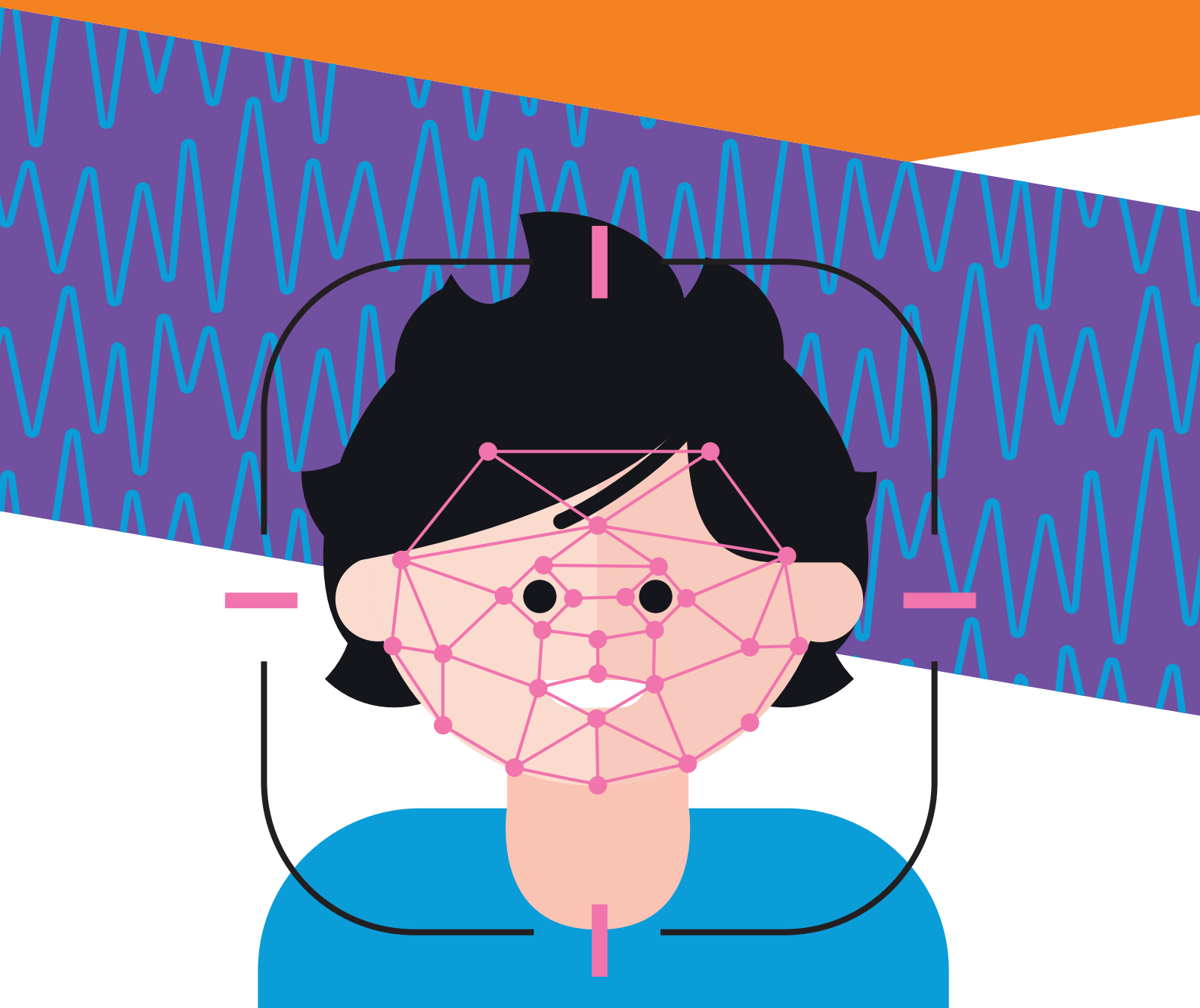
# Quand nos faits et gestes sont observés en permanence

**Synthèse de l'étude de TA-SWISS « Reconnaissance automatisée de la voix, de la parole et du visage : défis techniques, juridiques et sociétaux »**



TA-SWISS, Fondation pour l'évaluation des choix technologiques et centre de compétence des Académies suisses des sciences, entend mener une réflexion sur les répercussions – opportunités et risques – de l'utilisation de nouvelles technologies.

La synthèse se base sur une étude scientifique réalisée pour le compte de TA-SWISS par un groupe de projet interdisciplinaire composé de membres du Fraunhofer-Institut für System- und Innovationsforschung ISI à Karlsruhe (Allemagne) et de l'Université de Fribourg (Suisse), sous la direction générale de Dr Murat Karaboga. Cette synthèse présente les principaux résultats et les recommandations de l'étude sous forme condensée et s'adresse à un large public.

**Automatisierte Erkennung von Stimme, Sprache und Gesicht: Technische, rechtliche und gesellschaftliche Herausforderungen**

Murat Karaboga, Nula Frei, Frank Ebbers, Sophia Rovelli, Michael Friedewald, Greta Runge

TA-SWISS, Stiftung für Technologiefolgen-Abschätzung (éd.)
vdf Hochschulverlag an der ETH Zürich, 2022.
ISBN : 978-3-7281-4140-8

L'étude est disponible en libre accès : www.vdf.ch

La présente synthèse peut également être téléchargée gratuitement : www.ta-swiss.ch

# La reconnaissance de la voix, de la parole et du visage : introduction

**Installées un peu partout dans l'espace public, les caméras recueillent un riche matériau destiné à alimenter les programmes de reconnaissance de la voix, de la parole et du visage. Utilisées à bon escient, ces technologies peuvent contribuer à renforcer la sécurité individuelle et publique en aidant les autorités à retrouver des personnes disparues ou à surveiller des individus suspects. Mais les systèmes d'identification automatique peuvent également être détournés de leur usage et pousser les gens à aligner leur comportement aux normes sociétales. Cela a de multiples conséquences, non seulement pour les individus, mais aussi pour la démocratie.**

Les enceintes intelligentes (ou haut-parleurs intelligents) sont désormais chose courante dans bon nombre de foyers en Suisse et les gens ont pris l'habitude de déverrouiller leur smartphone à l'aide de la technologie Face-ID. La reconnaissance de la voix, de la parole et du visage simplifie de nombreuses actions du quotidien mais ces technologies ne sont pas toujours fiables : la reconnaissance faciale, notamment, identifie les femmes et les personnes à la peau foncée avec moins de précision que les hommes blancs. Toutefois, si les progrès se poursuivent au rythme actuel, il est probable qu'un niveau élevé de fiabilité technique sera atteint dans les années à venir – même si le risque de résultats erronés reste grand lorsque les données vocales et faciales sont utilisées pour tirer des conclusions sur l'état émotionnel d'une personne, ou sur sa santé physique ou mentale.

## Quelques opportunités …

Grâce aux enceintes intelligentes, de nombreux appareils peuvent être commandés par la voix, ce qui rend la multiplication de télécommandes superflue. Les assistants virtuels simplifient la vie quotidienne : tandis qu'ils enregistrent des entrées dans le calendrier et règlent la chaîne stéréo ou l'éclairage par commande vocale, l'utilisatrice ou l'utilisateur garde ses deux mains libres.

Dans le domaine médical, les systèmes de reconnaissance de la voix, de la parole et du visage pourraient contribuer au dépistage précoce d'affections graves comme la maladie de Parkinson ou d'Alzheimer, la dépression ou le burn-out. Des travaux de recherches sont également en cours sur des programmes qui, grâce à la reconnaissance faciale, sont capables d'identifier certaines pathologies rares que peu de professionnels de la santé ont l'occasion de croiser.

Dans certaines circonstances, repérer et suivre les actions d'individus suspects peut permettre de démasquer à temps leurs intentions criminelles. Selon les cas, les applications de reconnaissance de la voix, de la parole et du visage peuvent donc renforcer la sécurité publique.

## … et quelques risques

Les données vocales et faciales sont des données biométriques qui livrent beaucoup d'informations sur une personne et ne changent plus guère au cours de la vie adulte. Une fois piratées, elles sont définitivement compromises.

Parce qu'elle fait sortir les gens de l'anonymat, la reconnaissance de la voix, de la parole et du visage, met en péril leur vie privée. Les enceintes intelligentes qui se trouvent dans le salon, par exemple, sont au cœur du foyer. En permanence en mode écoute, elles sont susceptibles d'enregistrer des informations confidentielles.

Les technologies de reconnaissance de la voix, de la parole et du visage ont tendance à accentuer l'inégalité de pouvoir entre la population et les services publics ou les entreprises du secteur privé qui y ont recours – à fortiori lorsque les citoyennes et citoyens ne savent pas quelles données les autorités ou les acteurs du secteur privé détiennent à leur sujet.

## Recommandations principales

Certaines applications de reconnaissance de la voix, de la parole et du visage particulièrement problématiques doivent être interdites, notamment la surveillance automatique en temps réel, les lunettes

connectées et tout autre dispositif discret qui permette de suivre les gens à leur insu. Les systèmes d'analyse de l'attention dans les écoles ne devraient pas non plus être autorisés.

Le recours à la reconnaissance de la voix, de la parole et du visage par la police ou d'autres autorités requiert un cadre légal explicite qui définit les mécanismes de garantie de l'État de droit et impose l'examen du caractère nécessaire de cette pratique.

