POUR LES CONCEPTEURS D'ASSISTANTS VOCAUX

En se focalisant sur les aspects logiciels, les développeurs d'assistants vocaux sont responsables des implémentations techniques qui vont régir le fonctionnement de ces derniers. Modalités d'activation, choix d'architecture, accès aux données, gestions des enregistrements, spécifications matérielles, etc., c'est par ces choix de conception que se matérialisent les possibilités de l'assistant. Afin de garantir aux utilisateurs maîtrise et contrôle sur leurs données, sept points de vigilance doivent être gardés à l'esprit.

Établir la transparence, fondement de la confiance

Comme pour tout traitement de données personnelles, le RGPD impose un devoir d'information des personnes concernées par les traitements mis en œuvre par des assistants vocaux (voir *Les notions clés du RGPD*, page 48).

Plus encore que pour les services accessibles en ligne, l'usage d'un assistant vocal impose aux utilisateurs de faire confiance à un dispositif qui donne parfois peu d'information sur son fonctionnement et dont les modalités de maîtrise ne sont aujourd'hui pas aussi intégrées que celles qui s'exercent sur un ordinateur ou un smartphone.

Nos conseils

- Être transparent sur le fonctionnement de l'assistant vocal et, notamment, sur les différentes étapes du traitement depuis la phase de collecte en passant par la retranscription de la voix sous forme de texte pour analyse et réponse à l'utilisateur.
- Délivrer cette information avant l'achat du dispositif, soit en portant les mentions d'information sur l'emballage, soit en mettant à disposition des futurs clients une note d'information.
- Rappeler cette information à la première utilisation du dispositif, qui pourrait être également proposée en version audio.
- Si de nombreuses informations doivent être données, prévoir une information par strates permettant de prioriser les éléments à présenter.
- Penser les interfaces pour que les utilisateurs puissent naviguer aisément dans les différentes strates d'information et trouver celles dont ils ont besoin à tout moment (lors de l'installation ou ultérieurement).
- Prévoir une présentation orale et compréhensible des conditions d'utilisation et des règles de protection de la vie privée, accessible en interrogeant l'assistant.
- Permettre à l'utilisateur de poser des questions sur les traitements de données personnelles de l'assistant vocal et fournir des réponses claires par oral.
- Si le concepteur a recours à une utilisation ultérieure des données destinée à améliorer ses propres services – par exemple en employant des personnes pour réaliser des écoutes et annotations des conversations enregistrées – informer spécifiquement et clairement l'utilisateur de cet usage, et indiquer dans l'interface de gestion de l'assistant les commandes et enregistrements qui ont fait l'objet d'une telle utilisation.

- Mettre par défaut le paramétrage du dispositif dans son fonctionnement le plus protecteur de la vie privée pour son utilisateur.
- Si l'assistant vocal nécessite des évolutions logicielles et/ou mises à jour importantes, prévoir un contact avec l'utilisateur pour l'en avertir et lui préciser la nature des changements et leurs conséquences.
- Si le concepteur de l'assistant fournit également un kit de développement d'application (ou SDK pour *Software Development Kit*), intégrer des fonctionnalités et outils logiciels permettant aux développeurs d'applications tierces d'appliquer l'impératif de transparence.

ZOOM SUR...

Données et design, pour des interfaces respectueuses de la vie privée

Avec la publication de son Cahier IP n°6 *La forme des choix*, la CNIL s'est intéressée à promouvoir l'émergence d'un design des interfaces plus responsable et respectueux des principes de protection des données[143]. À l'instar des questions juridiques et techniques, le design des interfaces doit désormais être au centre des préoccupations du régulateur, tout comme il est déjà au cœur des relations entre les individus et les fournisseurs de services.

