

Ohjeet



Äänikomennoilla ohjattavia virtuaaliavustajia koskevat ohjeet 02/2021

Versio 2.0

annettu 7. heinäkuuta 2021

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

Aiemmat versiot

Versio 2.0	7. heinäkuuta 2021	Ohjeiden hyväksyminen julkisen kuulemisen jälkeen
Versio 1.0	9. maaliskuuta 2021	Ohjeiden hyväksyminen julkista kuulemista varten

TIIVISTELMÄ

Äänikomennolla ohjattava virtuaaliavustaja, jäljempänä 'VVA' (*Virtual Voice Assistant*), on palvelu, joka ymmärtää äänikomentoja ja suorittaa niitä tarvittaessa muiden IT-järjestelmien kanssa. VVA:ita on tällä hetkellä saatavilla useimpiin älypuheliin ja taulutietokoneisiin, perinteisiin tietokoneisiin ja viime vuosina jopa itsenäisiin laitteisiin, kuten älykaiuttimiin.

VVA:t toimivat rajapintana käyttäjien ja heidän tietokonelaitteidensa sekä verkkopalveluiden, kuten hakukoneiden tai verkkokauppojen, välillä. Jotta VVA:t voisivat suorittaa tehtäviään, niillä on käytettävissään valtava määrä henkilötietoja, mukaan lukien kaikkien käyttäjien komennot (esim. selaus- tai hakuhistoria) ja vastaukset (esim. merkinnät kalenterissa).

Suurimman osan VVA-palveluista on suunnitellut vain muutama VVA-suunnittelija. VVA:t voivat kuitenkin yhdessä kolmansien osapuolten (VVA-sovellusten kehittäjien) ohjelmoimien sovellusten kanssa tarjota kehittyneempiä komentoja.

Jotta VVA toimisi oikein, se tarvitsee päätelaitteen, jossa on mikrofoneja ja kaiuttimia. Laite tallentaa puhetietoja ja muuta dataa, ja VVA siirtää ne etäpalvelimilleen.

VVA-palveluita tarjoavien rekisterinpitäjien ja henkilötietojen käsittelijöiden on siksi otettava huomioon sekä yleinen tietosuoja-asetus¹ että sähköisen viestinnän tietosuojadirektiivi².

Näissä ohjeissa yksilöidään joitakin olennaisimpia sääntöjen noudattamista koskevia haasteita ja annetaan asianomaisille sidosryhmille suosituksia niiden ratkaisuun.

Rekisterinpitäjien, jotka tarjoavat VVA-palveluja ruuduttomiin päätelaitteisiin, on edelleen ilmoitettava käyttäjille yleisen tietosuoja-asetuksen mukaiset tiedot VVA:n käyttöönoton tai asentamisen yhteydessä tai kun VVA-sovellusta käytetään ensimmäistä kertaa. Tämän vuoksi Euroopan tietosuojaneuvosto suosittelee, että VVA-palvelun tarjoajat/suunnittelijat ja ohjelmoijat kehittäisivät äänipohjaisia rajapintoja pakollisen tiedottamisen helpottamiseksi.

Tällä hetkellä kaikki VVA:t edellyttävät vähintään yhden käyttäjän rekisteröitymistä palveluun. Sisäänrakennetun ja oletusarvoisen tietosuojan velvoitteen mukaisesti VVA-palvelun tarjoajien/suunnittelijoiden ja ohjelmoijien olisi harkittava käyttäjien rekisteröitymisen tarpeellisuutta kunkin toiminnallisuutensa osalta.

Useiden VVA-suunnittelijoiden käyttäjätili yhdistää VVA-palvelun muihin palveluihin, kuten sähköpostiin tai videon suoratoistoon. Euroopan tietosuojaneuvosto katsoo, että rekisterinpitäjien olisi pidättäydyttävä kyseisistä käytännöistä, sillä niihin sisältyy pitkiä ja monimutkaisia tietosuojakäytäntöjä, jotka eivät ole yleisen tietosuoja-asetuksen avoimuusperiaatteen mukaisia.

Näissä ohjeissa tarkastellaan neljää yleisintä tarkoitusta, joita varten VVA:t käsittelevät henkilötietoja: pyyntöjen toteuttaminen, VVA:n koneoppimisen mallin parantaminen, biometrinen tunnistaminen ja profilointi yksilöityä sisältöä tai mainontaa varten.

¹ Euroopan parlamentin ja neuvoston asetusta (EU) 2016/679, annettu 27 päivänä huhtikuuta 2016, luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta, jäljempänä 'yleinen tietosuoja-asetus'.

² Euroopan parlamentin ja neuvoston direktiivi 2002/58/EY, annettu 12 päivänä heinäkuuta 2002, henkilötietojen käsittelystä ja yksityisyyden suojasta sähköisen viestinnän alalla (sähköisen viestinnän tietosuojadirektiivi), sellaisena kuin se on muutettuna direktiivillä 2006/24/EY ja direktiivillä 2009/136/EY, jäljempänä 'sähköisen viestinnän tietosuojadirektiivi'.

Siltä osin kuin rekisterinpitäjät käsittelevät VVA:n tietoja käyttäjän pyyntöjen täyttämiseksi eli käyttäjän pyytämän palvelun tarjoamiseksi, heidät vapautetaan ennakkosuostumuksen vaatimuksesta sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdan mukaisesti. Sitä vastoin tarvitaan sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdassa edellytetty suostumus, kun tietoja tallennetaan tai käytetään mihin tahansa muuhun tarkoitukseen kuin käyttäjien pyyntöjen toteuttamiseen.

Osa VVA-palveluista säilyttää henkilötietoja, kunnes käyttäjät pyytävät niiden poistamista. Tämä ei ole tallentamisen rajoittamisen periaatteen mukaista. VVA:n tulisi säilyttää tietoja vain niin kauan kuin se on tarpeen tietojenkäsittelyn tarkoitusten toteuttamista varten.

Jos rekisterinpitäjä saa tietoonsa (esimerkiksi laadunarviointiprosessien yhteydessä) henkilötietojen tahattoman keräämisen, sen olisi varmistettava, että kullakin tietojenkäsittelyn tarkoituksella on asianmukainen oikeusperusta. Muussa tapauksessa vahingossa kerätyt tiedot olisi poistettava.

VVA voi käsitellä useiden rekisteröityjen tietoja. Näin ollen VVA-palvelun tarjoajien/suunnittelijoiden olisi otettava käyttöön pääsynvalvontajärjestelmiä, joilla varmistetaan henkilötietojen luottamuksellisuus, eheys ja saatavuus. Jotkin perinteiset pääsynvalvontajärjestelmät, kuten salasanat, eivät kuitenkaan sovellu VVA-kontekstiin, koska ne olisi sanottava ääneen. Näissä ohjeissa esitetään joitakin tähän liittyviä näkökohtia. Ne sisältävät muun muassa osan, joka koskee erityisten tietoryhmien käsittelyä biometristä tunnistamista varten.

VVA-palvelun tarjoajien/suunnittelijoiden tulee huomioida, että käyttäjän puhetta kerätessä saatu nauhoite voi sisältää muiden henkilöiden ääntä tai tietoja, kuten taustamelua, jotka eivät ole tarpeen palvelun kannalta. VVA-suunnittelijoiden olisi siksi mahdollisuuksien mukaan harkittava sellaisen teknologian käyttöönottoa, joilla suodatetaan tarpeettomat tiedot ja varmistetaan, että vain käyttäjän ääni tallennetaan.

Arvioidessaan tarvetta tietosuojan vaikutustenarvioinnille, Euroopan tietosuojaneuvosto katsoo, että on hyvin todennäköistä, että VVA-palvelut kuuluvat niihin luokkiin ja edellytyksiin, joiden on todettu edellyttävän tietosuojan vaikutustenarviointia.

VVA-palveluja tarjoavien rekisterinpitäjien olisi varmistettava, että käyttäjät voivat käyttää rekisteröidyn oikeuksiaan helposti ymmärrettävillä äänikomennoilla. VVA-palvelun tarjoajien/suunnittelijoiden sekä sovellusten ohjelmoijien olisi palvelutapahtuman lopuksi ilmoitettava käyttäjille, että heidän oikeutensa on otettu asianmukaisesti huomioon, joko toteamalla tämän ääneen tai lähettämällä kirjallisen ilmoituksen käyttäjän matkapuhelimeen, käyttäjätilille tai muulle käyttäjän valitsemalle alustalle.

Sisällysluettelo

TIIVISTELMÄ	3
1 YLEISTÄ	7
2 TEKNOLOGIATAUSTA	8
2.1 Äänikomennoilla ohjattavien virtuaaliavustajien perusominaisuudet	8
2.2 VVA-ekosysteemin toimijat	9
2.3 Vaiheittainen kuvaus	9
2.4 Herätyslausekkeet	10
2.5 Äänikatkelmat ja koneoppiminen	11
3 Tietosuojaan liittyvät näkökohdat	12
3.1 Oikeudellinen kehys	12
3.2 Tietojenkäsittelyn ja sidosryhmien tunnistaminen	14
3.2.1 Henkilötietojen käsittely	14
3.2.2 Rekisterinpitäjien ja henkilötietojen käsittelijöiden suorittama käsittely	16
3.3 Avoimuus	17
3.4 Käyttötarkoituksen rajoittaminen ja oikeusperusta	21
3.4.1 Käyttäjien pyyntöjen toteuttaminen	22
3.4.2 VVA:n parantaminen kouluttamalla koneoppimisjärjestelmiä ja tarkistamalla äänitiedostot ja litteraatiot manuaalisesti	23
3.4.3 Käyttäjän tunnistus (puhetietojen avulla)	24
3.4.4 käyttäjäprofilointi sisällön tai mainonnan mukauttamiseksi.	24
3.5 Lapsia koskevien tietojen käsittely	26
3.6 Tietojen säilyttäminen.....	26
3.7 Turvallisuus.....	28
3.8 Erityisiä henkilötietoryhmiä koskeva käsittely	31
3.8.1 Erityisien tietoryhmien käsittelyyn liittyvät yleiset näkökohdat	31
3.8.2 Biometristen tietojen käsittelyyn liittyvät erityishuomiot	31
3.9 Tietojen minimointi	33
3.10 Vastuu.....	33
3.11 Sisäänrakennettu ja oletusarvoinen tietosuoja	34
4 Rekisteröidyn oikeuksien käyttömekanismit	34
4.1 Oikeus saada pääsy tietoihin.....	35
4.2 Oikeus tietojen oikaisemiseen	36
4.3 Oikeus tietojen poistamiseen.....	36
4.4 Oikeus siirtää tiedot järjestelmästä toiseen.....	37

5	Liite: Automaattinen puheentunnistus, puhesynteesi ja luonnollisen kielen käsittely	38
5.1	Automaattinen puheentunnistus	39
5.2	Luonnollisen kielen käsittely	39
5.3	Puhesynteesi	40

Euroopan tietosuojaneuvosto

ottaa huomioon luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta 27 päivänä huhtikuuta 2016 annetun Euroopan parlamentin ja neuvoston asetuksen (EU) 2016/679, jäljempänä 'yleinen tietosuoja-asetus', 70 artiklan 1 kohdan j ja e alakohdan,

ottaa huomioon ETA-sopimuksen sekä erityisesti sen liitteen XI ja pöytäkirjan 37, sellaisina kuin ne ovat muutettuina 6. päivänä heinäkuuta 2018 annetulla ETA:n sekakomitean päätöksellä N:o 154/2018³,

ottaa huomioon työjärjestyksensä 12 ja 22 artiklan,

ON ANTANUT SEURAAVAT OHJEET:

1 YLEISTÄ

1. Viimeaikaiset teknologiset edistysaskeleet ovat lisänneet huomattavasti äänikomennoilla ohjattavien virtuaaliavustajien; jäljempänä 'VVA' (*Virtual Voice Assistant*), tarkkuutta ja suosiota. VVA:ita on sisällytetty muun muassa älypuhelimiin, verkottuneisiin ajoneuvoihin, älykaiuttimiin ja älytelevisioihin. Tämä on antanut VVA:ille pääsyn henkilökohtaisiin yksityistietoihin, jotka voivat vahingoittaa yksilön oikeutta tietosuojaan ja yksityisyyteen, jos niitä ei käsitellä asianmukaisesti. Tästä syystä VVA:t ja niitä käyttävät laitteet ovat olleet erinäisten tietosuojaviranomaisten valvonnassa.
2. Puheeseen pohjautuvasta vuorovaikutuksesta on monia etuja: luonnollisuus (käyttäjän ei tarvitse erikseen opetella uutta taitoa), komennon toteuttamisnopeus ja toiminta-alueen laajentuminen, joka mahdollistaa sen, että tietoihin pääsee entistä nopeammin. Puheeseen luottaminen aiheuttaa kuitenkin myös vaikeuksia viestin tulkinnassa: äänisignaalin vaihtelevuus eri puhujien välillä, akustinen ympäristö, kielen epäselvyys jne.
3. Tehtävien sujuvoittaminen tai yksinkertaistaminen on käytännössä edelleen ensisijainen motivaatio varustautua VVA:lla. Tähän voi kuulua esimerkiksi puhelujen soittaminen tai niihin vastaaminen, ajastimen asettaminen jne., erityisesti silloin, kun käyttäjän kädet eivät ole vapaat. VVA-suunnittelijoiden tärkein soveltamiskohde on kotien automatisointi. He ovat mukana helpottamassa kotitalouksien elämää ehdottamalla tehtävien suorittamisen yksinkertaistamista (valon sytyttäminen, lämmityksen säätäminen, kaihdinten sulkeminen jne.) ja niiden keskittämistä yhdelle työkalulle, jonka etäaktivointi on helppoa. Henkilökohtaisen tai kotitalouskäytön lisäksi äänikomennot voivat olla hyödyksi ammatillisissa ympäristöissä, joissa on vaikea käsitellä tietokoneita tai hyödyntää kirjallisia komentoja (esim. teollisuustyöt).
4. Teoriassa äänirajapinnan tärkeitä hyötyjä voivat olla vammaiset tai riippuvuussuhteessa olevat henkilöt, joille perinteisten rajapintojen käyttö on ongelmallista. VVA voi tuoda tietoja ja tietotekniikkaan liittyviä resursseja heidän ulottuvilleen ja siten edistää osallistavaa toimintamallia, sillä äänen käyttö mahdollistaa joidenkin käyttäjäryhmien kokemien kirjoitettuun tekstiin liittyvien vaikeuksien voittamisen.

³ Viittauksilla 'jäsenvaltioihin' tarkoitetaan koko tässä asiakirjassa ETA:n jäsenvaltioita.

5. Lopuksi voidaan todeta, että myös terveysalalla on monia tilanteita, joissa voidaan hyödyntää vuorovaikutteisia agentteja, toimivatpa ne äänin tai eivät. Esimerkiksi covid-19-pandemian aikana otettiin käyttöön erilaisia puhelurobotteja, jotka tarjoavat käyttäjille esidiagnooseja. On ennakoitu, että ihmisten/avustajien vuorovaikutus voi vaikuttaa pitkällä aikavälillä koko potilaan hoitoprosessiin: ei pelkästään hyvinvoinnin ja ennaltaehkäisyyn vaan myös hoidon ja tuen kannalta.
6. Tällä hetkellä älypuhelimia on yli 3 miljardia. Niihin kaikkiin on sisällytetty jokin VVA, joista useimmat ovat oletusarvoisesti käytössä. Myös osa tietokoneiden yleisimmistä käyttöjärjestelmistä sisältävät VVA:n. Viimeaikainen älykaiuttimien suosio (vuonna 2019 niitä myytiin 147 miljoonaa⁴) tuo VVA:n miljooniin koteihin ja toimistoihin. Nykyiset VVA-mallit eivät kuitenkaan tarjoa tunnistamista tai pääsynvalvontajärjestelmiä oletusasetuksena.
7. Tämän asiakirjan tarkoituksena on antaa ohjeita yleisen tietosuojasetuksen soveltamiseen VVA:ita koskeissa asioissa.

2 TEKNOLOGIATAUSTA

2.1 Äänikomennoilla ohjattavien virtuaaliavustajien perusominaisuudet

8. VVA voidaan määritellä sovellusohjelmaksi, joka mahdollistaa luonnollisella kielellä käytävän suullisen vuoropuhelun käyttäjän ja laitteen välillä.
9. Luonnollisella kielellä on ihmiskielinen semantiikka. Kielen ominaisuuksista ja sanaston monipuolisuudesta riippuen sama ohje voidaan muotoilla monella eri tavalla, kun taas jotkin komennot saattavat vaikuttaa samanlaisilta, mutta liittyä kahteen eri objektiin. Näitä epäselvyyksiä ratkaistaan usein päättelymekanismeilla, esimerkiksi sen mukaan, mitä aiemmin on sanottu, milloin ja missä käsky on annettu, mitkä ovat käyttäjän mielenkiinnon kohteet jne.
10. VVA:n voi jakaa moduuleihin, jotka suorittavat eri tehtäviä: äänen sieppaus ja palautus, automaattinen puheen litterointi (puheen tekstiksi muuttaminen), automaattinen kielenkäsittely, dialogistrategiat, pääsy ontologioihin (tarvittavaan erikoisalaan liittyvät datajoukot ja järjestelmälliset käsitteet) ja ulkoiset tietämyksen lähteet, kielen tuottaminen, äänisyntetisaatio (tekstin puheeksi muuttaminen), jne. Käytännössä VVA:n on sallittava vuorovaikutus, jotta se voi toteuttaa tehtäviä (esim. "laita radio päälle", "sammuta valo") tai päästä käsiksi tietoon (esim. "millainen sää huomenna on?", "kulkeeko kello 7:43 junavuoro?"). Sillä on siis välittäjän ja järjestelijän rooli, jonka on tarkoitus helpottaa käyttäjän tehtävien suorittamista.
11. Käytännössä VVA ei ole älykaiutin, mutta älykaiuttimessa voi olla VVA. Näiden kahden sekoittaminen on yleistä, mutta jälkimmäinen on kuitenkin vain ensimmäisen aineellistuma. VVA voi olla käytössä älypuhelimessa, älykaiuttimessa, verkottuneessa kellossa, ajoneuvossa, kodinkoneessa, jne.
12. Tietojenkäsittelyn, johon VVA:n toiminta perustuu, organisointi saattaa vaatia useita tiedonvirtauskaavioita. Kaavioista on mahdollista eristää kolme pääobjektia:

⁴ Ks. esimerkiksi Hampurin tietosuoja- ja tietoviranomaisen lehdistötiedote 1.8.2019: <https://datenschutz-hamburg.de/pressemitteilungen/2019/08/2019-08-01-google-assistant>

Fyysinen tapahtuma: laitteistoelementti, jonka osa VVA on (älypuhelin, kaiutin, älytelevisio jne.) ja jossa on mikrofoneja, kaiuttimia sekä verkko- ja tietojenkäsittelykapasiteettia (tapauksen mukaan enemmän tai vähemmän kehittyntä).

Ohjelmistotapahtuma: osa, joka toteuttaa ihmisen ja koneen välistä vuorovaikutusta, ja jossa integroidaan automaattisen puheentunnistuksen, luonnollisen kielen käsittelyn, dialogin ja puhesynteesin moduulit. Tämän tapahtuman suorittaminen suoraan fyysisten laitteiden sisällä on mahdollista, mutta monissa tapauksissa se suoritetaan etänä.

Resurssit: ulkoiset tiedot, kuten sisältötietokannat, ontologiat tai yrityssovellukset, jotka antavat tietoa (esim. ”paljonko kello on Yhdysvaltain länsirannikolla”, ”lue sähköpostini”) tai mahdollistavat pyydetyn toiminnan konkreettisen toteuttamisen (esim. ”nosta lämpötilaa 1,5 °C”).