Des formations de base et des cours de perfectionnement pour toutes celles et ceux qui utilisent ces technologies sont nécessaires. Des centres de conseil ou points de contact doivent aussi être mis sur pied pour les personnes désireuses de se protéger des inconvénients liés aux systèmes de reconnaissance de la voix, de la parole et du visage, et qui souhaitent connaître et faire valoir leurs droits.

Les avantages et les inconvénients des technologies de reconnaissance de la voix, de la parole et du visage, tout comme les domaines où il faudrait les autoriser doivent faire l'objet d'un débat sociétal approfondi.

> L'étude sur les technologies de reconnaissance de la voix, de la parole et du visage a été réalisée par un groupe de projet composé de membres du Fraunhofer-Institut für System- und Innovationsforschung ISI à Karlsruhe (Allemagne) et de l'Université de Fribourg (Suisse), sous la direction générale de Murat Karaboga. La méthodologie de l'étude s'appuie sur des recherches bibliographiques approfondies, sur l'analyse des comptes rendus des médias, sur plusieurs discussions avec des citoyennes et citoyens au sein de groupes de réflexion ainsi que sur une enquête représentative en ligne auprès de 1000 personnes.

# L'anonymat en voie de disparition

**KITT Trans Am, la super-voiture culte de la série policière américaine Knight Rider du début des années 1980 obéissait déjà aux commandes vocales de son conducteur. Et dans les premiers films de science-fiction, il suffisait au personnage principal de faire scanner son visage pour accéder au cockpit du vaisseau spatial. Désormais, la reconnaissance de la voix, de la parole et du visage fait partie de notre quotidien.**

La technologie exerce souvent une influence inattendue sur l'évolution de la société. Ainsi, en juin 2022, le tribunal administratif de Göttingen en Allemagne a autorisé un couple de parents à changer le prénom de leur fille de quatre ans, Alexa, qui faisait l'objet de tellement de moqueries qu'elle présentait des troubles psychiques avérés. Son prénom étant le même que le nom de l'assistant virtuel d'Amazon qui permet de contrôler divers appareils par commande vocale, la fillette était constamment victime d'ordres stupides et de plaisanteries exaspérantes.

Que les futurs parents privilégient la musicalité, la tradition familiale ou la mode, le choix du bon prénom pour leur bébé n'est jamais laissé au hasard – même s'il est possible d'en changer lorsque les circonstances de la vie l'exigent. Il en va autrement de la voix, de la forme du visage ou même de la façon de parler : ces particularités physiques ne changent plus guère chez l'adulte. Une fois que ces caractéristiques essentielles, indissociables d'un individu, ont été mesurées, numérisées et stockées sous forme de données biométriques, elles deviennent disponibles et peuvent faire l'objet d'un traitement électronique ultérieur.

Les origines de la reconnaissance de la voix, de la parole et du visage remontent à une bonne cinquantaine d'années. Même si ces systèmes diffèrent sur certains détails techniques, ils ont en commun de reposer sur des données biométriques qui en disent long sur l'individu en question.

## Dragons de bureau et autres services utiles

La recherche s'est intéressée au traitement automatique du langage dès le début des années 1960 déjà, mais sans grand succès, les ordinateurs étant alors trop peu performants. Cela a changé avec l'apparition des super-calculateurs modernes qui ont permis à IBM, Philips et Dragon Systems de développer des programmes de dictée commerciaux adoptés dès les années 1990 dans les bureaux. Fonctionnant indépendamment d'un système d'exploitation spécifique,

le programme de dictée Dragon s'est imposé sur le marché. Après un bref entraînement qui permet au logiciel de se familiariser avec la locutrice ou le locuteur, le taux de reconnaissance atteint 98%. Dragon est plus précis pour la dictée d'un langage technique avec une terminologie limitée que pour une langue aux tournures littéraires et au vocabulaire riche.

Les assistants virtuels Siri, Google Assistant et Alexa sont liés à des plateformes spécifiques et intégrés dans les systèmes d'exploitation d'Apple et d'Android ou dans les enceintes intelligentes d'Amazon. En raison de leur bas prix, ces dernières ont joué un rôle clé dans la popularisation de la reconnaissance de la voix et de la parole : en 2017, les enceintes intelligentes d'Amazon ont représenté près de 80% des ventes mondiales. Depuis lors, l'offre s'est bien entendu diversifiée.

Le smartphone réagit à des ordres comme « Alexa, règle le réveil sur sept heures » ou « Hey Siri, montre-moi les prévisions météo pour Zurich » et exécute l'action souhaitée. Dans un foyer intelligent, si le téléphone portable est connecté à d'autres appareils, ces derniers peuvent également être commandés par la voix : « Hey Siri, active l'imprimante » ou « Alexa, allume la machine à café » fonctionnent alors tout aussi bien.

## Un visage formé de points de données

Les débuts de la recherche sur la reconnaissance faciale remontent également aux années 1960, bien que les progrès tangibles aient été plus tardifs que pour la reconnaissance vocale. En 1964, une première tentative de cartographier les yeux, le nez, et l'implantation des cheveux donnait des résultats erronés à la moindre inclinaison de la tête du sujet ou dans des conditions peu favorables d'éclairage. Ces résultats ne se sont améliorés qu'à partir de 1970, lorsque les performances accrues des ordinateurs ont permis d'intégrer d'autres caractéristiques telles que la forme des lèvres et la couleur des cheveux.

Mais la véritable percée a eu lieu au début des années 1990 avec un algorithme qui, plutôt que de se baser sur les mesures des caractéristiques anatomiques de différents visages, effectue une analyse statistique des principales composantes d'un grand ensemble de données d'images faciales. Au départ, cette approche avait un défaut majeur : les taux de reconnaissance des différents groupes de recherche étaient difficilement comparables, chaque groupe travaillant avec ses propres bases de données d'images. La situation ne s'est débloquée qu'après que le ministère de la Défense américain a commencé à constituer une grande base de données pour pouvoir comparer les différents algorithmes dans le cadre d'un programme sur la reconnaissance faciale lancé en 1993. Depuis, les programmes de reconnaissance faciale disponibles sur le marché concourent chaque année pour obtenir les meilleurs scores au Face Recognition Vendor Test : en 2020, 99 fabricants ont présenté un total de 189 algorithmes à cette compétition.

Les attaques terroristes du 11 septembre 2001 marquent un tournant dans l'utilisation de la reconnaissance faciale. Sous l'effet de ce traumatisme, les États-Unis ont imposé l'obligation de présenter un passeport biométrique pour entrer sur le territoire américain, même pour des séjours de courte durée. L'Union européenne et la Suisse ont suivi le mouvement et, avec l'introduction des images faciales biométriques standardisées à partir de 2006, ont posé les jalons de l'utilisation par l'État de la reconnaissance faciale.

## Identification et vérification de l'identité par la voix et la parole

En règle générale, les systèmes de reconnaissance de la voix, de la parole et du visage sont utilisés pour confirmer automatiquement l'identité d'une personne et, tout au plus, lui donner accès à un service particulier. La reconnaissance de la parole consiste en principe à exécuter les commandes sur un téléphone portable ou un PC non plus à l'aide du clavier, mais à l'aide de commandes vocales.

Il convient de faire la distinction, d'un point de vue technique, entre identification et vérification. Vérifier l'identité d'un individu est un processus relativement simple dans la mesure où il suffit de comparer les caractéristiques d'un fichier son ou image avec un échantillon d'enregistrement. Si le nouveau fichier sonore ou l'image correspondent à l'échantillon, l'identité de la personne est confirmée. Processus plus complexe, l'identification consiste à déterminer qui est un individu en comparant un fichier son ou image avec un grand nombre d'enregistrements provenant d'une base de données.

Reconnaissance de la voix et reconnaissance de la parole n'ont pas non plus la même acception. Un système qui saisit des mots – c'est-à-dire qui « comprend » leur signification – n'identifie pas nécessairement la voix ou la personne qui parle. La reconnaissance de la parole consiste en effet à reconnaître

le contenu d'une déclaration afin, le cas échéant, d'exécuter ensuite la commande correspondante. La reconnaissance de la voix, quant à elle, vise à identifier un individu sur la base de caractéristiques biométriques telles que le ton et le timbre de sa voix – ce qui permet par exemple aux collaborateurs ou collaboratrices d'une banque ou d'une assurance maladie de savoir si la personne au bout du fil est bien celle qu'elle prétend être.