À la suite de cette publication, la plateforme Données & Design a été lancée[144]. Celle-ci vise à créer des opportunités de collaboration et des espaces d'échange entre des designers pour co-construire des parcours respectueux de la vie privée. L'objectif est d'intégrer concrètement ces réflexions dans le travail quotidien des designers afin de les aider à argumenter leurs choix et à travailler en plus proche collaboration avec d'autres fonctions (chefs de produits, chefs de projets, départements juridiques, etc.) sur la protection des données personnelles.

Divers contenus expliquant et illustrant les points de la réglementation sur lesquels les designers peuvent agir sont mis à disposition. Dans les faits, la plateforme Données & Design est structurée autour d'approches complémentaires relatives à l'explication de concepts clés du RGPD (information des personnes, consentement et exercice des droits), à la mise à disposition d'études de cas et la création d'espaces d'échanges sur ces questions tant en ligne que lors de rencontres physiques. Si les travaux Données & Design ne s'adressent pas spécifiquement aux interfaces vocales, les éléments y figurant peuvent alimenter la réflexion sur les bonnes pratiques à mettre en œuvre avec des assistants vocaux.

Donner des moyens de contrôle aux utilisateurs

Autre exigence majeure portée par le RGPD, la mise en œuvre de moyens permettant aux personnes de maîtriser les usages qui sont faits de leurs données et d'exercer leurs droits de façon simple et effective. Ces modalités de contrôle et d'exercice doivent être adaptées à l'interface vocale de l'assistant.

143 - LINC, Cahier IP6 *La forme des choix,* janvier 2019, https://linc.cnil.fr/fr/cahier-ip6-la-forme-des-choix-0
144 - https://design.cnil.fr/

Nos conseils

Sur l'architecture logicielle

- Promouvoir des architectures logicielles respectueuses de la vie privée des utilisateurs par construction. Par exemple :
  - Pour les services ne nécessitant pas d'accès distant (réveil, pilotage de lumières, etc.) mettre en œuvre des traitements fonctionnant exclusivement de façon locale.
  - Pour minimiser l'exposition des données personnelles, mettre en œuvre au maximum les principes de l'informatique en périphérie (*edge computing*) pour que ne soit transférées sur des serveurs centralisés que les données strictement nécessaires.

Sur les modalités de paramétrage

- Permettre aux utilisateurs de gérer facilement leurs données (écouter, supprimer, détecter des usages anormaux et le cas échéant, récupérer),
  - à travers un tableau de bord accessible via un écran compagnon ;
  - directement par interrogation de l'assistant par la voix.
- Permettre à l'utilisateur de paramétrer finement les fonctionnalités accessibles via son assistant vocal. Par exemple, mettre à sa disposition une fonctionnalité de suppression automatique des informations dès lors que l'utilisateur a obtenu la réponse à sa demande, ou après un délai qu'il peut fixer.
- Prévoir des interfaces fluides et facile d'accès pour la gestion des applications tierces que celles-ci soient vocales et/ou via un écran compagnon et laisser la possibilité à l'utilisateur de désactiver les services et applications préinstallés.
- Envisager, en fonction de la criticité des applications possibles, la mise en place de modalités de filtrage à destination des jeunes enfants activables par leurs parents.

Sur les modalités d'enregistrement

- Offrir à l'utilisateur un moyen de désactiver physiquement le microphone du dispositif.
- Proposer à l'utilisateur une fonction d'activation manuelle, qui peut soit déclencher l'écoute des instructions soit activer une période définie d'attente du mot-clé d'activation.
- Indiquer à l'utilisateur par un signal sonore le début et la fin des périodes d'enregistrement.
- Proposer à l'utilisateur une commande vocale spécifique de désactivation de l'appareil (par exemple lorsqu'il y a des invités, etc.).
- Penser dès la conception la possibilité d'une utilisation par des personnes en situation de dépendance ou de handicap. Par d'exemple, un signal lumineux indiquant que l'appareil est en mode d'écoute actif n'est pas approprié pour des personnes mal-voyantes.

