NOVARA
CONSULTING GROUP

Ce document a été traduit de l'anglais par une machine et peut contenir des erreurs. L'original anglais se trouve à l'adresse https://www.novaracg.com/2026/07/27/the-consent-gap-undisclosed-data-collection-in-sign-language-ai-interpreting-systems/.

Le déficit de consentement : collecte de données non divulguée dans les systèmes d'interprétation en langue des signes par IA

Heather Grizzle · 27 juillet 2026

Sign language AI depends on video of Deaf users. What vendors disclose about collection, retention, training use and sharing is inconsistent or absent.

Sommaire

1. Résumé exécutif
2. Problème politique
3. Analyse institutionnelle
4. Implications en matière de gouvernance
5. Position politique de NCG
6. Considérations relatives à la mise en œuvre
7. Références

Résumé exécutif

L'intelligence artificielle en langue des signes, qu'elle soit conçue pour la reconnaissance, la traduction ou la génération basée sur des avatars, dépend structurellement de la capture vidéo ou biométrique des utilisateurs Deaf et malentendants. Cette dépendance crée un problème de gouvernance des données qui a fait l'objet d'un examen institutionnel bien moins approfondi que les défaillances linguistiques et de compréhension documentées dans les précédents numéros de cette série. Dans l'ensemble du paysage des fournisseurs que NCG a examiné dans le cadre du référentiel SLAT (Sign

Language Access Trust), la divulgation des données collectées, de leur durée de conservation, de leur éventuelle utilisation pour entraîner de futurs modèles, et de leur éventuel partage avec des tiers, est au mieux incohérente et au pire absente. Il ne s'agit pas d'affirmer qu'un fournisseur en particulier a agi de mauvaise foi. Les preuves publiques ne permettent pas encore d'attribuer une quelconque intention à une entreprise nommée, et ce numéro ne procède pas à une telle attribution. Il s'agit d'affirmer que le marché n'a pas encore développé une norme de divulgation adaptée à la sensibilité des données en jeu, et que cette lacune constitue une défaillance de gouvernance indépendante de la conduite d'un fournisseur en particulier. La position politique exposée ci-dessous soutient que l'adéquation du consentement et de la divulgation devrait être traitée comme un critère éliminatoire dans les achats, et non comme une considération secondaire abordée après l'évaluation fonctionnelle.

Problème politique

La capture en langue des signes est une capture biométrique. La vidéo des mains d'un signeur, de sa grammaire faciale et de sa position corporelle n'est pas une donnée accessoire, comparable à une phrase tapée soumise à un système textuel. Elle est unique et identifiante, elle encode par définition des informations sur le statut de handicap, et dans de nombreuses juridictions, elle correspond à la définition légale de données biométriques ou de catégorie spéciale soumises à une protection renforcée. La loi de l'Illinois sur la protection des informations biométriques (BIPA), par exemple, exige un avis écrit et un consentement avant la collecte d'identifiants biométriques, et impose une responsabilité indépendante de la preuve d'un préjudice réel. Le RGPD (Règlement général sur la protection des données) de l'Union européenne classe les données révélant des informations liées au handicap, ainsi que certaines données biométriques utilisées à des fins d'identification, comme des données de catégorie spéciale au titre de l'article 9, nécessitant une base légale explicite au-delà du consentement ordinaire. La CCPA (California Consumer Privacy Act) et la CPRA (California Privacy Rights Act) de Californie imposent leurs propres obligations de divulgation et de retrait pour les informations biométriques.

