

Science des données:

Informations Personnelles Identifiables (IPI)

December 2022

Introduction

- Nous discuterons de la façon dont l'intégrité des informations personnellement identifiables est maintenue en protégeant et en supprimant les données dans les ensembles de données
- Avec la prolifération des mégadonnées, plusieurs lois et réglementations nationales ont été mises en place pour répondre aux préoccupations croissantes concernant les IPI
- Il est important de protéger la vie privée des individus et de se conformer aux lois et règlements relatifs aux IPI

Définition des IPI

- Les informations personnelles identifiables (IPI) sont toutes les informations qui peuvent être utilisées pour identifier une personne.
- Cela inclut des éléments tels que le nom, l'adresse, le numéro de téléphone, l'adresse e-mail, le numéro de sécurité sociale et les informations financières d'une personne.
- Les IPI sont souvent collectées et utilisées par des organisations, telles que des entreprises et des organismes gouvernementaux, à diverses fins, telles que la vérification de l'identité, la fourniture de services et la réalisation de recherches.

Pourquoi protéger les IPI ?

- Il est important de protéger les IPI, car elles peuvent être utilisées pour le vol d'identité et d'autres activités frauduleuses si elles tombent entre de mauvaises mains.
- Par exemple, si quelqu'un obtient votre nom et votre numéro de sécurité sociale, il peut être en mesure d'ouvrir des comptes de crédit ou de contracter des prêts en votre nom.
- De nombreuses lois et réglementations, telles que le Règlement général sur la protection des données (RGPD) dans l'Union européenne et la California Consumer Privacy Act (CCPA) aux États-Unis, ont été mises en place pour protéger les IPI et donner aux individus le contrôle sur la façon dont leurs informations personnelles sont collectées, utilisées et partagées.

Protection des données IPI

Protection des données IPI

- Les méthodes de protection des données personnelles dans les ensembles de données comprennent : anonymisation, pseudonymisation, cryptage, contrôles d'accès, minimisation des données
- Ces méthodes ont toutes leurs avantages et leurs limites

Protection des données IPI : anonymisation

- Cela implique de remplacer les données IPI originales par un pseudonyme ou un identifiant généré aléatoirement qui ne peut pas être utilisé pour identifier une personne.
- Par exemple, vous pouvez remplacer le nom d'une personne par un numéro ou un code unique.
- L'anonymisation peut être utile pour préserver l'intégrité des données tout en protégeant la vie privée des individus, mais il peut être difficile d'inverser le processus et de réidentifier les individus si nécessaire.

Protection des données IPI : pseudonymisation

- Cela implique de remplacer les données IPI originales par un pseudonyme qui ne peut pas être utilisé pour identifier une personne, mais qui peut être relié aux données originales si nécessaire.
- Par exemple, vous pouvez remplacer le nom d'une personne par un nom d'utilisateur ou un alias généré de manière aléatoire.
- La pseudonymisation peut être utile pour préserver l'anonymat des individus tout en permettant l'analyse des données, et elle est généralement plus facile à inverser que l'anonymisation.

Protection des données IPI : chiffrement

- Cela implique de convertir les données IPI originales en un formulaire codé auquel on ne peut accéder qu'avec une clé secrète ou un mot de passe.
- Le chiffrement peut être utile pour protéger la confidentialité des données personnelles, mais il peut aussi prendre beaucoup de temps et nécessiter des ressources supplémentaires pour gérer les clés et déchiffrer les données si nécessaire.

Protection des données IPI : contrôles d'accès

- Cela implique de limiter l'accès aux données personnelles aux seules personnes qui ont une raison légitime d'y accéder.
- Cela peut être fait par l'utilisation de comptes d'utilisateurs, d'autorisations et d'autres mesures de sécurité.
- Les contrôles d'accès peuvent être efficaces pour protéger les données personnelles, mais ils peuvent également prendre beaucoup de temps à gérer et ne pas être suffisants à eux seuls pour protéger pleinement les données.

Protection des données IPI : minimisation des données

- Cela implique de collecter et de stocker uniquement la quantité minimale de données personnelles nécessaires à un objectif spécifique.
- Cela peut aider à réduire le risque d'accès non autorisé ou de divulgation de données personnelles, car il y a moins de données à protéger.
- Cependant, la minimisation des données peut également limiter l'utilité des données à certaines fins.

Suppression des données IPI

Suppression des données IPI

- Les méthodes de suppression des données IPI dans les ensembles de données comprennent le masquage des données, le nettoyage des données, l'anonymisation des données, l'agrégation des données, l'échantillonnage des données.
- Ces méthodes ont toutes leurs avantages et leurs limites.

Suppression des données IPI : masquage des données

- Cela implique de remplacer les données IPI originales par des données fausses ou fictives qui ne peuvent pas être utilisées pour identifier une personne.
- Par exemple, vous pouvez remplacer le nom d'une personne par un nom généré aléatoirement ou son adresse par une fausse adresse.
- Le masquage des données peut être utile à des fins de test ou de développement, mais les fausses données peuvent ne pas représenter avec précision les données d'origine, de sorte qu'elles peuvent ne pas convenir à l'analyse ou à d'autres fins.

