Wytyczne 02/2021 w sprawie wirtualnych asystentów głosowych

Wersja 2.0

Przyjęta 7 lipca 2021 r.

Translations proofread by EDPB Members.
This language version has not yet been proofread.

Historia wersji

Wersja 2.0	7 lipca 2021 r.	Przyjęcie wytycznych po konsultacjach publicznych
Wersja 1.0	9 marca 2021 r.	Przyjęcie wytycznych do konsultacji publicznych

STRESZCZENIE

Wirtualny asystent głosowy (VVA) to usługa, która rozumie polecenia głosowe i w razie potrzeby wykonuje je lub pośredniczy w kontaktach z innymi systemami informatycznymi. VVA są obecnie dostępni w większości smartfonów i tabletów, w tradycyjnych komputerach, a w ostatnich latach nawet w samodzielnych urządzeniach, takich jak inteligentne głośniki.

VVA działają jako interfejs między użytkownikami i ich urządzeniami komputerowymi a usługami online, takimi jak wyszukiwarki czy sklepy internetowe. Ze względu na swoją rolę VVA mają dostęp do ogromnej ilości danych osobowych, w tym wszystkich poleceń użytkownika (np. historii przeglądania lub wyszukiwania) i odpowiedzi (np. spotkań w terminarzu).

Zdecydowana większość usług VVA została zaprojektowana przez nielicznych projektantów VVA. VVA mogą jednak współpracować z aplikacjami zaprogramowanymi przez osoby trzecie (programistów aplikacji VVA) w celu obsługi bardziej zaawansowanych poleceń.

Do prawidłowego działania VVA potrzebuje urządzenia końcowego wyposażonego w mikrofony i głośniki. Urządzenie to przechowuje dane głosowe i inne dane, które bieżąco używane VVA przesyłają do zdalnych serwerów VVA.

Administratorzy danych świadczący usługi VVA i ich podmioty przetwarzające muszą zatem brać pod uwagę zarówno ogólne rozporządzenie o ochronie danych (RODO)[1], jak i dyrektywę o e-prywatności[2].

W niniejszych wytycznych określono niektóre z najistotniejszych wyzwań związanych z zapewnieniem zgodności i przedstawiono zalecenia dla odpowiednich zainteresowanych stron dotyczące sposobów radzenia sobie z nimi.

Administratorzy danych świadczący usługi VVA za pośrednictwem bezekranowych urządzeń końcowych nadal muszą dopełnić zgodnie z RODO obowiązku informacyjnego wobec użytkowników podczas konfigurowania lub instalowania VVA bądź korzystania z aplikacji VVA po raz pierwszy. W związku z tym zalecamy dostawcom/projektantom i programistom VVA opracowanie interfejsów opartych na głosie, aby ułatwić przekazywanie obowiązkowych informacji.

Obecnie wszyscy VVA wymagają rejestracji co najmniej jednego użytkownika w serwisie. Zgodnie z obowiązkiem uwzględnienia ochrony danych w fazie projektowania i domyślnej ochrony danych, dostawcy/projektanci i programiści VVA powinni rozważyć konieczność posiadania zarejestrowanego użytkownika dla każdej ze swoich funkcji.

Konta użytkowników stosowane przez wielu projektantów VVA łączą usługę VVA z innymi usługami, takimi jak poczta elektroniczna lub strumieniowa transmisja wideo. Europejska Rada Ochrony Danych (EROD) uważa, że administratorzy danych powinni powstrzymać się od takich praktyk, ponieważ wiążą się one z wykorzystaniem długich i złożonych polityk prywatności, które nie byłyby zgodne z zasadą przejrzystości RODO.

[1] Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679/UE z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (zwane dalej „RODO").

[2] Dyrektywa 2002/58/WE Parlamentu Europejskiego i Rady z dnia 12 lipca 2002 r. dotycząca przetwarzania danych osobowych i ochrony prywatności w sektorze łączności elektronicznej (dyrektywa o prywatności i łączności elektronicznej), zmieniona dyrektywą 2006/24/WE i dyrektywą 2009/136/WE (zwana dalej „dyrektywą o e-prywatności").

W wytycznych uwzględniono cztery najczęstsze cele, dla których VVA przetwarzają dane osobowe: realizacja poleceń, doskonalenie modelu uczenia maszynowego VVA, identyfikacja biometryczna oraz profilowanie na potrzeby spersonalizowanych treści lub reklam.

W zakresie, w jakim dane VVA są przetwarzane w celu realizacji poleceń użytkownika, tj. gdy jest to absolutnie niezbędne do świadczenia usługi zamówionej przez użytkownika, administratorzy danych są zwolnieni z wymogu uzyskania uprzedniej zgody na mocy art. 5 ust. 3 dyrektywy o e-prywatności. Natomiast zgoda taka, wymagana na mocy art. 5 ust. 3 dyrektywy o e-prywatności, byłaby konieczna do przechowywania lub uzyskania dostępu do informacji w celu innym niż realizacja polecenia użytkownika.

Niektóre usługi VVA przetrzymują dane osobowe do czasu, gdy użytkownicy zażądają ich usunięcia. Nie jest to zgodne z zasadą ograniczenia przechowywania. VVA powinni przechowywać dane nie dłużej niż jest to konieczne do celów, dla których dane osobowe są przetwarzane.

Jeżeli administrator danych dowie się (np. w wyniku procesów przeglądu jakości) o przypadkowym zbieraniu danych osobowych, powinien sprawdzić, czy istnieje ważna podstawa prawna dla każdego celu przetwarzania takich danych. W przeciwnym wypadku przypadkowo zebrane dane powinny zostać usunięte.

VVA mogą przetwarzać dane wielu podmiotów danych. Dostawcy/projektanci VVA powinni zatem wdrożyć mechanizmy kontroli dostępu w celu zapewnienia poufności, integralności i dostępności danych osobowych. Niektóre tradycyjne mechanizmy kontroli dostępu, takie jak hasła, nie nadają się jednak do zastosowania w kontekście VVA, ponieważ musiałyby być wypowiadane na głos. Wytyczne zawierają pewne uwagi w tym względzie, w tym sekcję poświęconą przetwarzaniu specjalnych kategorii danych do celów identyfikacji biometrycznej.

Dostawcy/projektanci VVA powinni wziąć pod uwagę, że podczas zbierania głosu użytkownika nagranie może zawierać głos innych osób lub dane takie jak odgłosy w tle, które nie są niezbędne do świadczenia usługi. W miarę możliwości projektanci VVA powinni zatem rozważyć wprowadzenie technologii filtrujących zbędne dane i zapewniających, że nagrywany jest tylko głos użytkownika.

Oceniając potrzebę przeprowadzenia oceny skutków dla ochrony danych, EROD uważa, że jest bardzo prawdopodobne, iż usługi VVA mieszczą się w kategoriach i warunkach określonych jako wymagające przeprowadzenia takiej oceny skutków.

Administratorzy danych świadczący usługi VVA powinni zapewnić użytkownikom możliwość korzystania z praw przysługujących osobom, których dane dotyczą, za pomocą łatwych do zrozumienia poleceń głosowych. Dostawcy/projektanci VVA, a także programiści aplikacji powinni na koniec procesu poinformować użytkowników, że ich prawa zostały należycie uwzględnione, głosowo lub poprzez dostarczenie pisemnego powiadomienia na telefon komórkowy, konto użytkownika lub w inny sposób wybrany przez użytkownika.

Spis treści

Europejska Rada Ochrony Danych

uwzględniając art. 70 ust. 1 lit. e) i j) rozporządzenia Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (zwanego dalej „RODO"),

uwzględniając Porozumienie EOG, a w szczególności jego załącznik XI i protokół 37, w brzmieniu zmienionym decyzją Wspólnego Komitetu EOG nr 154/2018 z dnia 6 lipca 2018 r.[3],

uwzględniając art. 12 i 22 swojego regulaminu wewnętrznego,

PRZYJMUJE NINIEJSZE WYTYCZNE:

1 OGÓLNE

1. Ostatnie postępy technologiczne znacznie zwiększyły dokładność i popularność wirtualnych asystentów głosowych (VVA). Urządzenia VVA zostały zintegrowane m.in. ze smartfonami, pojazdami podłączonymi do internetu, inteligentnymi głośnikami i inteligentnymi telewizorami. Dzięki tej integracji VVA uzyskali dostęp do informacji o charakterze prywatnym, które, jeśli nie są odpowiednio zarządzane, mogą naruszać prawa osób fizycznych do ochrony danych i prywatności. W związku z tym VVA i zintegrowane z nimi urządzenia znaleźli się pod obserwacją różnych organów ochrony danych.

2. Istnieje kilka zalet wykorzystania interakcji opartych na mowie, takich jak: naturalność interakcji, która nie wymaga szczególnego uczenia się od użytkowników, szybkość wykonywania poleceń i rozszerzenie pola działania, co może umożliwić szybszy dostęp do informacji. Jednak bazowanie na mowie niesie ze sobą również trudności w prawidłowej interpretacji komunikatu: zmienność sygnału dźwiękowego pomiędzy różnymi mówcami, środowisko akustyczne, wieloznaczność języka itp.

3. W praktyce główną motywacją do korzystania z VVA pozostaje płynność lub uproszczenie zadań. Może to dotyczyć np. wykonywania i odbierania połączeń, czy ustawiania zegara, zwłaszcza gdy użytkownicy nie mogą używać rąk. Automatyka domowa to główne zastosowanie, jakie proponują projektanci VVA. Proponując uproszczenie wykonywania zadań (włączanie światła, regulowanie ogrzewania, opuszczanie żaluzji itp.) i scentralizowanie ich za pomocą jednego narzędzia, które można łatwo uruchomić zdalnie, wpisują się w dyskurs jako ułatwienia domowe. Poza zastosowaniami osobistymi czy domowymi, komendy głosowe mogą zyskać zainteresowanie w środowiskach zawodowych, gdzie trudno jest obsługiwać narzędzia komputerowe i używać poleceń pisemnych (np. w pracy na linii produkcyjnej).

4. W teorii głównymi użytkownikami interfejsu głosowego mogłyby być osoby z niepełnosprawnościami lub niesamodzielne, dla których korzystanie z tradycyjnych interfejsów jest problematyczne. Wirtualny asystent głosowy może zapewnić łatwiejszy dostęp do informacji i zasobów komputerowych, a tym samym wspierać logikę sprzyjającą włączeniu

[3] Odniesienia do „państw członkowskich" w niniejszym dokumencie należy rozumieć jako odniesienia do „państw członkowskich EOG".

społecznemu, ponieważ wykorzystanie głosu umożliwia przezwyciężenie trudności związanych ze słowem pisanym, które mogą występować u niektórych klas użytkowników.

5. Wreszcie również zdrowie jest dziedziną, w której istnieje wiele przypadków zastosowania agentów konwersacyjnych, głosowych lub innych. Na przykład w czasie pandemii COVID-19 wdrożono różnych wirtualnych rozmówców, aby zaoferować użytkownikom dzwoniącym możliwość uzyskania wstępnej diagnozy. W dłuższej perspektywie niektórzy przewidują, że interakcje człowiek-asystent mogą mieć wpływ na cały proces opieki nad pacjentem: nie tylko w zakresie dobrostanu i zapobiegania, ale również leczenia i wsparcia.

6. Obecnie istnieje ponad 3 miliardy smartfonów i wszystkie z nich mają zintegrowanych VVA, z których większość jest domyślnie włączona. Niektóre z najbardziej rozpowszechnionych systemów operacyjnych w komputerach osobistych i laptopach również posiadają zintegrowanych wirtualnych asystentów. Niedawny wzrost popularności inteligentnych głośników (w 2019 r. sprzedano 147 milionów sztuk[4]) sprawia, że VVA trafiają do milionów domów i biur. Jednak obecne projekty VVA nie oferują domyślnie mechanizmów uwierzytelniania lub kontroli dostępu.

7. Niniejszy dokument ma na celu przedstawienie wytycznych dotyczących stosowania RODO w kontekście VVA.

2 INFORMACJE TECHNOLOGICZNE

2.1 Podstawowe cechy wirtualnych asystentów głosowych

8. VVA można zdefiniować jako aplikację programową, która umożliwia prowadzenie ustnego dialogu z użytkownikiem w języku naturalnym.

9. Język naturalny posiada semantykę właściwą dla języka ludzkiego. W zależności od właściwości języka i różnorodności leksyki, ta sama instrukcja może być sformułowana na wiele sposobów, a niektóre polecenia mogą wydawać się podobne, ale odnosić się do dwóch różnych obiektów. Do rozwiązywania tych niejednoznaczności często wykorzystuje się mechanizmy wnioskowania, np. w zależności od tego, co zostało wcześniej powiedziane, czasu, w którym polecenie zostało wydane, miejsca, zainteresowań danej osoby.

10. VVA można podzielić na moduły umożliwiające realizację różnych zadań: przechwytywanie i odtwarzanie dźwięku, automatyczna transkrypcja mowy (rozpoznawanie mowy), automatyczne przetwarzanie języka, strategie dialogowe, dostęp do ontologii (zbiorów danych i uporządkowanych pojęć związanych z daną dziedziną) i zewnętrznych źródeł wiedzy, generowanie języka, synteza głosu (zamiana tekstu na mowę) itp. Konkretniej rzecz ujmując, asystent powinien umożliwiać interakcję w celu wykonania czynności (np. „włącz radio", zgaś światło") lub umożliwienia dostępu do wiedzy (np. „jaka będzie jutro pogoda?", „czy jeździ pociąg o 7:43?"). Odgrywa więc rolę pośrednika i aranżera, który ma za zadanie ułatwić użytkownikowi realizację jego zadań.

11. W praktyce VVA nie jest inteligentnym głośnikiem, ale inteligentny głośnik może być wyposażony w asystenta głosowego. Te dwa urządzenia są powszechnie mylone, jednak to

[4] Zob. na przykład komunikat prasowy z dnia 1 sierpnia 2019 r. wydany przez Urząd Ochrony Danych i Informacji w Hamburgu: https://datenschutz-hamburg.de/pressemitteilungen/2019/08/2019-08-01-google-assistant

drugie jest tylko materialnym wcieleniem pierwszego. VVA może być wdrożony w smartfonie, inteligentnym głośniku, zegarku podłączonym do internetu, pojeździe, sprzęcie AGD itp.

12. Organizacja przetwarzania danych może obejmować wiele schematów przepływu informacji. Można wyodrębnić trzy główne elementy:

    Element fizyczny: element sprzętowy, z którym asystent jest zintegrowany (smartfon, głośnik, smart TV itp.) i który posiada mikrofony, głośniki oraz możliwości sieciowe i obliczeniowe (mniej lub bardziej rozwinięte w zależności od przypadku).

    Element oprogramowania: część realizująca interakcję człowiek-maszyna, która integruje moduły automatycznego rozpoznawania mowy, przetwarzania języka naturalnego, dialogu i syntezy mowy. Może być obsługiwana bezpośrednio w obrębie fizycznego sprzętu, ale w wielu przypadkach jest wykonywana zdalnie.

    Zasoby: dane zewnętrzne, takie jak bazy danych treści, ontologie lub aplikacje biznesowe, które dostarczają wiedzę (np. „podaj godzinę na zachodnim wybrzeżu Stanów Zjednoczonych", „przeczytaj moje e-maile") lub umożliwiają wykonanie żądanej czynności w konkretny sposób (np. „zwiększ temperaturę o 1,5°C").

13. VVA pozwalają na instalację komponentów lub aplikacji innych firm, które rozszerzają ich podstawowe funkcje. Każdy VVA inaczej nazywa te komponenty, ale wszystkie wymagają wymiany danych osobowych użytkowników między projektantem VVA a programistą aplikacji.

14. Chociaż większość VVA nie udostępnia fragmentu głosu programistom aplikacji, podmioty te nadal przetwarzają dane osobowe. Ponadto, w zależności od charakteru udostępnianych funkcji, programista aplikacji otrzymuje intencje i sloty, które mogą zawierać informacje wrażliwe, takie jak dane zdrowotne.

2.2 Podmioty działające w ekosystemie VVA

15. VVA może angażować dużą liczbę podmiotów i pośredników w całym łańcuchu realizacji. W praktyce można zidentyfikować do pięciu różnych podmiotów. W zależności od modeli biznesowych i wyborów technologicznych niektóre podmioty mogą jednak przyjmować kilka kombinacji ról, np. projektanta i integratora lub projektanta i programisty aplikacji:

    a. **Dostawca VVA (lub projektant):** odpowiedzialny za opracowanie VVA, projektuje i definiuje jego możliwości i domyślne funkcje: sposoby aktywacji, wybór architektury, dostęp do danych, zarządzanie rekordami, specyfikacje sprzętowe itp.

    b. **Programista aplikacji VVA:** podobnie jak w przypadku aplikacji mobilnych, tworzy aplikacje rozszerzające domyślne funkcje VVA. W tym celu konieczne jest przestrzeganie ograniczeń projektowych narzuconych przez projektanta.

    c. **Integrator:** producent podłączonych obiektów, który chce je wyposażyć w VVA. Powinien on przestrzegać wymagań określonych przez projektanta.

    d. **Właściciel:** odpowiedzialny za fizyczne przestrzenie, w których przebywają ludzie (miejsca zakwaterowania, środowiska zawodowe, wynajmowane pojazdy itp.); jego celem jest dostarczenie VVA swoim odbiorcom (ewentualnie z dedykowanymi aplikacjami).

e. **Użytkownik:** ostatnie ogniwo w łańcuchu wartości VVA, który może korzystać z niego na różnych urządzeniach (głośnik, telewizor, smartfon, zegarek itp.) w zależności od tego, jak i gdzie VVA został rozmieszczony i skonfigurowany.