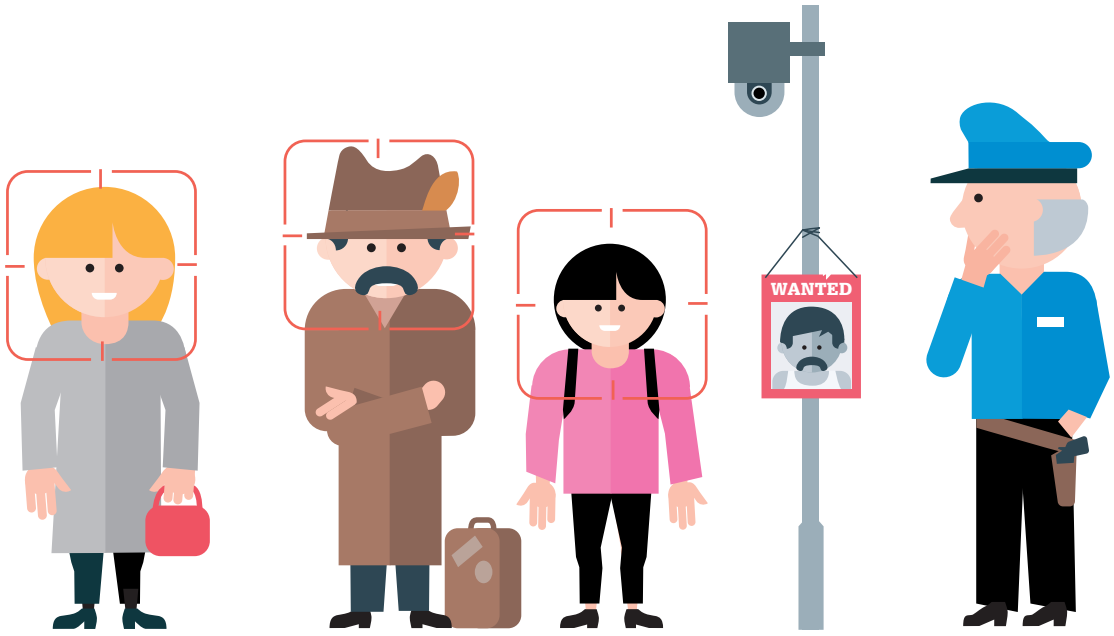
## Les données biométriques en disent bien plus long qu'une simple image

Les simples fichiers son ou image, assez faciles à manipuler, sont encore loin d'être des données biométriques. Ces dernières résultent au contraire d'un processus complexe de modélisation et de calcul. Par exemple, dans le cas de la méthode tridimensionnelle sur laquelle repose le système Face-ID d'un des principaux fabricants de téléphones portables, le smartphone projette 30 000 points de lumière infrarouge invisible sur le visage de sa ou son propriétaire en le scannant à l'aide d'un projecteur de points. Puis, sur cette base, le téléphone portable crée un modèle 3D qui identifie les structures et caractéristiques propres à ce visage. L'image infrarouge et l'image 3D sont converties en une formule mathématique et stockées sur la puce de l'appareil. À chaque déverrouillage, l'algorithme est exécuté quand le téléphone portable fait appel à cette formule.

De même, dans le cas de l'analyse de la voix, un algorithme saisit et traite plusieurs milliers de caractéristiques, dont la plupart ne sont même pas audibles pour l'oreille humaine et qui englobent bien plus que la tonalité, la mélodie de la voix, les pauses respiratoires et le rythme. C'est sur cette base que le logiciel crée une sorte d'empreinte acoustique, ou voiceprint en anglais.

Parce que les données biométriques sont étroitement liées à une personne et en disent long à son sujet, elles sont considérées comme sensibles et nécessitent une protection particulière. Dans un guide publié par les responsables suisses de la protection des données, les caractéristiques biométriques sont définies comme les « signes distinctifs physiques uniques et propres à une personne qui – du moins en théorie – peuvent toujours et partout être attribués sans équivoque à cette seule personne avec une certitude de presque 100% ».

En principe, l'utilisation par des particuliers de systèmes de reconnaissance de la voix, de la parole et du visage est contraire à la législation sur la protection sur la protection des données. En effet, il est difficile pour quiconque souhaite traiter ces données de faire valoir un intérêt prépondérant, car les données biométriques sont des données sensibles qui nécessitent une protection particulière. De plus, il est souvent quasiment impossible d'obtenir le consentement des personnes concernées puisque ces appareils doivent faire appel aux données avant de pouvoir fonctionner. Il arrive même parfois que, comme dans le cas de l'identification masquée, aucun consentement ne soit demandé, ce qui n'est pas autorisé par la législation sur la protection des données.



## La protection des données sous un angle plus large

L'identification automatisée des individus sur la base de leurs caractéristiques biométriques a très vite été la cible de critiques qui lui reprochent des résultats trop souvent erronés ou biaisés. Ainsi par exemple, la reconnaissance faciale identifie systématiquement moins bien les femmes et les personnes à la peau foncée que les hommes blancs – un biais qui s'explique notamment par le fait que les bases de données utilisées pour l'entraînement des algorithmes contenaient une vaste majorité d'images d'hommes blancs. L'analyse des données sonores n'est pas non plus à l'abri de distorsions, dans la mesure où, par exemple, une mauvaise qualité du son favorise à elle seule les erreurs d'interprétation.

Dans les années 1990, alors que l'utilisation des systèmes d'identification biométrique n'en était qu'à ses débuts, la protection des données était au centre des préoccupations. On craignait que le suivi automatisé des moindres faits et gestes des gens ne porte atteinte à leur vie privée et à leur liberté personnelle. Au fil du temps cependant, cette vision centrée sur l'individu est apparue trop étroite. Aujourd'hui, les spécialistes plaident plutôt pour une réflexion globale sur les conséquences des technologies susceptibles de permettre une surveillance automatisée. Car si leur utilisation représente clairement une menace pour la sphère privée, elle va aussi à l'encontre d'autres droits fondamentaux essentiels au bon fonctionnement d'une démocratie, comme la liberté de réunion ou la liberté d'expression.

## La pression de conformité menace les droits fondamentaux et la démocratie

Toute personne qui redoute d'être observée en permanence a tendance à s'adapter et à s'autocensurer. La pression exercée sur les individus pour qu'ils s'alignent sur le comportement du groupe va à l'encontre d'une législation qui, comme la Constitution fédérale suisse, rejette toute discrimination de personnes en raison de leur identité, de leur mode de vie et de leurs convictions, et qui accorde une grande importance à la liberté d'expression.

Du point de vue juridique, la protection des données reste un point de friction. En effet, les données biométriques révèlent des caractéristiques uniques propres à un individu et sont à cet égard des données personnelles particulièrement sensibles qui nécessitent une protection particulière. Le fait d'en faire usage risque de porter gravement atteinte aux droits fondamentaux individuels inscrits dans la Constitution, notamment le droit à la vie privée et la protection contre l'emploi abusif des données personnelles. Le traitement des données biométriques est encore compliqué par le fait que, en pratique, l'anonymisation est incompatible avec la plupart des systèmes de reconnaissance de la voix, de la parole et du visage. Enfin, la pression de conformité citée plus haut menace aussi les libertés d'expression, de réunion et d'association garanties par la Constitution.

Le traitement des données biométriques est notamment régi par le principe de proportionnalité qui prévoit que seules les données objectivement nécessaires à l'exécution d'une tâche spécifique doivent être collectées. La finalité doit également être vérifiée, c'est-à-dire que les données ne doivent être utilisées que pour le projet pour lequel elles ont été collectées. De plus, il faut veiller au caractère approprié d'une technologie : son utilisation n'est licite que si elle est adéquate pour atteindre l'objectif qu'elle poursuit. En outre, la transparence est de mise car les gens doivent savoir que leurs données biométriques seront collectées – et ils doivent avoir la possibilité d'accepter ou de refuser la collecte sans que cela leur porte préjudice. Enfin, le stockage sécurisé des données doit également être garanti afin d'éviter tout piratage ou accès illicite. Pour toutes ces raisons, l'utilisation de ces technologies doit être encadrée par une législation spécifique qui précise non seulement sa finalité, mais aussi ses limites et en particulier la restriction au strict nécessaire du traitement des données.

**Priorités différentes dans le monde de la recherche**

En Chine, les technologies de reconnaissance faciale font l'objet de recherches intensives depuis un certain temps déjà. En 2020, au moins deux fois plus d'articles scientifiques y étaient publiés sur ce thème qu'aux États-Unis ou en Inde. Dans l'ensemble de l'Europe – Suisse comprise –, le volume des publications équivaut à environ un tiers de la production chinoise. La situation est différente en ce qui concerne la reconnaissance vocale, où les États-Unis arrivent nettement en tête devant la Chine dont les efforts dans ce domaine ont pourtant été intensifiés depuis 2018. Il est frappant de constater l'importance que la Chine accorde à la recherche sur la reconnaissance de la voix, de la parole et du visage par rapport à d'autres domaines.

# Quand la technologie tend l'oreille

**Les enceintes intelligentes, ou haut-parleurs intelligents, ont fait leur apparition dans de nombreux bureaux et foyers. Les e-mails et lettres ne sont plus dactylographiés mais dictés à haute voix à un PC, les entrées de calendrier sont récitées à nos smartphones et les réveils sont réglés par commande vocale, tout comme la température de la pièce ou d'autres fonctions de l'habitat intelligent.**

Prononcer à haute voix une commande au lieu de la saisir au clavier ou sur un autre appareil présente des avantages évidents : les enceintes intelligentes remplacent une multitude de télécommandes. De plus, comme le maniement des différents appareils est plus aisé, cette technologie facilite aussi le quotidien de personnes physiquement handicapées.

## Enceintes intelligentes : beaucoup de monde à l'écoute

Sur le plan technique, les enceintes intelligentes se composent au minimum d'un microphone et d'un haut-parleur ainsi que d'une connexion à un fournisseur via un réseau WLAN et Internet pour les prestations et fonctionnalités basées sur le cloud et mises à la disposition de l'utilisatrice ou l'utilisateur.

Les fournisseurs se servent ensuite des enregistrements pour améliorer leurs services. Les données collectées sont stockées à différents endroits : sur le serveur du fournisseur, dans l'enceinte intelligente elle-même et, le cas échéant, sur le téléphone portable qui y est connecté.

