

Unlocking Data: Advancing Impact – French Summary and Transcript

(Note: This transcript was AI-generated using [NoteGPT](#). Transcripts may include inconsistencies or errors and some sentences have been condensed.)

Résumé

Le webinaire aborde le rôle des technologies d'amélioration de la vie privée (PET) dans l'accès aux données financières pour la santé publique, avec des commentaires d'experts du secteur.

Points forts

-  **Coopération Mondiale** : Le G7 soutient la promotion des technologies émergentes pour le partage de données.
-  **Vision de MasterCard** : Innovation dans les technologies de protection de la vie privée pour soutenir la confidentialité des données et l'inclusion financière.
-  **Confidentialité Différentielle** : Un standard d'or pour assurer la vie privée des individus tout en analysant les données.
-  **Perspectives Épidémiologiques** : Les données de transaction aident à modéliser le comportement en cas de pandémie et les zones à risque.
-  **Salles de Données Sécurisées** : Environnements sûrs pour analyser des données sensibles tout en préservant la confidentialité.
-  **Collaboration Intersectorielle** : Des équipes interdisciplinaires améliorent l'efficacité des PET.
-  **Portée Mondiale** : Les PET peuvent donner aux responsables de la santé locaux des informations en temps réel.

Aperçus Clés

-  **Approche Axée sur la Vie Privée** : Mettre l'accent sur une stratégie centrée sur la vie privée est crucial pour les organisations qui mettent en œuvre des PET, assurant conformité et confiance. Cette approche aide à traiter les préoccupations éthiques et réglementaires tout en maximisant l'utilité des données.
-  **Utilité des Données Financières** : Les données de transactions financières peuvent révéler des informations cruciales sur les comportements de consommation, aidant ainsi aux décisions de santé publique. Cette application montre le potentiel de l'analyse de données pour éclairer les politiques et la mise en œuvre de programmes.
-  **Collaboration Interdisciplinaire** : Impliquer des experts de divers domaines améliore l'efficacité des PET, comme on le voit dans les applications de

confidentialité différentielle. La collaboration peut mener à des solutions innovantes pour relever des défis sociétaux complexes.

-  Informations de Santé en Temps Réel : Les responsables de la santé locaux peuvent tirer parti des PET pour obtenir des informations sur les tendances pandémiques et prendre des décisions éclairées. Cette capacité soutient des interventions rapides et une allocation efficace des ressources.
-  Atténuation des Biais : Identifier et atténuer les biais dans les ensembles de données est essentiel pour créer des écosystèmes financiers équitables. Comprendre les données sous-jacentes est clé pour garantir des résultats justes dans les efforts d'inclusion financière.
-  Outils Innovants pour les Régions en Développement : Les pays en développement peuvent bénéficier des PET qui améliorent le partage de données tout en protégeant la vie privée. Cela ouvre des opportunités pour une prise de décision basée sur les données en matière de santé publique et de développement économique.
-  Engagement Communautaire : Rejoindre des initiatives comme la Confidentialité Différentielle Ouverte favorise un environnement collaboratif où les meilleures pratiques peuvent être partagées, menant à une adoption plus large des PET pour le bénéfice de la société.

Transcript

00:16

Bonjour, bon après-midi ou bonsoir, peu importe d'où vous nous rejoignez aujourd'hui. Je m'appelle Uyi Stewart, je suis directeur des données et de la technologie chez data.org. Je voudrais vous accueillir à notre webinaire qui se concentre sur l'utilisation des données pour avoir un impact. Bien que les technologies de protection de la vie privée existent depuis un certain temps, elles étaient surtout considérées comme des technologies de pointe, explorées par seulement quelques régulateurs et quelques entreprises du secteur privé. Mais les choses ont changé ces dernières années.

01:05

Au niveau mondial, en 2023, les autorités de protection des données et de la vie privée du G7 ont approuvé un plan qui vise à promouvoir le développement et l'utilisation de technologies émergentes, y compris des technologies qui peuvent instaurer la confiance dans le partage de données et protéger la vie privée. Chez data.org, on croit au potentiel transformateur des technologies qui améliorent la vie privée, qu'on va continuer à appeler « PETs » dans ce webinaire, pour débloquer des données financières pour la santé publique

tout en garantissant à la fois la vie privée et un impact positif. Aujourd'hui, on va explorer comment les PETs fonctionnent.

01:50

aborder les défis d'accès aux données tout en favorisant l'inclusion financière et en créant des écosystèmes équitables. Nous sommes ravis d'avoir trois experts internationaux dans notre panel aujourd'hui et j'aimerais les présenter. D'abord, John Derrico, vice-président chez MasterCard, responsable de l'approvisionnement en données et des technologies émergentes. Ensuite, Salil Vadhan, professeur d'informatique et de mathématiques appliquées, et directeur du programme Open DP à l'Université de Harvard. Et enfin, Shubham Kumar, qui est chercheur.

02:36

Je suis super excité de plonger dans la discussion d'aujourd'hui sur le pouvoir des données pour le bien social. Je voudrais commencer avec toi, John. Tu es à la tête de ce que fait Mastercard en matière de sourcing de données et des technologies d'impact. Mastercard a un rôle clé dans le sourcing de données et l'adoption des améliorations de la confidentialité. Alors, ma question pour toi, John, c'est quelle est la vision ou la stratégie de Mastercard pour cela ? Et comment ça garantit la confidentialité des données tout en débloquent les données des transactions financières pour soutenir la prise de décision en santé publique et d'autres domaines ?