Suppression des données IPI : nettoyage des données

- Cela implique d'identifier et de supprimer les données IPI d'un ensemble de données.
- Cela peut être fait manuellement, en examinant les données et en identifiant toutes les données IPI qui doivent être supprimées, ou automatiquement, en utilisant des outils logiciels qui peuvent analyser les données et identifier les données IPI.
- Le nettoyage des données peut prendre beaucoup de temps, mais c'est un moyen efficace de s'assurer que les données IPI sont supprimées d'un jeu de données.

Suppression des données IPI : anonymisation des données

- Cela implique de remplacer les données IPI originales par un pseudonyme ou un identifiant généré aléatoirement qui ne peut pas être utilisé pour identifier une personne.
- Par exemple, vous pouvez remplacer le nom d'une personne par un numéro ou un code unique.
- L'anonymisation des données peut être utile pour préserver l'intégrité des données tout en protégeant la vie privée des individus, mais il peut être difficile d'inverser le processus et de réidentifier les individus si nécessaire.

Suppression des données IPI : agrégation des données

- Cela implique de combiner plusieurs jeux de données en un seul jeu de données, puis de supprimer toutes les données IPI.
- Cela peut être utile pour préserver l'anonymat des individus tout en permettant l'analyse des données combinées.
- Cependant, il peut être difficile de faire correspondre et de fusionner avec précision les ensembles de données, et certaines informations peuvent être perdues au cours du processus.

Suppression des données IPI : échantillonnage des données

- Cela implique de sélectionner un échantillon aléatoire des données et de supprimer toutes les données IPI de l'échantillon.
- Cela peut être utile pour préserver l'anonymat des individus tout en permettant l'analyse des données, mais cela peut ne pas être représentatif de l'ensemble de données.

Conclusion

- Examiner et mettre à jour régulièrement vos politiques et procédures de protection des données pour vous assurer qu'elles sont efficaces et conformes à toutes les lois ou réglementations pertinentes.
- Utilisez des méthodes de stockage et de transfert de données sécurisées, telles que des serveurs cryptés ou des plateformes de partage de fichiers sécurisées.
- Veiller à ce que tous les employés et sous-traitants qui ont accès aux données personnelles reçoivent une formation sur les meilleures pratiques en matière de protection des données et soient conscients de leurs responsabilités.
- Mettre en œuvre des contrôles d'authentification et d'accès forts pour empêcher l'accès non autorisé aux données personnelles.
- Surveillez régulièrement vos systèmes et réseaux pour détecter toute faille de sécurité ou vulnérabilité potentielle.

Ressources des États-Unis:

- La Federal Trade Commission (FTC) a publié des lignes directrices sur la protection des informations personnelles
 - Vous pouvez trouver plus d'informations ici: <https://www.consumer.ftc.gov/articles/pdf-0003-protecting-personal-information>
- Le National Institute of Standards and Technology (NIST) a publié un cadre pour améliorer la cybersécurité des infrastructures critiques
 - Vous trouverez d'autres conseils sur la protection des renseignements personnels : <https://www.nist.gov/cybersecurity/framework>
- Le ministère américain de la Santé et des Services sociaux (HHS) a publié des lignes directrices sur la protection des informations personnelles sur la santé en vertu de la loi HIPAA (Health Insurance Portability and Accountability Act)
 - Vous pouvez trouver plus d'informations ici: <https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html>

Ressources pour l'Europe

- Le Règlement général sur la protection des données (RGPD) est une loi complète sur la protection des données qui s'applique à toutes les organisations qui traitent les données personnelles des individus dans l'Union européenne (UE). Elle énonce divers droits et obligations pour les individus et obligations pour les organisations, y compris la nécessité de protéger les données personnelles.
 - Vous trouverez plus d'informations sur le RGPD ici : https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations_en
- La directive sur les réseaux et les systèmes d'information (directive SRI) est une loi de l'UE qui s'applique à la protection des réseaux et des systèmes d'information dans l'UE. Elle comprend des dispositions sur la protection des données personnelles et s'applique aux organisations qui fournissent des services essentiels (par exemple, les soins de santé, l'énergie, les transports, etc.) ou des services numériques (par exemple, les marchés en ligne, les moteurs de recherche, les services d'informatique en nuage, etc.).
 - Vous trouverez plus d'informations sur la directive NIS ici: <https://ec.europa.eu/digital-single-market/en/network-information-systems-directive>

Ressources pour l'Afrique

- L'Union africaine (UA) a adopté la Convention sur la cybersécurité et la protection des données personnelles, qui énonce les principes de protection des données personnelles en Afrique.
 - Vous pouvez trouver plus d'informations sur la convention ici:
<https://au.int/en/treaties/convention-cyber-security-and-personal-data-protection>
- La Communauté de développement de l'Afrique australe (SADC) a adopté la Loi type de la SADC sur la protection des données, qui énonce les principes de protection des données personnelles dans la région de la SADC.
 - Vous pouvez trouver plus d'informations sur la loi type ici:
<https://www.sadc.int/model-law-data-protection/>