13. VVA:ihin on mahdollista asentaa niiden ydintoimintoja laajentavia kolmansien osapuolten komponentteja tai sovelluksia. Komponenttien nimet eroavat VVA:iden välillä, mutta ne kaikki välittävät käyttäjien henkilötietoja VVA-suunnittelijan ja sovelluskehittäjän välillä.
14. Vaikka suurin osa VVA:ista ei jaa äänikatkelmia sovelluskehittäjien kanssa, nämä toimijat käsittelevät silti henkilötietoja. Lisäksi sovelluksen kehittäjä saa toiminnon luonteesta riippuen käsiinsä aikomuksia ja aikavälejä, jotka voivat sisältää arkaluonteisia tietoja, kuten terveystietoja.

2.2 VVA-ekosysteemin toimijat

15. VVA:n toteutusketjussa saattaa olla mukana suuri määrä toimijoita ja välittäjiä. Käytännössä voidaan tunnistaa enintään viisi eri toimijaa. Liiketoimintamalleista ja teknologisista valinnoista riippuen osa toimijoista voi kuitenkin omaksua useita rooliyhdistelmiä, kuten suunnittelija ja integraattori tai suunnittelija ja sovelluskehittäjä:
 - a. **VVA-palvelun tarjoaja (tai suunnittelija):** vastaa VVA:n kehittämisestä, suunnittelee ja määrittelee sen mahdollisuudet ja oletustoiminnot: aktivoinnin modaliteetit, muodon valinta, tietojen saatavuus, tietueiden hallinta, laitteistomäärittelyt jne.
 - b. **VVA-sovelluskehittäjä:** luo sovelluksia, jotka laajentavat VVA:n oletustoimintoja mobiilisovellusten tapaan. Sovelluskehittäjän on noudatettava suunnittelijan asettamia kehitysrajoituksia.
 - c. **Integraattori:** verkottuneiden esineiden valmistaja, joka haluaa varustaa ne VVA:illa. Integraattorin olisi noudatettava suunnittelijan määrittelemiä vaatimuksia.
 - d. **Omistaja:** vastaa fyysisistä tiloista, joihin ihmiset tulevat (majoituspaikoista, ammattiympäristöistä, vuokra-autoista jne.), ja haluaa tarjota VVA:n asiakkailleen (mahdollisesti erityissovelluksilla varustettuna).
 - e. **Käyttäjä:** viimeinen linkki VVA-arvoketjussa, joka voi käyttää VVA:ta eri laitteissa (kaiutin, televisio, älypuhelin, kello jne.) sen mukaan, miten ja missä VVA on otettu käyttöön ja asennettu.

2.3 Vaiheittainen kuvaus

16. Jotta VVA voisi suorittaa toiminnon tai päästä käsiksi tietoon, toteutetaan seuraavat tehtävät:

- 1) VVA on valmiustilassa, kun laite (älypuhelin, kaiutin, ajoneuvo) otetaan käyttöön. Tarkemmin sanottuna se kuuntelee jatkuvasti. Ääntä ei kuitenkaan lähetetä vastaanottavasta laitteesta ennen kuin se havaitsee tietyn herätyslausekkeen, eikä se suorita mitään muuta toimintoa kuin herätyslausekkeen tunnistusta. Tätä varten VVA hyödyntää muutaman sekunnin puskuria (lisätietoa seuraavassa osassa).
- 2) Käyttäjä sanoo herätyslausekkeen ja VVA vertaa äänitiedostoa herätyslausekkeeseen. Jos ne vastaavat toisiaan, VVA avaa kuuntelukanaavan ja lähettää äänisisällön välittömästi.
- 3) Jos komento käsitellään etänä, avainsanan ääntäminen tarkistetaan usein palvelinpuolella toisen kerran ei-toivottujen aktivointien rajoittamiseksi.
- 4) Käyttäjä sanoo pyyntönsä, jonka VVA välittää suoraan VVA-palvelun tarjoajalle. Puhesekvenssi litteroidaan automaattisesti (puhe tekstiksi).
- 5) Käsky tulkitaan luonnollisen kielen käsittelyä hyödyntävällä tekniikalla. Viestin tarkoitukset poimitaan ja tietomuuttujat (aikavälit) tunnistetaan. Tämän jälkeen dialogin hallintajärjestelmä määrittää käyttäjän kanssa toteutettavan vuorovaikutusskenaarion tarjoamalla asianmukainen vastausmallin.
- 6) Jos komento sisältää kolmannen osapuolen sovelluksen tarjoaman toiminnon (taito, toiminta, pikakuvake tms.), VVA-palvelun tarjoaja lähettää sovelluksen kehittäjälle viestin aikeet ja tietomuuttujat (aikavälit).
- 7) Käyttäjän pyyntöön mukautettu vastaus tunnistetaan – ainakin lähtökohtaisesti, sillä jos VVA ei pysty tulkitsemaan pyyntöä oikein, se vastaa ”minulla ei ole vastausta kysymykseesi”. Tarvittaessa hyödynnetään etäresursseja: julkisia tietokantoja (verkkopohjainen tietosisältö jne.) tai todennuspalveluita (pankkitili, musiikkisovellus, verkko-ostoasiakkaan tili jne.), ja tietomuuttujat (aikavälit) täytetään palautetulla tiedolla.
- 8) Vastauslauseke luodaan ja/tai toimenpide tunnistetaan (kaihtimien sulkeminen, lämpötilan nostaminen, musiikin soittaminen, kysymykseen vastaaminen jne.). Lauseke syntetisoidaan (teksti puheeksi) ja/tai suoritettava toimenpide lähetetään suorittavaan laitteeseen.
- 9) VVA palaa valmiustilaan.

Huomaa, että vaikka suurin osa ääneen liittyvästä käsittelystä tehdään etäpalvelimilla, jotkin VVA-palvelun tarjoajat kehittävät järjestelmiä, jotka voivat suorittaa osan tästä käsittelystä paikallisesti⁵.

2.4 Herätyslausekkeet

17. VVA:n on oltava ”hereillä”, jotta sitä voi käyttää. Tämä tarkoittaa sitä, että avustaja siirtyy aktiiviseen kuuntelutilaan vastaanottaakseen pyyntöjä ja komentoja sen käyttäjältä. Vaikka VVA:n voi joissakin tapauksissa herättää myös fyysisellä toiminnalla (esim. painamalla nappia tai älykaiutinta jne.), lähes kaikkien markkinoilla olevien VVA:iden käyttäminen perustuu

⁵Tästä on raportoitu esimerkiksi osoitteessa <https://www.amazon.science/blog/alexa-new-speech-recognition-abilities-showcased-at-interspeech>

herätyslausekkeen tai -sanon (myös aktivointisana tai "hot word") havaitsemiseen ja sitä seuraavaan aktiiviseen kuuntelutilaan siirtymiseen.

18. VVA käyttää mikrofonia ja hienoisia laskennallisia ominaisuuksia tunnistukseen, onko avainsana sanottu. Tämä analyysi, jota VVA suorittaa käynnistymisestään lähtien, tehdään yksinomaan paikallisesti. Äänitallenteet käsitellään käskyn tulkitsemiseksi ja toteuttamiseksi vasta kun avainsana on tunnistettu. Useissa tapauksissa tämä tarkoittaa, että tallenteet lähetetään käsittelyn yhteydessä etäpalvelimiin internetin välityksellä. Avainsanojen tunnistus perustuu koneoppimisen tekniikoihin. Suurin haaste tällaisten menetelmien käytössä on se, että havaitseminen perustuu todennäköisyyksiin. Järjestelmä laskee siis jokaiselle sanalle tai lausekkeelle varmuuskertoimen siitä, onko avainsana todella lausuttu. Jos tämä kerroin osoittautuu ennalta määritettyä kynnysarvoa suuremmaksi, avainsana määritellään lausutuksi. Tällainen järjestelmä ei siis ole virheetön: joissakin tapauksissa aktivointia ei välttämättä havaita, vaikka avainsana on sanottu (virheellinen hylkäys), ja toisissa tapauksissa aktivointi voidaan havaita, vaikka käyttäjä ei ole sanonut avainsanaa (virheellinen hyväksyntä).
19. Käytännössä näiden kahden virhetyypin välillä olisi löydettävä hyväksyttävä kompromissi kynnysarvon määrittämiseksi. Koska avainsanan virheellinen havaitseminen saattaa kuitenkin johtaa äänitallenteiden lähettämiseen, on todennäköistä, että odottamattomia ja ei-toivottuja tiedonsiirtoja tapahtuu. Etäprosessointia toteuttavat VVA-palvelujentarjoajat käyttävät hyvin usein avainsanojen tunnistamiseen kaksivaiheista mekanismia: ensimmäinen läpivienti sulautetaan paikallisesti laitetasolla ja toinen tehdään etäpalvelimilla, joissa seuraava tietojenkäsittely tapahtuu. Tässä tapauksessa kehittäjät asettavat yleensä kynnyksen suhteellisen matalalle parantaakseen käyttäjäkokemusta ja varmistaakseen, että kun käyttäjä sanoo avainsanan, se tunnistetaan lähes aina – vaikka se johtaisikin "ylitunnistamiseen". Tätä seuraava palvelimen puolella tapahtuva toinen tunnistusvaihe on tiukempi.

2.5 Äänikatkelmat ja koneoppiminen

20. VVA:t hyödyntävät koneoppimismenetelmiä monenlaisten tehtävien suorittamiseen (avainsanojen tunnistus, automaattinen puheentunnistus, luonnollisen kielen käsittely, puhesynteesi jne.) ja näin ollen niiden käyttö edellyttää, että niillä on suuria datajoukkoja kerättävänä, valittavana ja merkittävänä jne.
21. Tiettyjen tilastollisten ominaisuuksien yli- tai aliedustus voi vaikuttaa koneoppimiseen perustuvien tehtävien kehittämiseen, mikä puolestaan heijastuu laskelmissa ja siten sen toimintatavassa. Tiedon laadulla on siis yhtä suuri merkitys koneoppimisprosessin tarkkuudelle kuin sen määrällä.
22. VVA-suunnittelijat saattavat haluta saada tietoa laitteen käytöstä todellisissa olosuhteissa – esimerkiksi äänikatkelmia – voidakseen parantaa VVA:n laatua ja käytössä olevia koneoppimismenetelmiä.
23. Olipa kyse oppimistietokannan laadun arvioinnista tai algoritmin käyttöönoton yhteydessä tehtyjen virheiden korjaamisesta, ihmisen osallistuminen tekoälyjärjestelmien opettamiseen ja koulutukseen on välttämätöntä. Tämä työn osa eli digitaalinen työ herättää kysymyksiä sekä työoloista että turvallisuudesta. Tähän liittyen tiedotusvälineet ovat myös kertoneet VVA-suunnittelijoiden ja alihankkijoiden välisistä tiedonsiirroista, jotka ovat tapahtuneet ilman tarvittavia tietosuojatakeita.

3 TIETOSUOJAAN LIITTYVÄT NÄKÖKOHDAT

3.1 Oikeudellinen kehys

24. VVA:ita koskeva asiaankuuluva EU:n oikeudellinen kehys on ensisijaisesti yleinen tietosuoja-asetus, koska henkilötietojen käsittely kuuluu VVA:n ydintehtäviin. Yleisen tietosuoja-asetuksen lisäksi myös sähköisen viestinnän tietosuojadirektiivissä⁶ asetetaan erityinen standardi kaikille toimijoille, jotka haluavat tallentaa tai käyttää tilaajan tai käyttäjän päätelaitteisiin tallennettuja tietoja ETA-alueella.
25. *"Päätelaitteiden"*⁷ määritelmän mukaisesti älypuhelimet, älytelevisiot ja vastaavat esineiden internetin laitteet ovat esimerkkejä päätelaitteista. Vaikka VVA:t itsessään ovat ohjelmistopalveluja, ne toimivat aina fyysisen laitteen, kuten älykaiuttimen tai älykkään television, kautta. **VVA:t käyttävät sähköisiä viestintäverkkoja päästäkseen käsiksi näihin fyysisiin laitteisiin, jotka ovat sähköisen viestinnän tietosuojadirektiivin tarkoittamia "päätelaitteita". Näin ollen sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdan säännöksiä sovelletaan aina, kun VVA tallentaa tietoja siihen liitettyyn fyysiseen laitteeseen tai käyttää jo tallennettuja tietoja.**⁸
26. Jotta edellä mainittujen käsittelytoimien jälkeen suoritettavat henkilötietojen käsittelytoimet, mukaan lukien päätelaitteissa olevien tietojen avulla saatujen henkilötietojen käsittely, olisivat lainmukaisia, niillä on myös oltava yleisen tietosuoja-asetuksen 6 artiklan mukainen oikeusperusta.⁹
27. Koska rekisterinpitäjän on pyytäessään sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdan mukaista suostumusta tietojen tallentamiseen tai käyttämiseen ilmoitettava rekisteröidylle kaikista käsittelyn tarkoituksista (eli "myöhemmästä käsittelystä") – mukaan lukien edellä mainittujen toimien jälkeinen käsittely – yleisen tietosuoja-asetuksen 6 artiklan mukainen suostumus on yleensä asianmukaisin oikeusperusta henkilötietojen myöhempää käsittelyä varten. Näin ollen suostumus muodostaa todennäköisesti oikeusperustan sekä jo tallennettujen tietojen tallentamiselle ja käyttämiselle että henkilötietojen käsittelylle edellä mainittujen käsittelytoimien jälkeen. Yleisen tietosuoja-asetuksen 6 artiklan noudattamista arvioitaessa olisi otettava huomioon, että käsittelyketjuun sisältyy erityisiä toimia, joita varten EU:n lainsäätäjä on pyrkinyt tarjoamaan lisäsuojaa.¹⁰ Rekisterinpitäjien on muun muassa otettava huomioon rekisteröityjen oikeuksiin kohdistuvat vaikutukset määrittäessään

⁶ Euroopan parlamentin ja neuvoston direktiivi 2002/58/EY, annettu 12 päivänä heinäkuuta 2002, henkilötietojen käsittelystä ja yksityisyyden suojasta sähköisen viestinnän alalla (sähköisen viestinnän tietosuojadirektiivi), sellaisena kuin se on muutettuna direktiivillä 2006/24/EY ja direktiivillä 2009/136/EY, jäljempänä 'sähköisen viestinnän tietosuojadirektiivi'.

⁷ Kilpailusta telepäätelaitemarkkinoilla 20 päivänä kesäkuuta 2008 annetun komission direktiivin 2008/63/EY 1 artiklan mukaan "päätelaitteella" tarkoitetaan "a) suoraan tai välillisesti yleisen televerkon rajapintaan liitettyä tietojen lähettämiseen, käsittelyyn tai vastaanottamiseen tarkoitettua laitetta; molemmissa tapauksissa (suora tai välillinen) liitäntä voidaan tehdä johdoin, optisella kuidulla tai sähkömagneettisesti; yhteys on välillinen, jos on sijoitettu laite päätelaitteen ja verkon rajapinnan väliin; | b) | satelliittimaa-asemalaitteita".

⁸ Ks. Euroopan tietosuojaneuvoston ohjeet 1/2020 12 kohta, jossa esitetään samankaltaiset perustelut verkottuneiden ajoneuvojen osalta (jäljempänä "Euroopan tietosuojaneuvoston ohjeet 1/2020"). Ks. Myös Euroopan tietosuojaneuvoston lausunto 5/2019 sähköisen viestinnän tietosuojadirektiivin ja yleisen tietosuoja-asetuksen välisestä vuorovaikutuksesta, erityisesti tietosuojaviranomaisten toimivallan, tehtävien ja valtuuksien osalta.

⁹ Ks. edellinen alaviite, 41 kohta.

¹⁰ Lausunto nro 5/2019, kohta 41.

asianmukaista laillista perustetta, jotta kohtuullisuusperiaatetta noudatettaisiin.¹¹ Lähtökohtana on, että rekisterinpitäjät eivät voi nojata tietosuoja-asetuksen 6 artiklaan pienentääkseen sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdassa säädettyä lisäsuojaa.

28. Kuten kohdassa 2.3 (vaiheet 2 ja 3) on todettu, nykyiset VVA:t tarvitsevat pääsyn VVA-laitteen tallentamaan äänidataan.¹² Näin ollen sovelletaan sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohtaa. Sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdan soveltaminen tarkoittaa sitä, että tietojen tallentaminen ja VVA:n jo tallentamien tietojen käyttö edellyttävät lähtökohtaisesti loppukäyttäjän ennakkosuostumusta¹³. Direktiivi sallii kuitenkin kaksi poikkeusta: ensinnäkin käytön tai tallentamisen, jonka ainoana tarkoituksena on toteuttaa viestinnän välittäminen sähköisissä viestintäverkoissa tai helpottaa sitä tai, toiseksi, joka on ehdottoman välttämätöntä sellaisen tietoyhteiskunnan palvelun tarjoamiseksi, jota tilaaja tai käyttäjä on erityisesti pyytänyt.
29. Toinen poikkeus ("ehdottoman välttämätöntä tilaajan tai käyttäjän nimenomaisesti pyytämän tietoyhteiskuntapalvelun tarjoamiseksi") antaisi VVA-palvelun tarjoajalle mahdollisuuden käsitellä käyttäjien tietoja käyttäjien pyyntöjen täyttämiseksi (ks. 3.4.1 kohdan 72 kohta) ilman sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdassa tarkoitettua suostumusta. Sitä vastoin **sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdassa edellytetty suostumus olisi tarpeen**, jotta tietoja voisi tallentaa tai käyttää **mihin tahansa muuhun tarkoitukseen kuin käyttäjien pyyntöjen toteuttamiseen** (esim. käyttäjäprofilointi). Rekisterinpitäjien olisi kytkettävä suostumus tiettyyn käyttäjään. Näin ollen rekisterinpitäjien olisi käsiteltävä rekisteröimättömiä käyttäjiä koskevia tietoja ainoastaan pyyntöjen toteuttamiseen.
30. VVA voi vahingossa siepata ääntä henkilöistä, joiden ei ollut tarkoitus käyttää VVA-palvelua. Ensinnäkin herätyslausekkeen muuttaminen on VVA:sta riippuen mahdollista jossain määrin. Henkilöt, jotka eivät ole tietoisia herätyslausekkeen muuttamisesta, voivat vahingossa käyttää päivitettyä herätyslauseketta. Toiseksi VVA voi havaita herätyslausekkeen vahingossa tai virheellisesti. On erittäin epätodennäköistä, että jompaakumpaa sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdassa säädetystä poikkeuksista voitaisiin soveltaa vahingossa tapahtuvaan aktivointiin. Lisäksi yleisen tietosuoja-asetuksen mukaisen suostumuksen on oltava *rekisteröidyn yksiselitteinen tahdonilmaisu*. Näin ollen on erittäin epätodennäköistä, että vahingossa tapahtuva aktivointi voitaisiin tulkita päteväksi suostumukseksi. Jos rekisterinpitäjät saavat tietoonsa (esimerkiksi automaattisen tai ihmisten suorittaman tarkastelun avulla), että VVA-palvelu on vahingossa käsitellyt henkilötietoja, heidän olisi varmistettava, että kullakin tällaisten tietojen käsittelyn tarkoituksella on voimassa oleva oikeusperusta. Muussa tapauksessa vahingossa kerätyt tiedot olisi poistettava.
31. Lisäksi todettakoon, että VVA:n käsittelemät henkilötiedot voivat olla luonteeltaan erittäin arkaluonteisia. Ne voivat sisältää henkilötietoja sekä sisällöltään (puhutun tekstin merkitys) että metatiedoiltaan (puhujan sukupuoli tai ikä jne.). Euroopan tietosuojaneuvosto

¹¹ Ks. Euroopan tietosuojaneuvoston ohjeet 2/2019 yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan b alakohdan perusteella tapahtuvasta henkilötietojen käsittelystä rekisteröidyille tarjottavien verkkopalvelujen yhteydessä, 2.0 versio, 8. lokakuuta 2019, 1 kohta.