Pour que l'enceinte intelligente puisse « entendre » le code d'activation vocal d'une commande, elle doit être en permanence en mode écoute. Dès que le code d'activation vocal est prononcé, l'enregistrement est déclenché puis transmis au serveur du fabricant. Le système de reconnaissance de la parole analyse la commande et renvoie le résultat à l'utilisatrice ou l'utilisateur via une sortie vocale synthétique automatisée.

Il existe une fonction que la plupart des gens ignorent, ou n'utilisent que rarement : sur de nombreux appareils, il est possible de supprimer les enregistrements audio via une app ou un site Internet. Du point de vue de la protection des données, cela doit être salué. En effet, certaines demandes ou commandes vocales sont délicates, par exemple la prise de rendez-vous chez son médecin ou chez une avocate spécialisée dans les divorces. Par ailleurs, outre les commandes vocales elles-mêmes,

les enceintes intelligentes enregistrent aussi d'autres données telles que le jour et l'heure, ce qui permet le cas échéant de tirer des conclusions sur les habitudes de l'individu concerné. Si d'autres personnes sont présentes, leur voix est également enregistrée, ce qui peut à son tour donner des indications sur la composition du ménage ou du cercle d'amis. En outre, si les demandes et commandes vocales sont révélatrices, le volume et le ton de la voix le sont tout autant : elles peuvent livrer des informations sur la condition physique d'une personne, notamment si elle est enrhumée ou si elle a trop bu et, souvent, tirer des conclusions sur son état d'esprit sur la base de ses rires joyeux ou de ses soupirs pesants.

## Authentification vocale au guichet automatisé

En Suisse, certaines banques et la Poste utilisent des enregistrements vocaux pour confirmer l'identité de leur clientèle, dans le but d'accroitre la sécurité et de gagner du temps. À la Banque Migros, la durée des appels a pu être réduite de 20% grâce à l'authentification vocale. Actuellement, ce procédé est surtout utilisé pour les opérations bancaires par téléphone mais il n'est pas exclu qu'il serve un jour lors d'entretiens conseil ou pour des systèmes d'assistance virtuels.

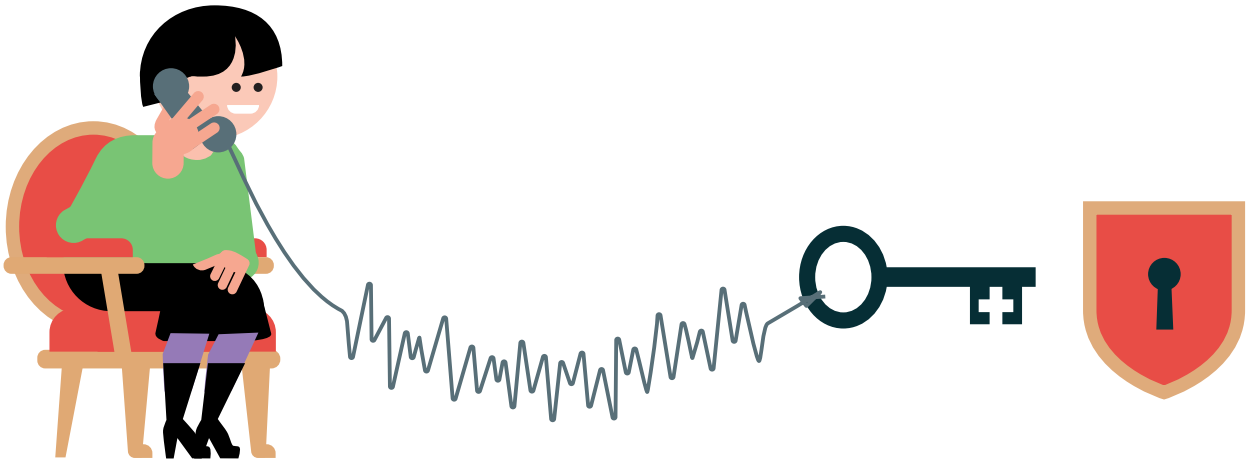
Avant que l'enregistrement de son appel ne commence, le client ou la cliente doit donner son accord et autoriser que celui-ci soit stocké et utilisé pour une authentification ultérieure. Chez PostFinance, il est aussi possible d'activer ou de refuser le recours à la reconnaissance vocale sur le portail web.

La question de savoir si l'identité d'une personne peut être certifiée sans équivoque par sa voix est controversée. Les adeptes de ce procédé estiment qu'une voix est extrêmement individuelle et considèrent donc la reconnaissance vocale biométrique comme fiable. Mais d'autres font remarquer que des journalistes et des pirates informatiques ont déjà réussi à déjouer les barrières d'accès acoustiques au moyen d'extraits sonores de vidéos YouTube et de programmes informatiques également utilisés pour créer des enregistrements sonores artificiels trompeurs – les hypertrucages, ou deep fakes.

Il est certain en revanche que l'authentification biométrique n'est sûre que si les données sont protégées. Une fois compromises, celles-ci ne peuvent plus servir à identifier une personne avec certitude : les empreintes vocales et faciales sont en effet étroitement et durablement liées à un individu et, contrairement à un mot de passe, elles ne peuvent pas être modifiées aisément.

## Échec à la criminalité

Certains individus ou organismes criminels se donnent beaucoup de mal pour accéder à des données sensibles par le biais de systèmes de reconnaissance vocale. C'est pourquoi les spécialistes recommandent aux banques et autres entreprises



qui traitent des données sensibles d'utiliser un mot de passe en plus de la reconnaissance vocale. Ce processus en deux étapes renforce la sécurité.

Avec leur multiplication, les enceintes intelligentes ont tendance à devenir des cibles pour les pirates informatiques qui peuvent tenter de s'emparer des données sensibles d'un individu en trompant le mécanisme d'authentification de l'appareil à l'aide d'enregistrements vocaux.

La recherche se penche actuellement sur les problèmes de sécurité des enceintes intelligentes. Une des approches consiste à supprimer de l'enregistrement les caractéristiques vocales qui ne sont pas nécessaires à l'interaction avec l'enceinte intelligente. Pour empêcher un accès non autorisé aux données, d'autres équipes scientifiques travaillent sur des programmes d'analyse de réseau qui détectent la transmission de fichiers audio sur Internet. Le but étant de pouvoir avertir les utilisatrices et utilisateurs lorsque leur code d'activation vocal a été déclenché.

**Retenue helvétique**
En 2020, Amazon occupait la première place sur le marché avec 22% de toutes les enceintes intelligentes vendues dans le monde et Google suivait de près avec 17%. En 2018, 18% de la population américaine utilisait des enceintes intelligentes, contre 10% en Allemagne. En comparaison, la Suisse fait preuve de retenue : en 2018, à peine 1% de la population suisse utilisait des enceintes intelligentes contre 3% un an plus tard. L'enquête représentative menée en octobre 2021 dans le cadre de l'étude de TA-SWISS a révélé que 63% des ménages n'en possédaient pas. Près de la moitié des utilisatrices et utilisateurs d'un tel appareil en ont fait l'acquisition au cours de l'année écoulée. Seuls 9% en possèdent un depuis plus de trois ans. Parmi les personnes interrogées qui n'utilisent pas d'enceintes intelligentes, 41% déclarent vouloir continuer de s'en passer. Elles invoquent pour raisons principales qu'elles ne voient pas d'utilité à ces appareils et qu'elles craignent pour la protection des données.

# Notre visage en guise de laisser passer

**Avec la pandémie de coronavirus, les systèmes de paiement et d'accès sans contact ont connu une forte progression. Déverrouiller son téléphone portable grâce à la technologie Face-ID d'Apple est aussi devenu très courant. Si les systèmes de reconnaissance faciale sur un smartphone personnel peuvent être pratiques, ils sont moins appréciés lors de matchs de football ou de concerts de rock. Ils peuvent même avoir des conséquences inquiétantes pour les individus.**

Les autorités, notamment les douanes et la police, ont depuis longtemps déjà recours à la reconnaissance faciale. Toutefois, en raison de la discrétion qui entoure le recours à cette technologie par la police, il reste difficile d'obtenir des informations sur la manière dont elle est utilisée et sur ses résultats.

Les systèmes utilisés par les services de police américains dans les années 1990 ont surtout marqué les esprits par leurs échecs. Lors d'un test pluriannuel à grande échelle à l'aéroport de Boston en 2003, ils n'ont donné aucune correspondance. En Floride non plus, l'utilisation d'un logiciel de ce type n'a pas été concluante, le système allant même, certains jours, jusqu'à ne fournir que des faux positifs. En 2006, un test réalisé à la gare de Mayence par l'Office fédéral allemand de la police criminelle a donné des résultats un peu meilleurs avec un taux de réussite de 30% – toutefois nettement inférieur aux 80% visés.

## Une utilisation modérée en Suisse

En Suisse, le contrôle facial est principalement utilisé à l'aéroport de Zurich, et ce sur une base volontaire : depuis 2020, et après une phase pilote de six mois à partir de l'automne 2017, les personnes qui voyagent par avion peuvent opter pour le contrôle automatisé du visage. Des panneaux d'information les guident vers le système de reconnaissance faciale, mais elles peuvent encore faire vérifier leurs documents personnels à un guichet conventionnel.

