

# CODE D'ÉTHIQUE

de la valorisation de données



**RECOMMANDATIONS**

## VUE D'ENSEMBLE

Le «Code d'éthique de la valorisation de données» se compose des documents suivants: 1) Vue d'ensemble; 2) Principes fondamentaux; 3) Recommandations; 4) Mise en œuvre; 5) Informations générales.

Le document **RECOMMANDATIONS** formule des recommandations concrètes pour une valorisation de données éthique, structurée sur les quatre phases du cycle de vie des données. Celles-ci sont regroupées en fonction des trois orientations éthiques de base et les trois valeurs procédurales. Pour chaque groupe, les intentions des recommandations sont explicitées avant que ne suivent les recommandations concrètes. Enfin, un exemple est donné pour chaque groupe afin d'illustrer à quoi pourrait ressembler la mise en œuvre concrète des recommandations. Des explications sur le cycle de vie des données et les orientations de base se trouvent dans le document «Principes fondamentaux». Le document est disponible en allemand, anglais, français et italien.

### MENTION LÉGALE

Le «Code d'éthique de la valorisation de données» a été élaboré par le groupe d'experts d'éthique des données de la Swiss Alliance for Data-Intensive Services. Équipe de rédaction: Markus Christen, Christoph Heitz, Tom Kleiber, Michele Loi (direction de la rédaction). Traduction française: Jean-Gabriel Piguet. Graphiques: Ana Nicolasa Caduff. Statut: 2020.

© Swiss Alliance for Data-Intensive Services, 2020.

ISBN 978-3-9522703-1-8; [www.data-service-alliance.ch/codex](http://www.data-service-alliance.ch/codex)

Licence: Attribution 4.0 International (CC BY 4.0).

# TABLE DES MATIERES

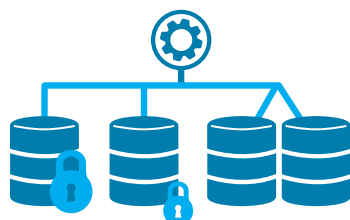


## ÉTAPE 1

### Génération et acquisition des données

Orientations éthiques fondamentales  
Valeurs procédurales

page 4



## ÉTAPE 2

### Stockage et gestion des données

Orientations éthiques fondamentales  
Valeurs procédurales

page 10



## ÉTAPE 3

### Analyse de données

Orientations éthiques fondamentales  
Valeurs procédurales

page 16



## ÉTAPE 4

### Utilisation des connaissances

Orientations éthiques fondamentales  
Valeurs procédurales

page 22

# ÉTAPE 1: GÉNÉRATION ET ACQUISITION DES DONNÉES

## ORIENTATIONS ÉTHIQUES FONDAMENTALES

### PRÉVENTION DES DOMMAGES

Les recommandations relatives au principe fondamental de prévention des dommages visent à garantir qu'aucun dommage n'est causé à l'utilisateur par la manière dont les données sont collectées. Des questions sur le préjudice éventuel peuvent se poser si l'interaction avec l'utilisateur est conçue en utilisant les résultats de la recherche comportementale de telle manière que cela encouragera l'utilisateur à partager davantage de données sur lui-même ; c'est ce qu'on appelle les «nudges» ou «coups de pouce». Si la conception d'une interface utilisateur est délibérément conçue pour faire une communication d'informations par l'utilisateur contraire à ses intérêts, on appelle cela un «dark pattern» ou «interface manipulatrice».



#### Recommandations

- Évitez d'exploiter l'inertie, le manque d'attention ou d'autres faiblesses psychologiques des personnes pour obtenir d'elles plus d'informations personnelles que nécessaire pour fournir votre service.
- Vérifiez la conception du site et l'expérience de l'utilisateur pour déceler les éventuelles interfaces manipulatrices et supprimez-les.
- Si vous utilisez des coups de pouce, vous devez prouver à l'avance ou ultérieurement qu'ils servent également les intérêts de l'utilisateur (par exemple, accroître la cybersécurité, améliorer le service fourni).

### EXEMPLE - ÉVITER LES INTERFACES TROMPEUSES

Les interfaces trompeuses sont faites pour tromper l'utilisateur et – dans le pire des cas – le conduire à une action qu'il ne voulait pas effectuer au départ. Pour déclencher ce «comportement non voulu», des modèles de comportement intelligents sont exploités et les utilisateurs sont manipulés par des boutons intelligemment placés ou des menus déroulants trompeurs. Un exemple est celui d'une entreprise qui voulait vendre une police d'assurance correspondante à un client après qu'il a réservé un voyage. L'option «pas d'assurance» est bien cachée sous la lettre «P» dans une liste alphabétique des pays dans lesquels l'assurance doit s'appliquer. A contrario, les interfaces devrait faciliter leur utilisation en recourant à plus de logique. Cela implique que, dans ce cas précis, la décision initiale «oui» ou «non» est séparée des options possibles qu'un «oui» implique.



## JUSTICE

Les recommandations relatives au principe fondamental de justice devraient assurer un certain degré de réciprocité, de sorte que lorsque des données sont collectées, les personnes concernées se voient offrir des avantages proportionnels aux données demandées.

### Recommandations

- Faites clairement comprendre à vos clients la valeur de la contrepartie que vous recevez en divulguant leurs données.
- Si la collecte de certaines données résulte d'exigences légales, indiquez-le explicitement.



### EXEMPLE - DOSSIER ÉLECTRONIQUE DU PATIENT (DEP)

Un médecin ayant son propre cabinet s'est inscrit au système DPE afin que ses patients puissent être (encore) mieux traités à l'avenir. Si un patient a créé un DEP individuel, le médecin peut saisir les données relatives à la santé et au traitement du patient. Cependant, le patient U n'est pas sûr de vouloir ouvrir un DEP et craint que ses données ne soient utilisées à mauvais escient. Eu égard au principe fondamental de justice, le médecin peut convaincre le patient U comme suit:

- Il ou elle montre au patient le bénéfice individuel: grâce au DEP, l'accès aux documents relatifs au traitement est garanti à tout moment et les doubles emplois ou les traitements inutiles peuvent être évités.
- Il ou Elle indique clairement au patient qu'il peut déterminer qui a accès à quelles données de santé. Ce n'est qu'en cas d'urgence que le DEP peut être consulté sans le consentement du patient. Les données sont stockées de manière décentralisée uniquement sur des serveurs en Suisse, l'échange d'informations n'ayant lieu que par le biais d'une connexion sécurisée et tous les processus ont été spécialement certifiés.
- Enfin, il ou elle fait également référence à la possibilité pour les patients de mettre facilement leurs données de santé à la disposition de la recherche s'ils le souhaitent.

