprendrait les services internes qui ont une incidence sur les fonctionnaires fédéraux.</p>

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
ÉIA (Section sur les consultations)	<p>N° 1 – Modifier la question pour ajouter de nouvelles options aux listes pour les intervenants internes et externes : « Prévoyez-vous échanger avec l'un des groupes suivants? Intervenants internes (politique et planification stratégiques, gouvernance des données, politique relative aux programmes, etc.) : « Politique sur le numérique, Ressources humaines, Bureau du dirigeant principal des ressources humaines du SCT, Bureau de la dirigeante principale de l'information du SCT ».</p> <p>Intervenants externes (société civile, universités, industrie, etc.) : « Agents négociateurs, gouvernements d'autres administrations, organisations internationales, clients ou leurs représentants ».</p>	<p>Le SCT cherche à déterminer de nouveaux intervenants internes et externes pour tenir compte des cas d'utilisation qui touchent des services internes (qui peuvent être assujettis à la Directive modifiée) et mettre en évidence les types d'intervenants qui manquaient auparavant. Ces changements permettront aux ministères de mieux définir et faire participer les intervenants potentiellement pertinents au sein de leurs institutions respectives et à l'extérieur du gouvernement sur leurs projets d'automatisation.</p> <p>Non exhaustives, ces listes fournissent des exemples de bureaux et d'institutions que les ministères devraient consulter tôt dans le cycle de vie du projet ou du système pour assurer la conformité avec les lois et politiques applicables et pour déceler les meilleures pratiques et les leçons apprises dans d'autres organisations.</p>
ÉIA (À propos de la section sur les décisions)	<p>N° 1 – Modifier la question : « La décision concerne-t-elle l'une des catégories ci-dessous (cochez toutes celles qui s'appliquent) : » en ajoutant une nouvelle option à la liste : « Emploi (recrutement, embauche, promotion, évaluation du rendement) ».</p>	<p>L'option proposée définit l'emploi comme une catégorie clé de cas d'utilisation qui relèverait du domaine de la Directive modifiée, qui s'appliquerait à l'automatisation des services internes impliquant une prise de décision administrative. Le recrutement, l'embauche, la promotion et l'évaluation du rendement sont des exemples de ces services. Bien qu'il existe d'autres types de services internes susceptibles d'être assujettis à la Directive, le troisième examen cherche à donner la priorité à l'évaluation et à l'atténuation des risques découlant des contextes d'emploi.</p>
DPDA, section 5. (Portée)	<p>N° 2 – Modifier le paragraphe 5.2 : « La présente Directive s'applique à tout système, outil ou modèle statistique utilisé pour recommander ou prendre une décision administrative ou effectuer une évaluation connexe au sujet d'un client. »</p>	<p>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 assujetti 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 introduit le concept d'évaluation pour établir un seuil permettant d'évaluer si les cas d'automatisation partielle relèvent de la portée de la Directive. Les différentes manières dont un système peut effectuer une évaluation d'un client seront exposées dans l'orientation. Il convient de prendre note que la modification ne change pas la portée actuelle; elle ne fait que l'exprimer en termes plus appropriés.</p>

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, section 1 (Date d'entrée en vigueur)	N° 3 – Modifier le paragraphe 1.2 : « La Directive sera examinée tous les deux ans, et 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 que cause la disposition 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 assujetties 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 , les institutions fédérales et la société canadienne , 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 terme 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 : la définition de client dans cette politique tient compte fidèlement du large éventail de personnes – citoyens canadiens et autres – qui peuvent être soumis à des décisions administratives automatisées. Le SCT propose aussi d'ajouter une référence à la société canadienne pour mettre en évidence la dimension sociale des décisions automatisées, qui pourraient avoir des répercussions systémiques sur certaines collectivités ou sur le grand public canadien. Ensemble, ces changements communiquent l'engagement du gouvernement envers l'inclusion, renforcent la responsabilité démocratique, et assurent la cohérence avec la terminologie de la politique en place.
DPDA, section 6 (Exigences)	N° 5 – Ajouter un nouveau paragraphe 6.3 intitulée « 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.