2.3 Opis krok po kroku

16. Aby VVA mógł wykonać jakąś czynność lub uzyskać dostęp do informacji, należy wykonać szereg zadań:

1) Po podłączeniu do urządzenia (smartfonu, głośnika, pojazdu) VVA jest w stanie czuwania. Dokładniej rzecz ujmując, stale nasłuchuje. Jednak do momentu wykrycia określonego wyrażenia budzącego z urządzenia odbierającego głos nie jest transmitowany i nie są wykonywane żadne inne operacje poza wykrywaniem wyrażenia budzącego. W tym celu wykorzystywany jest kilkusekundowy bufor (więcej szczegółów w następnej sekcji).

2) Użytkownik wypowiada wyrażenie budzące, a VVA lokalnie porównuje dźwięk z wyrażeniem budzącym. Jeśli są one zgodne, VVA otwiera kanał nasłuchowy i zawartość audio jest natychmiast transmitowana.

3) W wielu przypadkach, jeśli przetwarzanie polecenia odbywa się zdalnie, po stronie serwera przeprowadzana jest druga kontrola wymowy hasła, aby ograniczyć niepożądane aktywacje.

4) Użytkownik podaje swoje polecenie, które jest na bieżąco przekazywane do dostawcy VVA. Sekwencja wypowiadanych słów jest następnie automatycznie poddawana transkrypcji (rozpoznawanie mowy).

5) Za pomocą technologii przetwarzania języka naturalnego (PJN) polecenie jest interpretowane. Wyodrębniane są intencje komunikatu i identyfikowane zmienne informacyjne (sloty). Menedżer dialogu jest następnie wykorzystywany do określenia scenariusza interakcji z użytkownikiem poprzez zapewnienie odpowiedniego schematu odpowiedzi.

6) Jeśli polecenie dotyczy funkcjonalności udostępnianej przez aplikację zewnętrzną (umiejętność, działanie, skrót itp.), dostawca VVA wysyła do programisty aplikacji intencje i zmienne informacyjne (sloty) komunikatu.

7) Określana jest odpowiedź dostosowana do żądania użytkownika – przynajmniej przypuszczalnie; odpowiedź „Nie mam odpowiedzi na twoje pytanie” jest dostosowaną odpowiedzią w przypadku, gdy VVA nie był w stanie prawidłowo zinterpretować polecenia. W razie potrzeby wykorzystywane są zdalne zasoby: publicznie dostępne bazy danych wiedzy (encyklopedia internetowa itp.) lub bazy danych wymagające uwierzytelnienia (konto bankowe, aplikacja muzyczna, konto klienta dla zakupów internetowych itp.), a zmienne informacyjne (sloty) są wypełniane uzyskaną wiedzą.

8) Tworzona jest fraza odpowiedzi lub identyfikowane jest działanie (opuszczanie żaluzji, podnoszenie temperatury, odtwarzanie utworu muzycznego, odpowiedź na pytanie itp.). Zdanie podlega syntetyzacji (rozpoznawanie mowy) albo czynność, która ma być wykonana, wysyłana jest do wybranego urządzenia.

9) VVA powraca do trybu czuwania.

Należy pamiętać, że choć obecnie większość przetwarzania głosowego jest wykonywana na zdalnych serwerach, niektórzy dostawcy VVA opracowują systemy, które mogłyby wykonywać część tego przetwarzania lokalnie[5].

2.4 Wyrażenia budzące

17. Aby móc korzystać z VVA, należy go „przebudzić". Oznacza to, że asystent przełącza się w tryb aktywnego nasłuchiwania, aby odbierać polecenia i komendy od swojego użytkownika. Niekiedy to przebudzenie można wywołać działaniem fizycznym (np. przyciśnięciem przycisku, przyciśnięciem inteligentnego głośnika). Prawie wszyscy VVA dostępni na rynku opierają się na wykrywaniu wyrażenia lub słowa budzącego, które przełącza ich na tryb aktywnego słuchania (znanego również jako słowo aktywujące lub słowo budzące/słowo-klucz).

18. W tym celu asystent wykorzystuje mikrofon i niewielkie możliwości obliczeniowe, aby wykryć, czy słowo kluczowe zostało wypowiedziane. Analiza ta, która trwa nieprzerwanie od momentu włączenia VVA, jest przeprowadzana wyłącznie lokalnie. Dopiero po rozpoznaniu słowa kluczowego nagrania audio są przetwarzane w celu interpretacji i wykonania polecenia, co w wielu przypadkach oznacza przesłanie ich do zdalnych serwerów przez internet. Wykrywanie słów kluczowych jest oparte na technikach uczenia maszynowego. Głównym wyzwaniem w stosowaniu takich metod jest to, że wykrywanie jest probabilistyczne. Tak więc, dla każdego wymawianego słowa lub wyrażenia, system dostarcza wynik pewności, czy dane słowo kluczowe zostało rzeczywiście wymówione. Jeśli wynik ten okaże się wyższy niż zdefiniowana wcześniej wartość progowa, uznaje się, że tak właśnie było. Taki system nie jest zatem wolny od błędów: w niektórych przypadkach aktywacja może nie zostać wykryta, mimo że słowo kluczowe zostało wypowiedziane (fałszywe odrzucenie), a w innych przypadkach aktywacja może zostać wykryta, mimo że użytkownik nie wypowiedział słowa kluczowego (fałszywa akceptacja).

19. W praktyce, w celu określenia wartości progowej, należy znaleźć akceptowalny kompromis pomiędzy tymi dwoma rodzajami błędów. Ponieważ jednak konsekwencją fałszywego wykrycia słowa kluczowego może być przesłanie nagrań audio, może dojść do nieoczekiwanych i niepożądanych transmisji danych. Bardzo często dostawcy VVA wdrażający zdalne przetwarzanie wykorzystują dwuetapowy mechanizm detekcji: pierwszy etap osadzony lokalnie na poziomie urządzenia i drugi wykonywany na zdalnych serwerach, gdzie odbywa się kolejne przetwarzanie danych. W takim przypadku programiści mają tendencję do ustawiania stosunkowo niskiego progu, aby poprawić wrażenia użytkownika i zapewnić, że gdy użytkownik wypowiada słowo kluczowe, jest ono prawie zawsze rozpoznawane – nawet jeśli oznacza to „nadmierne wykrywanie" – a następnie wdrażają drugi etap wykrywania po stronie serwera, który jest bardziej restrykcyjny.

2.5 Fragmenty treści głosowych (voice snippets) i uczenie maszynowe

20. VVA opierają się na metodach uczenia maszynowego w celu realizacji szerokiego zakresu zadań (wykrywanie słów kluczowych, automatyczne rozpoznawanie mowy, przetwarzanie języka naturalnego, synteza mowy itp.) i dlatego też potrzebują ogromnych zbiorów danych, które należy zebrać, wybrać, oznaczyć itp.

[5] Informacje na ten temat można znaleźć na przykład tutaj: https://www.amazon.science/blog/alexas-new-speech-recognition-abilities-showcased-at-interspeech

21. Zbyt wysoka lub zbyt niska reprezentacja pewnych cech statystycznych może wpłynąć na rozwój zadań opartych na uczeniu maszynowym i następnie znaleźć odzwierciedlenie w jego obliczeniach, a tym samym w sposobie funkcjonowania. Tak więc, w równym stopniu co ilość, również jakość danych odgrywa istotną rolę w wyrafinowaniu i dokładności procesu uczenia.

22. W celu podniesienia jakości VVA i udoskonalenia stosowanych metod uczenia maszynowego, projektanci VVA mogą chcieć żądać dostępu do danych związanych z użytkowaniem urządzenia w rzeczywistych warunkach – np. fragmentów treści głosowych – aby pracować nad ich udoskonaleniem.

23. Uczenie się i szkolenie systemów sztucznej inteligencji wymaga interwencji człowieka, niezależnie od tego, czy chodzi o kwalifikację uczącej się bazy danych, czy też o korektę błędów popełnionych podczas wdrażania algorytmu. Ta część pracy, znana jako praca cyfrowa, rodzi pytania dotyczące zarówno warunków pracy, jak i bezpieczeństwa. W tym kontekście media informacyjne doniosły również o transferach danych pomiędzy projektantami VVA a podwykonawcami, które rzekomo nie były objęte niezbędnymi gwarancjami ochrony prywatności.

3 ELEMENTY BEZPIECZEŃSTWA DANYCH

3.1 Ramy prawne

24. Właściwe ramy prawne UE dla VVA to przede wszystkim RODO, ponieważ przetwarzanie danych osobowych należy do podstawowych funkcji VVA. Oprócz RODO, również dyrektywa o e-prywatności[6] ustanawia określone standardy dla wszystkich podmiotów, które chcą przechowywać informacje znajdujące się na urządzeniach końcowych abonenta lub użytkownika w EOG lub uzyskiwać do nich dostęp.

25. Zgodnie z definicją „*urządzenia końcowego*”[7], smartfony, inteligentne telewizory i podobne urządzenia internetu rzeczy są przykładami urządzeń końcowych. Nawet jeśli VVA sami w sobie są usługami programowymi, zawsze działają za pośrednictwem urządzenia fizycznego, takiego jak inteligentny głośnik lub inteligentny telewizor. **VVA korzystają z sieci łączności elektronicznej, aby uzyskać dostęp do tych urządzeń fizycznych, które stanowią „urządzenia końcowe” w rozumieniu dyrektywy o e-prywatności. W związku z tym przepisy art. 5 ust. 3 dyrektywy o e-prywatności mają zastosowanie zawsze, gdy VVA przechowuje informacje lub uzyskuje do nich dostęp w podłączonym do niego urządzeniu fizycznym**[8].

[6] Dyrektywa 2002/58/WE Parlamentu Europejskiego i Rady z dnia 12 lipca 2002 r. dotycząca przetwarzania danych osobowych i ochrony prywatności w sektorze łączności elektronicznej (dyrektywa o prywatności i łączności elektronicznej), zmieniona dyrektywą 2006/24/WE i dyrektywą 2009/136/WE (zwana dalej „dyrektywą o e-prywatności”).

[7] Artykuł 1 dyrektywy Komisji 2008/63/WE z dnia 20 czerwca 2008 r. w sprawie konkurencji na rynkach końcowych urządzeń telekomunikacyjnych definiuje „*końcowe urządzenie*” jako „*a) urządzenie bezpośrednio lub pośrednio podłączone do interfejsu publicznej sieci telekomunikacyjnej w celu przesyłania, przetwarzania lub odbierania informacji; w obydwu przypadkach (połączenia bezpośredniego lub pośredniego), połączenie może być dokonane za pomocą przewodu, światłowodu lub elektromagnetycznie; połączenie jest pośrednie, jeżeli urządzenie jest umieszczone między końcowym urządzeniem telekomunikacyjnym a interfejsem sieci; b) urządzenia naziemnych stacji satelitarnych*”.

[8] Zob. wytyczne EROD nr 1/2020 pkt 12 zawierające podobne rozumowanie dotyczące podłączonych pojazdów (zwane dalej „wytycznymi EROD nr 1/2020”). Zob. również opinię EROD 5/2019 w sprawie wzajemnej zależności między dyrektywą o prywatności i łączności elektronicznej a RODO, w szczególności w zakresie właściwości, zadań i uprawnień organów ochrony danych.

26. Wszelkie operacje przetwarzania danych osobowych następujące po wyżej wymienionych operacjach przetwarzania, w tym przetwarzanie danych osobowych uzyskanych poprzez dostęp do informacji w urządzeniu końcowym, muszą również posiadać podstawę prawną na mocy art. 6 RODO, aby były zgodne z prawem[9].

27. Ponieważ administrator danych, starając się o zgodę na przechowywanie lub uzyskanie dostępu do informacji zgodnie z art. 5 ust. 3 dyrektywy o e-prywatności, będzie musiał poinformować osobę, której dane dotyczą, o wszystkich celach przetwarzania (czyli dalszym przetwarzaniu) – w tym o przetwarzaniu następującym po wyżej wymienionych operacjach – zgoda na mocy art. 6 RODO będzie zasadniczo najbardziej odpowiednią podstawą prawną do objęcia dalszego przetwarzania danych osobowych. Dlatego też zgoda będzie prawdopodobnie stanowić podstawę prawną zarówno dla przechowywania i uzyskiwania dostępu do informacji już przechowywanych, jak i przetwarzania danych osobowych w następstwie wyżej wymienionych operacji przetwarzania. Faktycznie, oceniając zgodność z art. 6 RODO, należy wziąć pod uwagę, że przetwarzanie jako całość obejmuje konkretne działania, w przypadku których prawodawca UE starał się zapewnić dodatkową ochronę[10]. Ponadto administratorzy danych muszą uwzględnić wpływ na prawa osób, których dane dotyczą, przy określaniu właściwej podstawy prawnej w celu poszanowania zasady rzetelności[11]. Wniosek jest taki, że art. 6 GDPR nie może być przywoływany przez administratorów w celu obniżenia dodatkowej ochrony przewidzianej w art. 5 ust. 3 dyrektywy o e-prywatności.

28. Jak wykazano w sekcji 2.3 (etapy 2 i 3), obecni wirtualni asystenci wymagają dostępu do danych głosowych przechowywanych przez urządzenie VVA[12]. Dlatego też zastosowanie ma art. 5 ust. 3 dyrektywy o e-prywatności. Zastosowanie art. 5 ust. 3 dyrektywy o e-prywatności oznacza, że przechowywanie informacji oraz dostęp do informacji już przechowywanych w VVA wymaga, co do zasady, uprzedniej zgody użytkownika końcowego[13], ale dopuszcza dwa wyjątki: po pierwsze, w celu wykonania lub ułatwiania transmisji komunikatu za pośrednictwem sieci łączności elektronicznej, lub po drugie, gdy jest to szczególnie niezbędne w celu dostarczania usługi społeczeństwa informacyjnego, wyraźnie zażądanej przez abonenta lub użytkownika.

29. Drugi wyjątek („gdy jest to szczególnie niezbędne w celu dostarczania usługi społeczeństwa informacyjnego, wyraźnie zażądanej przez abonenta lub użytkownika") pozwoliłby dostawcy usług VVA na przetwarzanie danych użytkowników w celu realizacji ich poleceń (zob. pkt 72 w sekcji 3.4.1) bez zgody przewidzianej w art. 5 ust. 3 dyrektywy o e-prywatności. Natomiast **zgoda taka, wymagana na mocy art. 5 ust. 3 dyrektywy o e-prywatności, byłaby konieczna** do przechowywania lub uzyskania dostępu do informacji **w celu innym niż realizacja polecenia użytkownika** (np. profilowanie użytkownika). Administratorzy danych musieliby przypisywać zgodę do konkretnych użytkowników. W związku z tym administratorzy danych powinni przetwarzać dane niezarejestrowanych użytkowników wyłącznie w celu realizacji ich poleceń.

[9] Ibidem, pkt 41.

[10] Opinia 5/2019, ust. 41.

[11] Wytyczne EROD 2/2019 w sprawie przetwarzania danych osobowych na podstawie art. 6 ust. 1 lit. b) RODO w kontekście świadczenia usług online na rzecz osób, których dane dotyczą, Wersja 2.0, 8 października 2019 r., ust. 1.

[12] Możliwe jest, że przyszłe urządzenia VVA przyjmą paradygmat przetwarzania brzegowego i będą w stanie świadczyć pewne usługi lokalnie. W takim przypadku konieczna będzie ponowna ocena możliwości zastosowania dyrektywy o e-prywatności.

[13] Zob. również wytyczne EROD 1/2020, pkt 14.

30. VVA mogą przypadkowo przechwytywać dźwięk osób, które nie zamierzały korzystać z usługi VVA. Po pierwsze, do pewnego stopnia i w zależności od VVA, można zmienić wyrażenie budzące. Osoby, które nie są świadome tej zmiany, mogą przypadkowo użyć zaktualizowanego wyrażenia budzącego. Po drugie, VVA mogą wykryć wyrażenie budzące przez pomyłkę lub w wyniku błędu. Jest bardzo mało prawdopodobne, aby którykolwiek z wyjątków przewidzianych w art. 5 ust. 3 dyrektywy o e-prywatności miał zastosowanie w sytuacji przypadkowej aktywacji. Ponadto zgoda zdefiniowana w RODO musi być jednoznacznym okazaniem woli osoby, której dane dotyczą. Dlatego jest bardzo mało prawdopodobne, aby przypadkowa aktywacja mogła być interpretowana jako ważna zgoda. Jeśli administratorzy danych dowiedzą się (np. poprzez przegląd automatyczny lub przegląd dokonany przez człowieka), że usługa VVA przypadkowo przetworzyła dane osobowe, powinni sprawdzić, czy istnieje ważna podstawa prawna dla każdego celu przetwarzania takich danych. W przeciwnym wypadku przypadkowo zebrane dane powinny zostać usunięte.

31. Ponadto należy zauważyć, że dane osobowe przetwarzane przez VVA mogą mieć bardzo wrażliwy charakter. Mogą one zawierać dane osobowe zarówno w treści (znaczenie wypowiadanego tekstu), jak i w metadanych (płeć lub wiek mówiącego itp.). EROD przypomina, że dane głosowe są z natury biometrycznymi danymi osobowymi[14]. W związku z tym, gdy takie dane są przetwarzane w celu jednoznacznej identyfikacji osoby fizycznej lub są z natury rzeczy lub w sposób oczywisty danymi osobowymi szczególnej kategorii, przetwarzanie musi mieć ważną podstawę prawną w art. 6 i musi mu towarzyszyć odstępstwo od art. 9 RODO (zob. sekcja 3.7 poniżej).