## AUTONOMIE

Les recommandations relatives au principe fondamental d'autonomie devraient donner aux utilisateurs un degré suffisant d'autodétermination informationnelle.

### Recommandations

- Fournissez à vos clients des outils faciles à utiliser, qui leur permettent de prendre des décisions concernant le consentement à l'utilisation des données.
- À cette fin, il convient d'élaborer des formulaires de consentement informé intuitifs et, si possibles, adaptables en fonction des choix des utilisateurs.
- Permettez aux personnes d'adapter ou de révoquer leur consentement plus tard.
- Si possible, permettez la modification des préférences concernant la finalité du traitement des données.



### EXEMPLE – DIFFÉRENTS NIVEAUX DE CONSENTEMENT

Une plateforme web fournit à des groupes de patients des informations sur la recherche médicale concernant des maladies spécifiques, des possibilités de réseautage social et des applications d'auto-surveillance. La plateforme recueille, analyse et commercialise ces données d'utilisateurs à des fins commerciales, mais aussi pour la création de recommandations personnalisées, qui sont cruciales pour la fonctionnalité et l'attrait de la plateforme. Le modèle commercial de la plateforme consiste à fournir des données anonymes et personnalisées sur les patients aux sociétés pharmaceutiques afin de recruter des patients pour tester de nouveaux médicaments et de collaborer avec l'industrie pharmaceutique et les universités pour mener des recherches comportementales (par exemple sur l'attitude des patients à l'égard des thérapies et de leurs effets secondaires).

- L'entreprise propose différents niveaux de consentement: les patients peuvent donner leur consentement à l'utilisation de leurs données pour les recommandations sans que les données personnelles ne soient divulguées à l'industrie.
- Les utilisateurs peuvent passer en revue leurs différents consentements et préférences à tout moment. Un résumé peut être facilement accessible via le tableau de bord du patient.
- Un utilisateur peut décider à tout moment de ne pas partager les données avec l'industrie – si les données anonymes déjà partagées ne sont pas affectées, les données personnelles ne sont plus partagées une fois que le consentement est révoqué par voie électronique.



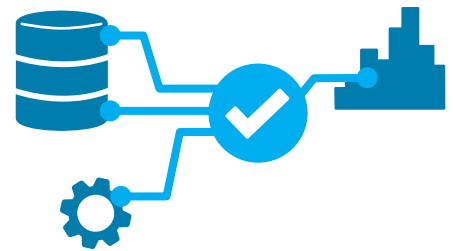
## VALEURS PROCÉDURALES

### CONTRÔLE

Les recommandations relatives au principe procédural fondamental de contrôle visent à garantir que toutes les informations pertinentes soient fournies pendant la production ou l'acquisition des données afin de garantir l'évaluation de l'utilisation éthique des données.

#### Recommandations

- Documentez de manière appropriée vos sources de données et toute restriction d'utilisation des données.
- Si d'autres entreprises utilisent vos données, fournissez-leur les informations dont elles ont besoin pour évaluer l'origine des données et les éventuelles restrictions d'utilisation.



### EXEMPLE - DES DONNÉES OUVERTES

#### POUR DE NOUVEAUX MODÈLES COMMERCIAUX

Un grand prestataire de services de transport décide de mettre les données collectées au sein de sa propre entreprise à la disposition de tiers pour un usage secondaire, gratuitement et sous une forme lisible par machine (Open Data). L'objectif est de soutenir les acteurs créatifs du marché dans le développement de nouvelles offres d'informations et de services attrayantes. On espère ainsi promouvoir de nouveaux modèles commerciaux avec des offres d'informations ou de services innovantes et, en fin de compte, accroître l'attrait de l'environnement de marché de l'entreprise. L'entreprise définit les conditions d'utilisation, telles que l'obligation de citer la source, l'obligation pour l'utilisateur des données de mettre à jour les données brutes (si de nouvelles données ont été mises à disposition par l'entreprise) afin d'éviter que les modèles commerciaux ne soient basés sur des données obsolètes, et les obligations de respecter la protection des données - c'est-à-dire notamment de s'abstenir de toute tentative d'extraire des données des informations désanonymisées sur les clients de l'entreprise.

## TRANSPARENCE

Les recommandations relatives au principe procédural fondamental de transparence visent à garantir que l'exigence de protection des données que constitue le consentement éclairé puisse être satisfaite d'une manière aussi simple, claire et accessible que possible pour les utilisateurs.



### Recommandations

- Trouvez des moyens simples d'expliquer quelles données vous collectez, votre politique de confidentialité et comment vous allez utiliser les données de vos clients (dans quel but et quels sont les résultats que vous en attendez).
- Utilisez des résumés et d'autres moyens pour expliquer les points clés de vos conditions générales (CG).
- Travaillez avec d'autres entreprises afin de créer de nouvelles normes pour des conditions générales plus lisibles et plus compréhensibles.

### EXEMPLE - PHOTOGRAPHE

Un photographe propose entre autres des séances de photos. Leurs conditions générales de vente (CGV) stipulent que les photos peuvent être mises à la disposition des clients sous forme numérique via un fournisseur de services en ligne externe afin que les clients puissent les télécharger à partir de là. Le photographe peut conserver les photos pendant un an après le traitement de la commande afin qu'il soit possible de les commander à nouveau et a également le droit d'utiliser les photos à ses propres fins, notamment pour son site web, ses affiches et autres. Les droits d'auteur des photos restent la propriété du photographe. Une cliente souhaite maintenant recevoir des informations sur les conditions générales. Le photographe peut réaliser ce souhait de la manière suivante:

- Il remet les termes et conditions au client avant la conclusion du contrat. Il résume les finalités pour lesquelles les données seront utilisées et s'engage à ne pas les utiliser à d'autres fins.
- Il indique au client qu'elle peut supprimer les photographies après la fin de la commande, mais que la cliente peut renoncer d'emblée à passer une nouvelle commande.
- Il demande explicitement au client s'elle peut utiliser ses photographies à ses propres fins publicitaires.
- Enfin, il s'assure également que le fournisseur externe de services en ligne respecte les mêmes règles en matière de traitement des données (notamment qu'il n'utilise pas les données à des fins personnelles) et en informe la cliente.





## RESPONSABILITÉ

Les recommandations relatives au principe procédural fondamental de responsabilité visent à garantir que vous pouvez utiliser de manière responsable les données collectées par des tiers.

### Recommandations

- Si votre produit dépend de données provenant d'autres entreprises, assurez-vous que ces données ont été recueillies de manière éthique.
- N'utilisez pas de données provenant de sources non fiables.
- N'obtenez des données que de partenaires qui font preuve de transparence quant à leurs pratiques de collecte de données et à toute restriction associée à l'utilisation des données.