¹² On mahdollista, että tulevaisuuden VVA-laitteet omaksuvat reunalaskennan paradigman ja pystyvät tuottamaan jonkin verran palvelua paikallisesti. Tällöin on tarpeen arvioida uudelleen sähköisen viestinnän tietosuojadirektiivin sovellettavuutta.

¹³ Ks. myös Euroopan tietosuojaneuvoston ohjeet 1/2020, 14 kohta.

muistuttaa, että puhutiedot ovat luonteeltaan biometrisiä henkilötietoja.¹⁴ Näin ollen, jos tällaisia tietoja käsitellään luonnollisen henkilön yksilöllistä tunnistamista varten tai jos ne kuuluvat luontaisesti henkilötietojen erityisryhmiin tai ne on määritelty sellaisiin kuuluviksi, tietojen käsittelyn oikeusperustan on oltava 6 artiklan mukainen ja siihen on liitettävä poikkeus yleisen tietosuoja-asetuksen 9 artiklasta (ks. jäljempänä 3.7 kohta).

3.2 Tietojenkäsittelyn ja sidosryhmien tunnistaminen

32. Kun otetaan huomioon, että VVA voi tarjota monenlaisia avustumahdollisuuksia niin monessa eri ympäristössä rekisteröidyn arjessa¹⁵, on huomioitava, että henkilötietojen käsittelyyn, johon myös eri sidosryhmät voivat vaikuttaa, on sovellettava huolellista harkintaa.

3.2.1 Henkilötietojen käsittely

33. Henkilötietojen suojan osalta voidaan havaita useita vakioita riippumatta siitä, millaista VVA-tyyppiä (esim. laitetyyppi, toiminnallisuudet, palvelut tai niiden yhdistelmä) rekisteröity voi käyttää. Nämä vakiot liittyvät kyseessä olevien henkilötietojen, rekisteröityjen ja tietojenkäsittelyn moninaisuuteen.

Henkilötietotyyppien moninaisuus

34. Yleisen tietosuoja-asetuksen 4 artiklan 1 kohdan mukainen henkilötietojen määritelmä sisältää useita erilaisia tietoja, ja sitä sovelletaan teknisesti neutraalissa yhteydessä kaikkiin tietoihin, jotka liittyvät *"tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön"*.¹⁶ Mikä tahansa rekisteröidyn vuorovaikutus VVA:n kanssa voi kuulua tämän määritelmän soveltamisalaan. Kuten kohdassa 2.4 todettiin, vuorovaikutuksen tapahduttua VVA:n toiminnan aikana voidaan käsitellä monenlaisia henkilötietoja.
35. Näin ollen ensimmäisestä pyynnöstä siihen liittyvään vastaukseen, toimeen tai seurantaan (esim. viikoittaisen hälytyksen laatiminen), jo ensimmäinen henkilötietojen syöttö tuottaa myöhempiä henkilötietoja. Tähän sisältyvät ensisijaiset tiedot (esim. tilitiedot, äänitallenteet, pyyntöhistoria), havaitut tiedot (esim. rekisteröidyn laitetiedot, toimintalokitiedot, verkkotoiminnot) sekä päätellyt tai johdetut tiedot (esim. käyttäjäprofilointi). VVA käyttää puhetta välittääkseen tietoja käyttäjien ja kaikkien liitettyjen palvelujen välillä (esim. hakukone, verkkokauppa tai musiikin suoratoistopalvelu), mutta toisin kuin muut välittäjät, VVA voi saada täyden pääsyn pyyntöjen sisältöön ja siten tarjota VVA-suunnittelijalle laajan valikoiman henkilötietoja käsittelyn tarkoituksista riippuen.

¹⁴ Yleisen tietosuoja-asetuksen 4 artiklan 14 kohdan määritelmän mukaan biometrisillä tiedoilla tarkoitetaan "luonnollisen henkilön fyysisiin ja fysiologisiin ominaisuuksiin tai käyttäytymiseen liittyvällä teknisellä käsittelyllä saatuja henkilötietoja, kuten kasvokuvia tai sormenjälkitietoja, joiden perusteella kyseinen luonnollinen henkilö voidaan tunnistaa tai kyseisen henkilön tunnistaminen voidaan varmistaa".

¹⁵ Esimerkiksi: kotona, ajoneuvossa, kadulla, työpaikalla tai missä tahansa muussa yksityisessä, julkisessa tai ammatillisessa tilassa tai näiden tilojen yhdistelmässä.

¹⁶ Lisäksi yleisen tietosuoja-asetuksen 4 artiklan 1 kohdassa täsmennetään että *"tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa erityisesti tunnistetietojen, kuten nimen, henkilötunnuksen, sijaintitiedon, verkkotunnistetietojen taikka yhden tai useamman hänelle tunnusomaisen fyysisen, fysiologisen, geneettisen, psyykkisen, taloudellisen, kulttuurillisen tai sosiaalisen tekijän perusteella"*.

36. VVA:ta käytettäessä käsiteltävien henkilötietojen moninaisuus tarkoittaa myös sellaisten henkilötietoryhmien moninaisuutta, joihin on syytä kiinnittää huomiota (ks. kohta 3.7). Euroopan tietosuojaneuvosto muistuttaa, että kun käsitellään erityisiä tietoryhmiä¹⁷, yleisen tietosuojasetuksen 9 artiklassa edellytetään, että rekisterinpitäjä yksilöi sekä pätevän poikkeuksen 9 artiklan 1 kohdan mukaiselle käsittelytoimien kieltämiselle että 6 artiklan 1 kohdan mukaisen pätevän oikeusperustan käyttäen 9 artiklan 2 kohdan mukaisesti yksilöityjä asianmukaisia keinoja. Nimenomainen suostumus voi olla yksi asianmukaisista poikkeuksista, jos suostumus on 6 artiklan 1 kohdan mukainen oikeusperusta. Asetuksen 9 artiklassa todetaan myös (yksityiskohtaisesti), että jäsenvaltiot voivat ottaa käyttöön lisäehtoja biometristen tai muiden erityisluokkien tietojen käsittelyyn.

Rekisteröityjen moninaisuus

37. VVA:ta käytettäessä henkilötietoja käsitellään ensimmäisestä VVA:n kanssa käydyistä vuorovaikutuksesta lähtien. Joidenkin rekisteröityjen osalta tämä tarkoittaa hetkeä, jolloin he ostivat VVA:n ja/tai loivat käyttäjätilin (rekisteröidyt käyttäjät). Muiden rekisteröityjen osalta tämä viittaa siihen hetkeen, kun he olivat tietoisesti ensimmäistä kertaa vuorovaikutuksessa sellaisen VVA:n kanssa, jonka joku toinen rekisteröity on ostanut ja/tai jolle tämä on luonut käyttäjän (rekisteröimättömät käyttäjät). Näiden kahden ryhmän lisäksi on olemassa kolmas luokka: satunnaiset käyttäjät – rekisteröityneet ja rekisteröimättömät – jotka tietämättään esittävät VVA:lle pyyntöjä (esimerkiksi sanovat oikean herätyslausekkeen tietämättä, että VVA on aktiivinen, tai sanovat muita sanoja, jotka VVA tunnistaa virheellisesti herätyslausekkeeksi).
38. Rekisteröityjen moninaisuudella viitataan niin yhden VVA:n useisiin käyttäjiin (esim. rekisteröidyt ja rekisteröimättömät käyttäjät, työtoverit, perheet, kouluyhteisö) kuin erityyppisiin käyttäjiin (esim. aikuinen, lapsi, vanhus tai vammainen). Vaikka VVA voi tarjota helpompaa vuorovaikutusta digitaalisen työkalun kanssa ja monia etuja joillekin rekisteröityjen ryhmille, on tärkeää ottaa huomioon kunkin rekisteröityjen ryhmän erityispiirteet ja VVA:n käyttöympäristö.

Tietojenkäsittelyn moninaisuus

39. Käsiteltävien tietojen määrään ja käsittelyn tyyppisiin vaikuttavat myös VVA:iden tuottamiseen käytetyt teknologiat. Mitä enemmän VVA tarjoaa palveluita tai ominaisuuksia ja on yhteydessä muihin muiden osapuolten hallinnoimiin laitteisiin tai palveluihin, sitä enemmän henkilötietojen käsittely ja käsittelyn uudelleenkohdentaminen lisääntyvät. Tästä johtuu edellisessä kohdassa kuvailtu automaattisten käsittelyjen moninaisuus. Automaation lisäksi joihinkin käsittelyihin voi osallistua myös ihmisiä. Näin on esimerkiksi silloin, kun käyttöön otetun teknologian toiminta hyötyy ihmisen toiminnasta. Ihmiset voivat esimerkiksi tarkistaa puheen litteraatioita tai jakaa henkilötietoja koskevia huomautuksia, joilla voidaan lisätä uusia malleja koneoppimisteknologiaan. Näin on myös silloin, kun ihmiset analysoivat henkilötietoja (esim. metatietoja) VVA:n tarjoaman palvelun parantamiseksi.

¹⁷ Yleisen tietosuojasetuksen 9 artiklan 1 kohdan määritelmän mukaan erityisiä henkilötietoryhmiä ovat tiedot, ”joista ilmenee rotu tai etninen alkuperä, poliittisia mielipiteitä, uskonnollinen tai filosofinen vakaumus tai ammattiliiton jäsenyys sekä geneettisten tietojen käsittely, biometristen tietojen käsittely henkilön yksiselitteistä tunnistamista varten tai terveyttä koskevien tietojen taikka luonnollisen henkilön seksuaalista käyttäytymistä ja suuntautumista koskevien tietojen käsittely on kiellettyä”.

3.2.2 Rekisterinpitäjien ja henkilötietojen käsittelijöiden suorittama käsittely

40. Yleisen tietosuoja-asetuksen mukaan rekisteröidyillä olisi oltava mahdollisuus ymmärtää ja tunnistaa eri toimijoiden roolit ja heidän olisi voitava olla yhteydessä kuhunkin sidosryhmään tai toimia niiden kanssa. Vaikka skenaariot voivat olla monimutkaisia tai kehittyviä, roolijaosta ei saisi olla haittaa rekisteröidylle. Sidosryhmät voivat roolejaan arvioidessaan noudattaa Euroopan tietosuojaneuvoston ohjeita 7/2020 rekisterinpitäjän ja henkilötietojen käsitteistä yleisessä tietosuoja-asetuksessa.¹⁸
41. Kuten 15 kohdassa todetaan, tärkeimpien sidosryhmien rooleiksi voidaan määrittää palvelun tarjoaja tai suunnittelija, sovelluskehittäjä, integraattori, omistaja tai niiden yhdistelmä. Eri skenaariot ovat mahdollisia riippuen roolijaosta sidosryhmien liikesuhteesta, käyttäjän pyynnöstä, henkilötiedoista, tietojenkäsittelytoimista ja niiden tarkoituksista. Sidosryhmien olisi selkeästi päätettävä ja tiedotettava rekisteröidylle edellytyksistä, joilla kukin niistä toimii ja noudattaa yleisen tietosuoja-asetuksen mukaisia rekisterinpitäjien, yhteisrekisterinpitäjien tai henkilötietojen käsittelijöiden rooleja.¹⁹ Kukin niistä voi ottaa yhden tai useamman roolin, koska ne voivat olla yhden tietojenkäsittelyn aikana yksittäisiä rekisterinpitäjiä, yhteisrekisterinpitäjiä tai tietojen käsittelijöitä, mutta suorittaa toista roolia toisen tietojenkäsittelyn aikana.
42. Yleisesti tarkasteltuna suunnittelija saattaa toimia rekisterinpitäjänä määrittäessään käsittelyn tarkoituksia ja keinoja, mutta voi puuttua henkilötietojen käsittelijänä henkilötietojen käsittelyyn muiden osapuolten, kuten sovelluskehittäjän, puolesta. Näin ollen VVA-käyttäjän tietoja käsittelee usea rekisterinpitäjä: sovelluskehittäjä ja suunnittelija. On myös mahdollista, että suunnittelija, integraattori ja kehittäjä ovat ryhmitelty yhdeksi toimijaksi, joka toimii yksittäisenä rekisterinpitäjänä. Soveltuva pätevyys on joka tapauksessa määritettävä tapauskohtaisesti.

Esimerkki 1:

VVA-suunnittelija käsittelee käyttäjätietoja moneen tarkoitukseen, muun muassa VVA:n äänenymmärtämistaitojen ja pyyntöihin vastaamisen parantamiseen. Tässä tapauksessa – ja vaikka tämä tarkoitus voi johtaa kolmansien osapuolten toimittamien sovellusten käytöstä johtuvien tietojen käsittelyyn – rekisterinpitäjiä on vain yksi: VVA-suunnittelija, jonka puolesta ja jonka tarkoituksia varten käsittely suoritetaan.

Esimerkki 2:

Pankki tarjoaa asiakkailleen sovelluksen, jota käyttämällä he voivat tiedustella tiliensä hoitamiseen liittyviä asioita VVA:n kautta.

Henkilötietojen käsittelyssä on mukana kaksi toimijaa: VVA-suunnittelija ja pankkisovelluksen kehittäjä.

Tässä skenaariossa pankki on palvelun tuottava rekisterinpitäjä, koska se määrittelee sovellukseen liittyvät tarkoitukset ja keskeiset käsittelykeinot, jotka mahdollistavat vuorovaikutuksen VVA:n kanssa. Pankki tarjoaa erillisen sovelluksen, jonka avulla käyttäjä,

¹⁸ Euroopan tietosuojaneuvoston ohjeet 07/2020 rekisterinpitäjän ja henkilötietojen käsittelijän käsitteistä yleisessä tietosuoja-asetuksessa, V2.0, annettu 7. heinäkuuta 2021, jäljempänä 'ohjeet 7/2020'.

¹⁹ Yleinen tietosuoja-asetus, 12–14 artikla, 26 artikla.

pankin asiakas, voi hallita tilejään etäältä. Lisäksi se päättää käsittelytavoista valitsemalla asianmukaisen henkilötietojen käsittelijän, joka on VVA-suunnittelija ja jonka asiantuntemusta voidaan hyödyntää näiden käsittelytapojen määrittämisessä (se voi esimerkiksi hallinnoida ohjelmointialustaa, joka mahdollistaa kolmansien osapuolten sovellusten integroimisen VVA:han ja siten asettaa puitteet ja ehdot, joita sovelluskehittäjiä on noudatettava).

43. Rekisteröidyn on syytä huomata, että useat sidosryhmät voivat käsitellä samoja henkilötietoja, vaikka rekisteröity olettaisi, että vain VVA-palvelun tarjoaja osallistuisi niiden käsittelyyn. Kun rekisteröity toteuttaa henkilötietojensa koskevia toimia VVA-palvelun tarjoajan kanssa (esim. rekisteröidyn oikeuksien käyttö), tämä ei automaattisesti tarkoita, että tämä toimenpide koskisi toisen sidosryhmän käsittelemiä samoja henkilötietoja. Kun nämä sidosryhmät ovat itsenäisiä rekisterinpitäjiä, on tärkeää, että rekisteröidyille ilmoitetaan selkeästi käsittelyn eri vaiheista ja toimijoista. Tapauksissa, joihin liittyy yhteisrekisterinpitäjä, olisi lisäksi tehtävä selväksi, onko jokaisella rekisterinpitäjällä valmiudet noudattaa kaikkien rekisteröityjen oikeuksia tai keillä rekisterinpitäjistä on valmiudet kunkin oikeuden noudattamiseen.²⁰

Esimerkki 3:

Tässä skenaariossa VVA-suunnittelija haluaa käyttää pankin tarjoamaa palvelua varten kerättyjä ja käsiteltyjä tietoja äänentunnistusjärjestelmänsä parantamiseksi. VVA:n suunnittelija, joka käsittelee tietoja omiin tarkoituksiinsa, on tällöin kyseisessä käsittelyssä rekisterinpitäjän asemassa.

44. Koska henkilötietojen käsittelyketjuun voi kuulua monia sidosryhmiä, ja vastaavasti monia työntekijöitä, riskitilanteita voi esiintyä, jos asianmukaisia toimenpiteitä ja suojatoimia ei ole käytössä. Vastuu on rekisterinpitäjillä, ja siksi heidän olisi keskityttävä erityisesti henkilötietojen suojaamiseen valitsemalla asianmukaisia liikekumppaneita ja henkilötietojen käsittelijöitä, soveltamalla sisäänrakennettua ja oletusarvoista yksityisyyden suojaa,²¹ ottamalla käyttöön asianmukaisia, muun muassa turvallisuuteen liittyviä, yleisen tietosuojasetuksen välineitä, kuten tarkastuksia ja oikeudellisesti sitovia sopimuksia (esimerkiksi yleinen tietosuojasetuksen yhteisrekisterinpitäjiä koskeva 26 artikla tai henkilötietojen käsittelijöitä koskeva 28 artikla).
45. VVA-ekosysteemi on monimutkainen ekosysteemi, jossa monilla toimijoilla on mahdollisuus vaihtaa ja käsitellä henkilötietoja rekisterinpitäjänä tai henkilötietojen käsittelijöinä. On erittäin tärkeää määrittää kunkin toimijan rooli kunkin käsittelyn osalta ja noudattaa tietojen minimoinnin periaatetta myös tietojenvaihdon osalta.
46. Lisäksi rekisterinpitäjien olisi valvottava henkilötietojen siirtoa ja taattava tarvittava suojan taso koko käsittelyketjussa, erityisesti silloin, kun ne käyttävät ETA:n ulkopuolella sijaitsevia palveluja.

3.3 Avoimuus

47. Koska VVA:t käsittelevät henkilötietoja (esimerkiksi käyttäjien ääntä, sijaintia tai viestin sisältöä), niiden on noudatettava yleisen tietosuojasetuksen avoimuusvaatimuksia, joista

²⁰ Ohjeet 7/2020, 165 kohta.

²¹ Ks. Euroopan tietosuojaneuvoston ohjeet 4/2019 sisäänrakennetusta ja oletusarvoisesta tietosuojasta, versio 2.0, annettu 20. lokakuuta 2020.

säädetään 5 artiklan 1 kohdan a alakohdassa sekä 12 ja 13 artiklassa (selitetään johdanto-osan 58 kappaleessa). Rekisterinpitäjien on ilmoitettava käyttäjille henkilötietojen käsittelystä tiiviisti, avoimesti ja ymmärrettävästi helposti saavutettavissa olevalla tavalla.