3.2 Identyfikacja przetwarzania danych i zainteresowanych stron

32. Biorąc pod uwagę wielorakie możliwości usług, jakie VVA może świadczyć w tak wielu różnych środowiskach codziennego życia osoby, której dane dotyczą[15], warto zauważyć, że należy dokładnie rozważyć przetwarzanie danych osobowych, na które mogą mieć wpływ różne zainteresowane strony.

3.2.1 Przetwarzanie danych osobowych

33. Z punktu widzenia ochrony danych osobowych można zaobserwować kilka stałych elementów, niezależnie od rodzaju VVA (tj. rodzaju urządzenia, funkcji, usług lub kombinacji tych czynników), z których może korzystać osoba, której dane dotyczą. Stałe te odnoszą się do mnogości danych osobowych, osób, których dane dotyczą, i przetwarzania danych, o którym mowa.

Mnogość rodzajów danych osobowych

34. Definicja danych osobowych na mocy art. 4 ust. 1 RODO obejmuje szeroki zakres różnych danych i ma zastosowanie w kontekście neutralnym technologicznie do wszelkich informacji,

[14] Na mocy art. 4 ust. 14 RODO „dane biometryczne" oznaczają „dane osobowe, które wynikają ze specjalnego przetwarzania technicznego, dotyczą cech fizycznych, fizjologicznych lub behawioralnych osoby fizycznej oraz umożliwiają lub potwierdzają jednoznaczną identyfikację tej osoby fizycznej, takie jak wizerunek twarzy lub dane daktyloskopijne".

[15] Na przykład: w domu, w pojeździe, na ulicy, w pracy lub w innych prywatnych, publicznych lub zawodowych przestrzeniach lub w kombinacji tych przestrzeni.

które odnoszą się do *„zidentyfikowanej lub możliwej do zidentyfikowania osoby fizycznej”*[16]. Każda interakcja osoby, której dane dotyczą, z VVA może wchodzić w zakres tej definicji. Po wejściu w interakcję, w trakcie działania VVA można przetwarzać różne rodzaje danych osobowych, jak opisano w sekcji 2.4.

35. Od pierwszego polecenia do związanej z nim odpowiedzi, działania lub działań następczych (np. skonfigurowanie cotygodniowego powiadomienia), pierwsze wprowadzenie danych osobowych wygeneruje zatem kolejne dane osobowe. Obejmuje to dane pierwotne (np. dane konta, nagrania głosowe, historię poleceń), dane obserwowane (np. dane urządzenia, które odnoszą się do osoby, której dotyczą dane, dzienniki aktywności, działania internetowe), a także dane wywnioskowane lub pochodne (np. profilowanie użytkownika). VVA wykorzystują mowę, aby pośredniczyć między użytkownikami a wszystkimi połączonymi usługami (np. wyszukiwarką, sklepem internetowym lub serwisem transmisji strumieniowej muzyki), ale w przeciwieństwie do innych pośredników VVA mogą mieć pełny dostęp do treści poleceń, a w konsekwencji dostarczać projektantowi VVA wiele różnych danych osobowych w zależności od celów przetwarzania.

36. Mnogość danych osobowych przetwarzanych podczas korzystania z VVA odnosi się również do mnogości kategorii danych osobowych, na które należy zwrócić uwagę (zob. poniżej sekcja 3.7). EROD przypomina, że w przypadku przetwarzania szczególnych kategorii danych[17] art. 9 RODO wymaga od administratora danych wskazania ważnego zwolnienia z zakazu przetwarzania na mocy art. 9 ust. 1 oraz ważnej podstawy prawnej na mocy art. 6 ust. 1, przy użyciu odpowiednich środków określonych na mocy art. 9 ust. 2. Wyraźna zgoda może być jednym z właściwych odstępstw, jeżeli zgoda stanowi podstawę prawną na mocy art. 6 ust. 1. W art. 9 zwraca się również (szczegółowo) uwagę, że państwa członkowskie mogą wprowadzić dalsze warunki przetwarzania danych biometrycznych lub innych szczególnych kategorii.

Mnogość osób, których dane dotyczą

37. Podczas korzystania z VVA dane osobowe są przetwarzane od momentu pierwszej interakcji z VVA. Dla niektórych osób, których dane dotyczą, oznacza to zakup VVA lub konfigurację konta użytkownika (np. zarejestrowani użytkownicy). W przypadku innych osób, których dane dotyczą, odnosi się to do pierwszej świadomej interakcji z VVA innej osoby, której dane dotyczą, która zakupiła lub skonfigurowała VVA (tj. niezarejestrowani użytkownicy). Oprócz tych dwóch kategorii osób, których dane dotyczą, istnieje jeszcze trzecia: przypadkowi użytkownicy, którzy, zarejestrowani lub nie, nieświadomie kierują polecenia do VVA (np. wypowiadają prawidłowe wyrażenie budzące, nie wiedząc, że VVA jest aktywny, lub wypowiadają inne słowa, które VVA błędnie identyfikuje jako wyrażenie budzące).

38. Termin „mnogość osób, których dane dotyczą” odnosi się również do wielu użytkowników jednego VVA (np. urządzenie współdzielone przez zarejestrowanych i niezarejestrowanych

[16] Ponadto na mocy art. 4 ust. 1 RODO „*możliwa do zidentyfikowania osoba fizyczna to osoba, którą można bezpośrednio lub pośrednio zidentyfikować, w szczególności na podstawie identyfikatora takiego jak imię i nazwisko, numer identyfikacyjny, dane o lokalizacji, identyfikator internetowy lub jeden bądź kilka szczególnych czynników określających fizyczną, fizjologiczną, genetyczną, psychiczną, ekonomiczną, kulturową lub społeczną tożsamość osoby fizycznej*”.

[17] W art. 9 ust. 1 RODO zdefiniowano „szczególne kategorie danych osobowych”, a na mocy tego artykułu „*zabrania się przetwarzania danych osobowych ujawniających pochodzenie rasowe lub etniczne, poglądy polityczne, przekonania religijne lub światopoglądowe, przynależność do związków zawodowych oraz przetwarzania danych genetycznych, danych biometrycznych w celu jednoznacznego zidentyfikowania osoby fizycznej lub danych dotyczących zdrowia, seksualności lub orientacji seksualnej tej osoby*”.

użytkowników, przez współpracowników, w rodzinie, w szkole) oraz różnych rodzajów użytkowników w zależności od ich stanu (np. dorośli, dzieci, osoby starsze lub z niepełnosprawnościami). Choć VVA może ułatwić interakcję z narzędziem cyfrowym i przynieść wiele korzyści niektórym kategoriom osób, których dane dotyczą, ważne jest, by wziąć pod uwagę specyfikę każdej kategorii osób, których dane dotyczą, oraz kontekst korzystania z VVA.

Mnogość procesów przetwarzania danych

39. Technologie wykorzystywane do dostarczania VVA mają również wpływ na ilość przetwarzanych danych i rodzaje przetwarzania. Im więcej VVA oferuje usług lub funkcji i im bardziej jest połączony z innymi urządzeniami lub usługami zarządzanymi przez inne strony, tym bardziej wzrasta ilość przetwarzanych danych osobowych i przetwarzania wtórnego. Powoduje to mnogość procesów przetwarzania wykonywanych w sposób zautomatyzowany, jak opisano w sekcji 2. Oprócz środków zautomatyzowanych, w niektórych procesach przetwarzania mogą brać udział również ludzie. Dzieje się tak na przykład w przypadku, gdy wdrożona technologia wymaga interwencji człowieka, np. przegląd transkrypcji głosów na teksty lub dostarczanie adnotacji na danych osobowych, które mogą być wykorzystane do wprowadzenia nowych modeli w technologii uczenia maszynowego. Dzieje się tak również w przypadku, gdy człowiek analizuje dane osobowe (np. metadane) w celu poprawy usług świadczonych przez VVA.

3.2.2 Przetwarzanie danych przez administratorów danych i podmioty przetwarzające

40. Osoby, których dane dotyczą, powinny być w stanie zrozumieć i zidentyfikować role, jakie są zaangażowane w przetwarzanie danych, oraz powinny być w stanie skontaktować się z każdą z zainteresowanych stron lub podjąć wobec nich działania zgodnie z wymogami RODO. Podział ról nie powinien odbywać się ze szkodą dla osób, których dane dotyczą, nawet jeśli scenariusze mogą być skomplikowane lub podlegać zmianom. W celu oceny swoich ról zainteresowane strony odsyła się do wytycznych EROD nr 7/2020 dotyczących pojęć administratora danych i podmiotu przetwarzającego dane w RODO[18].

41. Jak wskazano w ust. 15, główne zainteresowane strony można zidentyfikować w ramach roli dostawcy lub projektanta, programisty aplikacji, integratora, właściciela lub kombinacji tych ról. Możliwe są różne scenariusze, w zależności od tego, kto co robi w ramach relacji biznesowych zainteresowanych stron, od polecenia użytkownika, danych osobowych, działań związanych z przetwarzaniem danych i ich celów. Powinny one wyraźnie zadecydować i poinformować osoby, których dane dotyczą, o warunkach, na jakich każda z nich będzie działać i spełniać wynikające z tego role administratorów danych, współadministratorów danych lub podmiotów przetwarzających dane, zgodnie z RODO[19]. Każdy z nich może pełnić jedną lub kilka ról, ponieważ może być jedynym administratorem danych, współadministratorem lub przetwarzającym dane w odniesieniu do jednego procesu przetwarzania danych, podczas gdy pełni inną rolę w odniesieniu do innego procesu przetwarzania danych.

42. Z perspektywy wysokiego szczebla projektant może działać jako administrator danych przy określaniu celów i środków przetwarzania, ale może interweniować jako przetwarzający dane przy przetwarzaniu danych osobowych w imieniu innych stron, takich jak programista aplikacji. Użytkownik VVA podlegałby zatem kilku administratorom danych: programiście aplikacji i

[18] Wytyczne EROD nr 07/2020 w sprawie pojęć administratora danych i podmiotu przetwarzającego w RODO, V2.0, przyjęte w dniu 7 lipca 2021 r. (zwane dalej „wytycznymi nr 7/2020").
[19] Artykuły 12-14, artykuł 26 RODO.

projektantowi. Możliwe jest również, że projektant, integrator i programista będą zgrupowani w jednym podmiocie działającym jako jedyny administrator danych. W każdym przypadku stosowne kwalifikacje należy ustalić na podstawie analizy poszczególnych przypadków.

Przykład 1:

Projektant VVA przetwarza dane użytkownika w wielu celach, w tym w celu poprawy umiejętności rozumienia głosu przez VVA i dokładnego odpowiadania na polecenia. Dlatego i pomimo tego, że cel ten może prowadzić do przetwarzania danych wynikających z korzystania z aplikacji dostarczanych przez osoby trzecie, istnieje tylko jeden administrator danych: projektant VVA, w którego imieniu i dla którego celów odbywa się przetwarzanie.

Przykład 2:

Bank oferuje swoim klientom aplikację, do której można się zwracać bezpośrednio za pośrednictwem VVA w celu zarządzania ich rachunkami.

W przetwarzanie danych osobowych zaangażowane są dwa podmioty: projektant VVA i programista aplikacji bankowej.

W przedstawionym scenariuszu bank jest administratorem danych w zakresie świadczenia usługi, ponieważ określa cele i podstawowe środki przetwarzania związane z aplikacją umożliwiającą interakcję z asystentem. W istocie oferuje on specjalną aplikację, która pozwala użytkownikowi, klientowi banku, na zdalne zarządzanie swoimi rachunkami. Ponadto decyduje o sposobach przetwarzania poprzez wybór odpowiedniego podmiotu przetwarzającego, który jest projektantem VVA i może odegrać ważną rolę, wspomagając swoją wiedzą fachową określanie tych sposobów (na przykład może obsługiwać platformę rozwojową, która umożliwia integrację aplikacji osób trzecich z VVA, a zatem określa ramy i warunki, których muszą przestrzegać twórcy aplikacji).

43. Jeśli chodzi o stronę osoby, której dane dotyczą, warto zauważyć, że kilka zainteresowanych stron może przetwarzać te same dane osobowe, nawet jeśli osoba, której dane dotyczą, nie spodziewa się, że w łańcuch przetwarzania zaangażowane są inne strony niż dostawca VVA. Zatem gdy osoba, której dane dotyczą, podejmuje działania wobec dostawcy VVA w odniesieniu do swoich danych osobowych (np. wykonywanie praw osoby, której dane dotyczą), nie oznacza to automatycznie, że działania te będą miały zastosowanie do tych samych danych osobowych, które są przetwarzane przez inną zainteresowaną stronę. Gdy te zainteresowane strony są niezależnymi administratorami danych, ważne jest, by osoby, których dane dotyczą, otrzymały jasną informację wyjaśniającą poszczególne etapy procesu przetwarzania i jego uczestników. Ponadto w przypadku wspólnego administrowania danymi należy wyjaśnić, czy każdy administrator jest właściwy do przestrzegania wszystkich praw osoby, której dane dotyczą, lub też który administrator jest właściwy w odniesieniu do którego prawa[20].

[20] Wytyczne nr 7/2020, ust. 165.

Przykład 3:

W tym scenariuszu projektant VVA chce wykorzystać dane zgromadzone i przetworzone na potrzeby usługi świadczonej przez bank w celu ulepszenia jego systemu rozpoznawania mowy. Projektant VVA, który przetwarza dane do własnych celów, będzie miał status administratora danych w odniesieniu do tego konkretnego procesu przetwarzania.

44. Ponieważ w łańcuch przetwarzania danych może być zaangażowanych wiele zainteresowanych stron i odpowiednio wielu pracowników, może dojść do ryzykownych sytuacji, jeśli nie zostaną zastosowane odpowiednie środki i zabezpieczenia. Administratorzy ponoszą za to odpowiedzialność i dlatego powinni skupić się na ochronie danych osobowych, w szczególności wybierając odpowiednich partnerów biznesowych i podmioty przetwarzające dane, stosując zasady domyślnej ochrony prywatności i ochrony prywatności już w fazie projektowania[21], wdrażając odpowiednie zabezpieczenia i inne narzędzia RODO, takie jak audyty i umowy prawne (np. art. 26 RODO w odniesieniu do współadministratorów lub art. 28 RODO w odniesieniu do podmiotów przetwarzających).

45. Ekosystem VVA jest ekosystemem złożonym, w którym potencjalnie wiele podmiotów może wymieniać i przetwarzać dane osobowe jako administratorzy danych lub podmioty przetwarzające. Niezwykle ważne jest wyjaśnienie roli każdego podmiotu w odniesieniu do każdego procesu przetwarzania oraz przestrzeganie zasady minimalizacji danych również w odniesieniu do wymiany danych.

46. Ponadto administratorzy danych powinni zachować czujność przy przekazywaniu danych osobowych i zagwarantować wymagany poziom ochrony w całym łańcuchu przetwarzania, w szczególności gdy korzystają z usług zlokalizowanych poza EOG.

3.3 Przejrzystość

47. Ponieważ VVA przetwarzają dane osobowe (np. głos użytkowników, ich lokalizację lub treść komunikatu), muszą spełniać wymogi RODO dotyczące przejrzystości, uregulowane w art. 5 ust. 1 lit. a) oraz art. 12 i art. 13 (uściślone w motywie 58). Administratorzy danych są zobowiązani do informowania użytkowników o przetwarzaniu ich danych osobowych w zwięzłej, przejrzystej, zrozumiałej formie i w łatwo dostępny sposób.

48. Nieprzekazanie niezbędnych informacji stanowi naruszenie obowiązków, które może mieć wpływ na legalność przetwarzania danych. Przestrzeganie wymogu przejrzystości jest koniecznością, ponieważ służy jako mechanizm kontroli nad przetwarzaniem danych i umożliwia użytkownikom korzystanie z ich praw. Właściwe informowanie użytkowników o tym, jak wykorzystywane są ich dane osobowe, utrudnia administratorom danych nadużywanie VVA do celów, które wykraczają daleko poza oczekiwania użytkowników. Na przykład opatentowane technologie mają na celu wnioskowanie o stanie zdrowia i stanach emocjonalnych na podstawie głosu użytkownika i odpowiednie dostosowanie świadczonych usług.

49. Spełnienie wymogów przejrzystości może być szczególnie trudne dla dostawcy usług VVA lub jakiegokolwiek innego podmiotu działającego w charakterze administratora danych. Biorąc

[21] Zob. Wytyczne EROD nr 4/2019 w sprawie domyślnej ochrony danych i ochrony danych już w fazie projektowania na mocy art. 25, Wersja 2.0 przyjęta w dniu 20 października 2020 r.

pod uwagę specyficzny charakter VVA, administratorzy danych napotykają kilka przeszkód, aby spełnić wymogi RODO dotyczące przejrzystości:

- **Wielu użytkowników:** administratorzy danych powinni informować wszystkich użytkowników (zarejestrowanych, niezarejestrowanych i przypadkowych), a nie tylko użytkownika korzystającego z VVA.
- **Złożoność ekosystemu:** jak wyjaśniono w sekcji poświęconej podstawom technologicznym, tożsamość i rola osób przetwarzających dane osobowe podczas korzystania z VVA nie jest dla użytkowników oczywista.
- **Specyfika interfejsu głosowego:** systemy cyfrowe nie są jeszcze przystosowane do interakcji wyłącznie głosowych, czego dowodzi niemal powszechne stosowanie ekranu uzupełniającego. Jednak dostosowanie się do interfejsu głosowego i możliwość jasnego i poprawnego informowania użytkownika za jego pomocą jest koniecznością.