### EXEMPLE - UTILISATION COMMERCIALE DES DONNÉES DES RÉSEAUX SOCIAUX

Une entreprise pharmaceutique obtient des données à partir d'une plateforme web qui offre des possibilités de réseautage social et d'auto-surveillance à des groupes de patients. L'entreprise pharmaceutique reçoit des informations sur les pratiques de collecte de données de la plateforme web. Les données sont-elles obtenues de telle sorte que l'objectif de la collecte de données et le lien avec l'industrie, y compris la monétisation des données, soit communiqués de façon non transparente et ouverte aux utilisateurs ? Si tel est le cas, les entreprises pharmaceutiques n'acquerront pas de données à partir de cette plateforme et n'utiliseront pas ses services.

## ÉTAPE 2: STOCKAGE ET GESTION DES DONNÉES

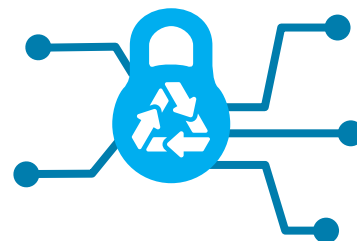
### ORIENTATIONS ÉTHIQUES FONDAMENTALES

#### PRÉVENTION DES DOMMAGES

Les recommandations relatives au principe fondamental de prévention des dommages doivent garantir qu'aucun dommage ne peut être occasionné par l'utilisation des données stockées.

Recommandations

- Mettez en œuvre des mesures de cybersécurité appropriées qui sont appliquées de manière cohérente à toutes les formes de stockage de données (par exemple, stockées dans le cloud ou de façon hybride).
- Élaborez des plans d'urgence pour minimiser l'étendue des dommages en cas de violation des données.
- Définissez des règles de suppression des enregistrements de données dans le but de réduire le risque de dommages causés par un accès non autorisé.



#### EXEMPLE - DONNÉES DE SANTÉ DANS LE CLOUD

Une entreprise d'analyse des données de santé développe des outils personnalisés de collecte et d'analyse de données pour les grands hôpitaux. La société fournit aux hôpitaux un service d'hébergement de toutes les données dans le cloud, ce qui permet aux hôpitaux de se concentrer sur les soins aux patients et d'améliorer les résultats.

Afin de garantir aux hôpitaux la cyber-sécurité nécessaire, la l'entreprise d'analyse des données met en œuvre les normes de cyber-sécurité les plus élevées et s'engage auprès des hôpitaux à les respecter (par exemple, les données ne sont accessibles qu'à certains employés de l'entreprise de logiciels avec une authentification à deux niveaux et il existe des cours de sensibilisation dans l'hôpital aux attaques d'ingénierie sociale).



## JUSTICE

Les recommandations relatives au principe fondamental de justice doivent permettre aux clients d'utiliser leurs données au profit de tiers et du bien commun.

### Recommandations

- Aidez vos clients à bénéficier efficacement du droit à la portabilité des données.
- Permettez à vos clients de recevoir et de réutiliser facilement leurs données (tant qu'elles n'ont pas été rendues anonymes) afin que leurs données puissent également être utilisées sous une forme appropriée pour des applications tierces.



### EXEMPLE - DON DE DONNÉES

Une entreprise qui collecte et traite les données personnelles de ses clients afin de leur fournir des services veut faciliter le don de leurs propres données par ses clients à des fins qui leur semblent raisonnables. Il peut s'agir, par exemple, de projets de recherche à but non lucratif ou de projets de l'administration publique. En réutilisant les données, l'entreprise veut également contribuer à la durabilité numérique. Les mesures suivantes pourraient être utiles à cet égard:

- Mise en place d'un système de gestion de l'information sur la plateforme de l'entreprise, avec lequel les clients de l'entreprise peuvent gérer eux-mêmes leurs données personnelles. Cela comprend notamment une option conviviale permettant de télécharger toutes les données ou certaines d'entre elles sous une forme lisible par une machine et de les exporter dans d'autres systèmes.

Mise en place d'une interface technique avec d'autres systèmes de gestion de l'information et des plateformes de dépôt fiduciaire de données, afin que les clients de la société puissent mettre leurs données personnelles directement à la disposition d'autres prestataires de services.

## AUTONOMIE

Les recommandations relatives au principe fondamental d'autonomie visent à garantir que le stockage et la gestion des données ne présentent aucun risque pour la vie privée des clients.

### Recommandations

- Examinez si vos mesures techniques visant à protéger la vie privée de vos clients sont appropriées et proportionnées.
- Tenez-vous régulièrement informé de l'état de la technique dans les technologies concernées.
- Veillez à ce que les données anonymes stockées ne puissent pas être reliées aux personnes auxquelles elles se rapportaient initialement.
- Proposez à vos clients des options simples pour limiter le stockage de leurs données.



### EXEMPLE - VISUALISATION DE L'AUTODÉTERMINATION NUMÉRIQUE

Une entreprise a placé l'autodétermination numérique de ses utilisateurs au centre de ses activités commerciales et fournit un système de gestion des informations personnelles et une place de marché pour les données personnelles. La plateforme permet à ses utilisateurs de mettre leur profil de données ou des extraits choisis à la disposition d'entreprises individuelles, d'administrations ou même d'organisations à but non lucratif à un certain prix ou même gratuitement ; l'entreprise participe au prix payé. Les utilisateurs ont à tout moment une vue d'ensemble des données qui sont envoyées à quels destinataires et à quel prix et peuvent également annuler ces offres. La valeur ajoutée de la plateforme pour les utilisateurs consiste à renforcer leur autodétermination numérique grâce aux éléments suivants:

- Les données des utilisateurs, telles que les données sur les achats, la mobilité ou la santé, sont agrégées et visualisées. Par exemple, les utilisateurs peuvent vérifier si leurs achats étaient sains ou durables ou quel a été leur comportement en matière de mobilité au cours des derniers mois.
- Les données sont stockées sous forme cryptée et ne peuvent être consultées par l'entreprise elle-même.

Si un client souhaite transmettre une partie de son ensemble de données à des tiers, le système vérifie automatiquement le niveau de risque, c'est-à-dire si l'ensemble de données ainsi défini peut être retracé jusqu'à la personne d'origine. Si le risque est élevé, le client est mis en garde contre la divulgation et doit l'accepter explicitement.