48. Tarpeellisten tietojen toimittamatta jättäminen on kyseisten velvoitteiden rikkomista, ja se voi vaikuttaa tietojenkäsittelyn laillisuuteen. Avoimuusvaatimuksen noudattaminen on välttämätöntä, koska se toimii tietojenkäsittelyn valvontamekanismina ja antaa käyttäjille mahdollisuuden käyttää oikeuksiaan. Kun käyttäjille tiedotetaan asianmukaisesti siitä, miten heidän henkilötietojaan käytetään, rekisterinpitäjien on vaikeampaa käyttää VVA:ta väärin tarkoituksiin, jotka eivät vastaa käyttäjien odotuksia. Esimerkiksi patentoiduilla teknologioilla pyritään päättämään käyttäjän äänen perusteella hänen terveyden- ja tunnetilaansa ja mukauttamaan tarjottuja palveluja niiden mukaan.
49. Avoimuusvaatimusten noudattaminen voi olla erityisen vaikeaa VVA-palvelun tarjoajalle tai muulle rekisterinpitäjänä toimivalle taholle. VVA:iden erityisluonteen vuoksi rekisterinpitäjät kohtaavat useita esteitä noudattaessaan yleisen tietosuoja-asetuksen avoimuusvaatimuksia:
- J **Useat käyttäjät:** Rekisterinpitäjien olisi annettava tietoa kaikille käyttäjille (rekisteröidyille, rekisteröimättömille ja tahattomille käyttäjille), ei ainoastaan VVA:n käyttöönottoon osallistuville käyttäjille.
 - J **Ekosysteemin monimuotoisuus:** kuten Teknologiausta-osiassa todettiin, henkilötietoja käsittelevien toimijoiden identiteetit ja roolit jäävät usein epäselviksi VVA:iden käyttäjille.
 - J **Äänirajapinnan erityispiirteet:** digitaaliset järjestelmät eivät vielä sovellu vuorovaikutuksiin, jotka toimivat vain äänikomennoin. Tästä on todisteena se, että niihin on melkein systemaattisesti kytketty näyttö. Äänirajapintaan sopeutuminen ja se, että sen kautta pystytään kommunikoimaan käyttäjän kanssa selkeästi ja asianmukaisesti, on välttämätöntä.
50. VVA:ita voidaan pitää äärellisinä automaatteina, jotka käyvät läpi useita tiloja niiden tavanomaisen toiminnan aikana. Ne voivat kuunnella paikallisesti havaitakseen herätyslausekkeita tai olla vuorovaikutuksessa etäpalvelimen kanssa komennon ratkaisemiseksi, mutta ne voivat myös omaksua muita tiloja kontekstin mukaan (esimerkiksi, jos taustalla on luontoääniä) tai jos käyttäjä puhuu niille (ne voivat esimerkiksi puhua tunnistetulle tai tuntemattomalle käyttäjälle). Näissä tilanteissa käyttäjä kokee valitettavasti informaation epäsymmetriaa, sillä hän tuskin tietää kuunteleeko laite, ja vielä vähemmän hän tietää laitteen tilasta.
51. On erittäin suositeltavaa, että VVA-suunnittelijat ja -kehittäjät ryhtyvät riittäviin toimiin ratkaistakseen nämä epäsymmetriat, sillä tämä tekisi VVA:iden toiminnasta vuorovaikutteisempaa. Käyttäjille olisi ilmoitettava, missä tilassa laite milläkin hetkellä on. Läpinäkyvyyttä voidaan parantaa tekemällä dialogista laitteen ja ihmisen välillä näin vuorovaikutteisempaa (laite voisi esimerkiksi kuitata äänikomentojen vastaanoton jollakin tavalla) tai ilmoittamalla laitteen tilasta tietyin signaalein. Tämän toteuttamiseksi voidaan kokeilla useita eri vaihtoehtoja, kuten tiettyjen äänitunnisteiden ja näkyvien kuvakkeiden ja valojen tai näyttöjen käyttöä laitteissa.
52. Nämä ongelmakohdat ovat erityisen merkityksellisiä, kun otetaan huomioon käyttäjien moninaisuus ja haavoittuvassa asemassa olevien henkilöiden, kuten lasten, vanhusten tai näkö- ja kuulovammaisten käyttäjien, olemassaolo.

53. Edellä mainituista ongelmista kumpuaa kaksi tärkeää kysymystä: mikä on tehokkain tapa tiedottaa käyttäjiä ja milloin olisi sille sopiva aika? Näitä kysymyksiä tulisi tarkastella tarkemmin kahdessa eri tilanteessa riippuen siitä, onko VVA:lla vain yksi käyttäjä (esimerkiksi henkilökohtainen älypuhelin) vai mahdollisesti useita käyttäjiä (esim. älykodinkone). VVA-teknologiaa käyttämällä voidaan myös häivyttää rajaa näiden kahden eri tilanteen välillä esimerkiksi silloin, kun käyttäjällä on henkilökohtainen älypuhelin, joka liitetään autoon. Tässä tapauksessa älypuhelimien VVA, jonka voidaan kohtuudella olettaa olevan vain kyseisen käyttäjän käyttämä, "ulotetaan" muiden autossa olevien käyttöön.
54. Tällä hetkellä kaikki VVA:t on yhdistetty johonkin käyttäjätiliin ja/tai ne on määritetty sovelluksessa, joka edellyttää käyttäjätilin luomista. Kysymystä siitä, miten rekisterinpitäjät voisivat harkita VVA:n käyttöön ottamisen yhteydessä tietosuojaselosteesta tiedottamista käyttäjille, olisi käsiteltävä 29 artiklan mukaisen tietosuojatyöryhmän avoimuutta koskevien ohjeiden mukaisesti. Sovelluksista on annettava tarvittavat tiedot ennen lataamista verkkokaupassa.²² Näin tiedot annetaan mahdollisimman aikaisin ja viimeistään silloin, kun henkilötiedot saadaan. Jotkin VVA-palvelun tarjoajat sisällyttävät kolmannen osapuolen sovelluksia VVA:n oletusasetuksiin, jotta nämä sovellukset voivat suorittaa toimintojaan käyttämällä tiettyjä herätyslausekkeita. Tätä kolmannen osapuolen sovellusten käyttöönottostrategiaa hyödyntävien VVA:iden olisi varmistettava, että käyttäjät saavat tarvittavat tiedot myös kolmannen osapuolen toteuttamasta tietojenkäsittelystä.
55. Monet VVA-suunnittelijat vaativat kuitenkin VVA:n käyttöä varten käyttäjätilejä, jotka yhdistävät VVA-palvelun useisiin muihin palveluihin, kuten sähköpostiin, videoiden suoratoistoon tai ostoksiin. VVA-suunnittelijan päätös yhdistää tili moniin eri palveluihin edellyttää hyvin pitkiä ja monimutkaisia tietosuojakäytäntöjä. Tällaisten tietosuojakäytäntöjen pituus ja monimutkaisuus vaikeuttavat huomattavasti avoimuusperiaatteen noudattamista.

Esimerkki 4:

VVA-suunnittelija vaatii käyttäjiltään tilin VVA-palvelun käyttöä varten. Tämä käyttäjätili ei ole kytköksissä vain VVA-palveluun, vaan sitä voidaan käyttää muihin VVA-suunnittelijan tarjoamiin palveluihin, kuten sähköpostiin, pilvitallennuspalveluun ja sosiaaliseen mediaan. Käyttäjätilin luominen edellyttää 30 sivun pituisen tietosuojaselosteen lukemista ja hyväksymistä. Seloste sisältää tietoja kaikkien tiliin liitettävien palveluiden suorittamasta henkilötietojen käsittelystä.

VVA-suunnittelijan tässä tapauksessa toimittamien tietojen ei pitäisi katsoa olevan tiiviissä muodossa. Tietojen monimutkaisuus vähentää vaadittavaa läpinäkyvyyttä. Näin ollen VVA-suunnittelija ei täyttäisi yleisen tietosuoja-asetuksen 12 ja 13 artiklassa säädettyjä avoimuusvaatimuksia.

56. Vaikka vaadittavat tiedot annetaan yleensä kirjallisesti, yleisessä tietosuoja-asetuksessa sallitaan myös "muut keinot". Johdanto-osan 58 kappaleessa todetaan nimenomaisesti, että tiedot voidaan antaa sähköisessä muodossa esimerkiksi verkkosivuston kautta. Valittaessa asianmukaista menetelmää, jolla rekisteröidylle annettaisiin pääsy tietoon, olisi lisäksi otettava huomioon erityiset olosuhteet, kuten se, miten rekisterinpitäjä ja rekisteröity muuten ovat vuorovaikutuksessa keskenään.²³ Jos laitteessa ei ole näyttöä, rekisteröidylle voidaan

²² Asetukseen 2016/679 perustuvat ohjeet avoimuudesta, WP260 rev. 01, Euroopan tietosuojaneuvoston hyväksymät, jäljempänä 'Tietosuojatyöryhmän ohjeet WP260', 11 kohta.

²³ 29 artiklan mukaisen tietosuojatyöryhmän ohjeet WP260, 19 kohta.

antaa helposti ymmärrettävä linkki joko suoraan tai sähköpostitse. Jo olemassa olevat ratkaisut voisivat toimia esimerkkinä tiedotukselle, esimerkiksi puhelinpalvelukeskusten käytännöt, joissa soittajalle ilmoitetaan puhelun nauhoittamisesta ja asiakas ohjataan lukemaan tietosuojakäytännöt. Näyttömien VVA:iden rajoitteet eivät vapauta rekisterinpitäjää toimittamasta tarvittavia tietoja yleisen tietosuoja-asetuksen mukaisesti VVA:n tai VVA-sovelluksen asennuksen tai käytön yhteydessä. VVA-palveluntarjoajien ja kehittäjien tulisi kehittää puhepohjaisia rajapintoja pakollisen tiedon välittämisen helpottamiseksi.

57. VVA:t voisivat kiinnostaa erityisesti heikkonäköisiä käyttäjiä, koska ne tarjoavat vaihtoehtoisia vuorovaikutuskeinoja visuaaliseen tietoon perinteisesti perustuvien IT-palvelujen rinnalle. Yleisen tietosuoja-asetuksen 12 artiklan 1 kohdan mukaan tarvittavien tietojen antaminen suullisesti on mahdollista ainoastaan rekisteröidyn pyynnöstä, mutta ei oletusasetuksena. Näyttömiin VVA:ihin liittyy kuitenkin rajoitteita, joiden vuoksi tieto olisi automaattisesti annettava ääneen ja mahdollisesti täydennettävä kirjallisesti. Rekisterinpitäjien olisi annettava tarvittavat tiedot tiiviisti ja selkeästi, kun ne annetaan rekisteröidylle äänen keinoin. Lisäksi rekisteröityjen olisi voitava kuunnella viesti uudelleen²⁴.
58. Asianmukaisten toimenpiteiden toteuttaminen yleisen tietosuoja-asetuksen avoimuusvaatimusten noudattamiseksi on monimutkaisempaa, kun VVA-laitteella on useita muita käyttäjiä kuin laitteen omistaja. VVA-suunnittelijoiden on mietittävä, miten rekisteröidylle ja tahattomille käyttäjille ilmoitetaan asianmukaisesti, kun heidän henkilötietojaan käsitellään. Kun suostumus on käyttäjien tietojen käsittelyn oikeusperusta, käyttäjille on tiedotettava asianmukaisesti henkilötietojen käsittelystä, jotta suostumus olisi pätevä²⁵.
59. Noudattaakseen yleistä tietosuoja-asetusta rekisterinpitäjien olisi löydettävä keino tiedottaa henkilötietojen käyttämisestä niin VVA:n rekisteröidylle käyttäjille kuin rekisteröimättömille ja satunnaisille käyttäjille. Käyttäjille olisi ilmoitettava henkilötietojen käsittelystä mahdollisimman pian **ja viimeistään käsittelyn aikana**. Tämä ehto voi olla käytännössä erityisen vaikea täyttää.
60. Tietyistä yritysten erityispiirteistä ei myöskään saisi olla haittaa rekisteröityneille. Koska monet sidosryhmät ovat maailmanlaajuisia yrityksiä tai hyvin tunnettuja tietystä liiketoiminnasta (esimerkiksi televiestintä, sähköinen kaupankäynti, tietotekniikka, verkkotoiminta), tavan, jolla ne tarjoavat VVA-palveluita olisi oltava selkeä. Rekisteröityjen tulisi ymmärtää riittävien tietojen perusteella, onko heidän VVA:n käyttönsä liitetty muihin VVA-palvelun tarjoajan hallinnoimiin käsittelytoimiin (esimerkiksi televiestintään, sähköiseen kaupankäyntiin, tietotekniikkaan tai verkkotoimintaan) pelkän VVA:n käytön lisäksi.

Esimerkki 5:

VVA-suunnittelija tarjoaa myös sosiaalisen median alustan ja hakukoneen, ja edellyttää, että käyttäjä linkittää käyttäjätilinsä VVA:han. Yhdistämällä VVA:n käyttöön käyttötilin suunnittelija voi parantaa käyttäjiensä profiilia VVA:n käytön mukaan, asennetuilla sovelluksilla (tai ominaisuuksilla), tehtyjen tilausten perusteella jne. Näin ollen VVA-vuorovaikutus on käyttäjään liitetty uusi tiedonlähde. VVA-suunnittelijan olisi annettava

²⁴ Tietosuojatyöryhmän ohjeet WP260, 21 kohta.

²⁵ Yleinen tietosuoja-asetus, 4 artiklan 11 kohta.

käyttäjille sekä selkeät tiedot siitä, miten heidän tietojaan käsitellään kunkin palvelun osalta, että hallintalaitteet, joiden avulla käyttäjä voi valita, käytetäänkö tietoja profilointiin vai ei.

Suositukseset

61. Euroopan tietosuojaneuvosto suosittelee, että kun käyttäjille tiedotetaan käyttäjätilin tietosuojaselosteessa henkilötietojen VVA-käsittelystä ja tili on yhteydessä muihin itsenäisiin palveluihin (esim. sähköposti- tai verkko-ostoksiin), tietosuojaselosteessa olisi oltava selkeästi erillinen osio henkilötietojen VVA-käsittelystä.
62. Käyttäjien saamien tietojen on sisällettävä tarkat tiedot kaikesta suoritetusta tiedonkeruusta ja käsittelystä. Vaikka ääninäytteessä on joitakin metatietoja (esim. onko puhuja stressaantunut), ei ole automaattisesti selvää, analysoidaanko niitä. On ratkaisevan tärkeää, että rekisterinpitäjät ovat avoimia käsittelemiensä raakatietojen erityispiirteistä.
63. Lisäksi olisi aina oltava selvää, missä tilassa VVA on. Käyttäjien olisi pystyttävä määrittämään, kuunteleeko VVA kyseisellä hetkellä suljetulla virtapiirillään ja, etenkin, lähettääkö se tietoja taustaohjelmaansa. Näiden tietojen pitäisi olla myös vammaisten saatavilla, kuten värisokeiden ja kuurojen. Huomio olisi kiinnitettävä etenkin siihen, että VVA:ita voi käyttää niin, että laitteeseen ei tarvita katsekontaktia. Kaiken käyttäjäpalautteen, mukaan lukien tilamuutoksien, tulisi siis olla ainakin visuaalisessa ja akustisessa muodossa.
64. Erityistä huomiota olisi kiinnitettävä siihen, mahdollistavatko laitteet kolmannen osapuolen toimintojen lisäämisen (VVA-sovellukset). Vaikka käyttäjälle voidaan antaa joitakin yleisiä tietoja, kun hän lisää tällaisia toimintoja (kun otetaan huomioon, että se on käyttäjän valinta), laitteen normaalin käytön aikana eri ohjainten väliset rajat voivat olla paljon epäselvempiä. Toisin sanoen käyttäjälle ei välttämättä anneta riittävästi tietoa siitä, miten ja kenen toimesta heidän tietojaan käsitellään (ja missä määrin) kunkin pyynnön yhteydessä.
65. Kaikkien tallennettujen äänten käsittelyssä kerättyihin ja johdettuihin tietoihin perustuvien tietojen olisi myös oltava käyttäjien saatavilla yleisen tietosuoja-asetuksen 12 artiklan mukaisesti.
66. VVA-rekisterinpitäjien olisi tehtävä selväksi, minkälaista tietoa VVA voi saada ympäristöstään, kuten esimerkiksi muista huoneessa olevista ihmisistä, taustalla kuuluvasta musiikista, äänten käsittelystä lääketieteellisiin, markkinointi- tai muihin tarkoituksiin, lemmikkieläimistä jne.

3.4 Käyttötarkoituksen rajoittaminen ja oikeusperusta

67. VVA:lle esitettyjen äänikomentojen käsittelyn ilmeinen tarkoitus on pyynnön toteuttaminen. On kuitenkin usein muita tarkoituksia, jotka eivät ole niin ilmeisiä, kuten VVA:n luonnollisen kielen ymmärtämisen parantaminen kouluttamalla VVA-mallia koneoppimisen tekniikoilla. VVA:iden tavallisimpia henkilötietojen käsittelytarkoituksia ovat:
 -) käyttäjien pyyntöjen toteuttaminen
 -) VVA:n parantaminen kouluttamalla koneoppimismallia ja ihmisten suorittama tarkistus ja äänen litteraatioiden merkitseminen
 -) käyttäjän tunnistus (puhetietojen avulla)
 -) käyttäjäprofilointi sisällön tai mainonnan mukauttamiseksi.
68. Koska VVA:t toimivat välittäjinä ja johtuen siitä, miten ne on suunniteltu, ne käsittelevät monenlaisia henkilötietoja ja muita tietoja. Tämä mahdollistaa henkilötietojen käsittelyn

moniin tarkoituksiin, jotka ylittävät käyttäjien pyyntöihin vastaamisen ja jotka voivat jäädä heiltä huomaamatta. Analysoimalla VVA:lla kerättyjä tietoja voidaan tietää tai päätellä käyttäjien kiinnostuksen kohteita, aikatauluja, ajoreittejä ja tottumuksia. Tämä voisi mahdollistaa henkilötietojen käsittelyn ennalta arvaamattomiin tarkoituksiin (esim. mielipidelouhinta tai terveydentilan arviointi²⁶), mikä ylittäisi käyttäjien kohtuulliset odotukset huomattavasti.

69. Rekisterinpitäjien olisi määriteltävä selkeästi henkilötietojen käsittelyn tarkoitus kaikissa konteksteissa, joissa VVA:ita käytetään, jotta rekisteröidylle ei jäisi mitään epäselväksi (tarkoitukset voidaan esimerkiksi luokitella). Yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan mukaisesti henkilötiedot tulisi kerätä tiettyä nimenomaista ja laillista tarkoitusta varten, eikä niitä myöhemmin saa käsitellä näiden tarkoitusten kanssa yhteensopimattomalla tavalla.

3.4.1 Käyttäjien pyyntöjen toteuttaminen

70. VVA:n pääasiallinen käyttötarkoitus on antaa äänikomentoja, jotka VVA:n tai siihen liittyvän sovelluksen tai palvelun on suoritettava (esim. musiikin suoratoistopalvelu, kartoituspalvelu tai sähkölukko). Näin ollen saatetaan käsitellä käyttäjän ääntä ja mahdollisesti muita tietoja (esimerkiksi käyttäjän sijainti, kun pyydetään reittiä tiettyyn kohteeseen).

Esimerkki 6:

VVA:lla varustetun älyauton matkustaja pyytää reittiä lähimmälle huoltoasemalle. VVA käsittelee käyttäjän äänen ymmärtääkseen komennon ja auton sijainnin löytääkseen reitin, jonka se lähettää älykomponentille, joka näyttää reitin auton näytöllä.