03:25

Nouveaux cas, ok cool. D'abord, merci, je suis super content de pouvoir parler ici aujourd'hui et de partager quelques idées et stimuler des conversations. Pour donner un peu de contexte, Mard est une entreprise technologique mondiale dans le secteur des paiements et on fournit des solutions aux gens ainsi qu'aux organisations partout dans le monde. Une partie de la question que tu as mentionnée est vraiment liée à notre mission : c'est notre but de connecter un pouvoir économique numérique inclusif qui profite à tout le monde, partout. Pour nous, il s'agit de rendre les transactions sûres, simples et intelligentes.

03:59

Accessible maintenant avec ce contexte, notre leadership en technologie de la vie privée remonte à un bon moment car nous avons élargi notre profil de données depuis presque 60 ans pour rendre le commerce plus efficace. Ça a toujours été une question de sécurité, de sûreté et de confiance, qui sont vraiment les bases de beaucoup de notre travail ici avec les animaux de compagnie. Les animaux de compagnie font partie de notre vraie

innovation et ils soutiennent à la fois la vie privée, l'utilité, mais aussi la transparence. Et il y a vraiment trois aspects que nous pensons.

04:27

Concernant ton point sur la vision, la stratégie et la disponibilité, comme tu l'as mentionné, une partie de la vision, c'est de penser à comment utiliser les données pour l'innovation de manière fluide, tout en étant sûr qu'elles respectent la vie privée, mais aussi en étant automatisées avec des cadres flexibles. Ensuite, pour la stratégie, c'est un mot un peu galvaudé, mais c'est quand même super important : comment avoir l'accès à des outils qui démocratisent l'accès à un écosystème complet de solutions, plutôt que de se limiter à des solutions isolées. Et ensuite, en faisant ça, comment on s'y prend ?

05:01

On permet l'explicabilité et la nécessité de la gouvernance pour qu'on puisse tous travailler ensemble de manière flexible et rapide. Ensuite, il y a la disponibilité : comment on fait pour garantir la confidentialité tout en rendant les données utiles pour la recherche, par exemple pour des cas d'utilisation comme celui-ci. Mais aussi pour d'autres cas d'utilisation avec des contrôles spécifiques, en s'assurant que la vie privée est protégée grâce à une approche "privacy by design" via notre cadre de gouvernance qui adopte une perspective axée sur la confidentialité avec les contrôles nécessaires. C'est là qu'on envisage la vision.

05:35

Stratégie et disponibilité en lien avec notre mission. Merci beaucoup, John, pour avoir présenté de manière excellente la vision de Mastercard sur comment vous comptez utiliser les données pour soutenir l'innovation et démocratiser l'accès à travers l'écosystème de Pet. Laisse-moi maintenant passer à Salil. Salil, en plus de ton boulot de professeur d'informatique et de mathématiques appliquées, tu es aussi le directeur académique du projet open DP, un projet de logiciel open-source, et tu es reconnu mondialement comme un expert en technologies de protection de la vie privée.

06:15

"En gros, la protection de la vie privée différentielle, que certains considèrent comme la référence en la matière. Ma question pour toi, c'est pourquoi on dit ça et est-ce que tu peux expliquer comment cette protection de la vie privée permet la collaboration interdisciplinaire, surtout dans la santé publique et l'inclusion financière ? C'est super d'être ici sur le panel avec toi, Yui, John et Shubham, et j'ai hâte de discuter de ces

questions et d'entendre les interrogations du public plus tard. Donc, ouais, la protection de la vie privée différentielle, c'est vraiment un atout pour améliorer la confidentialité."

06:57

La technologie que je maîtrise est un domaine qui m'intéresse beaucoup. Le terme "DP" dont on a parlé se réfère à la "differential privacy", ou confidentialité différentielle en français. On est en train de développer des logiciels open source pour permettre et démocratiser, comme l'a dit John, l'utilisation de la confidentialité différentielle, surtout en tant que technologie d'amélioration de la vie privée. Maintenant, la confidentialité différentielle est parfois, et je pense à juste titre, qualifiée de norme d'or, mais ça ne veut pas dire qu'elle doit être en opposition avec d'autres technologies d'amélioration de la vie privée, mais plutôt que c'est...

07:37

Le standard d'or pour ce qu'il est censé résoudre, et je vais développer ça dans un instant, c'est que ça complète d'autres technologies de confidentialité qui fonctionnent avec la confidentialité différentielle pour s'attaquer à d'autres préoccupations concernant la vie privée et la confiance. Donc, la confidentialité différentielle s'intéresse à comment s'assurer que les données que tu partages réellement, les données que tu publies, les statistiques que tu balances, les modèles d'apprentissage automatique que tu entraînes, comment tu fais pour que tout ça ne reflète que la population à laquelle ça se rapporte.