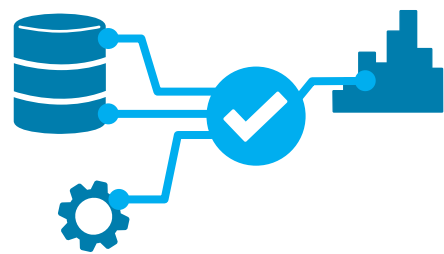
## VALEURS PROCÉDURALES

### CONTRÔLE

Les recommandations relatives au principe fondamental de contrôle visent à garantir qu'aucune information pertinente pour l'évaluation éthique de l'utilisation des données ne soit perdue lors de la gestion des données (par exemple, nettoyage des données, suppression des données, etc.).

#### Recommandations

- Documentez tous les changements et adaptations aux exigences de gestion des données lorsque les données sont déplacées vers des modèles différents, en particulier lorsque les services sont externalisés vers le cloud.
- Mettez en œuvre les exigences techniques qui assurent la traçabilité des modifications apportées aux données stockées.
- Réfléchissez à l'avance à ce que vous ferez des données des comptes inactifs ou des anciens clients.



### EXEMPLE - TRAÇABILITÉ DE L'UTILISATION DES DONNÉES

Une entreprise utilise les données sociodémographiques de ses clients ainsi que les données issues de son interaction avec eux (comportement de commande, correspondance, etc.) pour classer ses clients dans différents groupes à l'aide de modèles d'analyse de données et pour prévoir leur comportement futur. Ces modèles sont utilisés dans la gestion des relations avec les clients. En outre, les résultats des prévisions basées sur le modèle sont en partie partagés avec d'autres entreprises du même groupe d'entreprises. Ces données sont sujettes à des changements constants. Afin de pouvoir retracer à tout moment quelles données ont été utilisées pour un traitement ultérieur (analyses de données, modèles de prévision), chaque ensemble de données est enrichi de métadonnées décrivant la provenance des données, l'heure à laquelle les données ont été transmises aux analystes ou aux constructeurs de modèles, les restrictions d'utilisation liées à l'ensemble de données, etc. Cela permet de suivre chaque ensemble de données individuelles utilisées pour l'analyse ou la modélisation, tout en garantissant que les analystes et les créateurs de modèles utilisent les données de manière appropriée.

## TRANSPARENCE

Les recommandations relatives au principe fondamental de transparence visent à garantir que les clients ne soient pas seulement informés des aspects liés uniquement à la protection des données lors du traitement de leurs données. Le stockage et la gestion des données devraient plutôt être traitées d'une façon assez extensive pour susciter leur confiance.



### Recommandations

- Expliquez clairement à vos clients comment leurs données et accès sont protégés.
- S'il est nécessaire de stocker des données pendant une très longue période (par exemple pour des raisons juridiques), expliquez-le à vos clients au moment de la collecte des données.

### EXEMPLE - INFORMATIONS SUR LA CONSERVATION DES DONNÉES

La Loi fédérale sur la surveillance de la correspondance par poste et télécommunication oblige les fournisseurs de services de télécommunication en Suisse à stocker et à conserver les métadonnées des télécommunications de leurs clients pendant six mois. Cela devrait faciliter la clarification des données criminelles et la recherche des personnes disparues. Tout citoyen suisse a également le droit de consulter les données qui ont été conservées sur lui. Un fournisseur de télécommunications attire l'attention sur ce fait d'une manière appropriée et détaillée en indiquant de quelles données il s'agit:

- Qui était sur quel site web, quand et quelles applications ont été utilisées
- Avec qui, quand et à partir d'où la communication a eu lieu
- Qui s'est connecté à l'internet, quand et pendant combien de temps
- Qui a envoyé à qui un courrier électronique ou un autre message textuel, et quand
- l'endroit où se trouve actuellement l'utilisateur du téléphone portable ou l'historique de sa localisation au cours des six derniers mois



## RESPONSABILITÉ

Les recommandations sur la valeur procédurale de la responsabilité devraient garantir que les rôles de toutes les personnes impliquées dans la gestion des données ont été suffisamment clarifiés.

### Recommandations

- Définissez, au niveau de l'organisation de l'entreprise, qui est impliqué dans l'évaluation, la déclaration et la mise en œuvre des normes éthiques en matière de gestion des données.
- Définissez des règles claires pour l'accès aux données stockées (qui, quand, dans quelles conditions, suivi de l'accès, ...).



### EXEMPLE - LE RÔLE DE L'OFFICIER DE PROTECTION DES DONNÉES

Le rôle de l'officier de protection des données (OPD) s'est imposé dans de nombreuses organisations. L'OPD supervise l'utilisation des données afin de s'assurer qu'une organisation peut tirer le maximum de ses données tout en maintenant ses propres normes. Elles comprennent généralement l'ensemble du processus de gestion des données et la garantie de la qualité des données. Ces deux éléments sont abordés dans la présente recommandation. Bien que le bureau de l'officier de protection des données soit généralement établi non pas pour des raisons éthiques mais pour des considérations commerciales, il est néanmoins de son devoir d'établir et de contrôler les normes éthiques en matière de données, lorsque de telles normes sont établies.

## ÉTAPE 3: ANALYSE DE DONNÉES

### ORIENTATIONS ÉTHIQUES FONDAMENTALES

#### PRÉVENTION DES DOMMAGES

Les recommandations relatives au principe fondamental de prévention des dommages visent à garantir que le potentiel de risque des modèles soit identifié suffisamment tôt dans le processus d'analyse des données et de génération des modèles.

##### Recommandations

- Évaluez l'utilisation abusive prévisible de vos modèles.
- Pendant le processus d'élaboration du modèle, évitez notamment toute mesure qui pourrait conduire à la désanonymisation des données utilisées.
- Intégrez des mécanismes de contrôle dans vos modèles pour surveiller et limiter les dommages potentiels qui pourraient résulter de leur utilisation pratique ; cela s'applique notamment à la violation de la vie privée ou aux décisions incorrectes dues à des inexactitudes dans le modèle.
- Envisagez d'imposer des limites à la distribution du produit, le cas échéant.