71. Siltä osin kuin äänikomentojen käsittely edellyttää loppukäyttäjän päätelaitteisiin tallennettujen tietojen tallentamista tai käyttöä, on noudatettava sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohtaa. Vaikka 5 artiklan 3 kohtaan sisältyy yleinen periaate, jonka mukaan tällainen tallennus tai pääsy edellyttää loppukäyttäjän ennakkosuostumusta, siinä säädetään myös suostumusta koskevasta vaatimuksesta vapauttamisesta, jos se on ”ehdottoman välttämätöntä sellaisen tietoyhteiskunnan palvelun tarjoamiseksi, jota tilaaja tai käyttäjä on erityisesti pyytänyt”. Siltä osin kuin äänidataa käsitellään käyttäjän pyyntöjen täyttämiseksi, ne on vapautettu ennakkosuostumuksen vaatimuksesta.
72. Kuten edellä todettiin, kaikilla henkilötietojen käsittelytoimilla, jotka tehdään loppukäyttäjien päätelaitteissa olevan tiedon tallentamisen tai niihin pääsyn jälkeen, on oltava yleisen tietosuoja-asetuksen 6 artiklan mukainen oikeusperusta ollakseen laillisia.
73. VVA suorittaa kaksi peräkkäistä käsittelytoimintoa. Kuten edellä mainittiin, ensimmäinen edellyttää pääsyä VVA:han (ja siksi sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdan edellytykset on täytettävä). Sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdan ehtojen lisäksi tämä toinen vaihe edellyttää myös yleisen tietosuoja-asetuksen 6 artiklan mukaista oikeusperustaa.
74. Kun henkilö päättää käyttää VVA:ta, tämä tarkoittaa yleensä sitä, että alkuperäisen käyttäjän on ensin rekisteröitävä käyttäjätili VVA:n aktivoimiseksi. Toisin sanoen, tässä tilanteessa

²⁶ Eoghan Furey, Juanita Blue, “Alexa, Emotion, Privacy and GDPR”, konferenssiasiakirja, Human Computer Interaction -konferenssi, heinäkuu [2018].

käyttäjän ja VVA-rekisterinpitäjän välillä on sopimussuhde²⁷. Kun otetaan huomioon sopimuksen sisältö ja perustavoite, sopimuksen keskeisenä tarkoituksena on käyttää VVA:ta käyttäjän pyynnön täyttämiseen.

75. Mikä tahansa henkilötietojen käsittely, joka on tarpeen käyttäjän pyynnön toteuttamiseksi, voi näin ollen perustua sopimuksen täytäntöönpanon oikeusperustaan²⁸. Tällainen käsittely kattaa käyttäjän äänipyynnön tallentamisen, sen tekstiksi litteroimisen, sen tulkinnan, tiedot, jotka on vaihdettu vastauksen valmistelemiseksi tietolähteiden kanssa ja sen jälkeen käyttäjän pyynnön lopettavan lopullisen vastauksen äänen litteroinnin.
76. Sopimuksen täytäntöön paneminen voi olla oikeusperusta henkilötietojen käsittelemiselle koneoppimista käyttäen silloin, kun se on tarpeen palvelun tarjoamiseksi. Henkilötietojen käsittely koneoppimista käyttäen muihin tarkoituksiin, jotka eivät ole tarpeen, kuten palvelujen parantaminen, ei saisi perustua tähän oikeusperustaan.
77. Viimeisenä muttei vähäisimpänä voidaan todeta, että sopimuksen täytäntöönpanon oikeusperustaa ja yleisen tietosuoja-asetuksen mukaiseen suostumukseen perustuvaa oikeusperustaa ei pitäisi sekoittaa keskenään. Sopimuksen tekohetkellä eli sopimukseen suostumisen yhteydessä annettu suostumus on osa kyseisen sopimuksen voimassaoloa eikä viittaa yleisen tietosuoja-asetuksen mukaisen suostumuksen erityismerkitykseen²⁹.
78. Jos VVA:n käyttö ei edellytä, että käyttäjä luo VVA:lle käyttötilin ennen käyttöä, suostumus voisi olla mahdollinen oikeusperusta.

3.4.2 VVA:n parantaminen kouluttamalla koneoppimisjärjestelmiä ja tarkistamalla äänitiedostot ja litteraatiot manuaalisesti

79. Ihmisten puheessa on valtavasti aksentteja ja muuta vaihtelua. Vaikka kaikki VVA:t ovat heti käyttövalmiita, niiden suorituskyky voi parantua, kun niitä säädetään käyttäjän puheen ominaispiirteiden mukaisiksi. Kuten kohdassa 2.6 todetaan, tämä säätöprosessi perustuu koneoppimismenetelmiin ja koostuu kahdesta prosessista: VVA:n käyttäjiltä kerätyn uuden datan lisäämisestä koulutusaineistoon ja joidenkin pyyntöjen suorittamiseksi käsiteltyjen tietojen inhimillisestä tarkastelusta.

Esimerkki 7:

VVA:n käyttäjän on annettava kolme kertaa sama äänikomento, koska VVA ei ymmärrä sitä. Kolme äänikomentoa ja niiden litteraatiot välitetään ihmisille, jotka tarkistavat ja korjaavat litteraatiot. Äänikomennot ja tarkistetut litteraatiot lisätään VVA:n koulutusaineistoon VVA:n suorituskyvyn parantamiseksi.

80. Esimerkissä kuvattujen käsittelytoimien ei pitäisi katsoa olevan (välttämättä) *”tarpeen sopimuksen täytäntöönpanemiseksi”* yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan b alakohdassa tarkoitettulla tavalla, ja siksi ne edellyttävät yleisen tietosuoja-asetuksen 6 artiklan

²⁷ Edellyttäen, että *”sopimus on pätevä sovellettavan kansallisen sopimuslainsäädännön nojalla”*, ote ohjeista 2/2019 yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan b alakohdan perusteella tapahtuvasta henkilötietojen käsittelystä rekisteröidyille tarjottavien verkkopalvelujen yhteydessä, jatkossa *”ohjeet 2/2019”*, 26 kohta.

²⁸ Ohjeiden 2/2019 mukaisesti, joissa todetaan myös, että lausunto 06/2014 on edelleen merkityksellinen 6 artiklan 1 kohdan b alakohdan ja yleisen tietosuoja-asetuksen kannalta (ks. erityisesti lausunnon 6/2014 sivut 11, 16, 17, 18 ja 55).

²⁹ Ks. ohjeet 2/2019, 18, 19, 20, 21 ja 27 kohta.

sijaan toisen oikeusperustan. Pääasiallinen syy tähän on se, että VVA:t ovat heti ostettaessa käyttövalmiita, ja ne voivat jo toimia tavalla, joka on (välttämättä) tarpeen sopimuksen täytäntöönpanemiseksi. Euroopan tietosuojaneuvosto ei katso, että 6 artiklan 1 kohdan b alakohda on yleisesti asianmukainen laillinen peruste palvelun parantamista varten tai uusien toimintojen kehittämistä varten olemassa olevan palvelun sisällä. Useimmiten käyttäjä tekee sopimuksen hyödyntääkseen olemassa olevaa palvelua. Vaikka palvelun parantamista ja siihen tehtäviä muutoksia koskeva mahdollisuus saatetaan tavanomaisesti sisällyttää sopimusehtoihin, tällaista käsittelyä ei yleensä voida pitää objektiivisesti tarpeellisenä käyttäjän kanssa tehdyn sopimuksen täytäntöön panemiseksi.

3.4.3 Käyttäjän tunnistus³⁰ (puhetietojen avulla)

81. Puhetietojen käyttö käyttäjien tunnistamiseen edellyttää yleisen tietosuoja-asetuksen 4 artiklan 14 kohdassa määriteltyjen biometristen tietojen käsittelyä. Näin ollen rekisterinpitäjän on määriteltävä yleisen tietosuoja-asetuksen 9 artiklan mukainen poikkeus sen lisäksi, että se määrittelee yleisen tietosuoja-asetuksen 6 artiklan mukaisen oikeusperustan³¹.
82. Yleisen tietosuoja-asetuksen 9 artiklassa luetelluista poikkeuksista ainoastaan rekisteröityjen nimenomainen suostumus vaikuttaa soveltuvan tähän nimenomaiseen tarkoitukseen.
83. Koska tätä tarkoitusta varten on kuitenkin sovellettava yleisen tietosuoja-asetuksen 9 artiklan erityisiä oikeudellisia oloja, erityisten tietoryhmien käsittelyyn liittyvää lisätietoa on kohdassa 3.8.

3.4.4 käyttäjäprofilointi sisällön tai mainonnan mukauttamiseksi.

84. Kuten edellä mainittiin, VVA:illa on pääsy kaikkien äänikomentojen sisältöön myös silloin, kun ne on tarkoitettu kolmansien osapuolten tarjoamiin palveluihin. Tämä pääsyoikeus mahdollistaa sen, että VVA-suunnittelija voisi luoda erittäin tarkkoja käyttäjäprofiileja, joita voitaisiin käyttää yksilöllisten palveluiden tai mainosten tarjoamiseen.

Esimerkki 8:

Aina, kun VVA:n käyttäjä tekee Internet-haun, VVA lisää käyttäjäprofiiliin mielenkiinnonkohteista viestiviä nimiöitä. Kunkin uuden haun tulokset esitetään käyttäjälle järjesteltynä niin, että nämä nimiöt otetaan huomioon.

Esimerkki 9:

³⁰ Tunnistamisen käsite olisi periaatteessa erotettava varmentamisesta (todentamisesta). Tunnistaminen on yksi moneen -haku (1: N) ja -vertailu, ja se edellyttää periaatteessa tietokantaa, jossa on useita henkilöitä. Tästä poiketen, varmennusta varten toteutettu käsittely on yksi yhteen vertailua (1:1), ja sitä käytetään varmentamaan ja vahvistamaan biometrisellä vertailulla, onko henkilö sama henkilö kuin henkilö, jolta biometriset tiedot ovat peräisin. Euroopan tietosuojaneuvoston tietojen mukaan markkinoilla olevien VVA:iden toiminnot perustuvan yksinomaan kaiutinpohjaisten tunnistustekniikoiden käyttöön.

³¹ Yleisen tietosuoja-asetuksen mukaan pelkkä datan luonne ei aina riitä määrittämään, voidaanko sitä pitää erityisiin tietoryhmiin kuuluvana, koska valokuvien käsittely kuuluu ”biometristen tietojen määritelmän piiriin ainoastaan siinä tapauksessa, että niitä käsitellään erityisin teknisin menetelmin, jotka mahdollistavat luonnollisen henkilön yksilöllisen tunnistamisen tai todentamisen” (johdanto-osan 51 kappale). Sama perustelu koskee ääntä.

Aina kun VVA:n käyttäjä tekee ostoksen verkkokaupasta, VVA tallentaa ostotilauksesta tietueen. VVA-palvelun tarjoaja antaa kolmansille osapuolille mahdollisuuden kohdistaa VVA:n käyttäjälle mainoksia aiempien ostojen perusteella.

85. Sisällön mukauttaminen voi olla (mutta ei ole aina) VVA:n sisäänrakennettu ja odotettu ominaisuus. Se, voidaanko tällaista käsittelyä pitää VVA:n olennaisena osana, riippuu tarjotun palvelun luonteesta, keskimääräisen rekisteröidyn odotuksista, jotka koskevat sekä palvelua että tapaa, jolla palvelua mainostetaan käyttäjille, sekä siitä, voidaanko palvelu tarjota ilman räätälöintiä.³²
86. Jos räätälöinti tapahtuu sopimussuhteen yhteydessä ja osana loppukäyttäjän nimenomaista pyyntöä (ja käsittely rajoittuu siihen, mikä on ehdottoman välttämätöntä tämän palvelun tarjoamiseksi), tällainen käsittely voi perustua yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan b alakohtaan.
87. Jos käsittely ei ole ehdottomasti *"tarpeen sopimuksen täytäntöönpanemiseksi"* yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan b alakohdassa tarkoitetulla tavalla, VVA-palvelun tarjoajan on periaatteessa pyydettävä rekisteröidyn suostumus käsittelylle. Koska sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdassa edellytetään suostumusta tietojen tallentamiseen tai tietoihin pääsyyn (ks. edellä 28–29 kohta), yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan a alakohdan mukainen suostumus on periaatteessa myös asianmukainen oikeusperusta henkilötietojen käsittelylle näiden toimien jälkeen, koska oikeusperustan riippuminen rekisterinpitäjän oikeutetusta edusta saattaa tietyissä tapauksissa vaarantaa sähköisen viestinnän tietosuojadirektiivin 5 artiklan 3 kohdassa säädetyn lisäsuojan tason.
88. Mainontatarkoituksissa tehtävän käyttäjäprofiloinnin osalta on huomattava, että tätä tarkoitusta ei koskaan pidetä loppukäyttäjän nimenomaisesti pyytämänä palveluna. Näin ollen tätä tarkoitusta varten tapahtuvan käsittelyn yhteydessä olisi kerättävä käyttäjien suostumus järjestelmällisesti.

Suosituks

89. Käyttäjille olisi ilmoitettava henkilötietojen käsittelyn tarkoituksesta, ja tämän tarkoituksen olisi vastattava niitä odotuksia, joita heillä on ostamalleen laitteelle. VVA:n tapauksessa tämä tarkoitus – käyttäjän näkökulmasta – on selkeästi heidän äänikomentojensa käsittely, jonka ainoana tarkoituksena on tulkita heidän pyyntönsä ja vastata siihen merkityksellisesti (oli kyse sitten kyselyyn vastaamisesta tai muista reaktioista, kuten valokytimen kauko-ohjauksesta).
90. Kun henkilötietojen käsittely perustuu suostumukseen, jonka *"on oltava annettu yhtä tai useampaa erityistä tarkoitusta varten"* ja niin, että *"rekisteröidyllä on valinnan mahdollisuus kunkin tarkoituksen yhteydessä"*. Lisäksi rekisterinpitäjän, joka hakee suostumusta eri tarkoituksiin, olisi pyydettävä erillinen valinta kutakin tarkoitusta varten, jotta käyttäjät voivat antaa erityissuostumuksen tiettyjä tarkoituksia varten³³. Käyttäjien olisi esimerkiksi voitava erikseen hyväksyä tai olla hyväksymättä puheen litteraatioiden manuaalinen tarkastelu ja luokitteleminen ja puhetietojen käyttäminen käyttäjän tunnistamiseen/todentamiseen (ks. kohta 3.7).

³² Ks. myös Ohjeet 2/2019, 57 kohta.

³³ Ks. Euroopan tietosuojaneuvosto, Asetuksen 2016/679 mukaista suostumusta koskevat suuntaviivat 05/2020, hyväksytty 4. toukokuuta 2020, kohta 3.2.

3.5 Lapsia koskevien tietojen käsittely

91. Myös lapset voivat olla vuorovaikutuksessa VVA:iden kanssa tai luoda omia profiileja, jotka liittyvät aikuisten profiileihin. Jotkin VVA:t on upotettu laitteisiin, jotka on erityisesti suunnattu lapsille.
92. Kun käsittelyn oikeusperustana on sopimuksen täytäntöön paneminen, lapsien tietojen käsittelyn edellytykset riippuvat kansallisesta sopimusoikeudesta.
93. Kun käsittelyn oikeusperusta on suostumus ja yleisen tietosuoja-asetuksen 8 artiklan 1 kohdan mukaan lasten tietojen käsittely on laillista vain, *”jos lapsi on vähintään 16-vuotias”*. *”Jos lapsi on alle 16 vuotta, tällainen käsittely on lainmukaista vain siinä tapauksessa ja siltä osin kuin lapsen vanhempainvastuunkantaja on antanut siihen suostumuksen tai valtuutuksen.”* Näin ollen silloin, kun tietojen käsittelyn oikeusperusta on suostumus, vanhemmilta tai huoltajilta olisi pyydettävä nimenomainen lupa lapsen tietojen keräämiseen, käsittelemiseen ja tallentamiseen (esimerkiksi puhe- ja tekstitystietoja) yleisen tietosuoja-asetuksen noudattamiseksi.
94. VVA:ihin on tarjolla jonkin verran lapsilukkoja, mutta ne eivät ole käyttäjäystävällisiä nykyisissä muodoissaan (saattavat vaatia esimerkiksi kirjautumisen uuteen palveluun) tai niiden kapasiteetti on rajallinen. Rekisterinpitäjien olisi panostettava sellaisten keinojen kehittämiseen, joilla vanhemmat tai huoltajat voivat valvoa lastensa VVA:iden käyttöä.

3.6 Tietojen säilyttäminen

95. VVA:t käsittelevät ja tuottavat erilaisia henkilökohtaisia tietoja, kuten ääntä, litteraatioita, metadatasia tai järjestelmälokeja. Tällaisia tietoja voidaan käsitellä monenlaisiin tarkoituksiin, kuten palvelun tarjoamiseen, luonnollisen kielen käsittelyn parantamiseen, palvelun yksilöimiseen tai tieteelliseen tutkimukseen. Yleisen tietosuoja-asetuksen tietojen säilytyksen rajoittamista koskevan periaatteen mukaisesti riskienhallintaviranomaisten olisi säilytettävä tietoja enintään sen ajan, joka on tarpeen niitä tarkoituksia varten, joihin henkilötietoja käsitellään. Näin ollen tietojen sallitut säilytysajat olisi sidottava eri käsittelytarkoituksiin. VVA-palvelun tarjoajien tai kolmansien osapuolten, jotka tarjoavat palveluja VVA:iden kautta, olisi arvioitava kullekin datajoukolle ja käyttötarkoitukselle enimmäissäilytysajat.
96. Tietojen minimoinnin periaate liittyy läheisesti tietojen säilytyksen rajoittamista koskevaan periaatteeseen. Rekisterinpitäjien on rajoitettava tietojen säilytysajan lisäksi sekä tietojen tyyppiä että määrää.
97. Rekisterinpitäjien olisi kysyttävä itseltään muun muassa seuraavia kysymyksiä: Onko tarpeen tallentaa kaikki äänitteet tai kaikki litteraatiot, jotta tavoite X saavutetaan? Onko äänidatan säilyttäminen tarpeellista, kun litteraatio on tallennettu? Jos on, mihin tarkoitukseen? Kuinka kauan puhe- tai litterointitiedot ovat tarpeen kutakin tarkoitusta varten? Vastauksissa näihin ja muihin vastaaviin kysymyksiin määritellään säilytysajat, joiden pitäisi olla osa rekisteröityjen saatavilla olevia tietoja.
98. Jotkin VVA:t tallentavat oletusarvoisesti henkilökohtaisia tietoja, kuten äänikatkelmia tai litteraatioita määrittämättömäksi ajaksi antaen käyttäjille mahdollisuuden itse poistaa kyseiset tiedot. Henkilötietojen säilyttäminen toistaiseksi on säilytyksen rajoittamista koskevan periaatteen vastaista. Se, että rekisteröidylle annetaan mahdollisuus poistaa henkilötietonsa, ei poista rekisterinpitäjän vastuuta tietojen säilyttämisperiaatteen määrittelystä ja täytäntöönpanosta.

99. VVA-suunnittelussa on muistettava ottaa huomioon ohjaimet, joita käyttämällä käyttäjä voi poistaa henkilötietonsa laitteistaan ja kaikista etätallennusjärjestelmistä. Nämä ohjaimet voivat olla tarpeen käyttäjien erityyppisten pyyntöjen ratkaisemisessa, esimerkiksi poistamispyyntöjen tai aiemmin annetun suostumuksen peruuttamisessa. Tätä vaatimusta ei ole otettu huomioon joidenkin VVA:iden suunnittelussa.³⁴
100. Kuten muissakin yhteyksissä, rekisterinpitäjien on ehkä säilytettävä henkilötiedot todisteena käyttäjälle tarjotusta palvelusta täyttääkseen oikeudelliset velvoitteet. Rekisterinpitäjä voi säilyttää henkilötietoja tämän perusteella. Kyseisten oikeudellisten velvoitteiden noudattamiseksi säilytettävien tietojen määrän olisi kuitenkin oltava mahdollisimman pieni ja niitä olisi säilytettävä mahdollisimman vähän aikaa. Oikeudellisen velvoitteen noudattamiseksi säilytetyjä tietoja ei tietenkään pitäisi käyttää mihinkään muuhun tarkoitukseen ilman yleisen tietosuoja-asetuksen 6 artiklan mukaista oikeusperustaa.