50. VVA mogą być traktowani jako automaty skończone przechodzące przez szereg stanów podczas ich zwykłego funkcjonowania. Mogą nasłuchiwać lokalnie w celu wykrycia wyrażeń budzących lub wchodzić w interakcję ze zdalnym serwerem w celu wypełnienia polecenia, ale mogą przyjmować wiele innych stanów w zależności od kontekstu (np. jeśli w tle jest dźwięk otoczenia) lub użytkownika, który do nich mówi (np. mogą mówić do zidentyfikowanego lub nieznanego użytkownika). Niestety, sytuacje te odbywają się przy znacznej asymetrii informacji w stosunku do użytkownika, który praktycznie nie wie, czy urządzenie słucha, a tym bardziej w jakim stanie się znajduje.

51. Zalecane jest, aby projektanci i programiści VVA podjęli odpowiednie kroki w celu zrównoważenia tej asymetrii, czyniąc funkcjonowanie VVA bardziej interaktywnym. Użytkownicy powinni być informowani o statusie, w jakim aktualnie znajduje się urządzenie. To zwiększenie przejrzystości można osiągnąć zarówno poprzez uczynienie dialogu człowiek-maszyna bardziej interaktywnym (np. urządzenie może w jakiś sposób potwierdzać odbiór polecenia głosowego), jak i poprzez przekazywanie informacji o statusie maszyny za pomocą określonych sygnałów. Istnieje wiele opcji, które można zbadać w tym względzie, począwszy od stosowania określonych potwierdzeń głosowych i widocznych ikon lub diod, lub stosowania wyświetlaczy na urządzeniu.

52. Kwestie te są szczególnie istotne ze względu na mnogość użytkowników i obecność wśród nich kategorii osób szczególnie narażonych na zagrożenia, takich jak dzieci, osoby starsze lub użytkownicy z niepełnosprawnością audiowizualną.

53. Z powyższych kwestii wyłaniają się dwa ważne pytania: jaki jest najbardziej realny sposób informowania użytkowników i kiedy jest odpowiedni moment na informowanie ich? Kwestie te należy dokładniej zbadać w dwóch różnych sytuacjach, w zależności od tego, czy VVA ma tylko jednego użytkownika (np. osobisty smartfon), czy potencjalnie wielu użytkowników (np. inteligentne urządzenie domowe). Przy zastosowaniu technologii VVA może również dojść do odwrócenia tych dwóch podstawowych ustawień, np. gdy użytkownik posiada własny smartfon i podłącza go w samochodzie. VVA tego smartfonu, który powinien być używany tylko przez tego użytkownika, jest teraz rozszerzony na pozostałe osoby w samochodzie.

54. Obecnie wszyscy wirtualni asystenci głosowi są powiązani z kontem użytkownika lub są tworzeni przez aplikację, która tego wymaga. Kwestia określenia, w jaki sposób administratorzy danych mogliby rozważyć poinformowanie tych użytkowników o polityce prywatności podczas konfigurowania VVA, powinna zostać rozwiązana w sposób opisany w wytycznych Grupy Roboczej Art. 29 dotyczących przejrzystości. Aplikacje powinny udostępniać

niezbędne informacje w sklepie internetowym przed pobraniem[22]. W ten sposób informacje są podawane jak najwcześniej to możliwe, a najpóźniej w momencie pozyskiwania danych osobowych. Niektórzy dostawcy VVA włączają aplikacje innych firm do domyślnej konfiguracji VVA, dzięki czemu mogą one uruchamiać te aplikacje za pomocą określonych wyrażeń budzących. VVA stosujący tę strategię wdrażania aplikacji osób trzecich powinni zadbać o to, aby użytkownicy otrzymywali niezbędne informacje również na temat przetwarzania danych przez osoby trzecie.

55. Wielu projektantów VVA wymaga jednak kont użytkowników VVA, które łączą usługę VVA z wieloma innymi usługami, takimi jak poczta elektroniczna, strumieniowa transmisja wideo lub zakupy, a to zaledwie kilka z nich. Decyzja projektanta VVA o powiązaniu konta z wieloma różnymi usługami skutkuje koniecznością stosowania bardzo długich i złożonych polityk prywatności. Długość i złożoność takich polityk prywatności znacznie utrudnia spełnienie zasady przejrzystości.

> Przykład 4:
>
> Projektant VVA wymaga od swoich użytkowników posiadania konta, aby uzyskać dostęp do usługi VVA. To konto użytkownika nie jest specyficzne dla usługi VVA i może być używane do innych usług oferowanych przez projektanta VVA, takich jak e-mail, przechowywanie w chmurze i media społecznościowe. Aby utworzyć konto, użytkownicy muszą przeczytać i zaakceptować 30-stronicową politykę prywatności. Polityka ta zawiera informacje na temat przetwarzania danych osobowych przez wszystkie usługi, które mogą być powiązane z kontem.
>
> Informacji przekazanych przez projektanta VVA w tym przypadku nie można uznać za zwięzłe, a ich złożoność zmniejsza wymaganą przejrzystość. W związku z tym projektant VVA nie spełniałby wymogów przejrzystości określonych w art. 12 i 13 RODO.

56. Chociaż najbardziej powszechnym sposobem przekazywania niezbędnych informacji jest forma pisemna, RODO dopuszcza również „inne środki". W motywie 58 wyraźnie stwierdzono, że informacje mogą być przekazywane w formie elektronicznej, np. za pośrednictwem strony internetowej. Ponadto przy wyborze odpowiedniej metody informowania osób, których dane dotyczą, należy uwzględnić szczególne okoliczności, takie jak sposób, w jaki administrator danych i osoba, której dane dotyczą, kontaktują się ze sobą[23]. Opcją dla urządzeń bez ekranu mogłoby być podanie łatwego do zrozumienia linku, bezpośrednio lub w wiadomości e-mail. Jako przykład informacji mogłyby służyć już istniejące rozwiązania, np. praktyki centrów telefonicznej obsługi klienta polegające na informowaniu rozmówcy o tym, że jego rozmowa jest nagrywana i kierowaniu go do ich polityki prywatności. Ograniczenia VVA bez ekranów nie zwalniają administratora danych z obowiązku zapewnienia niezbędnych informacji zgodnie z RODO podczas konfigurowania VVA lub instalowania bądź korzystania z aplikacji VVA. Dostawcy i projektanci VVA powinni opracować interfejsy głosowe, aby ułatwić przekazywanie obowiązkowych informacji.

57. VVA mogą być bardzo interesujący dla użytkowników z dysfunkcją wzroku, ponieważ zapewniają alternatywne środki interakcji z usługami informatycznymi, które tradycyjnie opierają się na informacjach wizualnych. Zgodnie z art. 12 ust. 1 RODO udzielanie niezbędnych

[22] Wytyczne dotyczące przejrzystości na mocy rozporządzenia 2016/679, WP260, Wersja 01, zatwierdzone przez EROD (zwane dalej „Wytycznymi WP29 WP260"), ust. 11.

[23] Grupa Robocza Art. 29, Wytyczne WP260, pkt. 19.

informacji ustnie jest możliwe wyłącznie na żądanie osoby, której dane dotyczą, ale nie jako metoda domyślna. Ograniczenia związane z VVA bez ekranów wymagałyby jednak zautomatyzowanych środków ustnego przekazywania informacji, które mogłyby być uzupełnione środkami pisemnymi. Wykorzystując dźwięk do informowania osób, których dane dotyczą, administratorzy danych powinni przekazywać niezbędne informacje w sposób zwięzły i jasny. Ponadto osoby, których dane dotyczą, powinny mieć możliwość ponownego wysłuchania informacji[24].

58. Podjęcie odpowiednich środków w celu spełnienia wymogów RODO w zakresie przejrzystości jest bardziej złożone, gdy istnieje wielu użytkowników VVA innych niż właściciel urządzenia. Projektanci VVA muszą rozważyć, w jaki sposób odpowiednio informować niezarejestrowanych i przypadkowych użytkowników, kiedy ich dane osobowe są przetwarzane. Jeżeli podstawą prawną przetwarzania danych użytkowników jest zgoda, użytkownicy muszą być odpowiednio poinformowani, aby zgoda była ważna[25].

59. Aby spełnić wymogi RODO, administratorzy danych powinni znaleźć sposób na informowanie nie tylko zarejestrowanych użytkowników, ale także użytkowników niezarejestrowanych i przypadkowych użytkowników VVA. Użytkownicy ci powinni być informowani najwcześniej, jak to możliwe, **a najpóźniej w momencie przetwarzania danych.** Ten warunek może być szczególnie trudny do spełnienia w praktyce.

60. Pewne cechy szczególne przedsiębiorstwa również nie powinny być szkodliwe dla osób, których dane dotyczą. Ponieważ wiele zainteresowanych stron to przedsiębiorstwa globalne lub dobrze znane z konkretnej działalności gospodarczej (np. telekomunikacja, handel elektroniczny, technologie informacyjne, działalność internetowa), sposób, w jaki świadczą usługę VVA, powinien być jasny. Odpowiednie informacje powinny pozwolić osobom, których dane dotyczą, zrozumieć, czy korzystanie przez nie z VVA będzie powiązane z innymi czynnościami przetwarzania danych zarządzanymi przez dostawcę usług VVA (np. telekomunikacja, handel elektroniczny, technologie informacyjne lub działalność internetowa), poza ścisłym korzystaniem z VVA.

> Przykład 5:
>
> Aby korzystać ze jego asystenta, projektant VVA, który udostępnia również platformę mediów społecznościowych i wyszukiwarkę, wymaga od użytkownika powiązania jego konta z asystentem. Łącząc swoje konto z korzystaniem z VVA, projektant może w ten sposób wzmocnić profil swoich użytkowników poprzez możliwość korzystania z asystenta, zainstalowane aplikacje (lub umiejętności), złożone zamówienia itp. Interakcje z asystentem stanowią zatem nowe źródło informacji związanych z użytkownikiem. Projektant VVA powinien zapewnić użytkownikom jasne informacje o tym, w jaki sposób ich dane będą przetwarzane w przypadku każdej usługi oraz kontrole pozwalające użytkownikowi na wybór, czy dane będą wykorzystywane do profilowania, czy nie.

Zalecenia

61. Jeżeli użytkownicy są informowani o przetwarzaniu danych osobowych przez VVA za pomocą polityki prywatności konta użytkownika, a konto jest powiązane z innymi niezależnymi usługami (np. pocztą elektroniczną lub zakupami internetowymi), EROD zaleca, aby polityka

[24] Grupa Robocza Art. 29, Wytyczne WP260, pkt. 21.
[25] Art. 4 ust. 11 RODO.

prywatności zawierała wyraźnie oddzieloną sekcję dotyczącą przetwarzania danych osobowych przez VVA.

62. Informacje przekazywane użytkownikowi powinny odpowiadać dokładnemu procesowi zbierania i przetwarzania danych, który jest przeprowadzany. Chociaż próbka głosu zawiera pewne metadane (np. o poziomie stresu mówiącego), nie jest automatycznie jasne, czy taka analiza jest przeprowadzana. Niezwykle istotne jest, aby administratorzy zapewnili przejrzystość co do tego, jakie konkretne aspekty surowych danych przetwarzają.

63. Ponadto w każdym momencie powinno być oczywiste, w jakim stanie znajduje się VVA. Użytkownicy powinni być w stanie stwierdzić, czy VVA aktualnie nasłuchuje w swoim zamkniętym obiegu, a zwłaszcza czy przesyła informacje do swojego procesora końcowego. Informacje te powinny być również dostępne dla osób z niepełnosprawnościami, takimi jak ślepota na kolory (daltonizm), głuchota. Szczególną uwagę należy zwrócić na fakt, że VVA sugerują scenariusz użytkowania, w którym kontakt wzrokowy z urządzeniem nie jest konieczny. Dlatego wszystkie informacje zwrotne dla użytkownika, w tym zmiany stanu, powinny być dostępne przynajmniej w formie wizualnej i akustycznej.

64. Szczególną uwagę należy zwrócić na to, czy urządzenia pozwalają na dodawanie funkcjonalności innych firm (aplikacje do VVA). O ile pewne ogólne informacje można przekazać użytkownikowi, gdy to on dodaje taką funkcję (biorąc pod uwagę, że jest to jego wybór), o tyle podczas normalnego użytkowania urządzenia granice między różnymi zaangażowanymi administratorami danych mogą być znacznie mniej wyraźne, tj. użytkownik może nie być wystarczająco poinformowany, jak i przez kogo jego dane są przetwarzane (i w jakim zakresie) w konkretnym zapytaniu.

65. Wszystkie informacje o przetwarzaniu na podstawie danych zebranych i pochodzących z przetwarzania nagranego głosu powinny być również dostępne dla użytkowników zgodnie z art. 12 RODO.

66. Administratorzy VVA powinni w przejrzysty sposób informować o tym, jakiego rodzaju informacje VVA może uzyskać o swoim otoczeniu, np. o innych osobach w pomieszczeniu, muzyce grającej w tle, przetwarzaniu głosu z powodów medycznych lub marketingowych, zwierzętach domowych.

3.4 Ograniczenie celu i podstawa prawna

67. Przetwarzanie poleceń głosowych przez VVA ma oczywisty cel, jakim jest realizacja polecenia. Często jednak istnieją dodatkowe cele, które nie są tak oczywiste, jak np. poprawa zdolności VVA do rozumienia języka naturalnego poprzez szkolenie modelu VVA z wykorzystaniem technik uczenia maszynowego. Wśród najczęstszych celów przetwarzania danych osobowych przez VVA można wymienić:

- realizację poleceń użytkownika;
- ulepszanie VVA poprzez szkolenie modelu uczenia maszynowego oraz przeglądanie i oznaczanie transkrypcji głosowych przez człowieka;
- identyfikację użytkownika (z wykorzystaniem danych głosowych);
- profilowanie użytkowników w celu spersonalizowania treści lub reklam.

68. Ze względu na swoją rolę pośrednika i sposób, w jaki zostały zaprojektowane, VVA przetwarzają wiele różnych danych osobowych i nieosobowych. Pozwala to na przetwarzanie danych osobowych w wielu celach, które wykraczają poza odpowiadanie na polecenia

użytkowników i które mogłyby pozostać całkowicie niezauważone. Analizując dane zgromadzone za pośrednictwem VVA, można poznać lub wywnioskować zainteresowania, harmonogram dnia, trasy przejazdu lub zwyczaje użytkowników. Może to umożliwić przetwarzanie danych osobowych w nieprzewidzianych celach (np. analiza nastrojów lub ocena stanu zdrowia[26]), które znacznie wykraczałyby poza racjonalne oczekiwania użytkowników.

69. Administratorzy danych powinni jasno określić swoje cele w odniesieniu do kontekstu, w którym VVA jest wykorzystywany, tak aby były one wyraźnie zrozumiałe dla osób, których dane dotyczą (np. przedstawienie celów w kategoriach). Zgodnie z art. 5 ust. 1 RODO dane osobowe powinny być zbierane w konkretnych, wyraźnych i prawnie uzasadnionych celach i nieprzetwarzane dalej w sposób niezgodny z tymi celami.

3.4.1 Realizacja poleceń użytkownika

70. Głównym zastosowaniem VVA jest realizacja poleceń głosowych, które muszą zostać wykonane przez VVA lub powiązaną aplikację lub usługę (np. usługę strumieniowego przesyłania muzyki, usługę mapowania lub zamek elektroniczny). Głos użytkownika i ewentualnie inne dane (np. pozycja użytkownika w przypadku polecenia wyznaczenia trasy do określonego celu) mogą być zatem przetwarzane.

> Przykład 6:
>
> Pasażer inteligentnego samochodu wyposażonego w VVA wydaje polecenie wyznaczenia trasy do najbliższej stacji benzynowej. VVA przetwarza głos użytkownika, aby zrozumieć polecenie, oraz pozycję samochodu, aby znaleźć trasę i przesyła ją do inteligentnego komponentu, aby wyświetlić ją na ekranie samochodu.

71. W zakresie, w jakim przetwarzanie poleceń głosowych wiąże się z przechowywaniem lub dostępem do informacji przechowywanych w urządzeniach końcowych użytkownika końcowego, należy przestrzegać art. 5 ust. 3 dyrektywy o e-prywatności. Chociaż art. 5 ust. 3 zawiera ogólną zasadę, że takie przechowywanie lub dostęp wymaga uprzedniej zgody użytkownika końcowego, przewiduje również zwolnienie z wymogu uzyskania zgody, gdy „jest to szczególnie niezbędne w celu dostarczania usługi społeczeństwa informacyjnego, wyraźnie zażądanej przez abonenta lub użytkownika". O ile dane głosowe są przetwarzane w celu realizacji poleceń użytkownika, są one zwolnione z wymogu uzyskania uprzedniej zgody.

72. Jak wskazano wcześniej, wszelkie operacje przetwarzania danych osobowych następujące po przechowywaniu lub uzyskaniu dostępu do informacji w urządzeniu końcowym użytkowników końcowych muszą mieć podstawę prawną na mocy art. 6 RODO, aby były zgodne z prawem.

73. Na VVA odbywają się dwie kolejne operacje przetwarzania. Jak wspomniano powyżej, pierwsza z nich wymaga dostępu do VVA (a zatem muszą być spełnione warunki art. 5 ust. 3 dyrektywy o e-prywatności). Oprócz warunkówokreślonych w art. 5 ust. 3 dyrektywy o e-prywatności, ten drugi krok wymaga również podstawy prawnej na mocy art. 6 RODO.