08:19

Ils sont développés ou entraînés sur des données qui ne reflètent pas des informations spécifiques sur des individus. D'accord ? Donc, d'autres technologies visant à améliorer la confidentialité portent sur le processus de création de ce modèle d'une manière qui ne nécessite pas que différentes entités regroupent leurs données au même endroit, par exemple. Elles traitent d'autres problèmes de confiance et de la façon dont on peut travailler sur des données sensibles. Mais la protection différentielle de la vie privée concerne plutôt le produit final que l'on publie.

08:54

Tu mets à disposition comment tu fais pour t'assurer que ce n'est pas trop révélateur ? En gros, ça se fait en ajoutant des petites quantités de bruit aléatoire bien calibrées dans les calculs statistiques et dans les modèles d'apprentissage machine, pour que les infos sur des individus spécifiques soient floutées, tout en laissant passer des signaux concernant la population dans son ensemble. Et la raison, c'est que c'est la seule théorie systématique qu'on a pour réfléchir à ces types de statistiques.

09:33

des publications qui nous permettent d'être sûrs que tout ce qu'on publie dans l'ensemble ne peut pas être combiné pour reconstituer des informations sur des individus. Les tentatives passées de le faire de manière un peu improvisée ont montré qu'elles échouaient constamment. Donc, pourquoi la vie privée différentielle favorise-t-elle la collaboration interdisciplinaire ? La raison, c'est qu'on obtient le meilleur retour sur investissement quand c'est vraiment adapté à l'utilisation prévue des données, et ça va être très différent si...

10:08

C'est un, euh, un chercheur en santé publique qui étudie les données pour comprendre comment une maladie pourrait se transmettre ou l'impact d'un confinement, ou bien un économiste qui cherche à comprendre les relations entre les revenus des parents et les revenus futurs de leurs enfants pour saisir la mobilité économique. Et donc, il faut adapter la manière dont on introduit du bruit, quels statistiques on choisit de publier, à quel niveau de détail on les présente, lesquelles on priorise pour l'exactitude, euh, tout ça pour le faire de la meilleure façon possible.

10:51

Il est possible de tirer le meilleur parti des données tout en garantissant une protection forte de la vie privée. Ça nécessite que des experts en vie privée et en confidentialité différentielle collaborent avec les experts du domaine qui vont analyser les données. Merci, merci, Salil. Alors, John, tu nous as expliqué comment une organisation du secteur privé publie des données pour soutenir l'innovation, et Salil est venu parler de comment s'assurer que ces données publiées ne mettent en avant que les bonnes variables. J'aimerais qu'on ancre ça.

11:36

Bien sûr, Shubham, toi et tes collègues de Triple IT Camp êtes l'un des quatre finalistes de notre défi pour la santé publique, qui portait sur l'utilisation de données ouvertes transactionnelles, de la vie privée différentielle et de l'apprentissage machine pour informer des modèles épidémiologiques pour la santé publique en Colombie et, à terme, en Inde. Ma question est : peux-tu nous expliquer comment ces données ont été utilisées pour modéliser le comportement des pandémies et identifier les zones à risque ? Oui, bien sûr, merci et je suis...

12:25

C'est super d'être ici, alors euh, les technologies qui renforcent la confidentialité nous permettent d'accéder à une énorme quantité de données de transactions stockées par les banques et les systèmes de paiement, euh, avec leurs propriétés spatio-temporelles, euh, sans compromettre la vie privée des contributeurs. Euh, les données de transactions améliorées en matière de confidentialité peuvent nous donner des informations sur le comportement transactionnel en temps réel des gens. Euh, ce comportement transactionnel peut être utilisé pour améliorer l'analyse épidémiologique pendant les différentes phases d'une pandémie, euh, des gens.

13:08

Les comportements de consommation ont changé de manière drastique. Par exemple, on observe souvent un transfert des dépenses des biens de luxe vers des articles plus essentiels. Ces changements dans les habitudes de dépenses ont été utilisés pour modéliser le comportement pendant la pandémie. Une des caractéristiques particulières des données de transaction permet de visualiser les transactions à différents niveaux de codes postaux. Plus il y a de transactions dans une zone, plus il y a de chances d'interaction entre les gens.

13:54

Les gens sont donc plus nombreux, ce qui augmente le risque de propagation des maladies. Du coup, les zones économiques chaudes sont directement liées aux potentiels foyers de pandémie. C'est comme ça qu'on a essayé de modéliser notre comportement face aux pandémies et aux foyers, en utilisant les données de transaction. Merci, Shua. Alors, on a passé en revue un paysage qui est vraiment complet : de l'origine des données au traitement des données, jusqu'à l'utilisateur des données, qui dans ce cas est le SHBA. J'aimerais maintenant élever la conversation à un autre niveau.

14:36

Engagement ou détails, en regardant certains des défis et l'environnement pour mettre en œuvre cette technologie liée aux animaux de compagnie. J'aimerais commencer par toi, John, si ça ne te dérange pas. Pour une entreprise qui en est encore aux débuts de son parcours dans le domaine des animaux de compagnie, comment peut-elle équilibrer les aspects d'éthique, de risques, de valeur et de croissance dans cette approche ? Et j'ajouterais une autre question : comment Mastercard a-t-elle testé les cadres de gouvernance nécessaires et un ensemble de capacités intégrées pour les animaux de compagnie ?