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
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 déceler 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 extraits 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, section 6 (Exigences)	N° 7 – Ajouter un nouveau paragraphe sous 6.3 : « Réaliser une analyse comparative entre les sexes plus (ACS+) lors de l’élaboration du système décisionnel automatisé, conformément à ce qui est prévu à l’annexe C. ».	<p>La Directive actuelle n’exige pas explicitement que les ministères entreprennent une ACS+ pour leurs projets d’automatisation. Cependant, l’ÉIA demande aux utilisateurs s’ils ont entrepris une ACS+ pour les données recueillies ou utilisées par le système décisionnel automatisé. Comme elle applique uniquement la méthodologie de l’ACS+ aux données d’entrée, l’efficacité de cette question pour garantir des pratiques équitables dans le développement et l’utilisation de systèmes décisionnels automatisés semble limitée.</p> <p>L’établissement d’une exigence propre à l’ACS+ assurerait l’uniformité dans l’application de l’optique d’ACS+ à la prise de décisions automatisée et aiderait à favoriser les approches intersectionnelles dans la conception, le développement et l’utilisation de systèmes de décision automatisés. Les éléments proposés à l’annexe C aideraient à assurer une large application de la méthodologie aux données, aux systèmes, aux décisions et aux autres éléments d’un projet d’automatisation. Le SCT a proposé cette mesure pour les systèmes aux niveaux d’incidence II et IV considérant que les résultats de l’ACS+ seront vraisemblablement peu importants pour les systèmes de niveau I.</p>
DPDA, annexe C (Exigences par niveau d’incidence)	<p>N° 7 – Ajouter une nouvelle section sur l’analyse comparative entre les sexes plus. Les exigences dans ce domaine s’appliqueraient aux systèmes aux niveaux d’incidence II à IV : « Veillez à ce que votre analyse comparative entre les sexes plus aborde les enjeux suivants :</p> <ul style="list-style-type: none"> • les incidences du projet d’automatisation (incluant le système, les données et la décision) sur le sexe ou d’autres facteurs d’identité; • les mesures prévues ou actuelles pour faire face aux risques définis dans le cadre de l’analyse comparative entre les sexes plus. » 	Voir la justification des modifications à l’ACS+ (section 6).

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, annexe C (Exigences par niveau d'incidence)	<p>N° 8 – Modifier les mesures d'explication pour le niveau d'incidence I : « En plus de toute exigence juridique applicable, veiller à ce qu'une explication statique des résultats communs des décisions soit publiée. L'explication doit fournir une description générale des éléments suivants :-</p> <ul style="list-style-type: none"> • le rôle du système dans le processus décisionnel; • les données d'entrée, leur source et leur méthode de collecte; • les critères utilisés pour évaluer les données d'entrée 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. <p>Ces renseignements doivent être librement disponibles en langage simple par l'entremise de l'évaluation de l'incidence algorithmique et faciles à trouver par l'entremise du site Web d'un ministère. »</p>	<p>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</i> pour une République numérique 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.</p> <p>Les critères pour le niveau I sont conçus prenant en compte les attentes à l'égard des explications publiques. Le texte proposé précise que les explications des systèmes de niveau I doivent être publiées en langage clair dans le cadre de l'évaluation de l'incidence algorithmique (ÉIA) et faciles à trouver sur le site Web d'un ministère. Cela vise à renforcer la découvrabilité et l'accessibilité des explications des résultats de décisions communes. Comme pour tout renseignement contenu dans l'ÉIA, les explications publiques ne comprendraient aucuns renseignements personnels ou renseignements sensibles. Elles ne remplacent pas non plus les explications destinées à un client, que les ministères doivent toujours fournir pour répondre aux exigences juridiques et comme il est requis pour les niveaux d'incidence II à IV.</p>