74. Gdy osoba fizyczna podejmuje decyzję o korzystaniu z VVA, działanie to zasadniczo oznacza, że użytkownik początkowy musi najpierw zarejestrować konto, aby aktywować VVA. Inaczej

[26] Eoghan Furey, Juanita Blue, „Alexa, Emotion, Privacy and GDPR" („Alexa, emocje, prywatność i RODO"), referat z konferencji na temat interakcji człowiek-komputer „Human Computer Interaction Conference", lipiec 2018 r.

rzecz ujmując, sytuacja ta odnosi się do stosunku umownego[27] między zarejestrowanym użytkownikiem a administratorem VVA. Z uwagi na istotę i podstawowy cel umowy, głównym celem tej umowy jest korzystanie z VVA w celu realizacji polecenia użytkownika.

75. Wszelkie przetwarzanie danych osobowych, które jest niezbędne do wykonania polecenia użytkownika, może zatem opierać się na podstawie prawnej wykonania umowy[28]. Takie przetwarzanie obejmuje w szczególności przechwycenie polecenia głosowego użytkownika, jego transkrypcję na tekst, jego interpretację, informacje wymieniane ze źródłami wiedzy w celu przygotowania odpowiedzi, a następnie transkrypcję na wokalną odpowiedź końcową, która kończy polecenie użytkownika.

76. Wykonanie umowy może stanowić podstawę prawną do przetwarzania danych osobowych z wykorzystaniem uczenia maszynowego (ML), jeżeli jest to niezbędne do świadczenia usługi. Przetwarzanie danych osobowych z wykorzystaniem ML do innych celów, które nie są konieczne, takich jak poprawa jakości usług, nie powinno opierać się na tej podstawie prawnej.

77. Wreszcie nie należy mylić podstaw prawnych wykonania umowy i zgody w rozumieniu RODO. Zgoda udzielona na zawarcie umowy, tj. wyrażenie zgody na umowę, jest warunkiem ważności tej umowy i nie odnosi się do szczególnego znaczenia zgody na mocy RODO[29].

78. Gdy korzystanie z VVA nie wymaga wcześniejszego skonfigurowania konta użytkownika do VVA, zgoda może być możliwą podstawą prawną.

3.4.2 Ulepszanie VVA poprzez trening systemów ML oraz ręczne przeglądanie głosu i transkryptów

79. Istnieje wiele akcentów i odmian mowy ludzkiej. Chociaż wszyscy wirtualni asystenci głosowi są funkcjonalni już na początku użytkowania, ich wydajność można poprawić poprzez dostosowanie do specyficznych cech mowy użytkowników. Jak wspomniano w sekcji 2.6, proces dostosowywania opiera się na metodach uczenia maszynowego i składa się z dwóch procesów: dodawania do zbioru danych treningowych VVA nowych danych zebranych od użytkowników oraz przeglądu przez człowieka danych przetworzonych w celu realizacji części poleceń.

> Przykład 7:
>
> Użytkownik VVA musi trzykrotnie wydać to samo polecenie głosowe, ponieważ VVA go nie rozumie. Trzy polecenia głosowe i związane z nimi transkrypcje są przekazywane do osób dokonujących przeglądu w celu sprawdzenia i poprawienia transkrypcji. Polecenia głosowe i poprawione transkrypcje są dodawane do zbioru danych szkoleniowych VVA, aby poprawić jego wydajność.

80. Działania związane z przetwarzaniem danych opisane w przykładzie nie powinny być uważane za (ściśle) *„niezbędne do wykonania umowy”* w rozumieniu art. 6 ust. 1 lit. b) RODO, a zatem

[27] Pod warunkiem, że*„umowa jest ważna zgodnie z obowiązującymi krajowymi przepisami w zakresie zawierania umów”*, fragment Wytycznych nr 2/2019 w sprawie przetwarzania danych osobowych na mocy art. 6 ust. 1 lit. b) RODO w kontekście świadczenia usług internetowych osobom, których dane dotyczą („Wytyczne nr 2/2019”), ust. 26.

[28] Zgodnie z Wytycznymi nr 2/2019, które ponadto stwierdzają, że opinia nr 06/2014 obowiązuje w odniesieniu do art. 6 ust. 1 lit. b) i RODO (zob. w szczególności strony 11, 16, 17, 18 i 55 wspomnianej opinii 06/2014).

[29] Zob. Wytyczne nr 2/2019, odpowiednio ust. 18, 19, 20, 21 i 27.

wymagają innej podstawy prawnej niż art. 6 RODO. Głównym powodem jest to, że VVA są już funkcjonalni w momencie rozpoczęcia użytkowania i mogą już działać jako (ściśle) niezbędni do wykonania umowy. Europejska Rada Ochrony Danych uważa, że art. 6 ust. 1 lit. b) zasadniczo nie stanowi odpowiedniej i zgodnej z prawem podstawy przetwarzania do celów poprawy jakości usługi lub opracowania nowych funkcji w ramach istniejącej usługi. W większości przypadków użytkownik zawiera umowę na korzystanie z istniejącej usługi. O ile postanowienia umowy mogą rutynowo uwzględniać wprowadzanie ulepszeń i modyfikacji do usługi, takiego przetwarzania nie można uznać za obiektywnie niezbędne do wykonania umowy zawartej z użytkownikiem.

3.4.3 Identyfikacja użytkownika[30] (z wykorzystaniem danych głosowych)

81. Wykorzystanie danych głosowych do identyfikacji użytkownika oznacza przetwarzanie danych biometrycznych zgodnie z art. 4 pkt 14 RODO. W związku z tym administrator danych będzie musiał wskazać wyłączenie na mocy art. 9 RODO, oprócz wskazania podstawy prawnej na mocy art. 6 RODO[31].

82. Spośród wyłączeń wymienionych w art. 9 RODO jedynie wyraźna zgoda osoby, której dane dotyczą, wydaje się mieć zastosowanie do tego konkretnego celu.

83. Ponieważ jednak cel ten wymaga zastosowania szczególnego reżimu prawnego art. 9 RODO, dalsze szczegóły znajdują się w sekcji 3.8, dotyczącej przetwarzania szczególnych kategorii danych.

3.4.4 Profilowanie użytkowników w celu spersonalizowania treści lub reklam

84. Jak wspomniano powyżej, VVA mają dostęp do treści wszystkich poleceń głosowych, nawet jeśli są one skierowane do usług świadczonych przez osoby trzecie. Dostęp ten umożliwiłby projektantowi VVA skonstruowanie bardzo dokładnych profili użytkowników, które mogłyby być wykorzystane do oferowania spersonalizowanych usług lub reklam.

> Przykład 8:
>
> Za każdym razem, gdy użytkownik VVA dokonuje wyszukiwania w Internecie, VVA dodaje do profilu użytkownika etykiety sygnalizujące interesujące go tematy. Wyniki dla każdego nowego wyszukiwania są prezentowane użytkownikowi po uporządkowaniu z uwzględnieniem tych etykiet.
>
> Przykład 9:

[30] Z technicznego punktu widzenia, pojęcie identyfikacji należy odróżnić od weryfikacji (uwierzytelniania). Identyfikacja jest wyszukiwaniem i porównywaniem jeden do wielu (1: N) i wymaga zasadniczo bazy danych, w której znajduje się kilka osób. Przetwarzanie do celów weryfikacji jest natomiast porównaniem jeden do jednego (1:1) i służy do sprawdzenia i potwierdzenia za pomocą porównania biometrycznego, czy dana osoba jest tą samą osobą, od której pochodzą dane biometryczne. Zgodnie z wiedzą EROD, VVA na rynku opierają się wyłącznie na technologii identyfikacji mówcy.

[31] Zgodnie z RODO sam charakter danych nie zawsze wystarcza do ustalenia, czy kwalifikują się one jako szczególne kategorie danych, ponieważ *„fotografie są objęte definicją »danych biometrycznych« tylko w przypadkach, gdy są przetwarzane specjalnymi metodami technicznymi, umożliwiającymi jednoznaczną identyfikację osoby fizycznej lub potwierdzenie jej tożsamości*" (motyw 51). To samo rozumowanie odnosi się do głosu.

Za każdym razem, gdy użytkownik VVA dokonuje zakupu w serwisie handlu elektronicznego, VVA przechowuje zapis zamówienia. Dostawca VVA umożliwia osobom trzecim kierowanie do użytkownika VVA ukierunkowanych reklam na podstawie wcześniejszych zakupów.

85. Personalizacja treści może stanowić (ale nie zawsze stanowi) nieodłączny i oczekiwany element VVA. To, czy taki rodzaj przetwarzania można postrzegać jako nieodłączny element usługi VVA, zależeć będzie od dokładnego charakteru świadczonej usługi, oczekiwań przeciętnej osoby, której dane dotyczą, w świetle nie tylko warunków świadczenia usługi, ale również sposobu jej reklamowania użytkownikom, jak również od tego, czy usługę można świadczyć bez personalizacji[32].

86. W przypadku gdy personalizacja odbywa się w kontekście stosunku umownego i jako część usługi wyraźnie zleconej przez użytkownika końcowego (a przetwarzanie jest ograniczone do tego, co jest ściśle niezbędne do świadczenia tej usługi), takie przetwarzanie może opierać się na art. 6 ust. 1 lit. b) RODO.

87. Jeżeli przetwarzanie nie jest ściśle „*niezbędne do wykonania umowy*" w rozumieniu art. 6 ust. 1 lit. b) RODO, dostawca VVA musi zasadniczo uzyskać zgodę osoby, której dane dotyczą. W istocie, ponieważ zgoda będzie wymagana na mocy art. 5 ust. 3 dyrektywy o e-prywatności w przypadku przechowywania lub uzyskiwania dostępu do informacji (zob. pkt 28-29 powyżej), zgoda na mocy art. 6 ust. 1 lit. a) RODO będzie również co do zasady odpowiednią podstawą prawną przetwarzania danych osobowych w następstwie tych operacji, ponieważ opieranie się na uzasadnionym interesie mogłoby w niektórych przypadkach grozić naruszeniem dodatkowego poziomu ochrony przewidzianego w art. 5 ust. 3 dyrektywy o e-prywatności.

88. Jeśli chodzi o profilowanie użytkownika do celów reklamowych, należy zauważyć, że cel ten nigdy nie jest uważany za usługę wyraźnie zlecaną przez użytkownika końcowego. Dlatego w przypadku przetwarzania w tym celu należy systematycznie zbierać zgodę użytkowników.

Zalecenia

89. Użytkownicy powinni być informowani o celu przetwarzania danych osobowych, a cel ten powinien być zgodny z ich oczekiwaniami wobec urządzenia, które nabywają. W przypadku VVA celem tym – z punktu widzenia użytkownika – jest oczywiście przetwarzanie jego głosu wyłącznie w celu interpretacji jego polecenia i udzielenia sensownych odpowiedzi (czy to odpowiedzi na zapytanie, czy innych reakcji, takich jak zdalne sterowanie włącznikiem światła).

90. Gdy przetwarzanie danych osobowych odbywa się na podstawie zgody, „*zgoda taka powinna być udzielona w odniesieniu do jednego lub większej liczby konkretnych celów, a osoba, której dane dotyczą, powinna mieć wybór w odniesieniu do każdego z nich*". Ponadto „*administrator, który ubiega się o zgodę w różnych celach, powinien zapewnić możliwość odrębnego wyrażenia zgody dla każdego celu, aby umożliwić użytkownikom wyrażenie konkretnej zgody w konkretnych celach*"[33]. Na przykład użytkownicy powinni mieć możliwość odrębnego wyrażenia zgody lub odmowy zgody na ręczny przegląd i oznaczanie transkrypcji głosowych lub wykorzystanie ich danych głosowych do identyfikacji/uwierzytelniania użytkownika (zob. sekcja 3.7).

[32] Zob. również Wytyczne nr 2/2019, pkt 57.

[33] Zob. wytyczne EROD 05/2020 w zgody na mocy rozporządzenia 2016/679, przyjęte w dniu 4 maja 2020 r., sekcja 3.2.

3.5 Przetwarzanie danych dzieci

91. Dzieci mogą również wchodzić w interakcje z VVA lub tworzyć własne profile połączone z profilami dorosłych. Niektórzy VVA są wbudowani w urządzenia, które są przeznaczone specjalnie dla dzieci.

92. Gdy podstawą prawną przetwarzania danych jest wykonanie umowy, warunki przetwarzania danych dzieci będą zależały od krajowych przepisów dotyczących umów.

93. Gdy podstawą prawną przetwarzania jest zgoda i zgodnie z art. 8 ust. 1 RODO, przetwarzanie danych dzieci jest zgodne z prawem tylko w przypadku „*dziecka, które ukończyło 16 lat*. *Jeżeli dziecko nie ukończyło 16 lat, takie przetwarzanie jest zgodne z prawem wyłącznie w przypadkach, gdy zgodę wyraziła lub zaaprobowała ją osoba sprawująca władzę rodzicielską lub opiekę nad dzieckiem oraz wyłącznie w zakresie wyrażonej zgody*". W związku z tym, aby zachować zgodność z RODO, gdy podstawą prawną jest zgoda, należy uzyskać wyraźne pozwolenie od rodziców lub opiekunów na zbieranie, przetwarzanie i przechowywanie danych dzieci (głos, transkrypcje itp.).

94. Kontrole rodzicielskie są do pewnego stopnia dostępne, ale w obecnej formie nie są przyjazne dla użytkownika (np. konieczne jest zalogowanie się do nowej usługi) lub mają ograniczone możliwości. Administratorzy danych powinni zainwestować w opracowanie środków umożliwiających rodzicom lub opiekunom kontrolowanie korzystania przez dzieci z VVA.

3.6 Zatrzymywanie danych

95. VVA przetwarzają i generują szeroką gamę danych osobowych, takich jak głos, transkrypcje głosu, metadane czy logi systemowe. Tego typu dane mogą być przetwarzane w szerokim zakresie celów, takich jak świadczenie usług, doskonalenie PJN, personalizacja lub badania naukowe. Zgodnie z zasadą ograniczenia przechowywania danych na mocy RODO VVA powinni zatrzymywać dane nie dłużej niż jest to konieczne do celów, dla których dane osobowe są przetwarzane. Dlatego okresy zatrzymywania danych powinny być powiązane z różnymi celami przetwarzania. Dostawcy usług VVA lub osoby trzecie świadczące usługi za pośrednictwem VVA powinny ocenić maksymalny okres zatrzymywania danych dla każdego zbioru danych i celu.

96. Zasada minimalizacji danych jest ściśle powiązana z zasadą ograniczenia przechowywania danych. Administratorzy danych muszą nie tylko ograniczyć okres przechowywania danych, ale także rodzaj i ilość danych.

97. Administratorzy danych powinni zadać sobie m.in. następujące pytania: Czy konieczne jest przechowywanie wszystkich nagrań głosowych lub wszystkich transkrypcji, aby osiągnąć cel X? Czy konieczne jest przechowywanie danych głosowych po zapisaniu transkrypcji? Jeśli tak, to w jakim celu? Jak długo dane głosowe lub dane transkrypcji są niezbędne dla każdego celu? Odpowiedź na te i inne podobne pytania pozwoli określić okresy zatrzymywania danych, które powinny być jedną z informacji dostępnych dla osób, których dane dotyczą.

98. Niektórzy VVA domyślnie przechowują dane osobowe, takie jak fragmenty głosu lub transkrypcje, przez nieokreślony czas, zapewniając użytkownikom możliwość usunięcia takich danych. Zatrzymywanie danych osobowych na czas nieokreślony jest sprzeczne z zasadą ograniczenia przechowywania. Udostępnienie osobom, których dane dotyczą, środków umożliwiających usunięcie ich danych osobowych nie zwalnia administratora danych z obowiązku określenia i egzekwowania polityki zatrzymywania danych.

99. Projekt VVA musi uwzględniać kontrole użytkowników w celu usunięcia ich danych osobowych w ich urządzeniach i we wszystkich systemach zdalnego przechowywania. Kontrole te mogą być wymagane do rozpatrywania różnego rodzaju wniosków użytkowników, np. wniosku o usunięcie danych lub wycofania uprzednio udzielonej zgody. W projekcie niektórych VVA nie uwzględniono tego wymogu[34].

100. Podobnie jak w innych kontekstach, administratorzy danych mogą być zmuszeni do zatrzymania danych osobowych jako dowodu usługi świadczonej na rzecz użytkownika w celu wywiązania się z obowiązku prawnego. Administrator danych może zatrzymywać dane osobowe na tej podstawie. Zatrzymywane dane powinny jednak stanowić minimum niezbędne do wypełnienia takiego obowiązku prawnego i być przechowywane przez minimalny okres. Oczywiście dane zatrzymane w celu wypełnienia obowiązku prawnego nie powinny być wykorzystywane do żadnych innych celów bez podstawy prawnej na mocy art. 6 RODO.

> Przykład 10:
>
> Użytkownik kupuje telewizor za pośrednictwem usługi handlu elektronicznego za pomocą polecenia głosowego wydanego VVA. Nawet jeśli użytkownik zażąda później usunięcia swoich danych, dostawca lub programista VVA może zatrzymać pewne dane ze względu na swój prawny obowiązek przechowywania dowodów zakupu, wynikający z przepisów podatkowych. Dane przechowywane w tym celu nie powinny jednak wykraczać poza minimum niezbędne do wypełnienia obowiązku prawnego i nie mogą być przetwarzane do żadnych innych celów bez podstawy prawnej zgodnie z art. 6 RODO.