#### EXEMPLE - ÉVALUATION DE LA SÉCURITÉ DE L'IA

L'un des principaux acteurs dans le domaine de l'IA (Deep Mind) a développé un cadre théorique pour guider l'évaluation de la sécurité de ses produits. Le cadre se compose de trois éléments: (1) Spécification, que nous considérons cidessous comme un exemple de contrôle de l'étape 3 ; (2) Robustesse: identifier les environnements qui ne sont pas représentés dans l'ensemble de données et qui peuvent conduire à un mauvais comportement de l'IA (par exemple, un robot de nettoyage ménager testé dans une maison sans animaux peut être utilisé dans un bureau accueillant des animaux) et tester la possibilité d'un apport contradictoire ; (3) Assurance, qui comprend deux aspects: Systèmes d'inspection (statistiques sommaires, inspections automatisées et outils utilisés pour rendre les décisions relatives à l'IA explicables ou interprétables après utilisation) et mécanismes d'application (par exemple, interruptibilité, interrupteur fiable). Source: <https://deepmind.com/blog/article/specifying-ai-safety-problems>





## JUSTICE

Les recommandations relatives au principe fondamental de justice abordent en particulier le problème de la discrimination indirecte fondée sur des modèles et des prévisions fondées sur des données. La discrimination indirecte signifie que bien qu'une règle de décision s'applique à tous de la même manière, elle a un effet différent sur certains groupes de personnes (par exemple les femmes par rapport aux hommes) que sur d'autres et que cela n'est pas justifié. Ceci est particulièrement pertinent pour les modèles et les prévisions sur la base à partir desquels les décisions sont prises automatiquement ou qui soutiennent la prise de décision humaine (par exemple, les recommandations).



### Recommandations

- Vérifiez si un ensemble de données utilisé pour l'apprentissage machine est suffisamment diversifié pour éviter les biais dus à de mauvaises données de formation.
- Expliquez quel type de discrimination est prévu et intégré dans votre modèle, quelle discrimination indirecte non souhaitée a été supprimée de votre modèle, le cas échéant, et quelle forme de discrimination indirecte existe encore.
- Vérifiez l'existence d'une discrimination indirecte à l'aide de données de test.
- Développez une expertise interne sur la manière de traiter la discrimination indirecte qui peut résulter de l'apprentissage machine ; cela devrait également inclure des méthodes techniques (statistiques, informatique) pour éliminer ces injustices.
- Le cas échéant, introduisez des technologies pertinentes pour prévenir la discrimination indirecte.

### EXEMPLE - OUTIL DE DIAGNOSTIC BIAISÉ

Une équipe interdisciplinaire de scientifiques médicaux crée un modèle prédisant la réaction des patients à un traitement en se basant sur des données de génomique, d'épigénomique, de protéomique et d'autres domaines. Sur la base de ces données, le modèle fournit une prédiction sur l'efficacité d'un traitement donné pour un patient particulier. Le traitement peut sauver la vie du patient et a peu d'effets secondaires. Toutefois, en raison de contraintes de ressources, il est décidé de donner la priorité aux patients qui sont les plus susceptibles de bénéficier du traitement. Il s'avère que 90 % des hommes d'origine européenne, mais seulement 10 % des hommes d'origine africaine sont susceptibles de bénéficier d'un traitement, il y a donc une disparité entre les personnes à qui le traitement est recommandé. Cela soulève la question suivante: les personnes noires ont-elles une plainte valable contre cette inégalité ? Tout d'abord, on fait valoir que l'inégalité des chances de bénéficier d'un traitement justifie cette inégalité. À la suite d'une discussion éthique avec diverses parties prenantes, il est reconnu qu'un algorithme moins inégalitaire, où 70 % des blancs sont traités contre 40 % des noirs, est plus justifié, même s'il sacrifie le taux de réussite du traitement, parce qu'il existe un historique de biais dans le recrutement pour les essais cliniques qui explique en partie pourquoi des médicaments aussi efficaces ne sont pas disponibles pour les personnes noires, et parce que le taux de réussite global d'un tel prédicteur est suffisamment élevé pour être conforme à l'ensemble des directives des assurances sur la rentabilité des traitements. En revanche, ils sont conscients que le meilleur algorithme qui donne aux noirs et aux blancs une chance égale de recevoir un traitement se traduirait statistiquement par un taux de réussite prévisible trop faible et violerait les directives de rentabilité, de sorte qu'il ne serait pas remboursable.

## AUTONOMIE

Les recommandations relatives au principe fondamental d'autonomie doivent contribuer à garantir que les personnes concernées en dehors de l'entreprise puissent également procéder à l'évaluation des aspects éthiques de la construction de modèles si ces modèles ont une influence sur l'autonomie de ces acteurs.

### Recommandations

- Identifiez les questions éthiques pertinentes associées à l'utilisation de vos modèles basés sur les données ou les connaissances acquises à partir de ces données, en collaboration avec les parties prenantes concernées.
- Permettez aux parties prenantes de formuler des commentaires sur l'élaboration du modèle, notamment en ce qui concerne les objectifs intégrés dans le modèle (par exemple, l'objectif de classification ou la fonction de perte).



### EXEMPLE - WORLD CAFÉ AVEC LES PARTIES PRENANTES

Une entreprise suisse de télécommunications est en discussion avec ses principaux groupes de parties prenantes sur la base de sa stratégie de durabilité. Entre autres, 25 représentants de parties prenantes externes ont discuté des principaux problèmes de durabilité avec l'entreprise lors d'un World Café. Trois thèmes clés ont été identifiés:

- Informations sur les risques de la norme de communication mobile 5G.
- Formation des enseignants et des PME sur les questions de numérisation en créant des laboratoires de numérisation pour les PME et les écoles.
- Création d'incitations accrues pour les fournisseurs afin de réduire les émissions de CO2.

Les participants du World Café ont apprécié la possibilité d'échanger des informations directement avec l'entreprise. Sur la base des réactions positives reçues, l'entreprise a l'intention d'organiser d'autres événements de discussion impliquant ses parties prenantes pour compléter l'enquête réalisée auprès des parties prenantes. Ce type de dialogue entre les parties prenantes peut également être utilisé pour discuter de modèles spécifiques basés sur des données, par exemple le risque que ces modèles génèrent une discrimination.



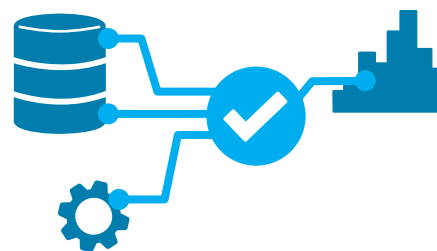
## VALEURS PROCÉDURALES

### CONTRÔLE

Les recommandations relatives à la valeur procédurale de contrôle visent à garantir que les aspects éthiques soient systématiquement documentés lors de l'analyse des données, et créent ainsi un socle de confiance.

#### Recommandations

- Montrez le caractère approprié des algorithmes utilisés par rapport à l'ensemble des données et à l'analyse spécifique visée ; cette justification doit inclure la documentation relative à l'évaluation des risques et aux mesures prises pour atténuer les risques à la lumière des principes de prévention des dommages, de justice et d'autonomie énoncés cidessus.
- Veillez à ce que les connaissances et les modèles générés puissent être rattachés aux données d'origine et à l'algorithme utilisé pour les créer.
- Si votre analyse est basée sur des données et/ou des modèles fournis par d'autres, assurez-vous que ces données ont été produites d'une manière qui respecte vos propres normes éthiques.