En plus de l'aéroport de Zurich, les polices d'Argovie et de Saint-Gall sont les seules connues à ce jour à avoir recours à la reconnaissance faciale en Suisse. D'autres services de police testent des systèmes de ce type ou font exécuter les tâches de reconnaissance faciale par des collaboratrices

et collaborateurs. La police cantonale de Bâle-Ville possède sept véhicules Tesla, chacun équipé de huit caméras. Bien qu'elles ne soient pas utilisées pour la reconnaissance faciale, il serait relativement aisé d'y intégrer le logiciel nécessaire.

Le groupe de projet de TA-SWISS a pu s'assurer auprès des deux polices cantonales qui recourent à la reconnaissance faciale que celles-ci s'efforcent de garantir une utilisation sûre et justifiable de cette technologie, tant sur le plan juridique que sur le plan éthique. Ainsi, les deux instances ont réalisé une analyse d'impact relative à la protection des données (AIPD) et, pour conserver les données en toute sécurité, les stockent sur des serveurs qui ne sont pas connectés à Internet ni à d'autres réseaux.

Mais l'étude de TA-SWISS relève aussi les points faibles de l'utilisation de la reconnaissance faciale par la police. Ni la police cantonale d'Argovie ni celle de Saint-Gall ne publient les résultats de leur AIPD et toutes deux ont renoncé à faire vérifier la performance et la sécurité de leurs systèmes par un organisme indépendant. La question de savoir quels offices ont accès à quelles données reste également obscure. De plus, sous prétexte qu'il n'y a pas de différence fondamentale entre la recherche d'images par un logiciel ou par l'œil humain, les deux polices cantonales arguent qu'il n'est pas nécessaire d'informer les gens lorsque leurs données sont traitées avec cette technologie.

## La reconnaissance faciale en temps réel, porte ouverte à la surveillance de masse ?

En Suisse, la surveillance de masse en temps réel n'existe pas. Cette pratique serait illicite et contraire à certains droits fondamentaux comme la liberté de réunion. De plus, elle constituerait une violation du principe de proportionnalité si, par exemple, tous les individus participant à une manifestation étaient surveillés dans le seul but de sanctionner les actes de vandalisme d'une partie d'entre eux. Pourtant, même en Suisse, la surveillance de masse suscite les plus intenses débats. La pandémie de coronavirus n'a pas non plus contribué à apaiser les craintes car plusieurs pays ont eu recours à des systèmes de reconnaissance faciale pour contrôler le respect des restrictions sanitaires.

Les groupes de discussion invités à débattre dans le cadre de l'étude de TA-SWISS ont appelé à s'opposer à la surveillance de masse étatique. Ils redoutent que l'État gagne trop de pouvoir sur les citoyennes et citoyens, ce qui entraînerait une perte de confiance à son égard. Ceci cache une crainte plus profonde : même si, dans un premier temps, la technologie de reconnaissance faciale est utilisée uniquement pour détecter des crimes graves contre la vie et l'intégrité physique, il n'est pas exclu que, une fois qu'elle aura fait ses preuves, son champ d'application soit étendu à des délits mineurs comme le vol à la tire, habituant peu à peu la population à une surveillance généralisée et permanente. Finalement, la Suisse risquerait de se retrouver dans une situation comparable à celle de la Chine, dont le système de crédit social à l'encontre des citoyennes et citoyens est souvent cité comme exemple dissuasif : les personnes qui traversent au rouge ou enfreignent d'autres prescriptions étatiques sont identifiées et s'exposent à des sanctions, comme le fait d'être désavantagées lors de recherche d'un emploi ou d'un logement.

## Une double approche, optique et acoustique, pour lutter contre le racisme

Tandis que, dans les pays européens, l'État fait preuve de retenue en matière de reconnaissance faciale, certains acteurs privés se montrent moins timorés. Par exemple, les clubs italiens de football envisagent d'utiliser la reconnaissance faciale et vocale pour lutter contre la xénophobie après que plusieurs matchs ont été interrompus suite aux insultes racistes répétées à l'encontre de joueurs à la peau foncée (cris de singe, slogans nazis et autres injures). La reconnaissance vocale pourrait permettre de détecter les propos racistes et la reconnaissance faciale d'identifier les coupables. Aucun système de ce type n'est encore utilisé dans les stades italiens à ce jour. À l'automne 2019 toutefois, Luigi De Siervo, directeur général de la plus haute ligue de football italienne, a déclaré être prêt à recourir à la reconnaissance faciale pour « attraper un par un les gens qui gâchent ce merveilleux sport ».

En Suisse, il n'existe aucune base légale permettant d'utiliser la reconnaissance vocale et faciale, ni dans les stades de sport, ni dans le cadre d'autres événements du secteur privé. Toutefois, une petite majorité des personnes interrogées dans le cadre de l'étude s'est prononcée en faveur de reconnaissance faciale dans les stades : dans ce contexte particulier, cette technologie a rencontré la plus grande acceptation, par ailleurs plutôt faible.

**Des directives claires sont nécessaires pour le recours à la reconnaissance faciale par la police**

L'enquête représentative menée dans le cadre de l'étude de TA-SWISS n'a pas révélé un rejet en bloc du recours à la reconnaissance faciale par la police : si un tiers des personnes interrogées s'est prononcé en faveur de son utilisation, un autre tiers a déclaré ne pas maîtriser suffisamment le sujet et un dernier petit tiers a indiqué refuser de l'utiliser. Pour ses partisans, cette technologie se prête en premier lieu à la recherche de personnes disparues, puis à la lutte contre le terrorisme. Celles et ceux qui s'opposent à la reconnaissance faciale le font avant tout par crainte d'une utilisation abusive des données et à cause de l'influence croissante de cette technologie sur la vie publique qui pourrait conduire à une surveillance de masse sans motif. Pour l'ensemble des personnes interrogées, seul un personnel habilité devrait être autorisé à utiliser la reconnaissance faciale, et chaque utilisation devrait être consignée et communiquée en toute transparence. De plus, une base légale spécifique est nécessaire, de même qu'un contrôle et une évaluation régulière de l'utilisation de cette technologie par des experts et expertes indépendants.

# Quand la technologie expose l'insondable

**Par son expérience et sa sensibilité, en observant les modifications du teint, de l'expression du visage et de la voix, une ou un médecin en apprend beaucoup sur l'état de santé de ses patientes et patients. Il paraît pourtant probable qu'un jour, les systèmes de reconnaissance de la voix, de la parole et du visage seront en mesure de détecter des maladies avant même les spécialistes de santé. Mais lorsqu'il s'agit de détecter les émotions, la technologie se heurte à des obstacles bien plus importants.**

« Pour les transgenres, l'app est vraiment utile pour entraîner la voix ou constater des changements », souligne un commentaire sur l'application Voice Pitch Analyzer. « Salut, je suis moi-même FtM (female to male) et j'utilise cette app parce que c'est vraiment intéressant de voir soi-même comment le ton de la voix change avec l'hormonothérapie », écrit une autre personne. Voice Pitch Analyzer est une app créée par Purr Programming qui peut être téléchargée gratuitement et utilisée pour analyser une voix, cette caractéristique révélant en général si la personne qui parle est une femme ou un homme. Les apps d'analyse et d'entraînement de la voix ont pour but d'aider les personnes transgenres à modifier leur voix parlée de manière à ce qu'elle se rapproche de l'objectif spécifique à leur sexe, c'est-à-dire qu'elle soit plus féminine ou plus masculine selon leur souhait.

Pour les performances sportives, le cycle de sommeil, le rythme cardiaque ou l'ovulation, il existe désormais sur Internet d'innombrables apps qui permettent de surveiller son propre corps. Il est possible qu'à l'avenir ces programmes recourent aussi davantage à la reconnaissance de la voix, de la parole et du visage pour évaluer l'état physique d'une personne. Cela fournirait des outils d'analyse supplémentaires au fameux « Dr. Google » que les gens consultent souvent en premier en cas de symptômes peu clairs.

À première vue inoffensives, les apps d'analyse vocale sur PC ou téléphone portable citées plus haut risquent de poser problème lorsque des versions plus sophistiquées apparaîtront. En effet, les progrès des outils de diagnostic vont de pair avec le risque de voir un public non averti dépassé s'il recourt de lui-même à un moyen de surveillance médicale via la caméra et le microphone d'un ordinateur. À l'heure actuelle toutefois, les outils de diagnostic performants basés sur la reconnaissance vocale ou faciale sont réservés aux spécialistes de la santé.

## En mission pour la médecine

Lorsque nous parlons, notre cerveau régule l'interaction de près d'une centaine de muscles. Une multitude de maladies peuvent être révélées par la voix, y compris par le biais de sons que l'oreille humaine n'entend même pas. Une équipe de recherche a ainsi étudié le son « aaah » émis par des personnes en bonne santé et des personnes atteintes de la maladie de Parkinson et ont trouvé dix caractéristiques acoustiques qui permettent de détecter cette pathologie avec une précision de près de 99%. Le choix des mots ou une difficulté à les trouver peuvent également indiquer la présence de la maladie d'Alzheimer. Les systèmes de reconnaissance vocale repèrent cette pathologie dans des environnements de test avec un taux de détection de 80 à 90%.

Depuis la pandémie de coronavirus, le son caractéristique de la toux que cette maladie provoque fait aussi l'objet de recherches afin de permettre un diagnostic aussi précoce que possible. Les affections cardiaques pourraient également être détectées à l'aide de la voix, certains schémas de fréquence vocale étant associés à de graves pathologies des artères coronaires.