101. Jak wspomniano w sekcji 2, zdolność VVA do rozpoznawania mowy zwiększa się poprzez szkolenie systemów uczenia maszynowego z wykorzystaniem danych użytkowników. Jeżeli użytkownicy nie wyrażą zgody lub wycofają swoją zgodę na wykorzystanie ich danych w tym celu, ich dane nie mogą być zgodnie z prawem wykorzystane do szkolenia kolejnych modeli i powinny zostać usunięte przez administratora danych, przy założeniu, że nie istnieje inny cel uzasadniający dalsze zatrzymywanie danych. Istnieją jednak dowody na to, że w niektórych modelach uczenia maszynowego może istnieć ryzyko ponownej identyfikacji.[35]

102. Administratorzy danych i podmioty przetwarzające dane powinni stosować modele, które nie ograniczają ich zdolności do zaprzestania przetwarzania danych, jeżeli osoba fizyczna cofnie swoją zgodę, ani nie powinni stosować modeli, które ograniczają ich zdolność do ułatwiania podmiotom danych korzystania z ich praw. Administratorzy i podmioty przetwarzające powinni stosować środki ograniczające w celu zmniejszenia ryzyka ponownej identyfikacji do akceptowalnego progu.

103. W przypadku gdy użytkownik wycofuje swoją zgodę, dane zebrane od użytkownika nie mogą być już wykorzystywane do dalszego szkolenia modelu. Niemniej jednak model uprzednio wytrenowany przy użyciu tych danych nie musi zostać usunięty. EROD podkreśla jednak, że istnieją dowody na to, że może istnieć ryzyko wycieku danych osobowych w niektórych modelach uczenia maszynowego. W szczególności liczne badania wykazały, że można

[34] Zob. pismo Amazona z 28 czerwca 2019 r. w odpowiedzi do amerykańskiego senatora Christophera Coonsa: https://www.coons.senate.gov/imo/media/doc/Amazon%20Senator%20Coons Response%20Letter 6.28.19 [3].pdf

[35] Veale Michael, Binns Reuben i Edwards Lilian 2018 „Algorithms that remember: model inversion attacks and data protection law" („Algorytmy, które pamiętają: ataki inwersji modelu a prawo ochrony danych") Phil. Trans. R. Soc. A.37620180083, doi: 10.1098/rsta.2018.0083

przeprowadzać ataki rekonstrukcyjne oraz ataki polegające na wnioskowaniu o przynależności, co pozwala osobom atakującym na odzyskanie informacji o osobach fizycznych[36]. Administratorzy danych i podmioty przetwarzające dane powinni zatem stosować środki ograniczające ryzyko w celu zmniejszenia ryzyka ponownej identyfikacji do akceptowalnego progu, aby upewnić się, że korzystają z modeli, które nie zawierają danych osobowych.

104. Osoby, których dane dotyczą, nie powinny być nakłaniane do wyrażenia zgody na przechowywanie ich danych w nieskończoność. Chociaż usunięcie przechowywanych danych głosowych lub transkrypcji może mieć wpływ na działanie usługi, wpływ ten należy wyjaśnić użytkownikom w jasny i wymierny sposób. Dostawcy usług VVA powinni unikać składania ogólnych oświadczeń o pogorszeniu jakości usługi po usunięciu danych osobowych.

105. Anonimizacja nagrań głosowych stanowi szczególne wyzwanie, ponieważ możliwa jest identyfikacja użytkowników na podstawie treści samej wiadomości i cech samego głosu. Prowadzone są jednak pewne badania[37] nad technikami, które mogłyby umożliwić usunięcie informacji sytuacyjnych, takich jak odgłosy tła, oraz anonimizację głosu.

Zalecenia

106. Z perspektywy użytkownika głównym celem przetwarzania jego danych jest wydawanie poleceń i otrzymywanie odpowiedzi lub wywoływanie działań, takich jak odtwarzanie muzyki lub włączanie i wyłączanie świateł. Po udzieleniu odpowiedzi na zapytanie lub wykonaniu polecenia, dane osobowe powinny zostać usunięte, chyba że projektant lub twórca VVA posiada ważną podstawę prawną do ich zatrzymania w określonym celu.

107. Przed rozważeniem anonimizacji jako sposobu na spełnienie zasady ograniczenia przechowywania danych, dostawcy i programiści VVA powinni sprawdzić, czy proces anonimizacji sprawia, że głos jest niemożliwy do zidentyfikowania.

108. Domyślne ustawienia konfiguracji powinny odzwierciedlać te wymagania poprzez domyślne ograniczenie do absolutnego minimum przechowywanych informacji o użytkowniku. Jeśli opcje te są przedstawiane jako część kreatora konfiguracji, ustawienia domyślne powinny to odzwierciedlać, a wszystkie opcje powinny być przedstawiane jako równe możliwości, bez wizualnej dyskryminacji.

109. Jeżeli w trakcie procesu przeglądu dostawca lub twórca VVA wykryje nagranie pochodzące z błędnej aktywacji, nagranie i wszystkie związane z nim dane powinny zostać natychmiast usunięte i nie powinny być wykorzystywane do żadnych celów.

3.7 Bezpieczeństwo

110. Aby bezpiecznie przetwarzać dane osobowe, VVA powinni chronić ich poufność, integralność i dostępność. Oprócz zagrożeń wynikających z elementów ekosystemu VVA, wykorzystanie głosu jako środka komunikacji stwarza nowy zestaw zagrożeń dla bezpieczeństwa.

---

[36] N. Carlini i in., „Extracting Training Data from Large Language Models" („Wyodrębnianie danych treningowych z dużych modeli językowych"), grudzień 2020 r.

[37] Zob. np. VoicePrivacy (https://www.voiceprivacychallenge.org), inicjatywę mającą na celu opracowanie rozwiązań w zakresie ochrony prywatności w technologii mowy.
Zob. również otwarte narzędzia anonimizacji głosu opracowane w ramach projektu badawczo-innowacyjnego H2020 COMPRISE: https://gitlab.inria.fr/comprise/voice_transformation.

111. VVA mają wielu użytkowników. Mogą pozwalać na korzystanie przez więcej niż jednego zarejestrowanego użytkownika, a każda osoba w ich otoczeniu może wydawać polecenia i korzystać z ich usług. Każda usługa VVA wymagająca poufności będzie wymagała jakiegoś mechanizmu kontroli dostępu i uwierzytelniania użytkownika. Bez kontroli dostępu każdy, kto jest w stanie wydawać polecenia głosowe VVA, może uzyskać dostęp, modyfikować lub usuwać dane osobowe użytkowników (np. pytać o odebrane wiadomości, adres użytkownika lub wydarzenia w kalendarzu). Wydawanie VVA poleceń głosowych nie wymaga fizycznej bliskości VVA, ponieważ można nimi manipulować, np. poprzez nadawanie sygnału[38] (przykładowo radiowego lub telewizyjnego). Niektóre ze znanych metod zdalnego wydawania poleceń VVA, takie jak fale laserowe[39] lub ultradźwiękowe (niesłyszalne)[40], nie są nawet wykrywalne przez ludzkie zmysły.

112. Uwierzytelnianie użytkownika może opierać się na jednym lub kilku z następujących czynników: coś, co się wie (np. hasło), coś, co się posiada (np. karta elektroniczna) lub coś, kim się jest (np. głosowy odcisk palca). Bliższe spojrzenie na te czynniki uwierzytelniania w kontekście VVA pokazuje, że:

    - uwierzytelnianie przy użyciu czegoś, co użytkownik zna, jest problematyczne. Sekret, który pozwoliłby użytkownikowi potwierdzić swoją tożsamość, powinien być wypowiadany na głos, a tym samym ujawnia się go każdemu w otoczeniu. Kanałem komunikacyjnym VVA jest otaczające powietrze, czyli rodzaj kanału, który nie może być zabezpieczony w sposób, w jaki są zabezpieczone tradycyjne kanały (np. poprzez ograniczenie dostępu do kanału lub zaszyfrowanie jego zawartości).

    - Uwierzytelnianie przy użyciu czegoś, co użytkownik posiada, wymagałoby od dostawców usług VVA tworzenia, dystrybucji „tokenów", które mogłyby być wykorzystywane jako dowód tożsamości, oraz zarządzania nimi.

    - Uwierzytelnianie przy użyciu czegoś, czym jest użytkownik, zakłada wykorzystanie danych biometrycznych w celu jednoznacznej identyfikacji osoby fizycznej (zob. sekcja 3.7 poniżej).

113. Konta użytkowników VVA są powiązane z urządzeniami, w ramach których świadczona jest usługa. Często to samo konto używane do zarządzania VVA jest wykorzystywane do zarządzania innymi usługami. Na przykład posiadacze telefonu komórkowego z systemem Android i głośnika Google Home mogą powiązać swoje konto Google z obydwoma urządzeniami i najprawdopodobniej to zrobią. Większość VVA nie wymaga ani nie oferuje mechanizmu identyfikacji lub uwierzytelniania, gdy urządzenie świadczące usługę VVA ma tylko jedno konto użytkownika.

114. Jeśli z urządzeniem związane jest więcej niż jedno konto użytkownika, niektórzy VVA oferują opcjonalną podstawową kontrolę dostępu w postaci numeru PIN bez rzeczywistego uwierzytelniania użytkownika. Niektórzy inni VVA mają możliwość wykorzystania rozpoznawania modelu głosu jako mechanizmu identyfikacji.

---

[38] X. Yuan i in., „All Your Alexa Are Belong to Us: A Remote Voice Control Attack against Echo" („Wszystkie Wasze Alexy należą do nas: atak zdalnej kontroli głosowej na Echo"), globalna konferencja telekomunikacji IEEE Global Communications Conference (GLOBECOM) 2018, Abu Dhabi, Zjednoczone Emiraty Arabskie, 2018, s. 1-6, doi: 10.1109/GLOCOM.2018.8647762.

[39] Zob. na przykład: https://lightcommands.com

[40] Zob. na przykład: https://surfingattack.github.io

115. Chociaż identyfikacja lub uwierzytelnienie użytkownika może nie być konieczne do uzyskania dostępu do wszystkich usług VVA, to z pewnością będzie konieczne w przypadku niektórych z nich. Bez mechanizmu identyfikacji lub uwierzytelniania każdy mógłby uzyskać dostęp do danych innych użytkowników i dowolnie je modyfikować lub usuwać. Na przykład każdy, kto znajdzie się w pobliżu inteligentnego głośnika, może usunąć listy odtwarzania innych użytkowników z usługi strumieniowego przesyłania muzyki, polecenia z historii poleceń lub kontakty z listy kontaktów.

116. Większość VVA ślepo ufa swoim sieciom lokalnym. Każde zagrożone urządzenie w tej samej sieci może zmienić ustawienia inteligentnego głośnika lub umożliwić instalację złośliwego oprogramowania lub przypisanie do niego fałszywych aplikacji/umiejętności bez wiedzy lub zgody użytkownika[41].

117. VVA, jak każde inne oprogramowanie, są narażeni na istnienie podatności oprogramowania. Jednakże, ze względu na koncentrację rynku VVA[42], każda podatność może dotknąć miliony użytkowników VVA. Jeśli VVA działają zgodnie z bieżącym projektem, nie wysyłają żadnych informacji do usługi chmury rozpoznawania mowy, dopóki nie zostanie wykryte wyrażenie budzące. Niemniej jednak podatności oprogramowania mogą pozwolić atakującemu na ominięcie ustawień VVA i środków bezpieczeństwa. Wówczas możliwe byłoby np. uzyskanie kopii wszystkich danych wysyłanych do chmury VVA i przesłanie ich do serwera kontrolowanego przez atakującego.

118. Dane legalnie przetwarzane lub uzyskiwane przez VVA pozwalają na zbudowanie dość dokładnego profilu ich użytkowników, ponieważ VVA znają lub mogą wywnioskować lokalizację, relacje i zainteresowania swoich użytkowników. VVA są coraz częściej obecni w domach i smartfonach użytkowników. Okoliczność ta zwiększa ryzyko masowego nadzoru i masowego profilowania. W związku z tym środki bezpieczeństwa mające na celu ochronę danych w czasie tranzytu i spoczynku, w urządzeniach i w chmurze, powinny odpowiadać tym zagrożeniom.

119. Rosnące wykorzystanie VVA w połączeniu z brakiem odpowiedniego wyważenia praw dostępu przez organy ścigania może wywołać efekt mrożący, który podważy podstawowe prawa, takie jak wolność słowa.

120. Organy ścigania, zarówno w UE[43], jak i poza nią[44], wyraziły już swoje zainteresowanie dostępem do fragmentów głosu przechwyconych przez VVA. Dostęp do danych przetwarzanych lub uzyskiwanych przez VVA w UE powinien być zgodny z obowiązującymi ramami prawnymi UE w zakresie ochrony danych i prywatności. W przypadku gdy niektóre państwa członkowskie rozważają wydanie szczegółowych przepisów ograniczających podstawowe prawa do

[41]Zob. na przykład Deepak Kumar i in., *„Skill Squatting Attacks on Amazon Alexa"* („Ataki typu skill squatting na Amazon Alexa") sympozjum USENIX Security Symposium, sierpień 2018 r., https://www.usenix.org/conference/usenixsecurity18/presentation/kumar
Security Research Labs, „*Smart Spies: Alexa and Google Home expose users to vishing and eavesdropping*" („Inteligentni szpiedzy: Alexa i Google Home narażają użytkowników na wyłudzanie danych i podsłuchiwanie"), listopad 2019 r., https://srlabs.de/bites/smart-spies

[42] Rynek VVA jest obecnie podzielony między mniej niż kilkunastu dostawców usług.

[43] Zob. na przykład: https://www.ft.com/content/ad765972-87a2-11e9-a028-86cea8523dc2

[44] Zob. na przykład: https://cdt.org/insights/alexa-is-law-enforcement-listening

prywatności i ochrony danych, takie ograniczenia powinny zawsze być zgodne z wymogiem określonym w art. 23 RODO[45].

121. Przegląd nagrań głosowych i powiązanych danych przez człowieka w celu poprawy jakości usług VVA jest powszechną praktyką wśród dostawców usług VVA. Ze względu na wrażliwy charakter danych przetwarzanych przez osoby dokonujące przeglądu oraz fakt, że proces ten jest często zlecany podmiotom przetwarzającym, niezwykle istotne jest wprowadzenie odpowiednich środków bezpieczeństwa.

Zalecenia

122. Projektanci VVA i twórcy aplikacji powinni zapewnić użytkownikom bezpieczne, najnowocześniejsze procedury uwierzytelniania.

123. Osoby dokonujące przeglądu powinny zawsze otrzymywać ściśle niezbędne dane opatrzone pseudonimem. Umowy prawne regulujące przegląd powinny wyraźnie zabraniać jakiegokolwiek przetwarzania, które mogłoby prowadzić do identyfikacji osoby, której dane dotyczą.

124. Jeżeli wzywanie służb ratunkowych jest świadczone jako usługa za pośrednictwem VVA, należy zagwarantować stabilny czas nieprzerwanej pracy (uptime)[46].

3.8 Przetwarzanie szczególnych kategorii danych

125. Jak wspomniano wcześniej, VVA mają dostęp do informacji o charakterze prywatnym, które mogą być chronione na mocy art. 9 RODO (zob. sekcja 3.7.1), takich jak dane biometryczne (zob. sekcja 3.7.2). Dlatego projektanci i twórcy VVA muszą dokładnie określić, w jakich przypadkach przetwarzanie wiąże się ze specjalnymi kategoriami danych.

3.8.1 Uwagi ogólne dotyczące przetwarzania specjalnych kategorii danych

126. VVA mogą przetwarzać specjalne kategorie danych w różnych okolicznościach:

- W ramach własnych usług, na przykład podczas zarządzania terminami wizyt lekarskich w terminarzach użytkowników.
- Gdy działają jako interfejs dla usług osób trzecich, dostawcy VVA przetwarzają treść poleceń. W zależności od rodzaju usługi zleconej przez użytkownika, dostawcy VVA mogą przetwarzać specjalne kategorie danych. Przykładem może być sytuacja, w której użytkownik wydaje VVA polecenie korzystania z aplikacji innej firmy służącej do monitorowania owulacji[47].
- Dane głosowe są wykorzystywane do celów jednoznacznej identyfikacji użytkownika, jak opisano poniżej.

3.8.2 Uwagi szczególne dotyczące przetwarzania danych biometrycznych

127. Niektórzy VVA są w stanie jednoznacznie identyfikować użytkowników jedynie na podstawie ich głosu. Proces ten jest znany jako rozpoznawanie modelu głosu. Podczas fazy rejestracji w

[45] Zob. również wytyczne EROD nr 10/2020 w sprawie zgody na mocy art. 23 RODO.

[46] Czas, w którym urządzenie lub usługa mogą pozostać bez nadzoru bez wystąpienia incydentów (crashing) lub konieczności ponownego uruchomienia (rebooting) w celach administracyjnych lub konserwacyjnych.

[47] Zob. na przykład produkt dostępny tutaj: https://www.amazon.com/Ethan-Fan-Ovulation-Period-Tracker/dp/B07CRLSHKY

systemie rozpoznawania mowy VVA przetwarza głos użytkownika w celu utworzenia modelu głosu (lub spektogramu). Podczas normalnego użytkowania VVA może obliczyć model głosowy dowolnego użytkownika i porównać go z zarejestrowanymi modelami, aby jednoznacznie zidentyfikować użytkownika, który wydał polecenie.

> Przykład 11:
>
> Grupa użytkowników konfiguruje VVA do korzystania z rozpoznawania modelu głosu. Następnie każdy z nich zapisuje swoje modele głosowe.
>
> Później jeden z użytkowników żąda od VVA dostępu do spotkań, które znajdują się w jego terminarzu. Ponieważ dostęp do terminarza wymaga identyfikacji użytkownika, VVA wyodrębnia model z głosu wydającego polecenie, oblicza jego model głosowy i sprawdza, czy pasuje on do zarejestrowanego użytkownika i czy ten konkretny użytkownik ma dostęp do terminarza.