Sur la gestion des comptes

- Permettre d'associer un ou plusieurs comptes personnels à l'assistant en fonction des utilisations possibles de celui-ci.
- Permettre de ne pas associer de compte ou d'associer un compte générique quand l'assistant est destiné à un lieu collectif ou public ou bien à un usage professionnel.
- Proposer un mode de navigation privée pour les actions ne nécessitant pas de s'authentifier, permettant à un utilisateur d'interagir sans qu'un compte soit associé, ni que soient conservées de traces de ces interactions.
- Si plusieurs comptes personnels sont associés à un même dispositif, mettre en œuvre des moyens d'authentification fiables pour le passage de l'un à l'autre et ainsi prévenir de possibles usurpations d'identité.

S'assurer du bon dimensionnement de la collecte de données

Par les interactions qu'ils ont avec eux, les utilisateurs d'assistants vocaux sont susceptibles de transmettre de nombreuses informations malgré eux. Qui plus est, suivant les modalités d'activation des dispositifs (par exemple suite à la prononciation d'un mot-clé), des enregistrements inopinés peuvent également advenir.

Nos conseils

- Déterminer des durées de conservation distinctes selon le type de données collectées. Par exemple, les données associées au compte utilisateur peuvent ainsi être conservées plus longtemps que les requêtes ponctuelles effectuées auprès de l'assistant vocal.
- Ne pas demander de création de compte utilisateur si l'assistant ne le nécessite pas, par exemple si sa fonction est de fournir des renseignements génériques ou de programmer des actions simples.
- Ne pas conserver les enregistrements occasionnés par une fausse activation ou, a minima, les identifier spécifiquement pour que l'utilisateur en soit averti.

Traiter des données biométriques

Certains assistants vocaux ayant vocation à être déployés dans des environnements partagés, des constructeurs proposent d'y associer des comptes pour chaque utilisateur (par exemple les différents membres d'un foyer). Une possibilité pour passer d'un compte à un autre est de se baser sur l'authentification ou l'identification du locuteur. Toutefois, celles-ci reposent sur l'exploitation de données biométriques - les gabarits ou modèles de voix – qui sont considérées comme des données sensibles au sens du RGPD. Pour rappel, celui-ci interdit le traitement de telles données, sauf certaines exceptions limitativement énumérées (article 9(2) - voir *Les notions clés du RGPD*, page 48).

Il est ainsi indispensable de s'assurer que le traitement des données biométriques est désactivé par défaut et conditionné à l'obtention du consentement explicite de chaque personne dont la voix est susceptible d'être ainsi analysée. Qui plus est, l'utilisateur doit disposer d'un mode d'authentification ou d'identification alternatif ne présentant pas de contraintes additionnelles afin qu'il dispose d'une véritable liberté de choix (voir Chapitre III *Cas d'usages : le RGPD en pratique*, page 46).

Nos conseils

- Pour les assistants vocaux non personnels, c'est-à-dire ceux qui sont utilisables par plus d'une personne ou disposés dans un espace partagé, prévoir un mot-clé spécifique ou une question aux personnes présentes et recueillir ainsi leur consentement pour déclencher une reconnaissance biométrique. Par exemple, l'utilisateur peut dire « authentification » ou bien l'assistant peut demander « souhaitez-vous être identifié ? » et attendre une réponse positive pour activer le traitement biométrique.
- Conserver le gabarit biométrique de l'utilisateur sous son contrôle exclusif et privilégier le stockage sur un support individuel, qui peut être le dispositif embarquant l'assistant.
- Réaliser les opérations d'authentification/identification en local, c'est-à-dire directement dans le dispositif embarquant l'assistant.

Satisfaire l'impératif de sécurité

Le RGPD précise que la protection des données personnelles nécessite de prendre des mesures techniques et organisationnelles appropriées afin de garantir un niveau de sécurité adapté aux risques. L'analyse de ces risques constitue donc une étape capitale qui doit être menée avant la conception de l'assistant vocal. Elle peut notamment prendre la forme d'un processus formalisé connu sous le nom sous le nom d'Analyse d'Impact relative à la Protection des Données (AIPD). Cette démarche, qui est obligatoire dans certains cas, et fortement conseillée dans d'autres, a été précisée par la CNIL dans de nombreux outils et méthodes[145] (voir également l'encadré page 56).