Esimerkki 10:

Käyttäjä ostaa television verkkokaupasta VVA:lle annetulla äänikomennolla. Vaikka käyttäjä pyytäisi tietojensa poistamista jälkikäteen, VVA-palvelun tarjoaja tai kehittäjä voi edelleen säilyttää joitakin tietoja ostotodisteiden säilyttämiseksi verosäännöksessä asetetun oikeudellisen velvoitteen perusteella. Tätä tarkoitusta varten tallennettujen tietojen määrä ei kuitenkaan saisi ylittää lakisääteisen velvoitteen noudattamisen edellyttämää vähimmäismäärää. Tietoja ei myöskään saa käsitellä muihin tarkoituksiin ilman yleisen tietosuoja-asetuksen 6 artiklan mukaista oikeusperustaa.

101. Kuten kohdassa 2 todetaan, VVA:n puheentunnistuskkyky paranee kouluttamalla koneoppimisjärjestelmiä käyttäjien tiedoilla. Jos käyttäjät eivät anna suostumustaan tai peruuttavat suostumuksensa tietojensa käyttöön tällaisissa tarkoituksissa, heidän tietojensa ei voida laillisesti käyttää enää uusien mallien kouluttamiseen. Jos oletetaan, että säilytyksen jatkamiselle ei ole muuta syytä, rekisterinpitäjän olisi poistettava tiedot. On kuitenkin näyttöä siitä, että joihinkin koneoppimisen malleihin saattaa liittyä uudelleentunnistamisriski.³⁵
102. Rekisterinpitäjien ja henkilötietojen käsittelijöiden olisi käytettävä malleja, jotka eivät rajoita heidän kykyään keskeyttää käsittely, jos henkilö peruuttaa suostumuksensa, eikä heidän myöskään olisi käytettävä malleja, jotka rajoittavat heidän kykyään edistää rekisteröityjen oikeuksia. Rekisterinpitäjien ja henkilötietojen käsittelijöiden olisi sovellettava lieventäviä toimenpiteitä uudelleentunnistamisriskin pienentämiseksi hyväksytyn kynnsarvon tasolle.
103. Mikäli käyttäjä peruuttaa suostumuksensa, käyttäjältä kerättyjä tietoja ei voida enää käyttää mallin jatkokoulutukseen. Mallia, jota on aiemmin koulutettu näitä tietoja käyttäen, ei kuitenkaan tarvitse poistaa. Euroopan tietosuojaneuvosto korostaa kuitenkin, että on olemassa todisteita siitä, että joidenkin koneoppimisen mallien käytössä voi olla henkilötietovuotojen riski. Useat tutkimukset ovat etenkin osoittaneet, että malleihin on mahdollista tehdä rekonstruktio- ja jäsenyyden päättely -hyökkäyksiä, jotka antavat hyökkääjälle mahdollisuuden saada tietoja yksilöistä.³⁶ Tästä syystä rekisterinpitäjien ja

³⁴ Ks. Amazonin 28 päivänä kesäkuuta 2019 päivätty kirje Yhdysvaltain senaattorille Christopher Coonsille: [https://www.coons.senate.gov/imo/media/doc/Amazon%20Senator%20Coons_Response%20Letter_6.28.19\[3\].pdf](https://www.coons.senate.gov/imo/media/doc/Amazon%20Senator%20Coons_Response%20Letter_6.28.19[3].pdf)

³⁵ Veale Michael, Binns Reuben ja Edwards Lilian, 2018 ”Algorithms that remember: model inversion attacks and data protection law” Phil. Trans. R. Soc. A.37620180083, doi: 10.1098/rsta.2018.0083

³⁶ N. Carlini ym., ”Extracting Training Data from Large Language Models” Joulukuu 2020.

henkilötietojen käsittelijöiden olisi sovellettava lieventämistoimenpiteitä, joilla he voivat pienentää uudelleentunnistamisriskiä hyväksytyn kynnysarvon tasolle sen varmistamiseksi, että ne käyttävät malleja, jotka eivät sisällä henkilötietoja.

104. Rekisteröityjä ei pitäisi suostutella säilyttämään tietojaan toistaiseksi. Vaikka tallennetun äänidatan tai litteraatioiden poistaminen saattaa vaikuttaa palvelun suorituskykyyn, käyttäjille olisi kuitenkin kerrottava tästä selkeästi ja mitattavissa olevalla tavalla. VVA-palveluntarjoajien tulee välttää yleisen tason lausuntoja palvelun heikkenemisestä henkilötietojen poistamisen jälkeen.
105. Äänitallenteiden anonymisointi on erityisen haastavaa, sillä käyttäjät voidaan tunnistaa viestin sisällön ja itse äänen ominaisuuksien avulla. Jonkin verran tutkitaan³⁷ kuitenkin tekniikoita, joilla voidaan poistaa tilannetietoja, kuten taustääniä, ja tehdä äänistä anonyymejä.

Suositukseset

106. Käyttäjän näkökulmasta päätarkoituksena on tiedustella ja vastaanottaa vastauksia ja/tai käynnistää toimenpiteitä, kuten musiikin soittaminen tai valojen kytkeminen päälle tai pois päältä. Kun kyselyyn tai komentoon on vastattu, henkilötiedot olisi poistettava, ellei VVA-suunnittelijalla tai -kehittäjällä ole pätevää oikeusperustaa niiden säilyttämiseksi tiettyä tarkoitusta varten.
107. Ennen kuin VVA-palvelujen tarjoajien ja kehittäjät harkitsevat tietojen tallentamisen rajoittamisen periaatteen toteuttamista anonymisoinnin keinoin, heidän olisi varmistettava, että ääni muutetaan anonymisointiprosessissa sellaiseksi, että sitä ei voi tunnistaa.
108. Toiminta-asetusten määrittämisen oletusasetusten pitäisi vastata näitä vaatimuksia siten, että oletuksena on, että tallennettuja käyttäjätietoja on mahdollisimman vähän. Jos nämä vaihtoehdot esitetään ohjatun asennustoiminnon osana, oletusasetuksen tulisi heijastaa tätä ja kaikki vaihtoehdot tulisi esittää yhtäläisinä mahdollisuuksina ilman visuaalista syrjintää.
109. Kun VVA-palvelun tarjoaja tai kehittäjä havaitsee tarkastelumenettelyn aikana nauhoituksen olevan peräisin virheellisestä aktivoinnista, nauhoitus ja kaikki siihen liittyvät tiedot on poistettava välittömästi eikä niitä saa käyttää mihinkään tarkoitukseen.

3.7 Turvallisuus

110. Jotta riskienhallintaviranomaiset voisivat käsitellä henkilötietoja turvallisesti, heidän olisi suojattava luottamuksellisuuttaan, koskemattomuuttaan ja käytettävyyttään. Joidenkin VVA-ekosysteemin tekijöistä aiheutuvien riskien lisäksi myös se, että VVA:n viestintäkeino on ääni, luo uudenlaisia turvallisuusriskejä.
111. VVA:lla voi olla monta käyttäjää. Ne voivat sallia useamman kuin yhden rekisteröidyn käyttäjän, ja kuka tahansa heidän ympärillään voi antaa komentoja ja käyttää palveluita. Kaikkiin luottamuksellisuutta edellyttäviin VVA-palveluihin liittyy jokin pääsynvalvontajärjestelmä ja käyttäjän todentaminen. Ilman pääsynvalvontaa kuka tahansa, joka voi antaa äänikomentoja VVA:lla, voi käyttää, muokata tai poistaa kenen tahansa käyttäjän henkilötietoja (esim. pyytää vastaanotettuja viestejä, käyttäjän osoitetta tai

³⁷ Ks. esimerkiksi VoicePrivacy (<https://www.voiceprivacychallenge.org>), aloite puheteknologian tietosuojaratkaisujen kehittämiseksi.

Ks. myös Horisontti 2020 ohjelmaan kuuluva tutkimus- ja innovointihankkeen COMPRISE:n puitteissa kehitetyt avoimen lähdekoodin äänenanonymisointityökalut https://gitlab.inria.fr/comprise/voice_transformation.

kalenteritapahtumia). Äänikomentojen antaminen VVA:lle ei edellytä sitä, että käyttäjä on fyysisesti lähellä VVA:ta, koska niitä voidaan ohjata esimerkiksi signaalilähetyksenä³⁸ (esim. radio tai TV). Ihmisen aistit eivät edes havaitse joitakin tunnettuja menetelmiä, joilla VVA-laitteille voidaan antaa etäkomentoja, kuten lasereita³⁹ tai äänettömiä ultraääniaaltoja⁴⁰.

112. Käyttäjän todentaminen voi perustua yhteen tai useampaan seuraavista tekijöistä: käyttäjän tietämiin asioihin (kuten salasana), käyttäjän omaisuuteen (kuten älykortti) tai käyttäjän olemukseen (kuten äänijälki). Näiden todentamistekijöiden tarkempi tarkastelu äänikomennoin ohjattavien VVA-kontekstissa osoittaa, että
-)] todennus, joka perustuu käyttäjän tietämiin asioihin, on ongelmallista. Salaisuus, jonka avulla käyttäjät voivat todistaa henkilöllisyytensä, olisi puhuttava ääneen ja paljastettava kaikille ympärillä oleville. VVA viestii ympäröivän ilman välityksellä, eli sitä ei voida vahvistaa perinteisten kanavien tapaan (esim. rajoittamalla pääsyä kanavaan tai salaamalla sen sisältö).
 -)] Todentaminen käyttäjän omistamien asioiden avulla edellyttäisi, että VVA-palvelun tarjoajat luovat, jakavat ja hallitsevat tunnistusvälineitä, joita voitaisiin käyttää henkilöllisyyden todistamiseen.
 -)] Todentaminen käyttäjän olemuksen avulla tarkoittaisi biometristen tietojen käyttöä luonnollisen henkilön yksilölliseen tunnistamiseen (ks. kohta 3.7 jäljempänä).
113. VVA-käyttäjätilit liittyvät laitteisiin, joissa palvelua tarjotaan. Usein samaa tiliä, jota käytetään VVA:n hallintaan, käytetään muiden palvelujen hallintaan. Esimerkiksi Android-matkapuhelimen ja Google Home -kaiuttimen omistajat voivat todennäköisesti liittää Google-tilinsä molempiin laitteisiin. Useimmat VVA:t eivät vaadi tai tarjoa tunnistus- tai todentamismekanismeja, jos VVA-palvelua tarjoavalla laitteella on vain yksi käyttäjätili.
114. Jos laitteeseen liittyy useampi kuin yksi käyttäjätili, jotkin VVA:t tarjoavat valinnaisen PIN-numeroon perustuvan peruspääsynvalvontatoiminnon ilman todellista käyttäjän todentamista. Joissakin muissa VVA:issa on mahdollisuus tunnistautua äänijäljellä.
115. Vaikka kaikkien VVA-palveluiden käyttö ei välttämättä edellyttäisikään käyttäjän tunnistusta tai todennusta, joissakin VVA-palveluissa se on pakollista. Ilman tunnistus- tai todentamismekanismeja kuka tahansa voi käyttää muiden käyttäjien tietoja ja muokata tai poistaa niitä mielensä mukaan. Kuka tahansa älykaiuttimen lähellä oleva voi esimerkiksi poistaa muiden käyttäjien soittolistoja musiikin suoratoistopalvelusta, komentoja komentohistoriasta tai yhteystietoja yhteystietoluettelosta.
116. Useimmat VVA:t luottavat sokeasti paikallisiin verkkoihinsa. Mikä tahansa samassa verkossa oleva vaarantunut laite voi muuttaa älykaiuttimen asetuksia, sallia haittaohjelmien asentamisen tai liittää siihen väärennettyjä sovelluksia/taitoja ilman käyttäjän tietoa tai suostumusta.⁴¹

³⁸ X. Yuan ym., "All Your Alexa Are Belong to Us: A Remote Voice Control Attack against Echo" 2018 kansainvälisen sähkö- ja elektroniikkainsinöörien järjestön Global Communications -konferenssi (GLOBECOM), Abu Dhabi, Yhdistyneet arabiemiirikunnat, 2018, s. 1–6, doi: 10.1109/GLOCOM.2018.8647762.

³⁹ Ks. esim. <https://lightcommands.com>

⁴⁰ Ks. esim. <https://surfingattack.github.io>

⁴¹ Ks. esimerkiksi Deepak Kumar ym., *Skill Squatting Attacks on Amazon Alexa*, USENIX Security Symposium, elokuu 2018, <https://www.usenix.org/conference/usenixsecurity18/presentation/kumar>

117. VVA:t, kuten muutkin ohjelmistot, ovat alttiita ohjelmiston haavoittuvuuksille. VVA-markkinoiden keskittymisen⁴² vuoksi tämä haavoittuvuus voi kuitenkin vaikuttaa miljooniin VVA-käyttäjiin. Jos VVA:t toimivat suunnitellusti, ne eivät lähetä mitään tietoa puheentunnistuksen pilvipalveluun, ennen kuin herätyslauseke havaitaan. Ohjelmiston heikkoudet voivat kuitenkin antaa hyökkääjän ohittaa VVA:n asentamis- ja suojaustoimenpiteet. Tällöin hyökkääjä voisi esimerkiksi saada kopion kaikista VVA-pilveen lähetetyistä tiedoista ja välittää ne palvelimelleen.
118. VVA:iden laillisesti käsittelemät tai niistä johdetut tiedot mahdollistavat melko tarkan käyttäjäprofiilin luomisen, koska VVA tietää tai voi päätellä käyttäjien sijainnin, suhteet, ja mieltymykset. VVA:ita on yhä enemmän käyttäjien kodeissa ja älypuhelimissa. Tämä lisää joukkovalvonnan ja -profiloinnin riskiä. Näin ollen turvallisuustoimenpiteiden, joilla turvataan sekä liikkuva että laitteissa ja pilvessä oleva data, olisi oltava sovitettuja näihin riskeihin.
119. VVA:iden lisääntyvä käyttö ja se, että lainvalvontaviranomaisten käyttöoikeuksia ei ole tasapainotettu tarpeeksi, voisivat yhdessä aiheuttaa sellaisia pelottavia vaikutuksia, jotka heikentäisivät perusoikeuksia, kuten sananvapautta.
120. Lainvalvontaviranomaiset sekä EU:ssa⁴³ että sen ulkopuolella⁴⁴ ovat jo ilmaisseet halukkuutensa päästä käsiksi VVA:iden kaappaamiin äänikatkelmiin. EU:ssa VVA:iden käsittelemien tai niistä johdettujen tietojen saannin olisi noudatettava EU:n nykyistä tietosuojaa ja yksityisyyden suojaa koskevaa sääntelykehystä. Jos jotkin jäsenvaltiot harkitsevat sellaisen erityislainsäädännön antamista, jolla rajoitettaisiin yksityisyyden suojaa ja tietosuojaa koskevia perusoikeuksia, kyseisten rajoitusten olisi aina oltava yleisen tietosuojasetuksen 23 artiklassa säädetyn vaatimuksen mukaisia⁴⁵.
121. Äänitallenteiden ja niihin liittyvien tietojen inhimillinen tarkastelu VVA-palvelujen laadun parantamiseksi on yleinen käytäntö VVA-palveluntarjoajien keskuudessa. Koska nämä ihmistarkastajat käsittelevät arkaluonteisia tietoja ja koska prosessi ulkoistetaan usein alihankintana henkilötietojen käsittelijälle, on erittäin tärkeää, että otetaan käyttöön oikeamittaiset turvatoimet.

Suosituksat

122. VVA-suunnittelijoiden ja sovelluskehittäjien olisi tarjottava käyttäjille uusimpia turvallisia todennusmenettelyjä.
123. Ihmistarkastajien on aina saatava ainoastaan tarpeelliset salanimetyt tiedot. Tarkastusta koskevissa oikeudellisissa sopimuksissa olisi nimenomaisesti kiellettävä käsittely, joka voisi johtaa rekisteröidyn tunnistamiseen.
124. Jos VVA tarjoaa hätäpuhelupalvelua, sen on taattava vakaa käyttöaika⁴⁶.

Security Research Labs, *Smart Spies: Alexa and Google Home expose users to vishing and eavesdropping*, marraskuu 2019, <https://srlabs.de/bites/smart-spies>

⁴² VVA-markkinat jakautuvat tällä hetkellä alle kymmenen palvelun tarjoajan kesken.

⁴³ Ks. esim. <https://www.ft.com/content/ad765972-87a2-11e9-a028-86cea8523dc2>.

⁴⁴ Ks. esim. <https://cdt.org/insights/alexa-is-law-enforcement-listening>.

⁴⁵ Ks. myös Euroopan tietosuojaneuvosto, Guidelines 10/2020 on restrictions under Article 23 GDPR.

⁴⁶ Aika, jonka laite tai palvelu voidaan jättää valvomatta ilman, että sen ohjelmisto kaatuu tai että se pitää käynnistää uudelleen hallinta- tai huoltotarkoituksessa.

3.8 Erityisiä henkilötietoryhmiä koskeva käsittely

125. Kuten edellä mainittiin, VVA:lla on pääsy intiimeihin tietoihin, joita voidaan suojata yleisen tietosuoja-asetuksen 9 artiklan nojalla (ks. kohta 3.7.1), kuten biometriset tiedot (ks. kohta 3.7.2). Tästä syystä VVA-suunnittelijoiden ja -kehittäjien on selvitettävä huolellisesti, missä tapauksissa käsittely edellyttää tietojen luokittelua erityisiin tietoryhmiin.

3.8.1 Erityisien tietoryhmien käsittelyyn liittyvät yleiset näkökohdat

126. VVA:t voivat käsitellä erityisiä tietoryhmiä erilaisissa olosuhteissa:

-)] Osana omia palveluitaan, esimerkiksi kun se hoitaa käyttäjän kalenterissa olevia terveysterveyspalveluvarauksia
-)] Kun VVA-palvelun tarjoajat toimivat rajapintana kolmannen osapuolen palveluille, ne käsittelevät komentojen sisältöä. VVA-palvelun tarjoajat voivat käsitellä erityisiä tietoryhmiä käyttäjän pyytämän palvelun tyypistä riippuen. Esimerkkinä tästä ovat tilanteet, joissa käyttäjä antaa VVA:lle komentoja käyttääkseen kolmannen osapuolen sovellusta ovulaationsa seuraamiseen.⁴⁷
-)] Kun puhetietoja käytetään käyttäjän yksilölliseen tunnistamiseen alla kuvatulla tavalla.

3.8.2 Biometristen tietojen käsittelyyn liittyvät erityishuomiot

127. Jotkin VVA:t pystyvät tunnistamaan käyttäjänsä jopa vain heidän äänensä perusteella. Tätä prosessia kutsutaan äänimallin tunnistamiseksi. Äänentunnistuksen ilmoittautumisvaiheessa VVA käsittelee käyttäjän ääntä ja luo äänimallin (tai äänijäljen). Tavallisen käytön aikana VVA voi laskea minkä tahansa käyttäjän äänimallin ja verrata sitä rekisteröityihin malleihin tunnistaa käyttäjän, joka on suorittanut komennon.

Esimerkki 11:

Ryhmä käyttäjiä asentaa VVA:n niin, että se käyttää äänimallin tunnistusta. Tämän jälkeen kukin heistä rekisteröi oman äänimallinsa.