Malgré ce paysage réglementaire, les produits d'IA en langue des signes destinés aux consommateurs s'appuient fréquemment sur des conditions générales d'utilisation génériques qui n'ont pas été rédigées en tenant compte des utilisateurs Deaf ni de la capture biométrique. Trois lacunes récurrentes apparaissent dans les instantanés de four-

nisseurs réalisés par NCG : premièrement, l'absence de divulgation distincte ou visible indiquant que la vidéo constitue une donnée biométrique soumise à des protections légales spécifiques ; deuxièmement, l'absence de déclaration claire quant à savoir si la vidéo capturée ou les caractéristiques dérivées sont conservées pour l'entraînement des modèles, et le cas échéant, pendant combien de temps et selon quels droits de suppression ; troisièmement, l'absence de divulgation dans un format accessible, ce qui signifie que le libellé du consentement lui-même n'est souvent pas traduit en ASL ni présenté dans un format utilisable par la population dont les données sont collectées. Ce dernier point n'est pas un détail technique. Un régime de divulgation qui exige une maîtrise de la lecture d'une seconde langue écrite pour comprendre ce qui est fait de ses propres données biométriques ne fonctionne pas comme un consentement éclairé pour une population dont la maîtrise de l'anglais écrit varie considérablement et pour laquelle l'ASL est la langue principale de compréhension.

Analyse institutionnelle

> La capture en langue des signes est une capture biométrique.

Les dynamiques institutionnelles à l'origine de cette lacune sont structurelles plutôt que délibérées. L'IA en langue des signes est une catégorie de produits jeune, et la plupart des fournisseurs sont de petites entreprises utilisant les mêmes modèles génériques de politique de confidentialité employés dans l'ensemble du secteur logiciel, adaptés a minima, voire pas du tout. Les processus d'achat au niveau institutionnel, quant à eux, ont historiquement évalué ces produits selon des critères fonctionnels : l'avatar restitue-t-il un rendu intelligible, le système de reconnaissance atteint-il une précision acceptable, le fournisseur emploie-t-il du personnel Deaf impliqué dans le développement. La gouvernance des données a été traitée comme une question de conformité juridique ou informatique abordée après la décision d'achat substantielle, lorsqu'elle est abordée du tout. Cet ordre des priorités constitue la véritable défaillance de gouvernance. Au moment où une lacune de gouvernance des données apparaît, que ce soit par une violation, une enquête réglementaire ou une critique publique de la communauté Deaf dont les données ont été concernées, la relation institutionnelle et l'engagement financier sont déjà établis, et le pouvoir de négociation pour exiger des mesures correctives s'en trouve d'autant plus affaibli.

Il existe également une question de provenance des données d'entraînement qui côtoie le problème du consentement mais qui s'en distingue analytiquement. Plusieurs systèmes d'IA en langue des signes ont été construits en partie à partir de corpus vidéo issus de contenus publiés publiquement, y compris du contenu créé par des utilisateurs et éducateurs Deaf en ASL sur des plateformes sociales. La question de savoir si cette utilisation relève du champ d'application des propres conditions générales de la plateforme, si elle constitue un usage loyal (fair use) au regard du droit d'auteur, et si les personnes dont les gestes apparaissent dans ces séquences ont été véritablement informées que leur contenu serait utilisé pour entraîner un système d'IA commercial, sont des questions ouvertes que NCG n'a pas vérifiées de manière indépendante pour un fournisseur en particulier et qu'elle n'affirme pas ici comme un fait établi. Le point pertinent pour la gouvernance est que les responsables des achats évaluant ces systèmes n'ont presque aucune visibilité sur la provenance des données d'entraînement, et que les normes de divulgation actuelles du secteur n'obligent pas les fournisseurs à la fournir.

Implications en matière de gouvernance

Une institution qui procède à l'achat d'une IA en langue des signes n'est pas simplement l'acquéreur d'un outil de traduction. Lorsque le système traite des vidéos en direct d'étudiants, de patients, d'employés ou de membres du public, l'institution est également un responsable ou un sous-traitant de données au sens de la plupart des cadres de confidentialité applicables, et elle hérite de l'exposition en matière de gouvernance liée aux pratiques de données de son fournisseur, que cette exposition ait été divulguée ou non au moment de l'achat. Une institution qui n'a pas vérifié les pratiques de collecte, de conservation et d'utilisation pour l'entraînement d'un fournisseur avant son déploiement ne peut pas affirmer à ses propres parties prenantes, aux régulateurs, ou aux personnes Deaf dont les données sont capturées, qu'elle a respecté ses propres obligations de protection des données. Cela est vrai indépendamment de la conduite réelle du fournisseur, car l'absence de divulgation vérifiée constitue en elle-même le défaut de gouvernance, indépendamment du fait qu'un fournisseur donné traite ou non les données de manière responsable.