15:21

Une salle de données propre, c'est un peu comme un gros sandwich d'éthique, de risque, de valeur et de croissance. Alors, si on y pense, pour commencer, ce qui a bien fonctionné pour nous, c'est d'adopter une approche axée sur la vie privée avec la stratégie PET et de comprendre ce qu'on cherchait à résoudre et pourquoi, plutôt que de commencer par le comment et les technologies. C'est super important. Ensuite, il s'agit de trouver le bon mélange de parties prenantes en fonction de vos données et de vos cas d'utilisation. Voilà la première partie.

15:57

On a commencé par ça et ensuite, en réfléchissant, on a aussi mis en place un groupe de gouvernance sur les animaux de compagnie. Ce groupe, qui regroupe des gens de différents secteurs, aide non seulement avec les évaluations dans les projets pilotes, mais aussi avec les étapes vers la production et le leadership éclairé qu'on doit avoir pour l'éducation, tant en interne qu'en externe. La deuxième partie, c'est que quand on pense à ce que débloque un cadre de gouvernance, ça prend en compte tous ces différents acteurs avec des mécanismes de contrôle spécifiques, mais ensuite, il y a des astuces.

16:31

Comment tu fais ça dans un processus flexible et répétable, tu sais, basé sur les données ? Donc, le cadre lui-même est axé sur les parties prenantes. Par exemple, comme on en parlait tout à l'heure, avec les différentes technologies, comment on pense à ça comme un écosystème de pets ? Et avec cet écosystème qui a différentes capacités, on peut l'utiliser pour plusieurs cas d'usage. Ce qui est vraiment cool, c'est qu'on peut voir où plusieurs pets peuvent être utilisés. Par exemple, une salle de données sécurisée peut être un moyen d'exécution, mais ça peut aussi...

17:04

Ce ne sera pas la seule façon, et d'autres technologies pourraient être nécessaires pour que ce cadre aide vraiment à cette approche axée sur la vie privée, ainsi qu'à impliquer les bonnes parties prenantes tout au long du processus, alors qu'on avance dans les évaluations et les différentes pistes. Et puis, en bossant sur la production, c'est super, j'adore le concept d'un écosystème de données personnelles. Je vais me tourner vers toi maintenant. Donc, en s'appuyant sur ce que John a dit, peux-tu expliquer comment les compromis historiques entre la vie privée et l'utilité sont abordés par le biais de la différenciation ?

17:43

La confidentialité est l'une des techniques principales, c'est super. Je suis content d'en parler. Ça fait écho à ce que j'ai dit en réponse à ta première question. Mais avant ça, pour rebondir sur ce que John a dit au sujet de la salle blanche, c'est un excellent exemple d'un cadre où ce que tu mets en place, donc la salle blanche, c'est ton environnement d'exécution qui offre certaines protections et limites, où les données peuvent circuler de manière contrôlée, et ça fait partie de la confidentialité différentielle.

18:25

C'est une des technologies qui te donne des outils pour savoir ce que tu peux vraiment calculer dans une salle blanche et ce qui peut traverser la limite entre la salle blanche et l'extérieur. Un super exemple de comment tout ça fonctionne ensemble dans cet écosystème de différentes technologies. Donc, la vie privée par rapport à l'utilité, comme je l'ai mentionné plus tôt, la confidentialité différentielle repose sur une des vérités fondamentales concernant la vie privée des données : tu ne peux pas tirer du lait des données.

19:00

Arbitrairement, on produit une quantité aléatoire d'analyses, on publie une quantité aléatoire de statistiques, on entraîne un nombre arbitraire de modèles de machine learning sur les mêmes données sans se mettre en danger face à de sérieuses vulnérabilités. Et ça a encore été montré à plusieurs reprises. Le Bureau du recensement des États-Unis a réalisé que la façon dont il publiait les tableaux lors du recensement de 2010 et auparavant était vulnérable à ça, parce qu'ils publiaient tellement de statistiques toutes basées sur le même ensemble de données collectées auprès de la population.

19:35

Ça arrive aussi dans l'industrie tech, et ce que la confidentialité différentielle offre, c'est une théorie quantitative pour réfléchir à combien de perte de vie privée on subit avec chaque publication supplémentaire. Ça te permet de peser combien de précision je peux publier, comment je publie les statistiques, quel bénéfice j'obtiens de ce modèle d'apprentissage automatique par rapport à l'impact sur la vie privée. Ça te permet de t'arrêter, de tracer une limite et de dire "voilà mon budget, je ne suis pas prêt à exposer les sujets des données, les personnes qui..."

20:12

Nous analysons ça par rapport à des risques de confidentialité plus importants, et c'est là que la vie privée différentielle nous aide à gérer ce compromis entre la vie privée et l'utilité, et à y penser de manière systématique. Merci, Salil, ça avance bien. Donc, John, tu nous as

présenté ce concept de l'écosystème des PET et tu as motivé l'idée d'une salle blanche. Cil, tu nous as maintenant mis en garde, du point de vue de la vie privée différentielle, qu'on ne peut pas, et j'adore ta phrase, qu'on ne peut pas exploiter les données de manière arbitraire sans s'exposer.