Enfin, la voix reflète l'état psychique d'une personne lorsqu'elle parle. Son tempo, son rythme, son ton et son volume révèlent une excitation, une anxiété, un abattement ou une manie. Dans le traitement de la dépression, la reconnaissance vocale est devenue la norme : un ton monotone et légèrement élevé par rapport à la voix habituelle peut indiquer une tendance suicidaire. Des caractéristiques vocales spécifiques ont également été mises en évidence pour le diagnostic de l'autisme et des troubles de déficit de l'attention ou d'hyperactivité.

La reconnaissance faciale est moins répandue en médecine que la reconnaissance vocale et la reconnaissance de la parole, mais elle progresse. Dans le cadre du projet européen SEMEOTICONS, une sorte de miroir équipé de différents capteurs et caméras a été développé. Ce dispositif est capable de reconnaître, outre les facteurs psychiques, la forme physique et l'état nutritionnel en analysant la couleur de la peau, les muqueuses, la répartition du tissu adipeux sous la peau et le modèle de transpiration. Le système développé à l'université de Bonn, appelé DeepGestalt, doit permettre de détecter des maladies rares ou des défauts génétiques à partir de la photo d'un visage. À cet effet, le logiciel compare l'expression du visage sur l'image avec de nombreuses photos de personnes atteintes de certaines maladies diagnostiquées. La base de données comprend aujourd'hui 17 000 photos de plus de 200 pathologies complexes. Plusieurs équipes de recherche travaillent sur d'autres systèmes permettant de détecter des maladies rares à l'aide d'images faciales.

L'un des obstacles à l'utilisation de la reconnaissance de la voix, de la parole et du visage dans le domaine médical réside dans la constitution d'une base de données fiable. Actuellement, il n'existe que peu de bases de données pour l'analyse de la voix. De plus, certains spécialistes doutent de la fiabilité et de la pertinence des caractéristiques vocales : la langue parlée est trop souvent altérée, notamment par un enrouement ou un rhume allergique. L'origine et la culture d'une personne influencent également le volume et le rythme de son discours. Il est clair que des recherches supplémentaires sur la reconnaissance de la voix, de la parole et du visage en médecine sont donc nécessaires.

L'utilisation des technologies de reconnaissance dans le respect du secret médical impose des exigences particulières en matière de sécurité de manipulation et de stockage des données dans le domaine des soins. Les apps d'autodiagnostic doivent être considérées comme des produits médicaux. Plutôt rares à l'heure actuelle, elles ne seront pas réglementées avant un certain temps.

## Sonder le monde des émotions

Les systèmes de reconnaissance de la voix, de la parole et du visage recèlent un potentiel dans le cadre de la recherche d'emploi, notamment en accélérant la procédure grâce à un logiciel d'aide à la présélection des candidatures particulièrement adaptées. Cette technologie pourrait aussi servir à tester l'impact que produit une personne et vérifier ainsi si elle correspond au poste à pourvoir.

Le secteur des assurances privées est également susceptible de tirer parti de cette technologie pour proposer des offres personnalisées à la clientèle. À l'instar de la médecine, les assurances pourraient exploiter les informations sur la forme physique et les caractéristiques corporelles obtenues à partir des données faciales et vocales. De manière générale, la détection des émotions devrait susciter un intérêt considérable dans différents cercles, à l'image des banques attirées par ses avantages potentiels en termes de marketing, de la police tentée de recourir à un détecteur de mensonges – comme cela a déjà été le cas dans le cadre du projet iBorderCtrl financé par l'UE – et des exploitants de stades espérant y trouver une aide pour détecter les supporters violents.

Comme pour les tests psychologiques ou l'analyse graphologique, l'analyse de l'image faciale ou de la voix requiert l'accord les personnes concernées. Ce consentement devrait aussi être donné de plein gré – une condition difficile à remplir compte tenu de l'asymétrie de pouvoir entre un employeur et un employé, ou entre un assureur et une assurée. L'utilisation de cette technologie en secret lors d'un entretien d'embauche ou de la conclusion d'une police d'assurance constituerait une atteinte à la personnalité et, notamment, une violation du principe de la bonne foi.

De manière générale, la détection des émotions suscite le scepticisme des spécialistes qui doutent de la fiabilité d'un logiciel en la matière, les émotions se traduisant de façon très différente selon les individus. Le simple fait que la culture d'une personne influence considérablement la manière dont elle manifeste ses émotions atteste de cette difficulté.

**Réserves à l'égard de l'analyse des émotions**
En Suisse, le recours à l'analyse de la voix, de la parole et du visage à des fins médicales, de détection des émotions ou d'analyse de l'attention à l'école suscite de grandes réserves. Si près d'un quart des personnes interrogées se déclarent sereines face à cette technologie lorsqu'elle sert à diagnostiquer des maladies physiques, 37% se disent très inquiètes, surtout si cela implique que les caisses maladie ont accès aux résultats. En ce qui concerne l'analyse des émotions, 65% des personnes interrogées indiquent craindre que le logiciel ne fonctionne pas correctement. Enfin, 62% des personnes interrogées considèrent que les émotions sont trop complexes pour qu'un logiciel puisse les détecter correctement.

## Bannir toute distraction à l'école

Dans de nombreuses sociétés, les performances des enfants en âge scolaire sont suivies de près. En effet, la réussite économique d'une nation dépend de la bonne formation de la relève. Différents pays, dont surtout les pays anglophones et la Chine, ont recours aux technologies d'analyse de l'attention dans les écoles. Ces dernières années, quelque 2,7 milliards de dollars par an ont été investis par les

écoles américaines dans des dispositifs de surveillance et de sécurité afin de se prémunir contre les fusillades et autres actes de violence. Bien que rien n'indique à ce jour que le niveau d'attention des enfants soit surveillé aux États-Unis, ces systèmes vidéo, une fois installés, peuvent relativement facilement être mis à niveau pour l'analyse de l'attention, et ce pour le coût modeste d'à peine 200 000 francs suisses par école.

Dans certaines écoles en Chine, on observe non seulement les élèves, mais aussi le corps enseignant. Aucun consentement n'est demandé, ni aux enfants ni à leurs parents. Les systèmes enregistrent notamment la durée pendant laquelle les élèves ont le regard dirigé sur le tableau et calculent un score d'attention que le personnel enseignant et les parents peuvent consulter. Le recours à cette technologie vise à améliorer la réussite scolaire : non seulement toute source extérieure de distraction est supprimée, mais l'enseignement se veut ainsi mieux adapté aux forces et aux faiblesses individuelles des élèves. En psychologie, on sait toutefois qu'un contrôle excessif peut aussi générer du stress, qui à son tour entrave l'apprentissage. De plus, un certain nombre de spécialistes doutent que l'analyse des expressions faciales puisse fournir des informations fiables sur le niveau d'attention d'un individu.

Les émotions des élèves ou des étudiants et étudiantes à l'université peuvent en principe être analysées sur n'importe quel système ou plateforme où des données d'images sont générées. La pandémie de coronavirus, en popularisant les plateformes d'apprentissage et de communication telles que Moodle, Google Classroom ou encore Zoom Meetings, a en même temps posé des jalons qui permettront d'introduire un jour la détection des émotions ou l'analyse de l'attention – d'autant plus que certaines écoles utilisent déjà de tels logiciels pour surveiller le comportement des candidates et candidats lors de tests.

## Tout le monde n'a pas forcément envie de connaître tout le monde

Les technologies de reconnaissance de la voix, de la parole et du visage mettent à portée de main ce que l'on appelle l'identification de toute personne : en s'équipant du matériel nécessaire, tout individu serait en mesure d'identifier quiconque croise son chemin. En 2014, l'entreprise Alphabet a fait un premier pas dans cette direction en lançant ses lunettes connectées Google Glass. L'idée derrière un mini-ordinateur sur le nez consiste à afficher sur les verres de lunettes des informations de toute nature sur les alentours, comme les bâtiments ou d'autres éléments du paysage ; les lunettes intelligentes se procurent elles-mêmes les données nécessaires sur Internet. Si le groupe a expressément exclu la reconnaissance faciale, les spécialistes de la protection des données ont fait remarquer que les lunettes connectées pouvaient relativement facilement être reliées aux programmes de reconnaissance faciale existants – et, à plus forte raison, que certaines entreprises misaient sur le fait qu'Alphabet renoncerait à l'interdiction de la reconnaissance faciale qu'elle s'était elle-même imposée.

En tout état de cause, il existe déjà des bases de données pour l'identification de personnes privées : à elle seule, l'application NameTag a rassemblé plusieurs millions d'images de visages et a annoncé qu'elle se procurerait dorénavant d'autres photos de portraits sur les réseaux sociaux. En 2015, Alphabet a cessé de vendre ses lunettes aux particuliers, notamment en raison des critiques persistantes et du désintérêt du public. Mais de nouvelles lunettes connectées performantes sont apparues depuis longtemps, comme les Ray-Ban Stories, développées en coopération avec Facebook. Ajouter une fonction de reconnaissance faciale à ces lunettes connectées n'est pas irréalisable pour des pirates informatiques. Avec l'aide d'une base de données de visages, il leur serait alors possible d'identifier n'importe quel individu en passant devant lui. L'avènement des technologies d'identification de toute personne mettrait un terme à l'anonymat dans les lieux publics, déjà mis à mal par l'omniprésence des smartphones.