### EXEMPLE - CONTRÔLE DES MODÈLES D'IA

L'un des principaux acteurs dans le domaine de l'IA (Deep Mind) a développé un cadre théorique pour guider l'évaluation de la sécurité de ses produits. Ce cadre comporte trois éléments, le premier élément, la spécification, met en œuvre la valeur du contrôle, tel que nous le définissons ici. Le concepteur distingue la «spécification idéale» (les souhaits du concepteur), la «spécification de conception» (par exemple la fonction de récompense utilisée dans un système d'apprentissage par renforcement) et la «spécification actuelle» (le comportement réel du système). Le premier élément de contrôle consiste à documenter clairement la spécification idéale (l'objectif que vous souhaitez atteindre) et la spécification de conception (la manière dont elle est traduite en code). Il utilise une méthode basée sur les données (par exemple, l'apprentissage par renforcement inverse et les scénarios simulés) pour identifier comment la spécification actuelle s'écarte de la spécification idéale, c'est-à-dire pour évaluer quand l'IA s'écarte des souhaits du créateur. Source: <https://deepmind.com/blog/article/specifying-ai-safety-problems> Une approche similaire est définie dans ABOUT ML, la méthodologie développée par Partnership for AI impliquant la transparence par la documentation. Voir: <https://www.partnershiponai.org/wp-content/uploads/2019/07/ABOUT-ML-v0-Draft-Chapter-2.pdf>

## TRANSPARENCE

Les recommandations relatives à la valeur procédurale de transparence visent à garantir que les utilisateurs des analyses et des modèles, ainsi que les experts et expertes en matière d'audit, soient en mesure de les évaluer en fonction de caractéristiques pertinentes sur le plan éthique. Cela s'applique en particulier aux limites des modèles, les informations pertinentes devant être adaptées au groupe cible.



### Recommandations

- Expliquez la logique de vos modèles et les décisions qui ont été incorporées dans la construction du modèle d'une manière que les utilisateurs du modèle peuvent comprendre.
- Faites connaître clairement les limites de vos modèles aux utilisateurs de vos modèles, par exemple, le caractère imprécis de certaines classifications.
- Expliquez aux utilisateurs les conséquences possibles de ces inexactitudes de modèle, comme la discrimination indirecte fondée sur des décisions basées sur un modèle.
- Documentez les mesures prises pour respecter vos normes éthiques en matière de modélisation basée sur les données ; par exemple, vos activités visant à éliminer ou à réduire la discrimination indirecte dans votre modèle, le cas échéant.
- Informez les utilisateurs de manière appropriée sur ces mesures ; le cas échéant, informez le public si les modèles sont utilisés dans des contextes décisionnels susceptibles d'affecter les droits de l'homme ; informés de façon spécifique les groupes concernés.

### EXEMPLE - INFORMATIONS ADAPTÉES AU GROUPE CIBLE SUR UN MODÈLE D'IA

Supposons qu'un informaticien crée un modèle d'apprentissage automatique avec des informations sur le demandeur d'un prêt afin de prédire s'il effectuera des paiements ponctuels sur une période de deux ans. Il utilise la prédiction d'apprentissage automatique pour décider si le propriétaire est éligible à une crédit. On peut distinguer trois types différents de destinataires d'une explication:

- 1) Les data scientists qui veulent comprendre le comportement du modèle dans son ensemble et pas seulement dans des cas spécifiques,
- 2) Les agent de crédit, qui doivent comprendre la décision d'accepter ou de rejeter la demande d'un demandeur, contrairement aux autres demandes,
- 3) Les clients qui veulent comprendre pourquoi il n'a pas droit à une ligne de crédit et ce qu'il pourrait faire pour revenir sur ce jugement.

Pour chaque demandeur, un type d'explication différent est sélectionné à l'aide du logiciel disponible dans la suite AI Explainability 360:

- 1) Un logiciel qui produit une approximation du produit de l'algorithme sous la forme de modèles directement interprétables, basés sur des règles, qui permettent une compréhension globale de son comportement ;
- 2) Les logiciels qui génèrent des profils des demandeur qui ont réussi ;
- 3) Un logiciel qui calcule des exemples qui indiquent au client quelle modification individuelle de son profil aurait changé la décision du modèle d'IA, le reste étant identique par ailleurs.

Source: page des tutoriels de l'AI Explicability 360 Toolkit, disponible sur <https://aix360.mybluemix.net>



## RESPONSABILITÉ

Les recommandations relatives à la valeur procédurale de responsabilité visent à garantir que, dans le processus complexe d'analyse des données et de génération de modèles, les responsabilités en matière de recommandations éthiques soient effectives.



### Recommandations

- Attribuez des responsabilités aux personnes participant à l'élaboration des modèles afin de garantir le respect des normes éthiques dans l'analyse des données et la production de connaissances et d'assurer que les responsabilités de chacun soient définies.
- Au niveau organisationnel, veillez à ce que les normes éthiques en matière d'analyse des données et de production de connaissances soient respectées et que les problèmes éthiques correspondants puissent être identifiés.

### EXEMPLE - CONSEIL DE GOUVERNANCE DES DONNÉES

Une société de médias, qui souhaite de plus en plus développer de nouveaux modèles commerciaux avec des données, a mis en place un conseil de gouvernance des données. Ce conseil se réunit au moins une fois par trimestre (si nécessaire également de manière ad hoc) et implique des représentants de tous les secteurs d'activité, en particulier les équipes actives dans le développement de produits et le service juridique. Ce conseil discute ouvertement des questions éthiques des nouvelles idées commerciales et sous la direction d'une personne qui a reçu une formation complémentaire sur les questions éthiques. Les résultats de la réunion sont consignés dans un procès-verbal. Cette «expérience d'apprentissage» commune sensibilise aux questions éthiques dans la vie quotidienne des entreprises. Dans les cas critiques, ces questions parviennent également à la direction et y sont discutées. Les enregistrements des discussions de groupe aident à présenter le problème le plus précisément possible, ce que la direction apprécie pour un traitement efficace du sujet

## ÉTAPE 4: UTILISATION DES CONNAISSANCES

### ORIENTATIONS ÉTHIQUES FONDAMENTALES

#### PRÉVENTION DES DOMMAGES

Les recommandations relatives à la prévention des dommages visent à garantir que le potentiel d'abus de leurs produits et services basés sur les données soit reconnu.