20:48

Des vulnérabilités sérieuses, donc laissez-moi reformuler ça comme je l'ai fait la première fois et venir vers toi, Shuba, pour te poser une question basée sur ta participation à notre défi. Avec des exemples concrets ou des cas d'usage, peux-tu nous donner un aperçu de comment les données de transactions financières peuvent vraiment soutenir la santé publique ? Ça montrerait la puissance de la prise de décision interdomaines, y compris la compréhension des relations entre des variables aléatoires ou la capacité à faire des prédictions sur l'avenir. Ouais, donc, les données de transactions...

21:33

Particulièrement dans la catégorie des dépenses de transport, on peut déduire la mobilité, qui est super utile pour prévoir de nouveaux cas. Donc, ces données financières peuvent être modélisées avec quelques paramètres supplémentaires pour prévoir de nouveaux cas. Maintenant, parlons d'un autre exemple : les données financières peuvent nous donner un aperçu de l'impact des politiques mises en place pendant la pandémie sur le comportement des gens normaux. Par exemple, pendant le Covid, à cause de la fermeture des grands...

22:19

L'établissement a un impact plus important sur l'expansion des gens que les normes de distanciation sociale. Maintenant, euh, en parlant d'un autre cas, euh, les groupes sociaux peuvent aussi être liés aux catégories de dépenses, et à travers ça, il y a un moyen, sur lequel nous travaillons encore, d'estimer les matrices de contact. Ça a une grande utilité en épidémiologie. Merci, c'est super. Euh, merci, Shuba. J'aimerais, euh, accueillir ceux qui nous ont rejoints et commencer à réfléchir à vos questions.

23:05

Alors, on est dans le dernier round de questions pour les panélistes, donc n'hésitez pas à utiliser le chat pour poser vos questions. John, voici ma troisième et dernière question pour toi : peux-tu nous donner des exemples de comment les insights des données de MasterCard pourraient créer une finance équitable ? Oui, comme je l'ai dit avant, notre mission est de connecter et de renforcer une économie numérique inclusive qui profite à

tout le monde, partout. On en a déjà parlé, mais il y a aussi la manière dont nous et d'autres entreprises pouvons tirer parti de ça.

23:49

Des capacités qui sont conçues pour penser de cette manière, identifier, quantifier mais aussi atténuer les biais dans les ensembles de données. C'est la base de ce que SE et d'autres vont faire avec ces ensembles de données. Ce n'est pas juste pour s'assurer que ces modèles ne produisent pas de biais, mais aussi pour évaluer et traiter si les ensembles de données eux-mêmes sont biaisés. Donc, commencer par cette base, ça nous fait vraiment réfléchir à comment tout ça, avec notre mission, peut soutenir des écosystèmes financiers équitables en examinant les données fondamentales elles-mêmes et ce qu'elles représentent.

24:25

On peut tous faire des efforts pour identifier et atténuer les biais, excellent, c'est super. Je vais me tourner vers toi, Salil, parce que pour atténuer ces biais, il faut qu'on fasse un vrai filtrage, non ? Alors, quels défis as-tu rencontrés en appliquant la DP dans différents secteurs et comment peut-on surmonter ces obstacles pour maximiser l'impact des données ? Ouais, super, chaque domaine où on essaie d'apporter des technologies pour améliorer la vie privée soulève de nouveaux défis.

25:13

Certains de ces problèmes sont similaires à ceux dont parlait John, d'autres concernent comment les gens dans un domaine particulier doivent analyser les données. Quel type de signal est le plus important à préserver et lequel est moins crucial ? Parce que faire ces choix, comme on l'a dit, fait partie de l'optimisation des technologies de protection de la vie privée, et en particulier de la confidentialité différentielle. On ne peut pas rendre les données disponibles pour toutes les utilisations sans risquer de les exposer.

25:47

Nous nous exposons à de sérieuses vulnérabilités, donc la façon de relever ce défi, c'est d'avoir des groupes comme celui de Shum et les autres participants dans le défi des animaux de compagnie qui soient créatifs et innovants pour développer les meilleures manières d'utiliser ce genre de données pour un bien social, pour la santé publique, pour comprendre quels signaux on peut exploiter et quelles sont les façons créatives de tirer des enseignements de tout ça. Et puis, l'autre chose que je voudrais juste conclure...

26:20

Avec notre expérience, euh, en déployant la vie privée différentielle et les PETs, à chaque nouveau secteur où on amène cette technologie, il y a toujours un peu de méconnaissance de cette technologie. Les lois, règlements et politiques des organisations ne sont pas encore adaptées pour prendre des décisions en tenant compte de cette technologie. Du coup, il peut y avoir pas mal de frictions saines qui rendent les premiers déploiements d'une technologie dans un nouveau domaine difficiles, et on a clairement rencontré ça, tu sais.

26:58

Toutes nos organisations essaient de collaborer sur le défi des animaux de compagnie à Harvard. Il y a beaucoup de prudence quand on traite des données sensibles, et c'est pour de bonnes raisons. Mais je pense que si on répète ce processus plusieurs fois, on accumule de l'expérience. D'autres dans le même secteur peuvent s'inspirer de ces exemples et des meilleures pratiques vont commencer à se développer, ce qui réduira les frictions à l'avenir. Merci, Salil. Pour ceux dans l'audience, je vous invite encore une fois à poser vos questions si vous en avez.