128. W powyższym przykładzie rozpoznawanie głosu użytkownika na podstawie modelu głosu jest równoznaczne z przetwarzaniem szczególnych kategorii danych osobowych w rozumieniu art. 9 RODO (przetwarzanie danych biometrycznych w celu jednoznacznego zidentyfikowania osoby fizycznej)[48]. Przetwarzanie danych biometrycznych w celu zidentyfikowania użytkownika, jak wymaga tego przykład, będzie wymagało wyraźnej zgody zainteresowanej osoby, której dotyczą dane (art. 9 ust. 2 lit. a) RODO). W związku z tym przy uzyskiwaniu zgody użytkowników administratorzy danych muszą spełniać warunki określone w art. 7 oraz wyjaśnione w motywie 32 RODO i powinni zaoferować alternatywną metodę identyfikacji w stosunku do biometrii, mając na uwadze dobrowolny charakter zgody.

129. Wykorzystując dane głosowe do identyfikacji biometrycznej lub uwierzytelniania, administratorzy danych są zobowiązani do zapewnienia przejrzystości w zakresie tego, gdzie stosowana jest identyfikacja biometryczna oraz w jaki sposób modele głosowe (modele biometryczne) są przechowywane i przekazywane w różnych urządzeniach. Aby spełnić ten wymóg przejrzystości, EROD zaleca przedstawienie odpowiedzi na następujące pytania:

- Czy aktywacja identyfikacji głosowej na jednym urządzeniu automatycznie aktywuje tę funkcję na wszystkich innych urządzeniach połączonych z tym samym kontem?
- Czy aktywacja identyfikacji głosowej rozprzestrzenia się poprzez infrastrukturę kontrolera VVA na urządzenia należące do innych użytkowników?
- Gdzie są generowane, przechowywane i dopasowywane modele biometryczne?
- Czy modele biometryczne są dostępne dla dostawców VVA, twórców oprogramowania lub innych podmiotów?

130. Gdy zarejestrowany użytkownik skonfiguruje VVA do identyfikacji głosu swoich użytkowników, głos niezarejestrowanych i przypadkowych użytkowników również będzie przetwarzany w celu ich jednoznacznej identyfikacji.

131. Wykrycie głosu właściwego mówcy wymaga bowiem również porównania go z głosem innych osób znajdujących się w pobliżu asystenta. Innymi słowy, funkcja rozpoznawania mówcy wdrożona w asystentach głosowych może wymagać zarejestrowania biometrii głosu osób rozmawiających w gospodarstwie domowym, aby umożliwić odróżnienie cech głosu

użytkownika od cech głosu osoby, która chce zostać rozpoznana. Skutkiem identyfikacji biometrycznej może być zatem poddawanie nieświadomych osób przetwarzaniu ich danych biometrycznych poprzez rejestrowanie ich modelu i porównywanie go z modelem użytkownika, który chce zostać rozpoznany.

132. Aby uniknąć takiego zbierania danych biometrycznych bez wiedzy osób, których dane dotyczą, a jednocześnie umożliwić rozpoznanie użytkownika przez asystenta, należy nadać priorytet rozwiązaniom opartym wyłącznie na danych użytkownika. Konkretnie oznacza to, że rozpoznawanie biometryczne jest aktywowane tylko przy każdym użyciu z inicjatywy użytkownika, a nie poprzez stałą analizę głosów słyszanych przez asystenta. Można na przykład podać konkretne słowo kluczowe lub pytanie do obecnych osób, aby uzyskać ich zgodę na uruchomienie przetwarzania biometrycznego. Na przykład, użytkownik może powiedzieć „identyfikacja" lub asystent może zapytać „czy chcesz zostać zidentyfikowany" i czekać na pozytywną odpowiedź, aby aktywować przetwarzanie biometryczne.

> Przykład 12:
>
> Jeśli użytkownik chce skonfigurować uwierzytelnianie biometryczne w celu uzyskania dostępu do pewnych chronionych danych, takich jak jego konto bankowe, asystent głosowy mógłby aktywować weryfikację głośnikową wyłącznie po uruchomieniu aplikacji i w ten sposób zweryfikować jego tożsamość.

Zalecenia

133. Modele głosowe powinny być generowane, przechowywane i dopasowywane wyłącznie na lokalnym urządzeniu, a nie na zdalnych serwerach.

134. Ze względu na wrażliwość modeli głosowych, należy starannie stosować standardy takie jak ISO/IEC 24745 oraz techniki ochrony modeli biometrycznych[49].

135. Jeżeli VVA wykorzystuje identyfikację biometryczną na podstawie głosu, dostawcy VVA powinni:

- upewnić się, że identyfikacja jest wystarczająco dokładna, aby w sposób wiarygodny powiązać dane osobowe z właściwymi osobami, których dane dotyczą;
- zapewnić, że dokładność jest podobna dla wszystkich grup użytkowników, sprawdzając, czy nie ma znaczącej tendencyjności w stosunku do różnych grup demograficznych.

3.9 Minimalizacja danych

136. Administratorzy powinni minimalizować ilość danych, które są zbierane bezpośrednio lub pośrednio oraz uzyskiwane w drodze przetwarzania i analizy, np. nie przeprowadzać żadnej analizy głosu użytkownika lub innych informacji dźwiękowych w celu uzyskania informacji o jego stanie psychicznym, ewentualnej chorobie lub okolicznościach życiowych.

[49] Zob. na przykład:
Jain, Anil & Nandakumar, Karthik & Nagar, Abhishek. (2008). „*Biometric Template Security*" („Bezpieczeństwo szablonów biometrycznych"). Dziennik EURASIP o postępach w przetwarzaniu sygnałów. 2008. 10.1155/2008/579416.
S. K. Jami, S. R. Chalamala i A. K. Jindal, „*Biometric Template Protection Through Adversarial Learning*" („Ochrona szablonów biometrycznych poprzez uczenie adwersaryjne") *2019 IEEE Międzynarodowa Konferencja Elektroniki Użytkowej (ICCE)*, Las Vegas, NV, Stany Zjednoczone, 2019, s. 1-6, doi: 10.1109/ICCE.2019.8661905.

137. Wprowadzić ustawienia domyślne, które ograniczają zbieranie lub przetwarzanie danych do minimalnej wymaganej ilości niezbędnej do świadczenia usługi.

138. W zależności od lokalizacji, kontekstu użytkowania i czułości mikrofonu, VVA mógłby zbierać dane głosowe osób trzecich jako część szumu tła podczas zbierania głosu użytkowników. Nawet jeśli szum tła nie zawiera danych głosowych, może zawierać dane sytuacyjne, które można przetwarzać w celu uzyskania informacji o uczestniku (np. lokalizacji).

Zalecenia

139. Projektanci VVA powinni rozważyć technologie usuwające szum tła, aby uniknąć nagrywania i przetwarzania głosów i informacji sytuacyjnych w tle.

3.10 Rozliczalność

140. W przypadku każdego przetwarzania, które opiera się na zgodzie, administratorzy danych są zobowiązani do udowodnienia zgody osób, których dane dotyczą, zgodnie z art. 7 ust. 1 RODO. Dane głosowe mogą być wykorzystywane do rozliczalności (np. w celu udowodnienia uzyskania zgody). Obowiązek zatrzymywania takich danych głosowych byłby wówczas podyktowany wymogami rozliczalności określonymi w odpowiednich przepisach szczegółowych.

141. Oceniając potrzebę przeprowadzenia oceny skutków dla ochrony danych, EROD określił kryteria[50], które mają być stosowane przez organy ochrony danych przy tworzeniu wykazów operacji przetwarzania wymagających obowiązkowej oceny skutków dla ochrony danych, oraz podał przykłady procesów przetwarzania, które prawdopodobnie będą wymagały takiej oceny. Jest bardzo prawdopodobne, że usługi VVA mieszczą się w kategoriach i warunkach określonych jako wymagające przeprowadzenia oceny skutków dla ochrony danych. Obejmuje to rozważenie, czy urządzenie może służyć do obserwacji, monitorowania lub kontrolowania osób, których dane dotyczą lub systematycznego monitorowania na dużą skalę zgodnie z art. 35 ust. 3 lit. c), użycia nowej technologii lub przetwarzania danych szczególnie chronionych i danych dotyczących szczególnie narażonych osób, której dane dotyczą.

142. Wszystkie czynności związane ze zbieraniem i przetwarzaniem danych muszą być udokumentowane zgodnie z art. 30 RODO. Dotyczy to również wszelkiego przetwarzania danych głosowych.

Zalecenia

143. Jeśli komunikaty głosowe mają być wykorzystywane do informowania użytkowników zgodnie z art. 13, administratorzy danych powinni publikować takie komunikaty na swojej stronie internetowej, aby były one dostępne dla użytkowników i organów ochrony danych.

3.11 Ochrona danych w fazie projektowania oraz domyślna ochrona danych

144. Dostawcy i programiści VVA powinni rozważyć konieczność posiadania zarejestrowanego użytkownika dla każdej z funkcji. O ile jasne jest, że do zarządzania terminarzem lub książką adresową konieczne jest posiadanie zarejestrowanego użytkownika, o tyle nie jest tak jasne, że wykonywanie połączeń telefonicznych lub wyszukiwanie w internecie wymaga posiadania zarejestrowanego użytkownika przez VVA.

[50] Grupa Robocza Art. 29 „Wytyczne dotyczące oceny skutków dla ochrony danych (wp248, wersja 01)" – zatwierdzone przez EROD.

145. Usługi, które nie wymagają identyfikacji użytkownika, nie powinny domyślnie kojarzyć żadnego z użytkowników zidentyfikowanych w VVA z poleceniami. Domyślne ustawienia VVA, przyjazne prywatności i ochronie danych, przetwarzałyby dane użytkowników tylko w celu realizacji ich poleceń i nie przechowywałaby ani danych głosowych, ani rejestru wykonanych poleceń.

146. Podczas gdy niektóre urządzenia mogą uruchamiać tylko jednego VVA, inne mogą wybierać spośród różnych VVA. Dostawcy VVA powinni opracować normy branżowe umożliwiające przenoszenie danych zgodnie z art. 20 RODO.

147. Niektórzy dostawcy VVA założyli, że ich VVA nie mogą usunąć wszystkich danych użytkownika, nawet na wniosek osoby, której dane dotyczą. Dostawcy VVA powinni zapewnić możliwość usunięcia wszystkich danych użytkowników na wniosek użytkownika zgodnie z art. 17 RODO.

4 MECHANIZMY KORZYSTANIA Z PRAW OSÓB, KTÓRYCH DANE DOTYCZĄ

148. Zgodnie z RODO, administratorzy danych świadczący usługi VVA muszą umożliwić wszystkim użytkownikom, zarejestrowanym i niezarejestrowanym, korzystanie z praw przysługujących osobom, których dane dotyczą.

149. Dostawcy i programiści VVA powinni ułatwić osobom, których dane dotyczą, kontrolę nad ich danymi przez cały okres ich przetwarzania, w szczególności ułatwić im korzystanie z prawa dostępu, sprostowania, usunięcia, prawa do ograniczenia przetwarzania oraz, w zależności od podstawy prawnej, przetwarzania, prawa do przenoszenia danych i prawa sprzeciwu.

150. Administrator danych powinien przekazać informacje o prawach osoby, której dane dotyczą, w momencie włączania VVA przez osoby, których dane dotyczą, a najpóźniej w momencie przetwarzania pierwszego polecenia głosowego użytkownika.

151. Biorąc pod uwagę, że głównym sposobem interakcji w przypadku VVA jest głos, projektanci VVA powinni zapewnić, aby użytkownicy, zarejestrowani lub nie, mogli wykonywać wszelkie prawa osób, których dane dotyczą, za pomocą łatwych do zrozumienia poleceń głosowych. Projektanci VVA, a także programiści aplikacji, jeśli są częścią rozwiązania, powinni na koniec procesu poinformować użytkowników, że ich prawa zostały należycie uwzględnione, głosowo lub poprzez dostarczenie pisemnego powiadomienia na telefon komórkowy, konto użytkownika lub w inny sposób wybrany przez użytkownika.

152. Wreszcie projektanci VVA i programiści aplikacji powinni przynajmniej wdrożyć specjalne narzędzia zapewniające skuteczny i efektywny sposób korzystania z takich praw. Powinni zatem zaproponować w swoich urządzeniach sposób wykonywania praw osób, których dane dotyczą, poprzez udostępnienie osobie, której dane dotyczą, narzędzi samoobsługowych, takich jak system zarządzania profilem[51]. Może to ułatwić skuteczną i szybką obsługę praw osób, których dane dotyczą, i umożliwić administratorowi danych włączenie mechanizmu identyfikacji do narzędzia samoobsługowego.

153. W odniesieniu do wykonywania praw osób, których dane dotyczą, w przypadku wielu użytkowników, gdy użytkownik, zarejestrowany lub niezarejestrowany, wykonuje jedno ze

[51] System zarządzania profilem jest rozumiany jako miejsce w systemie VVA, w którym użytkownicy mogą w dowolnym momencie zapisać swoje preferencje, dokonać modyfikacji i łatwo zmienić ustawienia prywatności.

swoich praw, powinien to robić bez uszczerbku dla praw innych użytkowników. Wszyscy użytkownicy, zarejestrowani i niezarejestrowani, mogą wykonywać swoje prawa tak długo, jak administrator danych nadal przetwarza dane. Administrator danych powinien ustanowić proces zapewniający wykonywanie praw osób, których dane dotyczą.

4.1 Prawo dostępu

154. Zgodnie z art. 12 ust. 1 RODO informacji na mocy art. 15 udziela się na piśmie lub w inny sposób, w tym w stosownych przypadkach – elektronicznie. W odniesieniu do dostępu do danych osobowych podlegających przetwarzaniu, art. 15 ust. 3 stanowi, że jeżeli osoba, której dane dotyczą, zwraca się o kopię drogą elektroniczną i jeżeli nie zaznaczy inaczej, informacji udziela się powszechnie stosowaną drogą elektroniczną. To, co można uznać za powszechnie stosowaną drogę elektroniczną, powinno opierać się na rozsądnych oczekiwaniach osób, których dane dotyczą, a nie na formacie używanym przez administratora danych w jego codziennej działalności. Osoba, której dane dotyczą, nie powinna być zobowiązana do zakupu określonego oprogramowania lub sprzętu komputerowego w celu uzyskania dostępu do informacji.

155. Administratorzy danych powinni zatem przesyłać na żądanie kopię danych osobowych, a w szczególności danych głosowych (w tym nagrań głosowych i transkrypcji), we wspólnym formacie czytelnym dla osoby, której dane dotyczą.

156. Podejmując decyzję o rodzaju formatu, w jakim powinny być dostarczone informacje na mocy art. 15, administrator danych musi pamiętać, że format ten powinien umożliwiać przedstawienie informacji w sposób zarówno zrozumiały, jak i łatwo dostępny. Administratorzy danych powinni również dostosować informacje do konkretnej sytuacji osoby, której dane dotyczą, składającej wniosek.

> Przykład 13:
>
> Administrator danych świadczący usługę VVA otrzymuje od użytkownika zarówno wniosek o dostęp, jak i wniosek o przeniesienie danych. Administrator danych postanawia przekazać informacje zarówno na mocy art. 15, jak i art. 20 w pliku PDF. W takim przypadku nie można uznać, że administrator danych prawidłowo rozpatruje oba wnioski. Plik PDF technicznie wypełnia obowiązki administratora danych na mocy art. 15, ale nie wypełnia obowiązków administratora danych na mocy art. 20[52].
>
> Należy zauważyć, że samo odesłanie użytkowników do historii ich interakcji z asystentem głosowym nie wydaje się umożliwiać administratorowi danych wypełnienia wszystkich obowiązków wynikających z prawa dostępu, ponieważ dostępne dane stanowią na ogół tylko część informacji przetwarzanych w kontekście świadczenia usługi.

157. Prawo dostępu nie powinno być wykorzystywane do łamania/obchodzenia zasad minimalizacji i zatrzymywania danych.

4.2 Prawo do sprostowania danych

158. Aby ułatwić sprostowywanie danych, użytkownicy, zarejestrowani lub nie, powinni mieć możliwość zarządzania swoimi danymi i ich aktualizowania w dowolnym momencie za pomocą

[52] Grupa Robocza Art. 29, Wytyczne dotyczące prawa do przenoszenia danych, zatwierdzone przez EROD, s. 18.

głosu bezpośrednio z urządzenia VVA, jak opisano powyżej. Ponadto w urządzeniu lub aplikacji należy wdrożyć narzędzie samoobsługowe, aby pomóc użytkownikom w łatwym sprostowywaniu ich danych osobowych. Użytkownicy powinni być powiadamiani głosowo lub pisemnie o aktualizacji.

159. Bardziej ogólnie rzecz ujmując, prawo do sprostowania ma zastosowanie do wszelkich opinii i wniosków[53] administratora danych, w tym profilowania, i powinno uwzględniać fakt, że zdecydowana większość danych jest wysoce subiektywna[54].