Myöhemmin käyttäjä pyytää VVA:lta tietoa kalenterissaan olevista kokouksista. Koska pääsy kalenteriin edellyttää käyttäjän tunnistamista, VVA poimii mallin pyynnön äänestä, laskee sen äänimallin ja tarkistaa, vastaako se rekisteröitynyttä käyttäjää ja onko kyseisellä käyttäjällä kalenterin käyttöoikeus.

128. Edellä olevassa esimerkissä se, että käyttäjän ääni tunnustetaan äänimallin perusteella, on yleisen tietosuoja-asetuksen 9 artiklassa tarkoitettua erityisten henkilötietoryhmien käsittelyä (biometristen tietojen käsittely henkilön yksiselitteistä tunnistamista varten).⁴⁸ Esimerkissä edellytetty biometristen tietojen käsittely käyttäjän tunnistamista varten edellyttää rekisteröidyn (tai rekisteröityjen) nimenomaista suostumusta (yleisen tietosuoja-asetuksen 9 artiklan 2 kohdan a alakohta). Näin ollen rekisterinpitäjien on käyttäjien suostumusta saadessaan noudatettava yleisen tietosuoja-asetuksen johdanto-osan 32 kappaleessa selitettyjä 7 artiklan ehtoja ja tarjottava jokin muu tunnistamismenetelmä vaihtoehtona biometrisille tiedoille, ottaen huomioon, että suostumuksen on oltava vapaasti annettu.

⁴⁷ Ks. esimerkiksi saatavilla oleva tuote: <https://www.amazon.com/Ethan-Fan-Ovulation-Period-Tracker/dp/B07CRLSHKY>

129. Kun puhetietoja käytetään biometriseen tunnistamiseen tai todentamiseen, rekisterinpitäjien on tehtävä selväksi, milloin biometristä tunnistusta käytetään ja miten äänijäljet (biometriset mallit) tallennetaan ja levitetään laitteesta toiseen. Avoimuusvaatimuksen täyttämiseksi Euroopan tietosuojaneuvosto suosittelee seuraaviin kysymyksiin vastaamista:
-) Aktivoiko äänitunnistamisen aktivointi yhdessä laitteessa automaattisesti äänitunnistamisen kaikissa muissa samaa tiliä käyttävissä laitteissa?
 -) Levittäytyykö äänitunnistamisen aktivointi VVA:n rekisterinpitäjän infrastruktuurin kautta muiden käyttäjien omistamiin laitteisiin?
 -) Missä biometriset mallit luodaan, tallennetaan ja sovitetaan yhteen?
 -) Ovatko biometriset mallit VVA-palvelun tarjoajan, kehittäjien tai muiden saatavilla?
130. Kun rekisteröity käyttäjä asentaa VVA:n tunnistamaan käyttäjiensä äänet, myös rekisteröimättömien ja tahattomien käyttäjien ääni käsitellään, jotta VVA pystyy yksilöimään ja tunnistamaan käyttäjän.
131. Oikean puhujan äänen tunnistamisprosessiin kuuluu myös äänen vertaaminen VVA:n läheisyydessä olevien ihmisten ääniin. Toisin sanoen äänikomennoin ohjattavissa VVA:issa toteutettava äänentunnistamistoiminto voi edellyttää samassa kotitaloudessa puhuvien äänen biometristen tietojen tallentamista, jotta käyttäjän ääniominaisuudet voidaan erottaa sen henkilön äänestä, joka haluaa tulla tunnistetuksi. Näin ollen biometrinen tunnistaminen voi johtaa siihen, että henkilöt joutuvat tietämättään biometrisen käsittelyn kohteiksi, kun VVA rekisteröi heidän mallinsa verratakseen sitä tunnistamista haluavan käyttäjän malliin.
132. Jotta voitaisiin välttää tällainen ilman rekisteröityjen tietoa tapahtuva biometristen tietojen kerääminen, ja jotta VVA voisi silti tunnistaa käyttäjän, etusijalle olisi asetettava ratkaisu, jotka perustuvat pelkästään käyttäjän tietoihin. Käytännössä tämä tarkoittaa sitä, että biometrisen tunnistamisen olisi aktivoiduttava ainoastaan käyttäjän aloitteesta, eikä avustajan kuulemien äänten jatkuvan analysoinnin seurauksena. Läsnä oleville henkilöille voitaisiin esimerkiksi antaa erityinen avainsana tai kysymys, jotta heiltä saataisiin suostumus biometrisen käsittelyn käynnistämiseen. Käyttäjä voi esimerkiksi sanoa ”tunnistus” tai VVA voi kysyä ”haluatko tunnistautua” ja odottaa myöntyvää vastausta biometrisen käsittelyn aktivoimiseksi.

Esimerkki 12:

Jos käyttäjä haluaa ottaa käyttöön biometrisen todentamisen tilanteissa, joissa hän haluaa päästä käsiksi tiettyihin suojattuihin tietoihin, kuten esimerkiksi pankkitiliinsä, VVA voisi aktivoida puhujan todentamisen vain, kun kyseinen sovellus käynnistetään, ja näin varmistaa käyttäjän henkilöllisyyden.

Suosituks

133. Äänimalleja saa luoda, tallentaa ja liittää ainoastaan paikalliseen laitteeseen, ei etäpalvelimiin.
134. Äänitunnisteiden herkkyyden vuoksi standardeja, kuten ISO/IEC 24745, ja biometrisen mallisuojausten⁴⁹ tekniikoita tulee soveltaa perusteellisesti.

⁴⁹ Ks. esim.

135. Jos VVA käyttää äänipohjaista biometristä tunnistusta, VVA-palveluntarjoajien olisi:
-) varmistettava, että tunnistaminen on riittävän tarkkaa, jotta henkilötiedot voidaan yhdistää luotettavasti oikeisiin rekisteröityihin.
 -) varmistettava, että kaikkien käyttäjäryhmien tunnistaminen on yhtä tarkkaa tarkistamalla, että eri väestöryhmiin ei kohdistu tiedostamattomia ennakoasenteita.

3.9 Tietojen minimointi

136. Rekisterinpitäjien olisi minimoitava tietojen määrä, jotka on kerätty suoraan tai välillisesti prosessoimalla ja analysoimalla, ja jättää analysoimatta esimerkiksi käyttäjän ääni tai muita äänitietoja, joista voidaan saada tietoa hänen henkisestä tilastaan, mahdollisesta sairaudestaan tai elämänolosuhteistaan.
137. Käyttöön otettavien oletusasetusten olisi rajoitettava tietojen kerääminen ja/tai käsittely vähimmäismäärään, joka tarvitaan palvelun tarjoamiseen.
138. Sijainnista, kontekstista ja mikrofoniin herkkyydestä riippuen VVA voi kerätä kolmansien osapuolten puhetietoja taustamelun mukana käyttäjien ääntä kerätessään. Vaikka taustamelu ei sisältäisi puhetietoja, se voi silti sisältää tilannetietoja, jota voitaisiin käsitellä käyttäjään liittyvien tietojen (esim. sijainti) pääättelemiseen.

Suosituks

139. VVA-suunnittelijoiden olisi harkittava sellaisten teknologioiden kehittämistä, jotka poistavat taustamelun, jotta vältetään taustäänien ja tilannetietojen kirjaaminen ja käsittely.

3.10 Vastuu

140. Suostumukseen perustuvan käsittelyn osalta rekisterinpitäjien on voitava todistaa rekisteröityjen suostumus yleisen tietosuojasetuksen 7 artiklan 1 kohdan mukaisesti. Puhetietoja voidaan käyttää vastuuvollisuuteen osoittamiseen (esim. suostumuksen osoittamiseen). Tällaisten puhetietojen säilyttämistä koskevat vaatimukset määräytyisivät asiaa koskevan erityislainsäädännön vastuuvollisuutta koskevien vaatimusten mukaan.
141. Arvioidessaan tietosuojan vaikutustenarvioinnin tarvetta Euroopan tietosuojaneuvosto määrittelee perusteet⁵⁰, joita tietosuojaviranomaisten on käytettävä laatiessaan luetteloita käsittelytoimista, jotka edellyttävät tietosuojan vaikutustenarviointia, ja antaessaan esimerkkejä käsittelyistä, jotka todennäköisesti edellyttävät tietosuojan vaikutustenarviointia. On hyvin todennäköistä, että VVA-palvelut kuuluvat niihin luokkiin ja olosuhteisiin, joiden on todettu tarvitsevan tietosuojan vaikutustenarviointia. Tähän sisältyy myös sen arviointi, voiko laite tarkkailla tai valvoa rekisteröityä tai suorittaa järjestelmällistä valvontaa laajamittaisesti 35 artiklan 3 kohdan c alakohdan mukaisesti, käyttää "uutta teknologiaa" tai käsitellä arkaluonteisia tietoja tai haavoittuvassa asemassa olevia rekisteröityjä koskevia tietoja.

Jain, Anil & Nandakumar, Karthik ja Nagar, Abhishek. (2008). "*Biometric Template Security*". EURASIP Journal on Advances in Signal Processing. 2008. 10.1155/2008/579416.

S. K. Jami, S. R. Chalamala ja A. K. Jindal, "Biometric Template Protection Through Adversarial Learning" *Biometric Template Protection Through Adversarial Learning* 2019 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 2019, s. 1–6, doi: 10.1109/ICCE.2019.8661905.

⁵⁰ 29 artiklan mukaisen tietosuojatyöryhmän ohjeet tietosuojaa koskevasta vaikutustenarvioinnista, WP248, rev.01, jonka Euroopan tietosuojaneuvosto on hyväksynyt.

142. Kaikki tietojen keruu ja käsittely on dokumentoitava yleisen tietosuoja-asetuksen 30 artiklan mukaisesti. Tämä koskee kaikkea puhetietojen käsittelyä.

Suosituks

143. Jos käyttäjille on tarkoitus toimittaa 13 artiklan mukaiset tiedot ääniviesteillä, rekisterinpitäjien olisi julkaistava kyseiset viestit verkkosivustollaan, jotta ne ovat käyttäjien ja tietosuojaviranomaisten saatavilla.

3.11 Sisäänrakennettu ja oletusarvoinen tietosuoja

144. VVA-palvelun tarjoajien ja kehittäjien olisi harkittava kunkin toiminnallisuutensa kohdalla, onko käyttäjien rekisteröityminen pakollista. Vaikka on selvää, että kalenterin tai osoitekirjan hallinnoimiseksi käyttäjän on oltava rekisteröity, ei ole niin selvää, että puhelun soittaminen tai Internet-haku edellyttäisi, että VVA:lla on rekisteröity käyttäjä.
145. Oletusarvoisesti palvelut, jotka eivät edellytä tunnistettua käyttäjää, eivät saisi liittää VVA-tunnistettuja käyttäjiä annettuihin komentoihin. Yksityisyyden suojaa ja tietosuojaa oletusarvoisesti edistävä VVA käsittelisi käyttäjien tietoja vain käyttäjien pyyntöjen suorittamista varten eikä tallentaisi puhetietoja eikä suoritettujen kommentojen rekisteriä.
146. Vaikka jotkin laitteet voivat käyttää vain yhtä VVA:ta, toiset voivat valita eri VVA:iden välillä. VVA-palveluntarjoajien olisi kehitettävä teollisuusstandardeja, jotka mahdollistavat tietojen siirtämisen järjestelmästä toiseen tietosuoja-asetuksen 20 artiklan mukaisesti.
147. Jotkin VVA-palvelun tarjoajat ovat väittäneet, että niiden VVA:t eivät voi poistaa käyttäjien kaikkia tietoja edes rekisteröidyn pyynnöstä. VVA-palveluntarjoajien olisi varmistettava, että käyttäjien kaikki tiedot voidaan poistaa käyttäjän pyynnöstä yleisen tietosuoja-asetuksen 17 artiklan mukaisesti.

4 REKISTERÖIDYN OIKEUKSIEN KÄYTTÖMEKANISMIT

148. Yleisen tietosuoja-asetuksen mukaisesti VVA-palveluja tarjoavien rekisterinpitäjien on sallittava kaikkien rekisteröityjen ja rekisteröimättömien käyttäjien käyttää rekisteröidyn oikeuksiaan.
149. VVA-tarjoajien ja -kehittäjien olisi tehtävä tietojensa valvonnasta helpompaa rekisteröidyille koko käsittelyjakson ajan ja erityisesti helpotettava heidän oikeuttaan tutustua tietoihin, oikaista niitä, poistaa ne, heidän oikeuttaan rajoittaa käsittelyä ja, käsittelyn oikeusperustasta riippuen, heidän oikeuttaan siirtää tietoja ja heidän oikeuttaan vastustaa henkilötietojen käsittelyä.
150. Rekisterinpitäjän olisi annettava tietoa rekisteröidyn oikeuksista silloin, kun rekisteröity ottaa VVA:n käyttöön, ja viimeistään silloin, kun ensimmäisen käyttäjän äänipyyntö käsitellään.
151. Koska VVA:iden pääasiallinen vuorovaikutuskeino on ääni, VVA-suunnittelijoiden olisi varmistettava, että rekisteröidyt tai rekisteröimättömät käyttäjät voivat käyttää mitä tahansa rekisteröidyn oikeuksia käyttämällä helposti noudatettavia äänikomentoja. VVA-suunnittelijoiden sekä sovelluskehittäjien, mikäli he ovat osa ratkaisua, tulee harjoitusprosessin lopuksi ilmoittaa käyttäjälle, että hänen oikeutensa on asianmukaisesti huomioitu, joko puhuen tai lähettämällä kirjallisen ilmoituksen käyttäjän matkapuhelimeen, tilille tai muulle käyttäjän valitsemalle alustalle.

152. VVA-suunnittelijoiden ja sovelluskehittäjien olisi vähintään otettava käyttöön erityisiä välineitä, jotka tarjoavat tehokkaan ja toimivan tavan käyttää kyseisiä oikeuksia. Heidän olisi siis suunniteltava laitteisiinsa tapoja rekisteröityjen oikeuksien käyttämiseen tarjoamalla rekisteröidylle välineitä, joilla he voivat itse hallita profiilinhallintajärjestelmää⁵¹. Tämä voisi helpottaa rekisteröidyn oikeuksien tehokasta ja oikea-aikaista käsittelyä ja antaa rekisterinpitäjälle mahdollisuuden sisällyttää tunnistusmekanismi itsepalveluvälineeseen.
153. Tapauksissa, joissa käyttäjiä on useita, käyttäjän – rekisteröidyn tai rekisteröimättömän – olisi käyttäessään jotain rekisteröidyn oikeuksistaan tehtävä niin rajoittamatta muiden käyttäjien oikeuksia. Kaikki rekisteröidyt ja rekisteröimättömät käyttäjät voivat käyttää oikeuksiaan niin kauan kuin rekisterinpitäjä käsittelee tietoja. Rekisterinpitäjän olisi otettava käyttöön menetelmä, jolla varmistetaan rekisteröityjen oikeuksien käyttö.

4.1 Oikeus saada pääsy tietoihin

154. Yleisen tietosuoja-asetuksen 12 artiklan 1 kohdan mukaan 15 artiklan mukaisen tiedonanto olisi annettava kirjallisesti tai muulla tavoin ja tapauksen mukaan sähköisessä muodossa. Käsittelyn kohteena oleviin henkilötietoihin pääsyn osalta 15 artiklan 3 kohdassa todetaan, että jos rekisteröity esittää pyynnön sähköisesti, tiedot on toimitettava sähköisesti mahdollisuuksien mukaan, paitsi jos rekisteröity toisin pyytää. Yleisesti käytettävän sähköisen muodon olisi perustuttava rekisteröityjen kohtuullisiin odotuksiin eikä rekisterinpitäjän päivittäisessä toiminnassaan käyttämään muotoon. Rekisteröidyn ei pitäisi olla velvollinen ostamaan tiettyjä ohjelmistoja tai laitteita päästäkseen tietoihin käsiksi.
155. Sen vuoksi rekisterinpitäjien olisi pyynnöstä lähetettävä jäljennös henkilötiedoista ja erityisesti äänidatasta (mukaan lukien äänitallenteet ja litteraatiot) rekisteröidyn luettavissa olevassa yleisessä muodossa.
156. Päättäessään, minkälaisessa muodossa 15 artiklassa tarkoitetut tiedot olisi toimitettava, rekisterinpitäjän on pidettävä mielessä, että tietojen muodon olisi mahdollistettava tietojen esittäminen tavalla, joka on sekä ymmärrettävä että helposti saatavilla. Rekisterinpitäjien olisi myös räätälöitävä tiedot pyynnön esittäneen rekisteröidyn erityistilanteeseen.

Esimerkki 13:

VVA-palvelua tarjoava rekisterinpitäjä vastaanottaa käyttäjältä sekä käyttöpyynnön että tiedon siirtoa koskevan pyynnön. Rekisterinpitäjä päättää toimittaa sekä 15 artiklan että 20 artiklan mukaiset tiedot PDF-tiedostossa. Tässä tapauksessa ei pitäisi katsoa, että rekisterinpitäjä käsittelee molempia pyyntöjä asianmukaisesti. PDF-tiedosto täyttää teknisesti rekisterinpitäjää koskevat 15 artiklan mukaiset velvoitteet, mutta ei 20 artiklan mukaisia rekisterinpitäjää koskevia velvoitteita.⁵²

On huomioitavaa, että pelkkä käyttäjien ohjaaminen VVA:n kanssa käydyn vuorovaikutuksen historiatietoihin ei näytä mahdollistavan rekisterinpitäjälle kaikkien velvollisuuksiensa, jotka perustuvat käyttäjän oikeuteen saada pääsy tietoihin, täyttämistä, koska saatavilla olevat tiedot edustavat yleensä vain osaa palveluntarjoamisen yhteydessä käsitellyistä tiedoista.

⁵¹ Profiilinhallintajärjestelmä ymmärretään VVA-järjestelmän sisäiseksi sijainniksi, jossa käyttäjät voivat milloin tahansa tallentaa mieltymyksensä, muokata alustaa ja muuttaa helposti yksityisyysasetuksiaan.

⁵² Tietosuojatyöryhmän suuntaviivat oikeudesta tietojen siirrettävyyteen - tietosuojaneuvoston hyväksymä, s. 18.

157. Tiedonsaantioikeutta ei pitäisi käyttää tietojen minimoinnin ja säilyttämisen periaatteiden kumoamiseen tai kiertämiseen.