27:29

Pour nous, John, je t'ai dit que c'était ma dernière question, mais j'ai menti, alors sois indulgent, je vais te poser une dernière question en rapport avec le dernier point de S. MasterCard a été un super partenaire pour le projet PETS sur lequel on collabore avec Harvard et d'autres. Donc, j'aimerais te demander, quelle est la suite pour MasterCard en termes de collaboration avec des partenaires pour utiliser les données de MasterCard à travers PETS ? D'accord, donc pour la suite, il y a plusieurs choses à considérer pour rester cohérent. D'abord, il faut qu'on continue l'innovation.

28:10

Le leadership dans ce domaine, tout en étant axé sur les pilotes et la disponibilité de ces solutions pour les entreprises, ainsi que pour les ONG et les organisations à but non lucratif. Ensuite, comment on fait pour rester en phase avec nos partenaires dans l'univers de la tech de la vie privée pour profiter de leurs services ? Après, il y a la question de la systématisation. Il y a plein de façons d'utiliser nos insights avec les clients et les partenaires, mais il faut être conscient qu'il doit y avoir une confiance réciproque, sinon ça peut poser problème. Donc, comment peut-on continuer à chercher à faire évoluer les contrôles ?

28:48

"C'était censé être contractuel jusqu'à ce que ça soit systématisé dans des livrables spécifiques pour les animaux de compagnie. Comme je l'ai dit, des choses comme les salles blanches rendent vraiment les données plus disponibles. Et puis, comme SE l'a déjà mentionné, il y a un niveau de transparence dans le domaine de la technologie de la vie privée qui est super important. Donc, comment continuer cette réflexion pour répondre aux enjeux de la vie privée, de la confidentialité, mais aussi de la transparence du résultat final ? C'est vraiment cette transparence qui compte."

29:23

où on veut lutter contre le fait que les animaux de compagnie puissent être compris de manière non superficielle, ce qui est parfois presque une forme de déni plausible ou un genre de, euh, j'emploie l'expression théâtre de la vie privée par rapport à ses véritables bénéfiques. C'est là que je mentionne que la transparence est super utile et que la compréhension des solutions elles-mêmes est géniale. En fait, je suis content de t'avoir posé cette question parce que maintenant je peux revenir à Shuba, tu sais, on peut imaginer que la transparence est importante dans, euh.

30:03

Mettre en place des outils pour les animaux de compagnie sur le terrain, dans les différentes organisations à travers le monde, surtout dans le Sud global. Ma question pour toi, Shba, c'est en lien avec le concept de localisme : comment ton outil va-t-il permettre d'obtenir des informations en temps réel pour les responsables de la santé locaux, comme ceux de Bota en Colombie ou de New Delhi en Inde, tout en garantissant la confidentialité des données ? Et surtout, quelles sont les leçons clés de ton projet qui peuvent être appliquées à l'échelle mondiale ? Ouais, donc notre outil utilise la confidentialité différentielle pour protéger...

30:54

Les contributeurs, euh, la vie privée, comme euh, Salil va, euh, bien définir la vie privée différentielle. En gros, avec ça, on ajoute un bruit calculé mathématiquement pour préserver la vie privée lors de la publication de statistiques. Dans notre cas, euh, on prend le marchand comme notre unité de vie privée, donc on sécurise les marchands au niveau des marchands. Maintenant, euh, les responsables de la santé locaux et le gouvernement peuvent euh, utiliser leurs propres données de transaction locales pour avoir un aperçu en temps réel des potentiels points chauds. Ils peuvent aussi euh, obtenir...

31:41

un aperçu de la façon dont la pandémie évolue, euh, à quel stade elle en est actuellement. Euh, il y a trois stades : le début, le pic et la phase de récupération. Grâce à ces informations, ils peuvent euh, prendre de meilleures décisions sur les politiques à mettre en place dans leur région. Maintenant, en ce qui concerne l'idée globale, ce qu'on a appris, c'est que c'est un outil très prometteur et puissant qui nous permet d'accéder à de gros volumes de données.

32:16

inaccessible à cause de préoccupations liées à la vie privée, et ça peut être utilisé pour créer des outils, des méthodes et des fonctions, y compris celles qu'on est en train de développer ici pour aider les épistémologues et les décideurs politiques. Merci, Shuba. Je vais rester sur le terrain, en parlant de ce concept de localisme, surtout en ce qui concerne les animaux de compagnie et son application pour ceux du Sud global ou dans les pays à revenu faible ou intermédiaire. Alors, Salil, ma question s'adresse à toi, si ça ne te dérange pas : quelles innovations, selon toi, vont faire bouger les choses ?

33:00

L'avenir des animaux de compagnie, surtout pour les pays en développement, est super important. C'est d'ailleurs l'une des raisons pour lesquelles on a lancé le projet Open DP. Ça se relie à ce que John a dit plus tôt sur la démocratisation des animaux de compagnie. Quand on a démarré le projet il y a cinq ans, les seules organisations qui utilisaient la confidentialité différentielle, c'étaient de grosses organisations bien dotées en ressources et avec beaucoup d'expertise, surtout dans les pays développés comme les États-Unis. Par exemple, le Bureau du recensement américain, qui a une énorme expertise statistique accumulée au fil des ans.