Position politique de NCG

La position de NCG est que l'adéquation de la gouvernance des données et de la divulgation en matière de confidentialité devrait fonctionner comme un critère éliminatoire dans les achats d'IA en langue des signes, évalué avant la performance fonctionnelle plutôt qu'après. L'incapacité d'un fournisseur à produire une divulgation claire, accessible en ASL, adaptée à la juridiction concernée, des données biométriques collectées, de leur durée de conservation, de leur éventuelle utilisation dans l'entraînement des modèles et des modalités, et des droits de suppression existants, devrait être considérée comme disqualifiante pour un déploiement institutionnel, quelle que soit la performance linguistique ou technique par ailleurs dans l'évaluation. Cette position est déjà opérationnalisée dans les domaines Gouvernance des données et confidentialité et Transparence et divulgation du référentiel SLAT, et ce numéro soutient que ces domaines méritent une pondération accrue par rapport à l'étalonnage actuel du référentiel, compte tenu de l'exposition légale à laquelle les institutions sont confrontées et de l'obstacle de compréhension que crée un langage de consentement non accessible pour la population concernée.

Considérations relatives à la mise en œuvre

Les responsables des achats évaluant une IA en langue des signes devraient exiger, comme condition de la réponse du fournisseur, une divulgation écrite de gouvernance des données couvrant les quatre éléments ci-dessus, demander que cette divulgation soit disponible au format vidéo ASL en plus de l'anglais écrit, et considérer le refus ou l'incapacité du fournisseur à produire cette divulgation comme une constatation défavorable indépendante de la performance fonctionnelle du système. Les institutions ayant déjà déployé ces systèmes devraient procéder à un examen rétrospectif plutôt que d'attendre le renouvellement du contrat, étant donné que l'exposition en matière de gouvernance s'accumule à partir de la date de déploiement plutôt qu'à partir de la date de découverte. Lorsque la provenance des données d'entraînement d'un fournisseur ne peut être vérifiée, les institutions devraient demander des garanties contractuelles concernant l'origine licite des données plutôt que de traiter l'absence de réponse comme une inconnue neutre.

LES QUATRE ÉLÉMENTS DE DIVULGATION

Quelles données biométriques sont collectées, pendant combien de temps elles sont conservées, si et comment elles sont utilisées pour l'entraînement des modèles, et quels droits de suppression existent.

Références

- Biometric Information Privacy Act, 740 ILCS 14 (Illinois).
- California Consumer Privacy Act of 2018, tel que modifié par la California Privacy Rights Act, Cal. Civ. Code § 1798.100 et suivants.
- Règlement (UE) 2016/679 (Règlement général sur la protection des données), art. 9.
- Federal Trade Commission Act, 15 U.S.C. § 45 (actes ou pratiques déloyaux ou trompeurs).
- Novara Consulting Group, *Référentiel Sign Language Access Trust (SLAT)*, 2026.

TIPS

Know something about a sign language AI system that the public should know? Send it to tips@novaracg.com. Every tip is read by me and kept confidential. If you work in the UK, writing to us is not by itself a protected disclosure under the Public Interest Disclosure Act 1998, so before you share employer information, check your options with Protect (protect-advice.org.uk), the free whistleblowing advice line.

Comment citer

Grizzle, H. M. (2026, July 27). Le déficit de consentement : collecte de données non divulguée dans les systèmes d'interprétation en langue des signes par IA. Novara Consulting Group. https://doi.org/10.5281/zenodo.23003818