Comme l'a montré le scandale de l'entreprise Clearview AI, les craintes pour la protection des données de ces systèmes – dits d'identification de toute personne – ne sont pas dénuées de fondement : cette société américaine fondée en 2017 avait récupéré sur Internet, en particulier sur les réseaux sociaux, des milliards de portraits photo pour en faire une vaste base de données et l'insérer dans son moteur de recherche de visages. Ce service a ensuite été proposé par l'entreprise à des organismes publics et privés du monde entier, sans que le public des pays concernés, ni même les propriétaires des images utilisées ou les personnes figurant sur ces images n'en soient jamais informées. Plusieurs plaintes ont été déposées contre l'entreprise, avec pour effet de restreindre l'accès au logiciel aux seules autorités de sécurité depuis 2020. La guerre en Ukraine a remis

cette société à la une des journaux lorsqu'il a été révélé que les autorités ukrainiennes utilisaient ce logiciel pour identifier les soldats russes tombés ou capturés et informer leurs familles. Les talibans ont également recours à la reconnaissance faciale en Afghanistan. En effet, lors de leur prise de pouvoir, des groupes de fondamentalistes se sont non seulement emparés de véhicules et d'armes appartenant aux forces armées occidentales, mais aussi de fichiers contenant des millions d'empreintes digitales et d'images faciales. Désormais, ils les utilisent dans leur chasse aux personnes ayant « collaboré » avec des organisations occidentales et aux autres « traîtres ».

**Rejet de l'analyse de l'attention et de l'identification de toute personne**
Selon l'enquête représentative, le recours à la reconnaissance de la voix, de la parole et du visage à l'école est rejeté en bloc par 56% des personnes interrogées. Parmi ces personnes, la résistance à l'identification de toute personne s'avère encore plus marquée (51% expriment de grandes réserves), à l'unisson de la tendance générale des groupes de discussion : un monde où toute personne serait identifiée en tout temps serait un monde « dans lequel elle ne voudrait pas vivre », déclare une participante. Une des craintes des participantes et participants est que l'identification de toute personne ne conduise à une augmentation et à une intensification du harcèlement (stalking). Celles et ceux qui n'ont pas exprimé d'inquiétudes à l'égard des technologies d'identification de toute personne dans l'enquête représentative justifient leur position notamment par le fait que l'anonymat a déjà disparu.

# Tenir Big Brother à l'écart : recommandations

**Les données biométriques sont hautement sensibles et nécessitent une protection particulière. Il est d'autant plus urgent de créer un cadre légal pour les technologies de reconnaissance de la voix, de la parole et du visage dans les situations où leur utilisation est particulièrement délicate – notamment le domaine de la médecine, de l'application de la loi, de l'octroi de crédits ou des assurances, et dans le monde du travail.**

Les enregistrements vocaux et les images faciales numérisées sont le reflet de caractéristiques physiques qui ne changent pratiquement plus pour une personne adulte. Une fois ces données biométriques compromises, il est impossible de rétablir leur intégrité. C'est pourquoi il est généralement recommandé de poser des limites à la collecte des données et de recourir à des procédures d'authentification multifactorielles, par exemple en complétant l'image faciale par un mot de passe.

## Interdire les applications à haut risque

La surveillance automatisée en temps réel par la reconnaissance de la voix, de la parole et du visage se heurte à un large rejet dans les démocraties occidentales. Elle se révèle également incompatible avec plusieurs droits fondamentaux inscrits dans la Constitution fédérale suisse et doit par conséquent être interdite, à l'instar du système de crédit social qui jauge le comportement de chaque citoyen et citoyenne sur la base d'une surveillance généralisée.

De même, il faut proscrire toute utilisation dans l'espace public de lunettes connectées et d'autres dispositifs techniques discrets qui permettent de surveiller des tierces personnes à leur insu grâce à la reconnaissance faciale.

Les systèmes d'analyse de l'attention à l'école doivent eux aussi être interdits. De plus, il faut prohiber le recours à des systèmes de prise de décision entièrement automatisés et basés sur la

reconnaissance de la voix, de la parole et du visage dans toute situation particulièrement délicate, notamment dans les hôpitaux, les banques et les assurances, comme dans le domaine de l'application de la loi ou au travail. A contrario, tout résultat provenant de systèmes décisionnels semi-automatisés doit systématiquement être examiné d'un œil critique et approuvé par des spécialistes qualifiés.

## Reporter la détection des émotions et des maladies

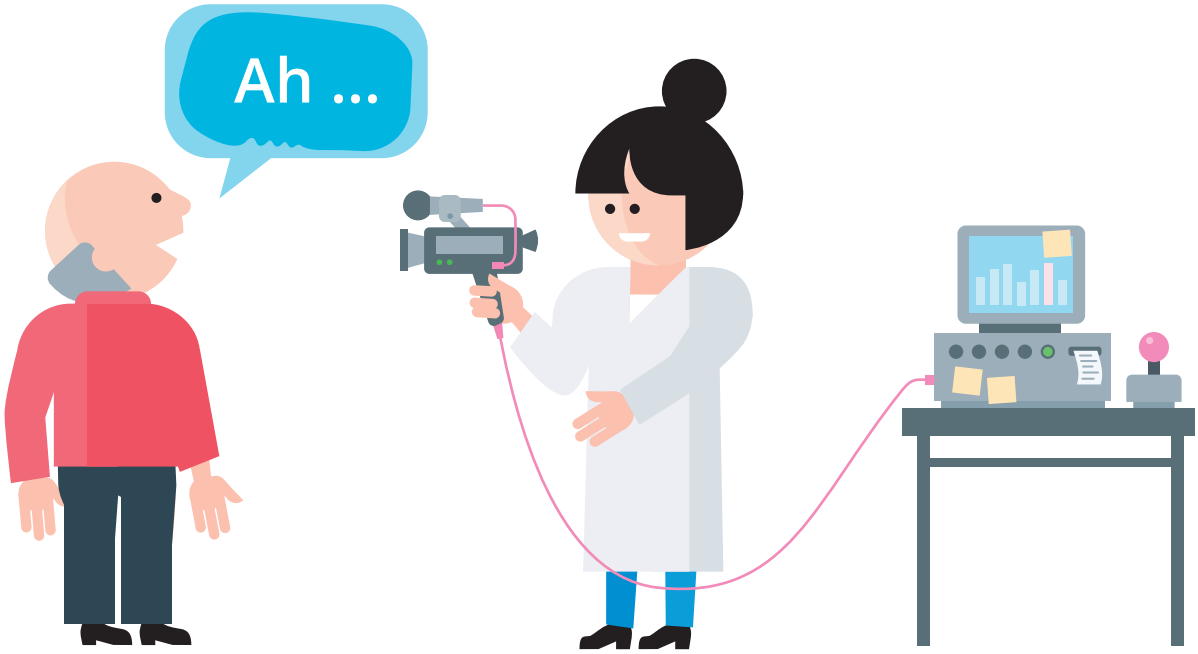
Tant que la fiabilité technique et organisationnelle de la détection des émotions et des maladies par les données faciales et vocales n'est pas assurée, ces technologies devraient être soumises à un moratoire dans certains domaines de la vie courante. La plus grande prudence s'impose notamment dans le domaine de l'application de la loi et des assurances. En particulier, analyser la voix de quelqu'un qui appelle un call-center ou présente sa candidature en vue de détecter des maladies ou des émotions doit être interdit, l'utilisation de ces données biométriques étant contraire au but prévu. Certaines applications bénéfiques pourraient être classées comme utilisation à haut risque et donc autorisées sous réserve de l'accord des autorités.

## Combler les lacunes légales, promouvoir la formation continue, soutenir les personnes concernées

Une base juridique explicite doit être mise en place pour l'utilisation des technologies de reconnaissance de la voix, de la parole et du visage par les services publics. Il convient d'une part de définir les mécanismes de garantie de l'État de droit et, d'autre part, d'assurer la vérification du caractère nécessaire de leur utilisation.

Les spécialistes qui recourent à la reconnaissance de la voix, de la parole et du visage, qui accèdent aux données collectées et qui partagent les résultats doivent recevoir une formation de base et suivre les cours de perfectionnement appropriés. Il faut également des directives et des services d'assistance pour veiller à ce que les opérateurs de ces systèmes respectent les principes de la protection des données.

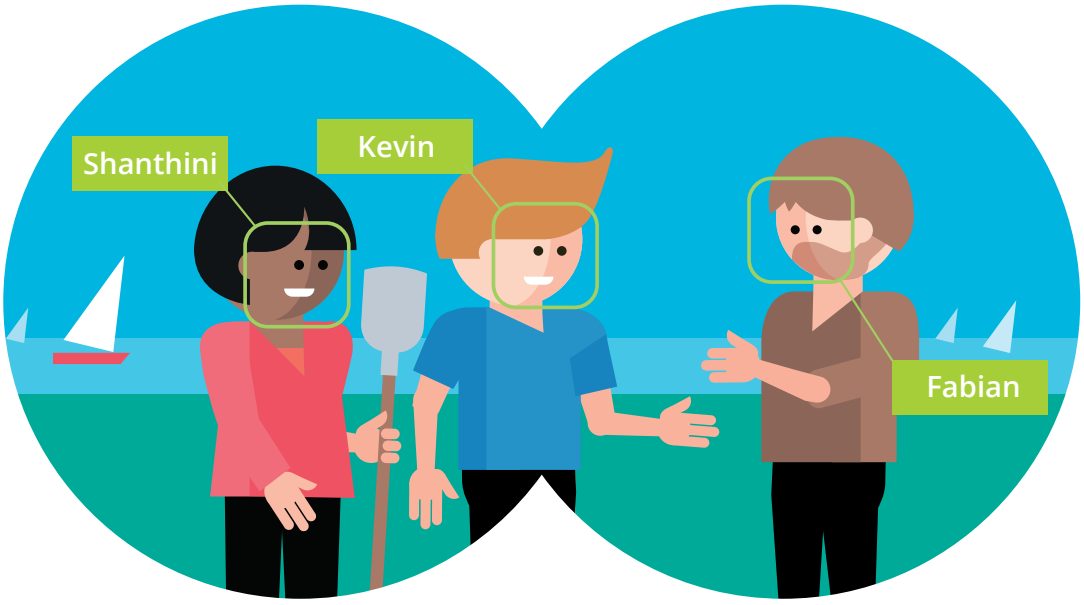
En cas de recours à des technologies de reconnaissance de la voix, de la parole ou du visage, une communication claire et transparente est nécessaire. Autant que possible, les personnes concernées doivent se voir proposer des alternatives qui n'impliquent pas d'inconvénients, comme un temps d'attente plus long ou des frais.