La mise en œuvre de tout traitement de données à caractère personnel implique donc une obligation de sécurité. Outre les mesures génériques qui peuvent être retrouvées dans le Guide sur la sécurité des données personnelles[146] et dans le Guide développeur[147] (présenté dans l'encadré page 75), des bonnes pratiques spécifiques aux assistants vocaux peuvent être précisées.

145 - CNIL, L'analyse d'impact relative à la protection des données (AIPD), https://www.cnil.fr/fr/RGPD-analyse-impact-protection-des-donnees-aipd
146 - CNIL, Le guide de sécurité des données personnelles, édition 2018 https://www.cnil.fr/sites/default/files/atoms/files/cnil_guide_securite_personnelle.pdf
147 - CNIL, La CNIL publie un guide RGPD pour les développeurs, janvier 2020, https://www.cnil.fr/fr/la-cnil-publie-un-guide-rgpd-pour-les-developpeurs

Nos conseils

- Choisir une modalité d'activation de l'assistant proportionnée au niveau de risque des services gérés par l'assistant vocal (par exemple un bouton dans certains cas).
- Mettre en œuvre un paramétrage raisonné de la détection du mot-clé : privilégier un taux bas de fausses acceptations pour éviter les déclenchements inopinés.
- Proposer à l'utilisateur de choisir son mot-clé, avec des conseils sur le choix : le mot-clé doit respecter certains critères (ne pas être utilisé trop fréquemment dans les discussions, ne pas être trop proche d'autres mots, etc.).
- Identifier les applications à risque et proposer pour celles-ci des mesures de sécurité comme l'authentification à deux facteurs (par exemple via une validation à effectuer suite à un envoi de courriel ou de SMS).
- Réaliser une Analyse d'Impact relative à la Protection des Données et la mettre régulièrement à jour afin d'assurer que les mesure techniques et organisationnelles prises sont en adéquation avec les risques que le traitement de données fait peser sur les personnes (pour plus de précisions sur l'AIPD, voir l'encadré page 56).
- Prévoir des mécanismes d'information et d'alerte de l'utilisateur en cas de dysfonctionnement de l'assistant ou d'activités inhabituelles, notamment en cas de violations de données[148].

Organiser son écosystème applicatif

Comme présenté précédemment, certains assistants – et notamment les plus répandus auprès du grand public – se positionnent comme une plateforme pour héberger des applications tierces. Ces modes de fonctionnement doivent s'accompagner de mesures spécifiques et d'une attention renforcée sur le partage des données.

Nos conseils

- Lorsqu'un acteur tiers utilise les ressources technologiques mises à disposition pour le développement de son application, définir contractuellement les règles applicables en matière de confidentialité et de respect de la vie privée de manière suffisamment claire et précise.
- Préciser la chaîne de responsabilité impliquant le concepteur de l'assistant et le développeur de l'application.
- Accompagner les développeurs d'application dans la mise en œuvre d'un service sécurisé, comme des API d'authentification génériques et de présentation d'information adaptée au dispositif.
- Limiter le nombre d'applications disponibles par défaut au strict nécessaire et favoriser l'installation des applications à l'initiative de l'utilisateur, par exemple via un magasin (*store*), plutôt qu'une mise à disposition directe depuis l'assistant, sans sélection préalable ni information spécifique.
- Lorsque l'assistant accède directement à une application tierce, ne partager aucune donnée personnelle avec le tiers sans une information claire de la personne.
- Mettre en œuvre une politique de validation des applications déposées dans le magasin (store) et régulièrement vérifier celui-ci, notamment en surveillant la présence d'applications aux noms très proches d'applications légitimes.
- Offrir à l'utilisateur des outils de contrôle granulaire pour toutes les applications installées et les données accédées par celles-ci, notamment les données sensibles ou révélatrices de sa vie privée (données de santé ou biométriques, géolocalisation, historique de recherche, etc.). Ces contrôles devraient aussi pouvoir être temporaires (accorder un accès une seule fois ou pour une durée limitée).
- Si des protocoles d'autorisation sont mis en œuvre pour permettre à une application tierce d'accéder à un service, s'assurer que les jetons (*tokens*) permettant l'authentification des utilisateurs ont une durée de vie limitée et raisonnable et qu'ils soient facilement révocables.