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, annexe C (Exigences par niveau d'incidence)	<p>N° 8 (suite) – Modifier les mesures d'explication pour les niveaux d'incidence II à IV : « En plus de toute exigence juridique applicable, veiller à ce qu'une explication significative soit fournie au client avec toute décision qui a conduit à un refus de prestation ou de service, ou qui impliquait une mesure réglementaire. L'explication doit informer le client dans un langage simple sur les éléments suivants :</p> <ul style="list-style-type: none"> • le rôle du système dans le processus décisionnel; • les données d'entraînement et sur les 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; • la justification de la décision administrative. <p>Une description générale de ces éléments devrait également être disponible dans l'évaluation de l'incidence algorithmique et facile à trouver dans le site Web d'un ministère. »</p>	<p>Les critères d'explications proposées pour les niveaux II à IV permettent de mieux distinguer les explications publiques et privées. Celles relatives aux systèmes des ces niveaux doivent être destinées aux clients afin de leur permettre de comprendre et de contester une décision, et de respecter les obligations juridiques procédurales. Ces explications concernent généralement une personne en particulier (qui est soumise à une décision automatisée) et sont donc susceptibles d'impliquer des droits procéduraux.</p> <p>Les critères diffèrent principalement de ceux proposés pour le niveau I en ce sens qu'ils obligeraient les ministères à articuler le raisonnement derrière une décision. Ce critère vise à veiller à ce que les clients comprennent pourquoi, compte tenu des résultats d'un système (et éventuellement du jugement d'un agent), une décision a été prise d'une certaine manière. (Étant donné que ces justifications concernent toujours une personne en particulier, ce critère n'est pas proposé pour les exigences de niveau I, qui concernent les explications publiques.)</p> <p>S'appuyant sur le langage proposé pour les explications de niveau I, le SCT cherche également à veiller à ce qu'une version publique des explications fournies aux clients soit rédigée et publiée dans le cadre de l'ÉIA. Fournir aux intervenants publics une « description générale » des mêmes éléments que ceux énoncés dans les critères renforcerait la transparence et la responsabilité algorithmiques. Une telle disposition améliorerait également l'efficacité de l'ÉIA en soi, qui demanderait aux utilisateurs d'expliquer à un large public comment leurs systèmes prennent ou appuient des décisions, élargissant davantage les possibilités de dialogue public sur l'utilisation de l'IA au gouvernement fédéral.</p>

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
ÉIA (À propos de la section sur les données)	N° 8 – Ajouter la question suivante : « Veillez décrire les données d’entrée recueillies et utilisées par le système, leur source et leur méthode de collecte. [texte libre] »	Cette question est tirée des critères d’explication proposés, que le SCT cherche à intégrer dans l’ÉIA. Voir les justifications des mises à jour des exigences d’explication (annexe C).
ÉIA (Facteur opérationnel / Section sur les incidences positives)	<p>N° 9 – Ajouter une nouvelle série de questions sur les raisons de l’automatisation :</p> <ul style="list-style-type: none"> • « À quels besoins des utilisateurs le système doit-il répondre et comment va-t-il les satisfaire? Si possible, décrivez comment les besoins des utilisateurs ont été définis. [texte libre] ». • « Dans quelle mesure le système répondra-t-il efficacement aux besoins des utilisateurs? [Légèrement efficace; Modérément efficace; Très efficace] ». • « Veillez décrire les améliorations, les bénéfices ou les avantages que vous attendez de l’utilisation d’un système automatisé, lesquels pourraient inclure des indicateurs de programme pertinents et des objectifs de rendement. [texte libre] ». • « Veillez décrire comment vous vous assurerez que le système se limite à répondre aux besoins des utilisateurs identifiés ci-dessus? [texte libre] ». • « D’autres processus non automatisés ont-ils été envisagés? [Oui/Non] ». • « Si l’on a envisagé des processus non automatisés, pourquoi l’automatisation a-t-elle été considérée comme l’option préférée? [texte libre] ». • « Quelles seraient 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] ». • « Veillez décrire tout avantage public attendu du système. [texte libre] » 	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 ou fournir des avantages publics. 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.
ÉIA (Section sur le profil de risque)	N° 9 – Ajouter une nouvelle question : « L’utilisation du système posera-t-elle des risques importants pour les personnes en situation de handicap? [Oui/Non] Si oui, veuillez décrire les risques liés à l’accessibilité et toute mesure d’atténuation prévue ou existante. [texte libre] ».	La question proposée intègre les considérations d’accessibilité dans l’ÉIA en veillant à ce que les ministères identifient, évaluent et atténuent les risques de l’automatisation pour les personnes en situation de handicap. Cela permet aux ministères de se conformer aux exigences de la <i>Loi canadienne sur l’accessibilité</i> et complète les mesures d’ACS+ proposées pour renforcer davantage l’équité et l’inclusion dans la prise de décision automatisée.