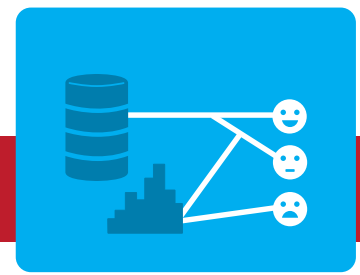
##### Recommandations

- Ne distribuez pas de produits et services basés sur des données avant d'avoir étudié dans quelle mesure ils pourraient être utilisés à mauvais escient par des utilisateurs avertis ou non avertis.
- Examinez dans quelle mesure vous pouvez vérifier si les produits ou services basés sur des données que vous développez sont effectivement utilisés de manière abusive par des tiers (par exemple, si un logiciel est utilisé à des fins autres que celles pour lesquelles il a été conçu et testé).
- Vérifiez si l'utilisation de votre produit pourrait conduire à la désanonymisation des données ou pourrait être utilisée pour collecter des données sensibles à partir de données personnelles ne posant pas de problème en tant que telles.



#### EXEMPLE - INTERACTION ABUSIVE AVEC UN BOT

Une entreprise utilise l'intelligence artificielle pour développer un robot d'information (bot) qui, à l'instar de l'ancien service d'information téléphonique, est conçu pour répondre à des questions générales. Le bot apprend de l'interaction avec les utilisateurs à quel point ils sont satisfaits des réponses ; il a été formé à l'aide de nombreuses informations accessibles au public telles que Wikipédia. Pour les questions sur des sujets controversés, le bot fournit des réponses qui incluent tous les aspects de la controverse afin de maintenir la neutralité. Il est vite apparu que le bot fournit de plus en plus de réponses « neutres » même lorsque les informations sont effectivement sécurisées : par exemple, il laisse ouverte la possibilité que la terre soit un disque. La raison en est que les partisans des théories du complot ont exprimé en grand nombre leur insatisfaction quant aux réponses du bot, ce qui a influencé l'apprentissage du bot. Les développeurs avaient omis de vérifier les interactions abusives systématiques avec le bot. Au moins, il a été déterminé que les abus avaient eu lieu et que des contre-mesures pouvaient être prises.



## JUSTICE

Les recommandations relatives au principe fondamental de justice doivent contribuer à garantir que l'utilisation de produits et services basés sur des données n'ait pas d'effets sociaux indésirables tels que la discrimination indirecte, la stigmatisation ou l'exclusion de groupes de personnes. Le devoir éthique d'évaluer et de traiter les impacts sociaux indirects est proportionnel aux capacités financières et technologiques d'une entreprise et à l'impact des produits et services basés sur les données sur la société.



### Recommandations

- Évitez d'utiliser des produits et services basés sur des données dont les effets discriminatoires indirects n'ont pas été évalués et quantifiés s'il existe un risque de discrimination indirecte significative.
- Évitez d'utiliser des produits et services basés sur des données qui pourraient par inadvertance stigmatiser des groupes entiers de personnes ou nuire à leur réputation.
- Évaluez si les produits et services basés sur les données que vous utilisez excluent les groupes vulnérables et évitez cette exclusion si possible.
- Surveillez les systèmes de décision automatisés ou les systèmes de recommandation pour repérer toute discrimination involontaire, même si le produit ou sa base de donnée a été testé en ce qui concerne la discrimination indirecte, car les données testées peuvent avoir des propriétés statistiques différentes des données réelles et de nouvelles formes de discrimination peuvent donc apparaître.

### EXEMPLE - CONFLITS DE VALEURS DANS LES APPLICATIONS COVID

De nombreuses applications européennes de suivi numérique des contacts, telles que Covid en Suisse, Immuni en Italie et l'application Corona Warn en Allemagne, ont opté pour une architecture décentralisée mise en œuvre par une API Google/Apple. Cela limite considérablement les données qui peuvent être collectées pour une analyse plus approfondie. Cette situation et l'architecture décentralisée font qu'il est très difficile d'évaluer si certains groupes sociaux sont touchés de manière très inégale par les notifications et les fausses notifications. Si l'application était rendue obligatoire ou si elle était associée à des avantages sociaux importants (par exemple l'utilisation des transports publics), le manque de connaissance des groupes qui sont notifiés et l'inégalité du taux de faux positifs signifieraient que des éléments importants d'équité dans son utilisation seraient ignorés. Ainsi, l'application, dans sa forme conçue, protège la vie privée, mais rend plus difficile l'obtention de la justice. Si la connaissance de l'injustice ne peut être générée par une application obligatoire de suivi des contacts décentralisés, il y a une tension entre la justice et la vie privée. Par conséquent, si l'application devient obligatoire, Google/Apple devrait modifier le protocole pour permettre la collecte systématique de données anonymes afin d'établir des corrélations entre les taux de notification et les notifications faussement positives avec les caractéristiques de sexe, d'ethnicité ou de revenu, identifiant ainsi les groupes qui sont vulnérables et ce de façon disproportionnée.

## AUTONOMIE

Les recommandations relatives au principe fondamental d'autonomie visent à garantir que le retour d'informations sur les éventuels effets négatifs de vos produits et services basés sur les données est évalué à sa juste mesure.

### Recommandations

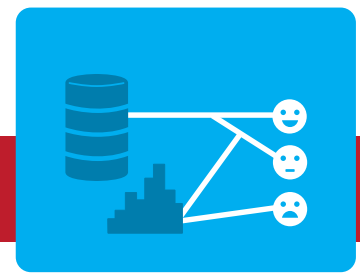
- Mettez en place des processus permettant aux parties concernées de signaler les résultats négatifs de vos produits et d'en assurer le suivi.
- Suivez les débats d'experts relatifs aux applications respectives de vos produits qui fournissent des preuves basées sur les connaissances des experts respectifs de déviations systématiques dans les processus de prise de décision.