148 - CNIL, Notifier une violation de données personnelles, https://www.cnil.fr/fr/notifier-une-violation-de-donnees-personnelles

Encadrer l'utilisation des données pour l'amélioration des technologies

Les assistants vocaux, comme d'autres objets connectés, peuvent faire remonter des données personnelles vers leurs concepteurs à des fins d'amélioration des services. Il peut s'agir de données techniques utilisées à des fins statistiques (par exemple, des informations relatives à l'utilisation d'une ampoule connectée et à sa durée de vie). Dans le cas spécifique des assistants vocaux, la question de l'amélioration du produit peut également impliquer des traitements de données issues des commandes vocales, tels que les enregistrements audio ou leurs retranscriptions textuelles.

En effet, ceux-ci mettent en œuvre des algorithmes d'intelligence artificielle dont les performances sont directement corrélées à des jeux de données utilisés pour l'apprentissage de modèles statistiques. Par conséquent, il peut être légitime de souhaiter accéder à des données relatives à l'utilisation en conditions réelles du dispositif pour travailler à son amélioration. En pratique, il peut par exemple s'agir pour le concepteur de l'assistant vocal de recourir à des salariés en interne ou à ceux d'une entreprise sous-traitante pour procéder à l'écoute et annotation des enregistrements vocaux afin que ceux-ci puisse permettre l'amélioration des modèles. Une grande attention doit être portée aux modalités de mise en œuvre de telles utilisations des données.

Il convient ainsi de respecter les droits de personnes concernées en ne traitant aucune donnée pour cette finalité sans s'assurer de leur bonne information et de la base légale de leur traitement Qui plus est, ces personnes doivent être dans la capacité de savoir clairement et à tout moment si leurs données sont utilisées à cette fin et de s'y opposer facilement.

Nos conseils

- Mettre en œuvre des mesures de sécurité fortes : restriction de l'accès aux données des enregistrements vocaux aux seuls salariés habilités à écouter les conversations, authentification forte, mesure de traçabilité renforcé, blocage de l'extraction des enregistrements audio, etc.
- Ne pas fournir d'informations relatives à l'utilisateur de l'assistant en accompagnement de l'enregistrement et en particulier, ne pas mettre en correspondance les enregistrements et fichiers transcrits avec d'autres données susceptibles d'être collectées (identifiants de l'appareil, localisation, comptes associés, etc.).
- Mettre en œuvre des mesures d'altération des informations présentes dans les enregistrements audio en procédant par exemple :
  - À une modification des caractéristiques du locuteur (timbre, pitch, prosodie, etc.) de manière irréversible ;
  - À une suppression/offuscation des informations contenues dans les enregistrements et retranscriptions (noms et prénoms, adresse, etc.).
- Échantillonner le message concerné en plusieurs parties qui seront analysées par des personnes différentes.
- Limiter la durée des écoutes des conversations des utilisateurs à quelques secondes par échantillon.
- S'assurer que les appareils à partir desquels les échantillons des conversations sont écoutées ne sont pas systématiquement les mêmes.
- En cas de recours aux salariés d'une entreprise sous-traitante, prévoir dans le contrat de sous-traitance toutes les garanties nécessaires en matière de sécurité (nécessité d'inclure une clause de confidentialité dans les contrats de travail du personnel concerné, modalités d'accès aux locaux et aux données, processus d'habilitation des personnes, durées de conservation des données, etc.).