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, section 6 (Exigences)	N° 10 – Modifier le paragraphe 6.3.4 : « Consulter l’expert qualifié approprié pour demander une évaluation du système décisionnel automatisé et publier l’ensemble de l’examen ou un résumé en langage clair des constatations avant la production du système , conformément à ce qui est prévu à l’annexe C. ».	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 décisionnels 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 examens par les pairs publiés (ou leur sommaire) 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).
DPDA, annexe C (Exigences par niveau d’incidence)	<p>N° 10 – Modifier les mesures d’examen par les pairs pour les niveaux d’incidence II et III : « Consulter au moins l’un des experts suivants et publier l’examen complet ou un résumé des constatations en langage clair sur un site Web du gouvernement du Canada : »; « OU Publier les spécifications du système décisionnel automatisé dans une revue à comité de lecture. Lorsque l’accès à l’examen publié est restreint, veiller à ce qu’un résumé en langage simple des conclusions soit 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 : « Consulter au moins deux des experts suivants et publier l’examen complet ou un résumé des constatations en langage clair sur un site Web du gouvernement du Canada : »; « Publier les spécifications du système décisionnel automatisé dans une revue à comité de lecture. Lorsque l’accès à l’examen publié est restreint, veiller à ce qu’un résumé en langage simple des conclusions soit librement disponible. »</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 l’ensemble de l’examen par les pairs ou un résumé des constatations sur un site Web du gouvernement. 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.

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, section 6 (Exigences)	<p>N° 11 – 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° 11 (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>	
DPDA, section 6 (Exigences)	<p>N° 12 – Modifier le paragraphe 6.1.1 : « Effectuer et publier les résultats définitifs d'une évaluation de l'incidence algorithmique avant la production de tout système décisionnel automatisé. »</p>	<p>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 AIA 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.</p>

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
ÉIA (Section sur l'évaluation de l'incidence)	<p>N° 1, 2 et 8 – Remplacer les deux premières questions actuellement dans cette section de l'ÉIA par :</p> <ul style="list-style-type: none"> • « Lequel des énoncés suivants décrit le mieux le type d'automatisation que vous envisagez? <ul style="list-style-type: none"> • Automatisation complète (le système prendra une décision administrative) • Automatisation partielle (le système contribuera à la prise de décisions administratives en appuyant un agent par l'entremise d'évaluations, de recommandations, de décisions intermédiaires ou d'autres extrants.) <p>Veillez décrire le rôle du système dans le processus décisionnel. [texte libre] »</p> <p>Ajouter les questions suivantes :</p> <ul style="list-style-type: none"> • « Veillez décrire les critères utilisés pour évaluer les données des clients et les opérations effectuées pour les traiter. [texte libre] » • « Veillez décrire l'extrait produit par le système et tous renseignements pertinents nécessaires pour l'interpréter dans le contexte de la décision administrative. [texte libre] » • « Le système effectuera-t-il une évaluation ou une autre opération qui, autrement, ne serait pas effectuée par un humain? [Oui/Non] Dans l'affirmative : Veillez décrire les fonctions pertinentes du système. [texte libre] » • « Les incidences que la décision aura sur l'égalité, la dignité et l'autonomie des fonctionnaires fédéraux seront probablement : [incidence faible à nulle, incidence modérée, incidence élevée, incidence très élevée] Veillez justifier votre réponse (selon l'option sélectionnée ci-dessus) [texte libre] ». <p>Modifier la question : « Le système effectuera-t-il des décisions ou des évaluations qui exigent du jugement ou de la discrétion? »</p>	<p>Il est nécessaire d'aborder les redondances et les contradictions dans les questions de l'ÉIA sur le rôle d'un système dans un processus décisionnel : les deux premières questions de cette section sont non seulement redondantes, mais associent également à tort une automatisation partielle à un niveau plus élevé de risque que l'automatisation complète. Les mises à jour proposées répondent à cette préoccupation conformément à la proposition du SCT de modifier le langage utilisé pour décrire la portée de la Directive. Le SCT propose une question consolidée qui définit le type d'automatisation poursuivi et demande aux utilisateurs de décrire plus en détail le rôle de leur système dans la prise de décisions. Cette description ferait partie de l'explication publique requise à l'annexe C de la Directive modifiée. Voir la justification des modifications des exigences d'explication sous l'annexe C.</p> <p>Les mises à jour dans cette section introduisent également de nouvelles questions évaluant si un système effectuerait de nouvelles évaluations qui pourraient ne pas être réalisables dans un processus non automatisé, et toute incidence possible sur les fonctionnaires fédéraux (conformément aux mises à jour proposées à l'annexe B). La question sur les nouvelles évaluations vise à définir de nouveaux processus qui peuvent ne pas faire partie d'un processus décisionnel existant qui, en raison de leurs propres risques, doivent être déterminés dans le cadre de l'ÉIA.</p> <p>La question relative aux incidences sur les fonctionnaires fédéraux vise à rendre compte de l'application de la Directive modifiée aux services internes touchant les fonctionnaires fédéraux. Cela correspond au domaine d'incidence proposée à l'annexe B de la Directive. Consultez la justification de cette modification pour plus d'informations.</p> <p>Comme la demande d'articuler le rôle du système dans le processus décisionnel, les questions concernant les données d'entrée et de sortie proviennent des critères d'explication proposés. Voir la justifications des modifications des exigences d'explication sous l'annexe C.</p>