### EXEMPLE - FORUM DES CITOYENS

Deep Mind a fait équipe avec la Royal Society for the encouragement of Arts, Manufactures and Commerce (RSA) pour promouvoir un engagement public significatif concernant l'impact de l'IA dans le monde réel. Le RSA a mis en place un jury citoyen pour enquêter sur l'utilisation de l'IA dans la prise de décision. Des questions générales sur l'IA ont été abordées, par exemple le type d'engagement dont le public a besoin dans l'IA, ou les lignes rouges des citoyens qui sont pertinentes avant le déploiement et dans la phase de conception (voir l'étape 3 - Autonomie). Afin de mettre en œuvre le principe d'autonomie décrit ici, la même méthodologie (forum des citoyens) pourrait être utilisée comme processus pour discuter des aspects négatifs de l'introduction d'une technologie après son déploiement. Source: <https://www.thersa.org/action-and-research/rsa-projects/economy-enterprise-manufacturing-folder/tech-and-society/forum-for-ethical-ai>

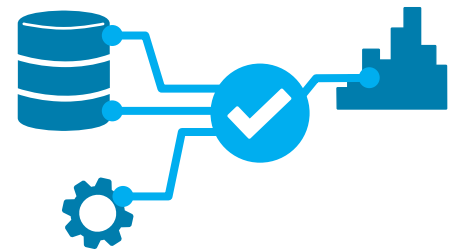




## VALEURS PROCÉDURALES

### CONTRÔLE

Les recommandations relatives à la valeur procédurale de contrôle visent à garantir, à un niveau général, que les valeurs propres de l'entreprise sont correctement reflétées dans les produits et services basés sur des données que l'entreprise entend offrir sur le marché (on y trouve des mesures de contrôle directement liées aux trois principes d'orientations de base).



Recommandations

- Déterminez les priorités en matière d'éthique des données pour votre entreprise, qui résultent de vos obligations en matière de valeur et de votre mission en tant qu'entreprise, et déduisez-en les exigences correspondantes pour vos produits et services basés sur les données.
- Évaluez si les normes éthiques ont été respectées à toutes les étapes du traitement des données qui a mené à la réalisation de votre produit ou service (qu'il soit effectué au sein de votre entreprise ou par d'autres) et si elles sont compatibles avec la mission, les valeurs et l'image publique de l'entreprise.

### EXEMPLE - PROBLÈMES D'APPLICATION DES DIRECTIVES SUR L'IA

L'ONG AlgorithmWatch a commencé à compiler des directives pour une «IA éthique» dans un répertoire. Toutefois, sur les plus de 160 lignes directrices qui figurent désormais dans cet inventaire mondial des lignes directrices éthiques de l'IA (2020), seule une fraction (10) dispose de mécanismes de surveillance et d'application adéquats. Un examen régulier de ces lignes directrices, comme dans le cas d'AlgorithmWatch, et une discussion critique ultérieure pourraient créer une pression publique pour rendre contrôlables les exigences résultant des normes d'éthique des données.

## TRANSPARENCE

Les recommandations relatives à la valeur procédurale de transparence visent à garantir que vous fournissez des informations adéquates sur votre respect des recommandations énoncées dans le présent code et à contribuer ainsi au maintien de pratiques éthiques en matière de données.

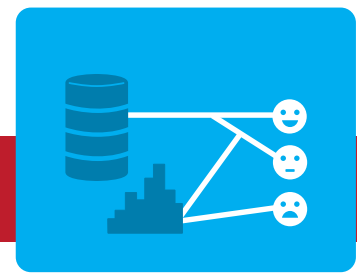


### Recommandations

- Communiquez vos normes et pratiques éthiques à tous ceux qui peuvent être affectés par les effets de votre produit.
- Si votre entreprise produit des données, des modèles ou des produits pour des tiers, fournissez toutes les informations nécessaires à l'utilisation éthique de ces données/modèles/produits ; par exemple, si des tiers forment un modèle sur la base des données que vous fournissez, donnez suffisamment d'informations sur l'ensemble des données pour éviter toute partialité non intentionnelle du modèle.

### **EXEMPLE - REVENTE DE CONNAISSANCES GÉNÉRÉES PAR UN MODÈLE**

Une entreprise utilise les données pour générer par l'analyse de nouvelles informations, qui sont ensuite mises à la disposition d'autres entreprises pour un traitement ultérieur. L'entreprise a précisé les fins auxquelles elle peut utiliser les données qu'elle utilise et les a stockées dans des métadonnées. Les lignes directrices de l'entreprise en matière d'éthique des données définissent des valeurs fondamentales telles que la non-discrimination, le respect de la personnalité, l'autodétermination, la transparence, la responsabilité, etc. En particulier, l'entreprise s'est engagée à assurer la traçabilité des données, c'est-à-dire à ce que l'origine et la finalité des données restent connues même au sein de la chaîne de valeur. Ces principes sont joints aux informations vendues dans un simple «supplément». Grâce à des enquêtes occasionnelles, la société s'assure que les entreprises tierces qui basent leurs produits sur les résultats et les modèles de la société sont informées de la finalité des données utilisées à l'origine par la société.



## RESPONSABILITÉ

Les recommandations relatives à la valeur procédurale de la responsabilité visent à garantir que la direction générale de l'entreprise puisse assumer la responsabilité du produit final de manière appropriée.

### Recommandations

- Faites preuve de responsabilité en tant que membre de la direction en ne permettant pas la dilution de la responsabilité dans le processus complexe de développement de produits et services basés sur les données.
- Mettez en œuvre des processus pour garantir le respect des normes éthiques en ce qui concerne la mise à disposition de la base de données.
- Ne créez pas de boucs émissaires pour nier la responsabilité des entreprises en cas de problème éthique avec vos bases de données.
- Soyez conscient du fait que les normes éthiques appliquées à votre base de données peuvent être fortement influencées par les normes éthiques des autres entreprises avec lesquelles vous interagissez (par exemple, parce qu'elles ont recueilli les données que vous utilisez).
- Évaluez la norme éthique appropriée pour la base de données en tenant compte de toutes ces influences et assumez la responsabilité du produit final.



### EXEMPLE - RESPONSABLE DE LA SÉCURITÉ DES SYSTÈMES D'INFORMATION

Une partie de la responsabilité de la direction consiste à assurer la sécurité des données (c'est un élément de l'étape 2). Le rôle du responsable de la sécurité des systèmes d'information (RSSI), qui a maintenant été introduit dans de nombreuses entreprises, peut servir d'exemple. Lors de la mise en place d'une telle fonction, il convient d'éviter ce qui suit:

- Manque d'autorité du RSSI, de sorte que le RSSI ne peut s'acquitter de sa tâche que de manière formelle, mais pas de manière efficace.
- Installation du RSSI comme bouc émissaire à sacrifier lorsqu'une erreur est commise.

Inversement, les caractéristiques suivantes d'un RSSI peuvent contribuer à son succès:

- Se concentrer sur la gestion du changement et la transformation, si une entreprise souhaite se développer de manière substantielle
- Se concentrer sur les questions opérationnelles et techniques afin de développer des solutions qui fonctionnent pour l'opération en cours
- Se concentrer sur les questions de réglementation et de conformité lorsque celles-ci sont d'une grande importance pour l'entreprise

Ces différenciations sont utiles non seulement pour le sous-domaine de la sécurité de l'information, mais aussi pour les autres domaines mentionnés aux étapes 1, 3 et 4.