33:49

"Ça fait un siècle ou plus d'expérience en matière d'évitement de divulgation, avec des grosses entreprises tech comme Google, Apple et Microsoft, bien sûr, qui ont une énorme expertise technique. Si d'autres organisations, même celles dans le secteur technologique, voulaient essayer d'adopter ces technologies, elles devaient recommencer à zéro, comme partir d'une feuille blanche. Elles devaient même refaire les bases elles-mêmes, et c'est super difficile d'y arriver, sans parler de tout le côté éducatif pour comprendre ce que c'est."

34:27

La valeur du POS que tu peux obtenir avec cette technologie, qu'est-ce qu'elle protège et comment l'utiliser en toute sécurité ? Désolé, je ne veux pas être brusque, mais John a une

urgence et il doit partir. C'est le monde dans lequel on vit en ce moment. Je m'excuse, l'équipe, merci beaucoup. J'ai hâte de la suite de cet événement et de rendre plus de données et de capacités disponibles, et de terminer le reste de ce défi à l'avenir. Merci beaucoup à tous, j'apprécie vraiment l'implication de chacun.

34:57

À l'intérieur, merci John, euh, au revoir John S. Je suis désolé d'interrompre, mais euh, il devait s'absenter. Ouais, pas de souci, je vais conclure rapidement. Prenez votre temps. Je ne vois pas encore de question du public, ok, mais c'est important parce que je viens de jeter un œil à la liste des participants à ce webinaire et il y a pas mal de gens du Sud global. Je suis sûr qu'ils se demandent : "Est-ce encore une technologie du Nord global ? Quel est l'intérêt et la pertinence des animaux de compagnie ?"

35:31

Dans mon entourage, c'est pour ça que je te pose cette question. Oui, oui, donc, euh, où je voulais en venir, c'est que, euh, au moins une partie du problème, c'est d'avoir des outils qui sont, euh, fiables, en open source et qui sont, euh, facilement adoptables par des organisations qui, euh, n'ont pas forcément d'expérience préalable avec ces technologies et qui n'ont pas les ressources pour embaucher une équipe afin de créer leur propre solution de confidentialité différentielle sur mesure. Voilà, et en plus de ça, il y a un large éventail de ressources disponibles.

36:16

Des supports éducatifs qui permettent aux décideurs de comprendre la valeur de tout ça pour les politiques des organisations, afin de donner des conseils sur les meilleures pratiques sur comment utiliser cette technologie, ce sont quelques-unes des choses qu'on essaie de faire avec le projet de confidentialité différentielle ouverte. Bien sûr, juste produire les matériaux, ce n'est pas suffisant, il faut aussi un engagement actif. Il y a un groupe qui essaie de promouvoir son utilisation.

36:52

Les Technologies de protection de la vie privée au niveau mondial sont représentées par le laboratoire des Technologies de protection de la vie privée des Nations Unies, le UN PET Lab. On collabore avec eux, et ça implique des représentants d'organisations statistiques nationales du monde entier. Il y a des expériences et des prototypes en cours de création, aussi bien dans le Nord que dans le Sud. On espère que c'est un exemple de ce qui peut se passer dans d'autres secteurs pour favoriser l'adoption à l'échelle mondiale et impliquer le monde entier.

37:29

Profitez de ces technologies maintenant, merci Salil. En fait, il y a une question du public qui fait écho à ce que tu viens de dire. J'aimerais la lire et commencer avec toi, Salil, et ensuite voir si Shubham a des idées. Alors, Salil, voici la question de Colin Rice : je suis curieux de savoir comment l'utilisation des animaux de compagnie, surtout avec des données intersectorielles, est améliorée ou compliquée par l'émergence de réglementations nationales sur la confidentialité des données intersectorielles, par rapport à celles spécifiques à un secteur.

38:09

Des réglementations de vie privée spécifiques, super ! C'est vraiment un des espoirs, que les technologies qui protègent la vie privée puissent nous aider à franchir les barrières actuelles, celles qui bloquent le flux des données, et qui nous empêchent de profiter des analyses qu'on pourrait faire en combinant des données de différents pays à l'échelle mondiale. Un des prototypes test qui a été réalisé par le laboratoire de l'ONU dont j'ai parlé plus tôt portait sur un cas de données de commerce international.

39:01

Une computation conjointe sur des données venant de plusieurs pays différents, c'était juste un prototype de test, pas sur des données sensibles réelles, mais plutôt comme une preuve de concept que c'est faisable. Donc, construire la technologie, c'est juste la première étape. Ensuite, il faut que les régulateurs et les décideurs soient à l'aise pour établir les meilleures pratiques, et idéalement, il faudrait ajuster les lois et les règlements pour que les technologies de protection de la vie privée soient utilisées de manière appropriée.

39:31

Ils doivent fournir les protections nécessaires pour que les données puissent circuler plus facilement pour ce genre d'utilisation transfrontalière, ça a beaucoup de sens. Merci, Salil. Voici une question pour toi, Shubham, et merci d'être patient. Ça dit : j'aimerais savoir quelles autres applications il pourrait y avoir pour utiliser les données financières afin de tirer des insights ou des résultats en santé publique, surtout dans des contextes plus petits qu'une pandémie. Ouais, donc il peut y avoir plein d'applications des données financières, mais on a besoin de deux choses basiques.