Modifications proposées à la Directive

Instrument et section	Modification proposée (texte provisoire)	Justification
DPDA, section 10 (Références)	<p>N° 1, 7 et 9 – Modifier la section 10.1 : « Loi canadienne sur l’accessibilité »</p> <p>Modifier la section 10.2 : « Politique sur la gestion des personnes »</p>	<p>Les ajouts proposés à la section Références visent à compléter trois modifications: l’élargissement de la portée de la Directive pour couvrir les services internes, l’exigence de remplir une ACS+ lors de l’élaboration d’un système décisionnel automatisé et l’introduction d’une question relative aux incidences sur l’accessibilité dans l’ÉIA.</p> <p>La Politique sur la gestion des personnes régit l’organisation et la gestion de l’effectif de la fonction publique fédérale. Elle établit des règles pour la prestation de services internes dans le domaine des ressources humaines, dont beaucoup relèveraient du champ d’application de la Directive modifiée (dans la mesure où elles comprennent des décisions administratives). Ces exigences devraient être prises en compte dans les évaluations de l’incidence des systèmes déployés dans ce domaine (p. ex., pour appuyer le recrutement, l’embauche ou la gestion du rendement). La Politique est donc pertinente pour l’adoption efficace des mesures de la Directive modifiée.</p> <p>La <i>Loi canadienne sur l’accessibilité</i> vise à obtenir un Canada sans obstacles d’ici 2040. La Loi profite à tous les Canadiens, en particulier aux personnes en situation de handicap, grâce à la reconnaissance, à l’élimination et à la prévention proactive des obstacles à l’accessibilité dans une gamme de domaines, y compris les technologies de l’information et des communications (TIC). L’exigence proposée de remplir une ACS+ et la nouvelle question d’ÉIA concernant l’incidence possible d’un projet d’automatisation sur les personnes en situation de handicap s’inspirent des obligations juridiques établies dans la Loi. La pertinence de la Loi comme point de référence dans la Directive va toutefois au-delà de ces mesures. La Loi donne la priorité à l’accessibilité dans la conception et la mise en œuvre de programmes et de services, ce qui en fait une source de droit pertinente pour toute utilisation possible de systèmes décisionnels automatisés dans la prestation de services.</p>

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

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 – automne	2022 – Hiver 2023
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 les différentes équipes au 3^e de façon continue. 	<ul style="list-style-type: none"> Affiner les recommandations et les modifications. Consulter les ministères, le Commissariat à la protection de la vie privée du Canada, les agents des services des ressources humaines, et les agents négociateurs. Collaborer avec des chercheurs, la société civile et d'autres gouvernements. 	<ul style="list-style-type: none"> Démarrer le processus de gérance intégrée des politiques du BDPI. Demander l'approbation des comités supérieurs. Demander l'approbation de la DPI du GC et du secrétaire. Publier la directive modifiée. Appuyer les ministères dans la mise en conformité (permanent).