Il faut également mettre sur pied des services d'assistance auxquels peuvent s'adresser les personnes désireuses de se protéger des inconvénients liés aux systèmes de reconnaissance de la voix, de la parole et du visage, et qui veulent connaître et faire valoir leurs droits.

Il est recommandé aux développeurs de dispositifs et d'apps de reconnaissance de la voix, de la parole et du visage de veiller, dans la mesure du possible, à stocker et à traiter les données collectées uniquement dans les appareils eux-mêmes. Cela renforcerait à la fois la protection des données et le droit des utilisatrices et utilisateurs à disposer de leurs propres données.

Enfin, il est important de stimuler le dialogue sociétal sur les avantages et les inconvénients des technologies de reconnaissance de la voix, de la parole et du visage, ainsi que sur les domaines où il faudrait les autoriser et ceux où leur utilisation doit être limitée. Le grand public devrait aussi être encouragé à s'interroger sur l'importance de la reconnaissance de la voix, de la parole et du visage pour la vie en société et pour le fonctionnement d'une démocratie – notamment dans la mesure où ces technologies augmentent la pression sociale et favorisent un regard moralisateur sur les citoyennes et citoyens ou sur leurs actions. La présente étude est une invitation à mener ces réflexions.



## Groupe d'accompagnement

- Dr Bruno Baeriswyl, expert en matière de protection des données, membre du comité directeur de TA-SWISS, président du groupe d'accompagnement
- Dominik Brumm, Head of Development, Cubera
- Dr Volker Dellwo, Institut für Computerlinguistik, Universität Zürich
- Dr Jean Hennebert, informatique et systèmes de communication, Université de Fribourg
- Dre Anna Jobin, sociologue, Alexander von Humboldt Institut für Internet und Gesellschaft
- Dre Annett Laube, Technik und Informatik, Berner Fachhochschule
- Dr Klaus Scherer, Swiss Center for Affective Sciences, Université de Genève
- Remo Schmidlin, juriste, Lenz & Staehelin
- Dr Thomas Vetter, Departement Mathematik und Informatik, Universität Basel
- Patrick Walder, Amnesty International Suisse

## Gestion du projet chez TA-SWISS

- Dre Elisabeth Ehrensperger, directrice
- Dre Christina Tobler, responsable de projet (2020 – 2021)
- Dre Laetitia Ramelet, responsable de projet (2022)

## Impressum

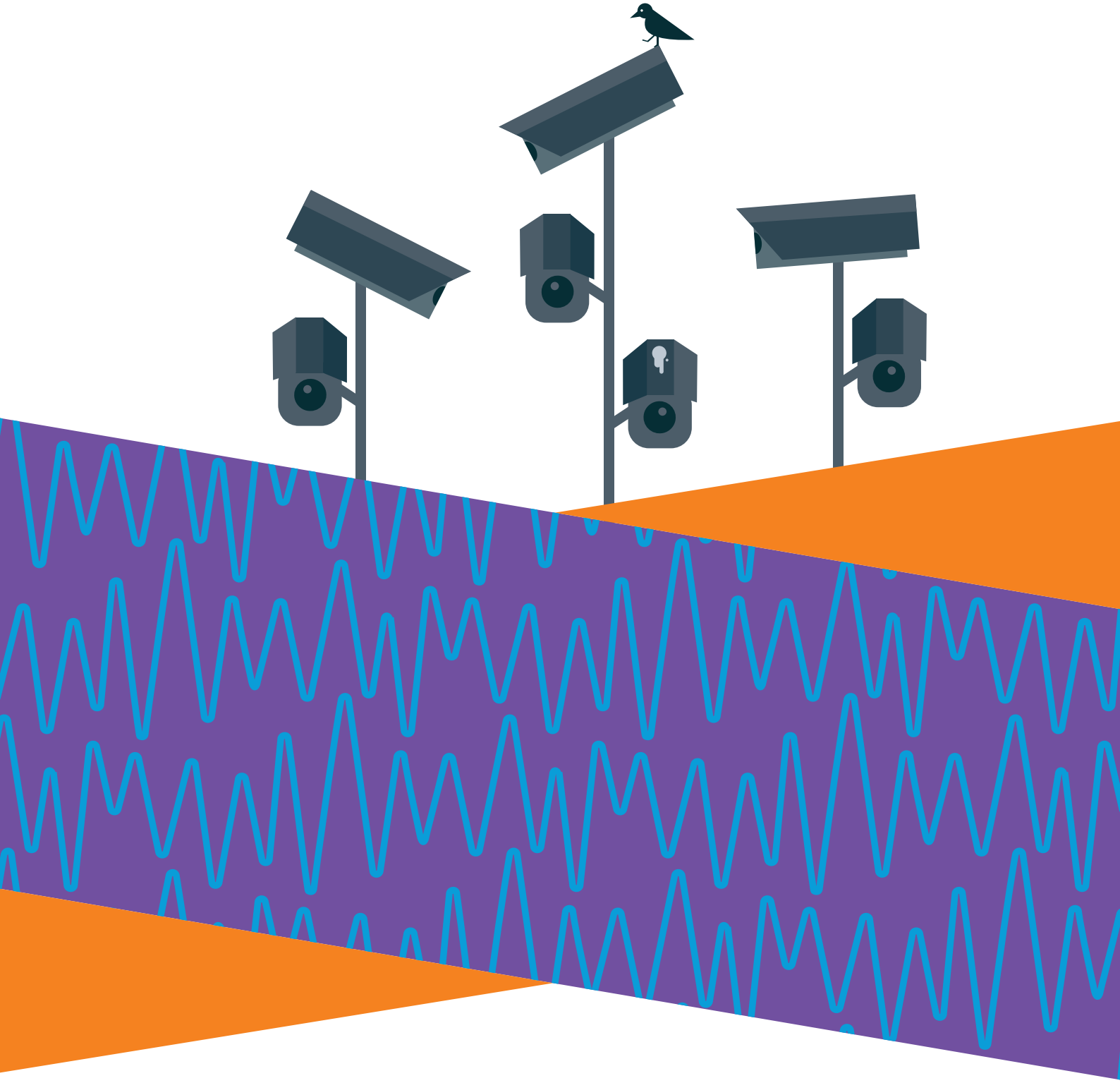
Quand nos faits et gestes sont observés en permanence
Synthèse de l'étude «Automatisierte Erkennung von Stimme, Sprache und Gesicht: Technische, rechtliche und gesellschaftliche Herausforderungen»
TA-SWISS, Berne 2022
TA 79A/2022

Rédaction : Dre Lucienne Rey, TA-SWISS, Berne
Traduction : pro-verbial, Zurich
Production : Dre Laetitia Ramelet et Fabian Schluep, TA-SWISS, Berne
Mise en page et illustrations : Hannes Saxer, Berne
Impression : Jordi AG – Das Medienhaus, Belp

**TA-SWISS – Fondation pour l'évaluation des choix technologiques**

Souvent susceptibles d'avoir une influence décisive sur la qualité de vie des gens, les nouvelles technologies peuvent en même temps comporter des risques nouveaux, qu'il est parfois difficile de percevoir d'emblée. La Fondation pour l'évaluation des choix technologiques TA-SWISS s'intéresse aux avantages et aux risques potentiels des nouvelles technologies qui se développent dans les domaines « biotechnologie et médecine », « société de l'information » et « mobilité / énergie / climat ». Ses études s'adressent tant aux décideurs du monde politique et économique qu'à l'opinion publique. TA-SWISS s'attache, en outre, à favoriser par des méthodes participatives, l'échange d'informations et d'opinions entre les spécialistes du monde scientifique, économique et politique et la population. TA-SWISS se doit, dans toutes ses projets sur les avantages et les risques potentiels des nouvelles technologies, de fournir des informations aussi factuelles, indépendantes et étayées que possible. Il y parvient en mettant chaque fois sur pied un groupe d'accompagnement composé d'experts choisis de manière à ce que leurs compétences respectives couvrent ensemble la plupart des aspects du sujet à traiter.

La fondation TA-SWISS est un centre de compétence des Académies suisses des sciences.



TA-SWISS
Fondation pour l'évaluation
des choix technologiques
Brunngasse 36
CH-3011 Berne
info@ta-swiss.ch
www.ta-swiss.ch

membre des
académies suisses
des sciences