40:24

Alors, pour que ça fonctionne, il faut d'abord qu'il y ait une tendance dans les données financières ou dans les catégories de dépenses liées à cet événement. Si la catégorie de

dépense concerne cet événement, alors tu peux vraiment en profiter. La deuxième condition, c'est qu'il doit y avoir suffisamment de données. Comme on va ajouter du bruit pour que ce soit privé de manière différentielle, il faut que les données soient assez volumineuses pour que même avec ce bruit, on puisse toujours capter la tendance de l'ensemble de la population.

40:55

niveau donc ouais, il peut y avoir plein de cas d'utilisation mais il faut que ça se reflète sur une grande population, excellent. Alors, t'as des idées là-dessus aussi ? Je sais que c'était adressé à Shuan, mais dans ton expérience, surtout du côté DP, quelles autres applications t'as vues, surtout dans le travail qu'on fait avec toi, où les données financières peuvent vraiment aider à obtenir des insights plus profonds, pas seulement en lien avec la pandémie, mais dans des contextes à plus petite échelle ? Ouais, c'est une super question.

41:45

Excellente question, euh, alors, laissez-moi réfléchir un moment et si j'ai une idée à ajouter, je vous le dirai. Pendant que vous pensez, laissez-moi en poser une à Shubham, euh, en essayant de conclure ici sur la même lancée. Alors, comment les gouvernements locaux ou les responsables de la santé peuvent-ils profiter des outils que vous avez développés, euh, en Inde ? Plus précisément, comment cela peut-il être appliqué et utilisé en Inde par les ministères de la Santé ? Oui, donc, les gouvernements locaux peuvent utiliser notre outil pour obtenir des données en temps réel.

42:31

un aperçu des zones chaudes, euh, détecter les foyers d'infection, euh, donc, euh, en utilisant ça, ils peuvent ajuster leurs politiques sur, euh, le confinement ou décider s'ils doivent, euh, imposer un confinement dans certaines zones et s'ils doivent le lever, euh, un autre, euh, usage, c'est de voir comment le comportement des gens change à cause de la pandémie, comment la pandémie progresse, donc, euh, ils peuvent visualiser, euh, l'évolution de la pandémie dans différentes zones et voir dans certaines zones si, euh, les gens se rétablissent plus lentement ou pas.

43:13

Que ce soit dans certains domaines, les gens récupèrent de manière plus coûteuse pour pouvoir ajuster leurs politiques en conséquence. Merci ! Alors, comme il ne nous reste qu'une minute, Salil, je vais te redonner la parole, mais je vais ajouter une autre question : quel est ton appel à l'action ? Super, d'accord. Pour répondre à la question précédente sur

d'autres utilisations des données de transactions financières, un exemple qui a beaucoup de potentiel est très similaire à une utilisation récente d'un logiciel d'ODP.

43:54

Microsoft s'est penché sur la fracture économique entre différentes régions. Ce qu'ils ont étudié, c'est la littératie numérique : comment les applications que les utilisateurs utilisent sont liées à la situation économique de leur région. Est-ce que les gens dans des zones plus aisées utilisent des applications différentes sur leurs ordinateurs ou leurs téléphones par rapport à ceux dans des zones de pauvreté ou moins favorisées ? On peut imaginer aller au-delà de ça.

44:41

Des applications numériques dans plein de domaines d'achats et d'activités économiques, et comprendre comment ça varie d'une région à l'autre, entre les différentes tranches démographiques et classes socio-économiques. Ce qu'on veut, c'est que les gens s'impliquent. Par exemple, op DP est une communauté que tu peux rejoindre. Tu peux venir sur notre Slack, notre liste de diffusion, et il y a d'autres communautés qui essaient de faire avancer les choses et de les amener vers de nouvelles applications. En bossant ensemble, on peut les amener vers de nouveaux secteurs et de nouvelles utilisations.

45:19

Dans des cas où la société en bénéficie, les technologies vont progresser et on va acquérir plus d'expérience pour les utiliser plus efficacement pour le bien commun. Merci beaucoup, Salil. On arrive à la fin du temps imparti, mais une autre question vient d'arriver. Je tiens à remercier Salil, Shubham et John qui était avec nous plus tôt, pour avoir partagé des idées aussi intéressantes. Personnellement, je pense plus que jamais que les technologies qui protègent la vie privée ont un énorme potentiel pour soutenir l'inclusion financière et...

46:00

Prise de décision en santé publique. Je vais encourager tout le monde qui est avec nous aujourd'hui, comme Salil l'a dit dans son appel à l'action, à continuer d'explorer des communautés qui offrent des possibilités de collaboration interdisciplinaire, afin que nous puissions débloquer des données pour avoir un impact sur tous les cas d'utilisation critiques. C'est le travail que nous faisons chez data.org. Je vous invite à vous inscrire à notre newsletter et à nous suivre en ligne. Encore un grand merci à tous pour votre participation et aux intervenants. Passez une bonne après-midi.

46:36

"Bonjour et bonne journée à tous, au revoir."