4.2 Oikeus tietojen oikaisemiseen

158. Tietojen oikaisun helpottamiseksi rekisteröityjen tai rekisteröimättömien käyttäjien olisi voitava milloin tahansa hallita ja päivittää tietojaan äänellä suoraan VVA-laitteessa edellä kuvatulla tavalla. Lisäksi laitteessa tai sovelluksessa olisi otettava käyttöön itsepalvelutyökalu, jonka avulla käyttäjät voivat helposti oikaista henkilötietojaan. Käyttäjille on ilmoitettava päivityksestä ääneen tai kirjallisesti.
159. Yleisemmin oikeus tietojen oikaisuun koskee rekisterinpitäjän mielipiteitä ja päätelmiä⁵³, myös profilointia, ja siinä olisi otettava huomioon, että valtaosa tiedoista on hyvin subjektiivisia.⁵⁴

4.3 Oikeus tietojen poistamiseen

160. Rekisteröityjen ja rekisteröimättömien käyttäjien olisi voitava milloin tahansa poistaa heitä koskevat tiedot VVA-laitteesta tai mistä tahansa VVA-laitteeseen integroidusta itsepalvelutyökalusta äänikäskyllä. Tässä suhteessa rekisteröity voi poistaa henkilötiedot yhtä helposti kuin ne on annettu. Puhetietojen anonymisoinnin luontaisten vaikeuksien sekä kerättyjen, havaittujen ja pääteltyjen rekisteröidyn henkilötietojen laajan kirjon vuoksi⁵⁵ oikeutta tietojen poistamiseen ei tässä yhteydessä pystytä juuri mahdollistamaan anonymisoimalla henkilötietoja. Koska yleinen tietosuoja-asetus on teknologianeutraali ja teknologia kehittyä nopeasti, ei kuitenkaan voida sulkea pois sitä mahdollisuutta, että oikeus tietojen poistamiseen voidaan tulevaisuudessa toteuttaa anonymisoimalla.
161. Joissakin tapauksissa, ilman kolmannen osapuolen näyttölaitetta tai muuta mahdollisuutta näyttää tallennettuja tietoja (esim. mobiilisovellus tai tabletilaite), on vaikeaa esikuunnella tallennettuja raitoja ja arvioida ehdotusten merkityksellisyyttä. VVA:n käytön helpottamiseksi sen käyttäjille pitäisi olla yleisesti tarjolla koontinäyttö (tai sovellus), jota käyttäjät voivat käyttää pyyntöhistoriansa poistamiseen ja jolla he voisivat mukauttaa laitetta tarpeidensa mukaiseksi.⁵⁶
162. Tietojen käsittelyn osalta ja erityisesti silloin, kun rekisteröidyt suostuvat äänitallenteiden litterointiin, jota palvelun tarjoaja käyttää parantaakseen palvelujaan, VVA-palvelun tarjoajien olisi käyttäjän pyynnöstä voitava poistaa alkuperäinen äänitallenne ja siihen liittyvät henkilötietojen litteraatiot.
163. Rekisterinpitäjän olisi varmistettava, ettei tietojen käsittelyä enää voi tapahtua tietojen poistamisoikeuden käyttämisen jälkeen. On hyvä huomioida, että aiempien toimien osalta oikeus tietojen poistamiseen voi törmätä joihinkin oikeudellisiin ja teknisiin rajoituksiin.

⁵³ Euroopan unionin tuomioistuin on vahvistanut, että lausunnot ja päätelmät voidaan katsoa henkilötiedoiksi, ja todennut, että termi ”kaikenlaiset tiedot” sisältää henkilötietojen määritelmässä ”sekä objektiiviset että subjektiiviset tiedot, jotka ilmaistaan mielipiteen tai arvion muodossa, edellyttäen, että ne ’koskevat’ kyseessä olevaa henkilöä” – Asia C-434/16 *Peter Nowak vastaan tietosuojavaaltuutettu*, ECLI:EU:C:2017:994 [34].

⁵⁴ Getting Data Subject Rights Right, Tietosuoja-aktijoiden artikkeli Euroopan tietosuoja-neuvostolle, marraskuu 2019.

⁵⁵ Tietosuoja-työryhmä (2014), lausunto 5/2014 anonymisointitekniikoista, annettu 10 huhtikuuta 2014.

⁵⁶ ”Assistants vocaux et enceintes connectées, l’impact de la voix sur l’offre et les usages culturels et médias”, ranskalainen ”Conseil Supérieur de l’Audiovisuel”, toukokuu 2019.

Esimerkki 14:

Jos käyttäjä on ennen poistopyyntöä tehnyt verkko-ostoksen VVA:n avulla, VVA-tarjoaja voi poistaa verkko-ostokseen liittyvän äänitteen ja varmistaa, ettei sitä enää käytetä tulevaisuudessa. Ostotapahtuma on kuitenkin edelleen voimassa, samoin kuin sähköisen kaupankäynnin verkkosivuston käsittelemä tilauksen äänitiedosto tai litteraatio (tässä poikkeus perustuu sähköisen kaupankäynnin verkkosivuston oikeudelliseen velvoitteeseen).

Jos käyttäjä on ennen poistopyyntöä lisännyt tietyn kappaleen soittolistalleen VVA:n avulla, VVA-tarjoaja voi poistaa suullisen pyynnön, mutta eivät tällaisen pyynnön aiempia seurauksia (poistaminen ei vaikuta käyttäjän soittolistaan).

164. Jos samoja henkilötietoja käsitellään eri käsittelytarkoituksiin, rekisterinpitäjien olisi edellä esitetyn perusteella tulkittava poistopyynnot selkeäksi merkiksi lopettaa tietojen käsittely kaikissa tarkoituksissa, joita oikeudelliset poikkeukset eivät koske.

Yleisen tietosuojasetuksen 21 artiklan 1 kohdassa säädettyjen edellytysten mukaisesti VVA-palveluntarjoajien oikeutettujen etujen perusteella käsittelemiä tietoja ei saisi vapauttaa tietojen poistamisoikeudesta, sillä rekisteröidyt eivät kohtuudella odota henkilötietojen jatkokäsittelyä.

4.4 Oikeus siirtää tiedot järjestelmästä toiseen

165. VVA-palvelun tarjoajien suorittama tietojenkäsittely kuuluu tietojen siirrettävyyden piiriin, koska käsittelytoimet perustuvat pääasiassa rekisteröidyn suostumukseen (6 artiklan 1 kohdan a alakohdan mukaisesti tai 9 artiklan 2 kohdan a alakohdan nojalla, kun on kyse erityisistä henkilötietoryhmistä) tai sopimukseen, jonka osapuolena rekisteröity on 6 artiklan 1 kohdan b alakohdan mukaisesti.
166. Käytännössä oikeuden tietojen siirtämiseen järjestelmästä toiseen pitäisi helpottaa vaihtamista eri VVA-palvelun tarjoajien välillä. Erityisesti kun VVA toimii digitaaliympäristössä, ja kun rekisteröidyn ääni tallennetaan sovellukseen tai alustalle, oikeus siirtää tiedot järjestelmästä toiseen olisi myönnettävä kaikille rekisteröidyn toimittamille henkilötiedoille. Lisäksi rekisterinpitäjän olisi tarjottava käyttäjille itsepalveluväline, jota käyttämällä heillä on mahdollisuus hakea henkilötietonsa suoraan omalta käyttäjäalueeltaan. Käyttäjien tulisi pystyä käyttämään tätä oikeuttaan myös äänikomennolla.
167. VVA-palvelun tarjoajien ja kehittäjien olisi annettava rekisteröidyille laaja määräysvalta itseään koskeviin henkilötietoihin, jotta he voivat siirtää henkilötietojaan VVA-palvelun tarjoajalta toiselle. Rekisteröidyn olisi näin ollen saatava rekisterinpitäjälle toimittamansa henkilötiedot jäsennellyssä, yleisesti käytetyssä ja koneellisesti luettavassa muodossa sekä keinoin⁵⁷, jotka

⁵⁷ Ks. esimerkkinä 29 artiklan mukaisen työryhmän oikeutta tietojen siirtämiseen järjestelmästä toiseen koskevien ohjeiden perustelut, s. 17:

"Teknisellä tasolla rekisterinpitäjien olisi tarkasteltava ja arvioitava kahta toisiaan täydentävää erilaista tapaa siirrettävien tietojen tarjoamiseksi rekisteröityjen tai muiden rekisterinpitäjien saataville:

- siirrettävien tietojen koko tietoaaineiston (tai koko tietoaaineistosta poimittujen useiden otteiden) suora siirto
- automaattinen väline, joka mahdollistaa olennaisten tietojen poimimisen.

Rekisterinpitäjät saattavat pitää parempana toiseksi mainittua keinoa tapauksissa, joihin liittyy monimutkaisia ja suuria tietoaaineistoja, koska se mahdollistaa minkä tahansa osan poimimisen tietoaaineistosta, joka on rekisteröidyn pyynnön kannalta olennainen, saattaa auttaa minimoimaan riskin ja voi mahdollistaa tietojen synkronointimekanismien käytön (esimerkiksi rekisterinpitäjien välisen säännöllisen tietoliikenteen yhteydessä).

pyrkivät vastaamaan tietojen siirrettävyyttä koskeviin pyyntöihin (kuten lataustyökalut ja sovellusrajapinnat)⁵⁸. Kuten ohjeissa, jotka koskevat oikeutta siirtää tiedot järjestelmästä toiseen, todetaan, jos kyseessä on suuri tai monimutkainen henkilötietojen keruu, rekisterinpitäjän olisi esitettävä yleiskatsaus ”*tiiviisti esitetyssä, läpinäkyvässä, helposti ymmärrettävässä ja saatavilla olevassa muodossa selkeällä ja yksinkertaisella kielellä*” (ks. yleisen tietosuoja-asetuksen 12 artiklan 1 kohta) siten, että rekisteröidyillä olisi aina oltava selkeät tiedot siitä, mitä tietoja on ladattava tai toimitettava toiselle rekisterinpitäjälle kunkin käyttötarkoituksen osalta. Rekisteröityjen olisi esimerkiksi voitava käyttää ohjelmistosovelluksia tiettyjen tietojen helppoon löytämiseen, tunnistamiseen ja käsittelyyn.

168. Tämän oikeuden pitäisi antaa käyttäjälle mahdollisuus hakea henkilökohtaista käyttöönsä varten tiedot, jotka hän on välittänyt äänellään (esim. ääniperusteisen vuorovaikutuksen historiatiedot) ja erityisesti käyttäjätilinsä luomisen yhteydessä (esim. nimi ja etunimi).
169. Jotta rekisteröityjen oikeutta voitaisiin soveltaa täysimääräisesti digitaalisten sisältöjen sisämarkkinoilla, etenkin VVA-suunnittelijoiden ja sovelluskehittäjien olisi kehitettävä yleisiä koneluettavia muotoja, jotka edistävät tiedon tallennusmuotojen yhteentoimivuutta VVA-järjestelmien välillä⁵⁹, mukaan lukien äänidatan vakiomuodot. Teknologioiden olisi oltava jäsennellyssä muodossa sen varmistamiseksi, että uusi rekisterinpitäjä voi helposti ja täysimääräisesti käyttää uudelleen henkilötietoja, myös äänidataa⁶⁰.
170. VVA-palveluntarjoajien olisi toimitettava yleisesti käytetyissä avoimissa tiedostomuodoissa (esim. mp3, wav, csv, gsm jne.) henkilötietoja sekä asianmukaisia metatietoja, joita käytetään vaihdettujen tietojen merkityksen tarkkaan kuvaamiseen.⁶¹

5 LIITE: AUTOMAATTINEN PUHEENTUNNISTUS, PUHESYNTESI JA LUONNOLLISEN KIELEN KÄSITTELY

171. Signaalinkäsittelyn teoreettisen perustan, erityisesti Claude Shannonin informaatio- ja näytteenottoteorioiden, jälkeen automaattinen puheenkäsittely on muodostunut insinööritieteiden perusosa-alueeksi. Fysiikan (akustiikka, aaltojen eteneminen), sovelletun matematiikan (mallinnus, tilastointi), tietotekniikan (algoritmit, oppimistekniikat) ja humanististen tieteiden (havainnointi, päättely) risteyksissä puheenkäsittely on nopeasti jaettu lukuisiin tutkimuskohteisiin: puhujan tunnistaminen ja todentaminen, automaattinen puheentunnistus, äänisyntesi, tunteiden havaitseminen jne. Viimeisten noin viidentoista

Se voi olla parempi keino varmistaa yhteensopivuus ”uuden” rekisterinpitäjän kanssa ja toimisi hyvänä käytäntönä, jolla vähennetään alkuperäisen rekisterinpitäjän osalta yksityisyydensuojaa koskevia riskejä.”

⁵⁸ Tämän osalta: Tietosuojatyöryhmä (2016), oikeutta tietojen siirtämiseen järjestelmästä toiseen koskevat ohjeet – Euroopan tietosuojaneuvoston hyväksymät, s. 1.

⁵⁹ Tältä osin: yleisen tietosuoja-asetuksen johdanto-osan 68 kappale; 29 artiklan mukaisen tietosuojaryhmän ohjeet: oikeus siirtää tiedot järjestelmästä toiseen, Euroopan tietosuojaneuvoston hyväksymä, s. 17.

⁶⁰ ”Tähän liittyen johdanto-osan 68 kappaleessa kannustetaan rekisterinpitäjiä kehittämään yhteentoimivia muotoja, jotka mahdollistavat tietojen järjestelmästä toiseen siirtämisen, luomatta kuitenkaan rekisterinpitäjille velvoitetta hyväksyä tai ylläpitää tietojenkäsittelyjärjestelmiä, jotka ovat teknisesti yhteensopivia. Yleisessä tietosuoja-asetuksessa kielletään kuitenkin rekisterinpitäjiä luomasta esteitä siirtämiselle” – Tietosuojaryhmä, Oikeutta tietojen siirtämiseen järjestelmästä toiseen koskevat ohjeet – Euroopan tietosuojaneuvosto, s. 5.

⁶¹ Euroopan tietosuojaneuvosto kannustaa voimakkaasti alan sidosryhmiä ja ammattijärjestöjä tekemään yhteistyötä yhteisten yhteentoimivien standardien ja muotojen laatimiseksi, jotta voidaan täyttää tietojen siirtämistä järjestelmästä toiseen koskevan oikeuden vaatimukset.

vuoden aikana, ala on ottanut suuria edistysaskelia, mihin on vaikuttanut monta tekijää: parannellut menetelmät, merkittävä laskentakapasiteettien kasvu ja se, että saatavilla olevaa tietoa on enemmän.

5.1 Automaattinen puheentunnistus

172. Automaattinen puheentunnistus (tunnetaan myös äänentunnistuksena) sisälsi ennen kolme eri vaihetta, joiden tavoitteena oli: 1) määrittää akustisen mallin avulla, mitkä foneemit on sanottu; 2) määrittää foneettisen sanaston avulla mitkä sanat on sanottu; 3) litteroida kielimallin avulla sanasarja (lause), joka on todennäköisimmin sanottu. Nykyään syväoppimisen (koneoppimistekniikka) ansiosta monet järjestelmät tarjoavat ”päästä päähän” automaattista puheentunnistusta. Näin vältetään tarve käydä läpi kolmen eri mallin monimutkainen koulutus ja samalla suorituskyky tulosten ja käsittelyajan suhteen paranee. Lähes kaikki suuret digialan toimijat tarjoavat nyt omia automaattisen puheentunnistuksen sovelluksia, joita ohjelmointirajapintajärjestelmät voivat helposti käyttää, mutta olemassa on myös avoimen lähdekoodin järjestelmiä (esimerkiksi DeepSpeech⁶² ja Kaldi⁶³).

5.2 Luonnollisen kielen käsittely

173. Luonnollisen kielen käsittely on kielitieteiden, tietotekniikan ja tekoälyn monitieteinen ala, jonka tavoitteena on luoda luonnollisen kielen käsittelyn välineitä erilaisiin sovelluksiin. Tutkimuksen ja sovellusten aloja on useita: syntaktinen analyysi, konekääntäminen, automaattinen tekstintuottaminen ja yhteenveto, oikeinkirjoituksen tarkistus, kysymyksiin vastaavat järjestelmät, tekstinlouhinta, nimettyjen entiteettien tunnistaminen, tunteiden analysointi jne. Käytännössä luonnollisen kielen käsittelyn tavoitteena on antaa tietokoneille mahdollisuus lukea, ymmärtää ja johtaa merkityksiä ihmiskielistä. Luonnollisen kielen käsittelyn sovellusten kehittäminen on haastavaa, koska tietotekniset työkalut vaativat perinteisesti ihmisiä olemaan vuorovaikutuksessa niiden kanssa käyttäen virallista täsmällistä, yksiselitteistä ja erittäin jäsenneiltyä ohjelmointikieltä. Ihmisten puhe ei kuitenkaan aina ole täsmällistä. Se on usein moniselitteistä, ja sen kielellinen rakenne voi riippua monista monimutkaisista muuttujista, kuten slangista, alueellisista murteista ja sosiaalisesta kontekstista.
174. Kaksi päätekniikkaa, joita käytetään luonnollisen kielen käsittelyn yhteydessä, ovat syntaksin ja semantiikan analyysi. Syntaksi on kieliopillisesti järkevä sanajärjestys lauseessa. Luonnollisen kielen käsittely käyttää syntaksia kieliopillisiin sääntöihin perustuvan kielen merkityksen arviointiin. Syntaksitekniikoita ovat esimerkiksi lauseenjäsenitys (lauseen kieliopillinen analyysi), segmentointi sanoiksi (joka jakaa suuren tekstipätkän yksiköihin), virkkeiden jakaminen (joka asettaa lauserajat suuriin teksteihin), morfologinen segmentointi (joka jakaa sanat ryhmiin) ja kantasanaahaku (joka jakaa taivutetut sanat kantasanoiksi). Semantiikka koskee sanojen käyttöä ja merkitystä. Luonnollisen kielen käsittely soveltaa algoritmeja ymmärtääkseen lauseiden merkityksen ja rakenteen. Tekniikoita, joita luonnollisen kielen käsittely käyttää semantiikan kanssa, ovat yksiselitteistäminen (joka johtaa asiayhteyteen perustuvan sanan merkitykseen), nimettyjen kohteiden tunnistus (joka määrittää sanoja, jotka voidaan luokitella ryhmiin) ja luonnollisen kielen luominen (joka käyttää tietokantaa sanojen taustalla olevan semantiikan määrittämiseen). Aiemmat luonnollisen kielen käsittelyn lähestymistavat olivat sääntöperusteisia. Niissä kerrottiin yksinkertaisille koneoppimisen algoritmeille, mitä sanoja ja ilmauksia tekstistä on etsittävä, ja

⁶² <https://github.com/mozilla/DeepSpeech>

⁶³ <https://github.com/kaldi-asr/kaldi>

annettiin tarkkoja vastauksia kyseisten ilmausten ilmetessä. Nykyiset luonnollisen kielen käsittelyn nykyiset lähestymistavat perustuvat sen sijaan syvään oppimiseen, tietynlaiseen tekoälyyn, joka tutkii ja käyttää tietomalleja ohjelman ymmärtämisen parantamiseksi.

5.3 Puhesynteesi

175. Puhesynteesi on ihmisen puheen keinotekoista tuottamista. Puhesynteesiä on toteutettu pääasiassa liittämällä yhteen tietokantaan tallennettuja ääniyksiköitä. Tässä menetelmässä valitaan aiemmin foneemeiksi, tavuiksi ja sanoiksi litteroiduista äänitteistä äänipätkät, jotka vastaavat sanoja, joita VVA haluaa ääntää. Äänipätkät järjestellään niin, että niistä muodostuu ymmärrettävä lause, jolla on luonnollinen merkitys. Vaihtoehtoisesti puhesyntetisaattori voi hyödyntää ääniväylän mallia ja muita ihmisäänen ominaisuuksia mallintaakseen äänen parametreja, kuten intonaatiota, rytmiä ja äänenväriä, generatiivisilla tilastollisilla malleilla (kuten WaveNet⁶⁴, Tacotron⁶⁵ tai DeepVoice⁶⁶) ja luoda täysin synteettisen äänituotoksen.

⁶⁴ Aaron van den Oord et Sander Dieleman, *WaveNet: A generative model for raw audio*, Deepmind-blogi, syyskuu 2016, <https://deepmind.com/blog/article/wavenet-generative-model-raw-audio>

⁶⁵ Yuxuan Wang, *Expressive Speech Synthesis with Tacotron*, Google AI -blogi, maaliskuu 2018, <https://ai.googleblog.com/2018/03/expressive-speech-synthesis-with.html>

⁶⁶ *Deep Voice 3: 2000-Speaker Neural Text-to-Speech*, Baidu Research -blogi, lokakuu 2017 <http://research.baidu.com/Blog/index-view?id=91>