4.3 Prawo do usunięcia danych

160. Użytkownicy, zarejestrowani lub nie, powinni mieć możliwość, w dowolnym momencie, głosowo z urządzenia VVA lub z narzędzia samoobsługowego zintegrowanego z dowolnym urządzeniem powiązanym z VVA, usunięcia dotyczących ich danych. W tym względzie osoba, której dane dotyczą, może usunąć dane osobowe równie łatwo, jak je przekazała. Ze względu na nieodłączne trudności związane z anonimizacją danych głosowych oraz dużą różnorodność danych osobowych zebranych od osoby, której dane dotyczą[55], zaobserwowanych u niej i o niej wywnioskowanych, w tym kontekście trudno wyegzekwować prawo do usunięcia danych poprzez anonimizację zbiorów danych osobowych. Ponieważ RODO jest neutralne technologicznie, a technologia szybko ewoluuje, nie można jednak wykluczyć, że prawo do usunięcia danych może zostać urzeczywistnione poprzez anonimizację.

161. W niektórych przypadkach, bez ekranu osoby trzeciej lub możliwości wyświetlenia przechowywanych danych (np. aplikacja mobilna lub urządzenie typu tablet), trudno jest mieć podgląd zarejestrowanych ścieżek, ocenić trafność sugestii. Pulpit (lub aplikacja) szeroko dostępny dla użytkowników w celu ułatwienia korzystania z niego powinien być dostarczany wraz z asystentem głosowym, aby można było usunąć historię zadanych zapytań i dostosować narzędzie do potrzeb użytkownika[56].

162. W przypadku każdego przetwarzania danych, a w szczególności, gdy zarejestrowane osoby, których dane dotyczą, wyrażają zgodę na transkrypcję nagrań głosowych i wykorzystanie ich przez dostawcę w celu poprawy jakości jego usług, dostawcy VVA powinni, na żądanie użytkownika, mieć możliwość usunięcia pierwotnego nagrania głosowego, jak również wszelkiej związanej z nim transkrypcji danych osobowych.

163. Administrator danych powinien zapewnić, że po skorzystaniu z prawa do usunięcia danych nie będzie można już ich przetwarzać. W odniesieniu do poprzednich działań prawo do usunięcia danych może podlegać pewnym ograniczeniom prawnym i technicznym, w szczególności.

---

[53] Fakt, że opinie i wnioski mogą być kwalifikowane jako dane osobowe potwierdził TSUE, który zauważył, że termin „wszelkie informacje" w definicji danych osobowych obejmuje informacje, które są „zarówno obiektywne, jak i subiektywne, w postaci opinii czy oceny, a jedynym warunkiem, które muszą one spełniać, jest to, aby »dotyczyły« one danej osoby"- sprawa C-434/16 *Peter Nowak/Data Protection Commissioner* ECLI:EU:C:2017:994 [34].

[54] Prawidłowe egzekwowanie praw osób, których dane dotyczą, wniosek do EROD złożony przez pracowników naukowych zajmujących się ochroną danych, listopad 2019 r.

[55] Grupa Robocza Art. 29, Opinia 05/2014 w sprawie technik anonimizacji, przyjęta 10 kwietnia 2014 r.

[56] „Assistants vocaux et enceintes connectées, l'impact de la voix sur l'offre et les usages culturels et médias", francuska „Conseil Supérieur de l'Audiovisuel", maj 2019 r.

Przykład 14:

Jeżeli przed złożeniem wniosku o usunięcie użytkownik dokonał zakupu internetowego za pomocą swojego VVA, dostawca VVA może usunąć nagranie głosu dotyczące zakupu internetowego i uniemożliwić jego dalsze wykorzystanie w przyszłości. Zakup będzie jednak nadal skuteczny, podobnie jak głosowe zamówienie lub pisemna transkrypcja przetwarzana przez stronę internetową (tutaj wyłączenie wynika z obowiązku prawnego strony handlu elektronicznego).

Podobnie w przypadku, gdy przed żądaniem usunięcia użytkownik dodał za pomocą VVA określony utwór do swojej listy odtwarzania, dostawcy VVA będą mogli usunąć polecenie głosowe, ale nie skutki tego żądania w przeszłości (usunięcie nie będzie miało wpływu na listę odtwarzania użytkownika).

164. W oparciu o powyższe, w przypadku przetwarzania tych samych danych osobowych w różnych celach przetwarzania, administratorzy danych powinni interpretować wnioski o usunięcie danych jako wyraźny sygnał do zaprzestania przetwarzania danych we wszystkich celach, które nie są prawnie wyłączone.

    Zgodnie z warunkami określonymi w art. 21 ust. 1 RODO dane przetwarzane na podstawie uzasadnionych interesów dostawców VVA nie powinny stanowić wyłączenia prawa do usunięcia danych, w szczególności dlatego, że osoby, których dane dotyczą, nie oczekują w sposób uzasadniony dalszego przetwarzania ich danych osobowych.

4.4 Prawo do przenoszenia danych

165. Przetwarzanie danych przez dostawców VVA wchodzi w zakres przenoszenia danych, ponieważ operacje przetwarzania opierają się głównie na zgodzie osoby, której dane dotyczą (na mocy art. 6 ust. 1 lit. a) lub na mocy art. 9 ust. 2 lit. a), jeżeli chodzi o szczególne kategorie danych osobowych), lub na umowie, której stroną jest osoba, której dane dotyczą, na mocy art. 6 ust. 1 lit. b).

166. W praktyce prawo do przenoszenia danych powinno ułatwić zmianę między różnymi dostawcami VVA. W przypadku VVA działających w środowisku cyfrowym, w szczególności gdy głos osoby, której dane dotyczą, został zarejestrowany w aplikacji lub na platformie, prawo do przenoszenia danych powinno być zapewnione w odniesieniu do wszystkich danych osobowych przekazanych przez osobę, której dane dotyczą. Ponadto administrator danych powinien zaoferować użytkownikom możliwość bezpośredniego wyszukania ich danych osobowych w obszarze użytkownika będącego narzędziem samoobsługi. Użytkownicy powinni mieć również możliwość korzystania z tego prawa za pomocą poleceń głosowych.

167. Dostawcy i programiści VVA powinni zapewnić osobom, których dane dotyczą, szeroką kontrolę nad dotyczącymi ich danymi osobowymi, aby umożliwić im przenoszenie danych osobowych od jednego dostawcy VVA do drugiego. Osoby, których dane dotyczą, powinny zatem otrzymywać swoje dane osobowe przekazane administratorowi danych w ustrukturyzowanym, powszechnie używanym i nadającym się do odczytu maszynowego formacie, a także za pomocą środków[57], które są wykorzystywane do odpowiedzi na wnioski o

[57] Zob. jako ilustrację rozumowanie Grupy Roboczej Art. 29 w Wytycznych dotyczących prawa do przenoszenia danych, zatwierdzonych przez EROD, s. 16:

przenoszenie danych (takich jak narzędzia pobierania i interfejsy programowania aplikacji)[58]. Jak stwierdzono w wytycznych w sprawie prawa do przenoszenia danych, w przypadku dużego lub złożonego zbioru danych osobowych, co może mieć miejsce w tym przypadku, administrator danych powinien przedstawić przegląd „*w zwięzłej, przejrzystej, zrozumiałej i łatwo dostępnej formie, jasnym i prostym językiem*" (zob. art. 12 ust. 1 RODO) w taki sposób, aby osoby, których dane dotyczą, zawsze miały jasne informacje, jakie dane należy pobrać lub przekazać innemu administratorowi danych w związku z danym celem. Na przykład osoby, których dane dotyczą, powinny być w stanie korzystać z aplikacji oprogramowania, aby łatwo identyfikować, rozpoznawać i przetwarzać konkretne dane.

168. Prawo to powinno umożliwiać użytkownikowi odzyskanie na własny użytek szczególnie danych, które przekazał za pomocą głosu (np. historii interakcji głosowych), i w ramach tworzenia jego konta użytkownika (np. nazwisko i imię).

169. W celu pełnego zastosowania tego prawa osób, których dane dotyczą, w kontekście jednolitego rynku cyfrowego, projektanci VVA i programiści aplikacji powinni opracować wspólne formaty odczytu maszynowego, które ułatwią interoperacyjność formatu danych między systemami VVA[59], w tym standardowe formaty danych głosowych. Technologie powinny być skonstruowane w taki sposób, aby zapewnić łatwe i pełne ponowne wykorzystanie przetwarzanych danych osobowych, w tym danych głosowych, przez nowego administratora danych[60].

170. Jeśli chodzi o format, dostawcy VVA powinni dostarczać dane osobowe przy użyciu powszechnie stosowanych otwartych formatów (np. mp3, wav, csv, gsm) wraz z odpowiednimi metadanymi stosowanymi w celu dokładnego opisania znaczenia wymienianych informacji[61].

„Na poziomie technicznym administratorzy danych powinni zbadać i ocenić dwa różne i uzupełniające się sposoby udostępniania danych podlegających przenoszeniu osobom, których dane dotyczą, lub innym administratorom danych:

- bezpośrednie przesłanie całego zbioru danych zawierającego dane podlegające przenoszeniu (lub kilku fragmentów części globalnego zbioru danych);

- zautomatyzowane narzędzie umożliwiające wyodrębnianie istotnych danych.

Drugi sposób może być preferowany przez administratorów danych w przypadkach obejmujących złożone, obszerne zbiory danych, ponieważ umożliwia on wyodrębnienie jakiejkolwiek części zbioru danych, która jest istotna dla osoby, której dane dotyczą, w kontekście jej żądania, może pomóc zminimalizować ryzyko oraz potencjalnie umożliwia wykorzystanie mechanizmów synchronizacji danych (np. w kontekście regularnej komunikacji między administratorami danych). Może być to lepszy sposób zapewnienia zgodności w przypadku »nowego« administratora oraz stanowiłby on dobrą praktykę w zakresie ograniczania zagrożeń dla prywatności ze strony pierwotnego administratora danych".

[58] W tym zakresie: Grupa Robocza Art. 29, Wytyczne dotyczące prawa do przenoszenia danych, zatwierdzone przez EROD, s. 1.

[59] W tym względzie: motyw 68 RODO; Grupa Robocza Art. 29, Wytyczne dotyczące prawa do przenoszenia danych, zatwierdzone przez EROD, s. 17.

[60] „*W tym zakresie w motywie 68 zachęca się administratorów danych do opracowywania interoperacyjnych formatów, które umożliwiają przenoszenie danych, nie nakładając jednak na administratorów obowiązku prowadzenia lub wprowadzenia kompatybilnych technicznie systemów przetwarzania. W RODO nie zakazano jednak administratorom tworzenia barier utrudniających przesyłanie.*" – Grupa Robocza Art. 29, Wytyczne dotyczące prawa do przenoszenia danych, zatwierdzone przez EROD, s. 5.

[61] EROD zdecydowanie zachęca zainteresowane strony z branży i stowarzyszenia handlowe do współpracy nad wspólnym zestawem interoperacyjnych standardów i formatów w celu spełnienia wymogów prawa do przenoszenia danych.

5 ZAŁĄCZNIK: AUTOMATYCZNE ROZPOZNAWANIE MOWY, SYNTEZA MOWY I PRZETWARZANIE JĘZYKA NATURALNEGO

171. Podążając za teoretycznymi podstawami przetwarzania sygnałów, w szczególności za teoriami informacji i próbkowania Claude'a Shannona, automatyczne przetwarzanie mowy stało się fundamentalnym elementem nauk inżynieryjnych. Na styku fizyki (akustyka, propagacja fal), matematyki stosowanej (modelowanie, statystyka), informatyki (algorytmy, techniki uczenia się) i nauk o człowieku (percepcja, rozumowanie), przetwarzanie mowy zostało szybko podzielone na liczne przedmioty badań: identyfikacja i weryfikacja mówcy, automatyczne rozpoznawanie mowy, synteza głosu, detekcja emocji itp. W ciągu ostatnich piętnastu lat nastąpił znaczący postęp w tej dziedzinie, do którego przyczyniły się różne czynniki: udoskonalone metody, znaczny wzrost mocy obliczeniowej oraz większa ilość dostępnych danych.

5.1 Automatyczne rozpoznawanie mowy (ASR)

172. Automatyczne rozpoznawanie mowy (znane również jako transkrypcja mowy na tekst) obejmowało kiedyś trzy odrębne etapy mające na celu: 1) określenie, które fonemy zostały wypowiedziane przy użyciu modelu akustycznego; 2) określenie, które słowa zostały wypowiedziane przy użyciu słownika fonetycznego; 3) transkrypcję sekwencji słów (zdania), które najprawdopodobniej zostały wypowiedziane przy użyciu modelu językowego. Obecnie, dzięki postępowi, jaki umożliwiło głębokie uczenie (technika uczenia maszynowego), wiele systemów oferuje automatyczne rozpoznawanie mowy typu end-to-end. Pozwala to uniknąć konieczności przechodzenia przez skomplikowany trening trzech różnych modeli, oferując jednocześnie lepszą wydajność pod względem wyników i czasu przetwarzania. Niemal wszyscy główni gracze w branży cyfrowej oferują obecnie własne implementacje ASR, które mogą być łatwo wykorzystane przez systemy API, ale istnieją również systemy otwarte (na przykład DeepSpeech[62] lub Kaldi[63]).

5.2 Przetwarzanie języka naturalnego (PJN)

173. Przetwarzanie języka naturalnego to multidyscyplinarna dziedzina obejmująca lingwistykę, informatykę i sztuczną inteligencję, której celem jest tworzenie narzędzi do przetwarzania języka naturalnego dla różnorodnych zastosowań. Obszary badań i zastosowań są liczne: analiza składniowa, tłumaczenie maszynowe, automatyczne generowanie i streszczanie tekstów, sprawdzanie pisowni, systemy odpowiadania na pytania, eksploracja tekstów, rozpoznawanie nazw własnych, analiza nastrojów itp. Konkretnie celem PJN jest zapewnienie komputerom zdolności do czytania, rozumienia i wydobywania znaczenia z języków ludzkich. Rozwój aplikacji PJN stanowi wyzwanie, ponieważ narzędzia komputerowe tradycyjnie wymagają od człowieka interakcji z nimi w języku programowania, który jest formalny, czyli precyzyjny, jednoznaczny i wysoce ustrukturyzowany. Natomiast ludzka mowa nie zawsze jest precyzyjna. Jest ona często wieloznaczna, a struktura językowa może zależeć od wielu złożonych zmiennych, w tym slangu, dialektów regionalnych i kontekstu społecznego.

174. Analiza składniowa i semantyczna to dwie główne techniki wykorzystywane w PJN. Składnia to układ słów w zdaniu, nadający mu sens gramatyczny. PJN wykorzystuje składnię do oceny znaczenia języka w oparciu o reguły gramatyczne. Techniki składniowe obejmują parsowanie

[62] https://github.com/mozilla/DeepSpeech
[63] https://github.com/kaldi-asr/kaldi

(analiza gramatyczna zdania), segmentację słów (która dzieli duży fragment tekstu na jednostki), łamanie zdań (które umieszcza granice zdań w dużych tekstach), segmentację morfologiczną (która dzieli słowa na grupy) i stemming (który dzieli słowa z fleksją na formy źródłowe). Semantyka obejmuje użycie i znaczenie słów. PJN stosuje algorytmy w celu zrozumienia znaczenia i struktury zdań. Techniki, które PJN wykorzystuje w odniesieniu do semantyki, obejmują rozróżnianie znaczenia słów (które określa znaczenie słowa na podstawie kontekstu), rozpoznawanie nazw własnych (które określa słowa, które można podzielić na grupy) oraz generowanie języka naturalnego (które wykorzystuje bazę danych do określenia semantyki słów). Podczas gdy wcześniejsze podejścia do PJN obejmowały podejścia oparte na regułach, gdzie proste algorytmy uczenia maszynowego były informowane o słowach i frazach, których należy szukać w tekście i otrzymywały konkretne odpowiedzi, gdy te frazy się pojawiały, obecne podejścia do PJN opierają się na głębokim uczeniu, rodzaju sztucznej inteligencji, która bada i wykorzystuje wzorce w danych w celu poprawy zrozumienia programu.

5.3 Synteza mowy

175. Synteza mowy jest sztucznym generowaniem ludzkiej mowy. Synteza mowy jest realizowana głównie poprzez konkatenację jednostek wokalnych, które są przechowywane w bazie danych. Technika ta polega na tym, że ze wszystkich nagrań aktora, które zostały wcześniej zapisane w postaci fonemów, sylab i słów, wybiera się cegiełki dźwiękowe, które odpowiadają słowom, które mają być wymawiane przez VVA, i składa się je jedna po drugiej, aby utworzyć zrozumiałe zdanie z naturalną dykcją. Alternatywnie syntezator mowy może zawierać model traktu głosowego i inne cechy ludzkiego głosu w celu modelowania parametrów głosu, takich jak intonacja, rytm i barwa, za pomocą generatywnych modeli statystycznych (takich jak WaveNet[64], Tacotron[65] lub DeepVoice[66]) i tworzenia całkowicie syntetycznego głosu wyjściowego.

[64] Aäron van den Oord et Sander Dieleman, *„WaveNet: A generative model for raw audio”* („WaveNet: Model generatywny surowego dźwięku”), blog Deepmind, wrzesień 2016 r., https://deepmind.com/blog/article/wavenet-generative-model-raw-audio

[65] Yuxuan Wang, *„Expressive Speech Synthesis with Tacotron”* („Ekspresyjna synteza mowy w Tacotronie”), blog Google AI, marzec 2018 r., https://ai.googleblog.com/2018/03/expressive-speech-synthesis-with.html

[66] *„Deep Voice 3: 2000-Speaker Neural Text-to-Speech”* („Deep Voice 3: 2000 - Przetwarzanie tekstu na mowę neutralne wobec mówcy”), blog Baidu Research, październik 2017 r. http://research.baidu.com/Blog/index-view?id=91