Personnes-ressources

Secrétariat du Conseil du Trésor du Canada

Bureau de la dirigeante principale de l'information

ai-ia@tbs-sct.gc.ca

Annexe

Questions clés identifiées dans le 3^e examen

- **Portée** : La portée de la directive 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 directive 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 directive 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 directive 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 directive. 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 directive ne tiennent pas compte des préjugés 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.
- **Inclusion** : Bien qu'elle nécessite des essais de biais, la directive ne prescrit aucun cadre ou méthode pour examiner l'inclusion dans l'automatisation. Cela limite le potentiel de la directive à garantir une prise de décision juste et équitable.
- **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. L'exigence ne prescrit pas non plus une approche de publication.

Questions clés identifiées dans le 3^e examen

- **Raisons de l'automatisation** : La directive 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 directive 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 directive 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 sur la formation 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 décisions
- 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
- Assurer une intervention humaine

Recours

Corriger

- Options de recours pour contester les décisions


Aperçu des exigences de la directive

- Publiée en 2019, la directive 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 directive 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 directive officialise la responsabilisation algorithmique en maintenant des sous-ministres adjoints (SMA) qui supervisent les projets d'automatisation pertinents et qui sont chargés de se conformer aux exigences de la Politique.
- Les institutions fédérales assujetties à la directive 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 directive

- 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

- 
- 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.

L'automatisation des aspects du processus d'établissement des faits ou d'analyse peut influencer les décisions ultérieures.

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-responsable-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.

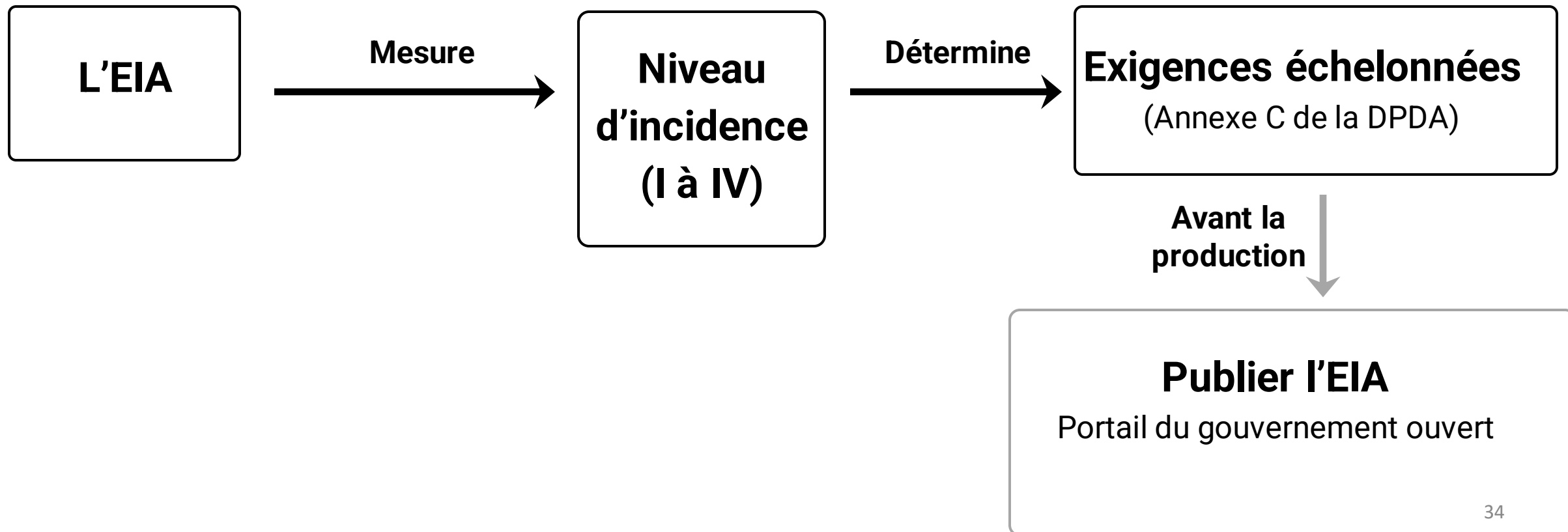
The impacts that the decision will have on the rights or freedoms of individuals will likely be:

Little to no impact

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

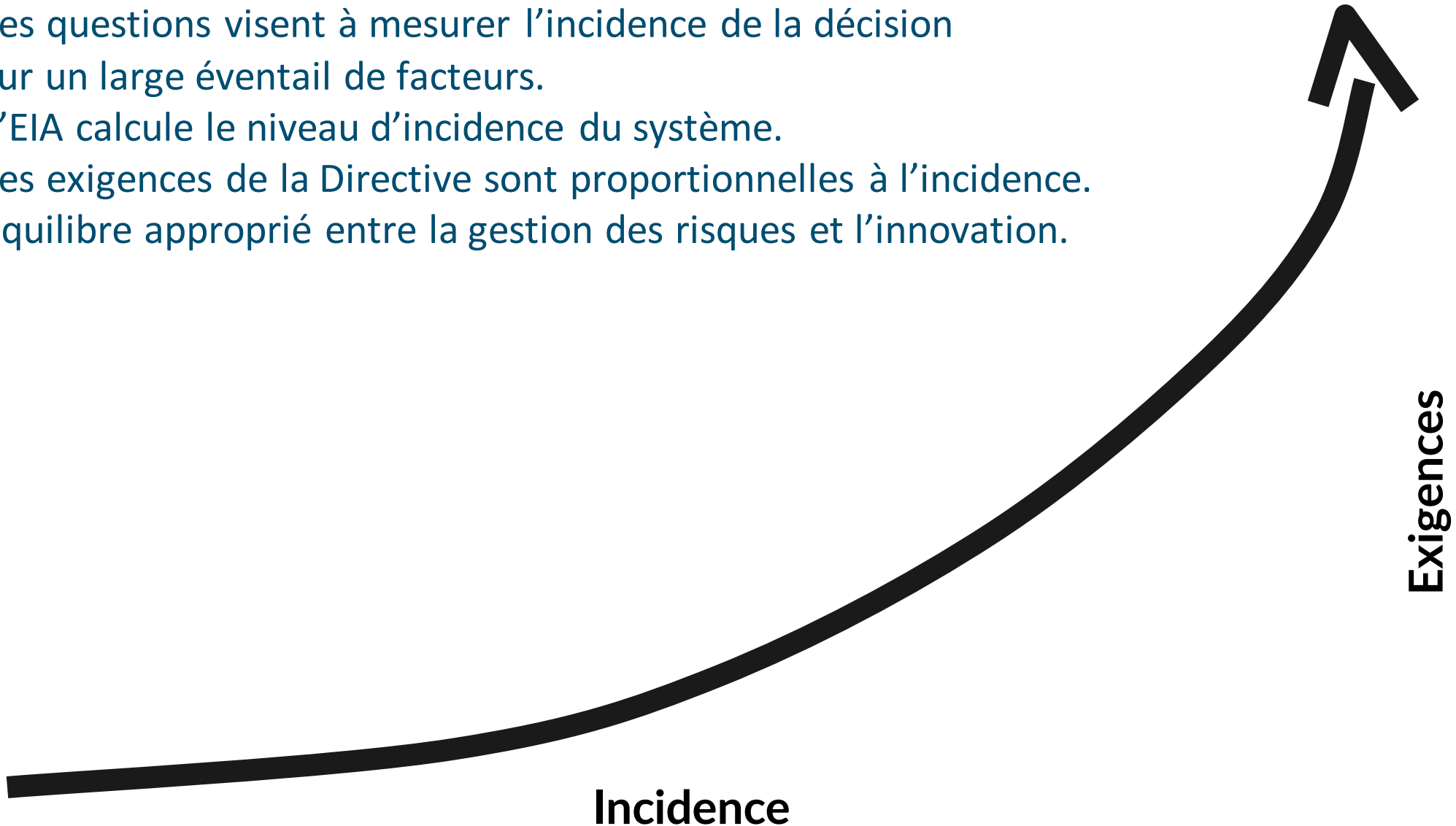
Impact Level: 1 Current Score: 20 Raw Impact Score: 20 Mitigation Score: 0

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 de formation ou un lien vers les données de formation 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 directive, si sa portée devait être élargie pour couvrir les services internes.