

# Donera tal-kampanjen – Jämviktstest för berättigat intresse

Den personuppgiftsansvarige ska använda ett jämviktstest för att bedöma om ett berättigat intresse kan läggas till grund för behandlingen av personuppgifter. Nedan finns en skriftlig beskrivning av hur vi har gjort jämviktstestet för Donera prat-kampanjen och de relaterade beslutsstegen. Jämviktstestet har utförts i samarbete med Svenska litteratursällskapet i Finland rf., Helsingfors universitet och Yleisradio Oy. Testet kommer att utföras på nytt och beskrivningen nedan kommer att uppdateras i enlighet med detta om behandlingens syfte, karaktär eller sammanhang ändras.

## 1. Är ett berättigat intresse den lämpligaste grunden för behandling?

### Val av behandlingsgrund

Det ska finnas en laglig grund för behandling av personuppgifter. Det har i projektet ansetts eftersträvansvärt att det finns en gemensam grund för så mycket som möjligt av behandlingen i nätverket av operatörer, så att kriteriet inte behöver bedömas och inhämtas separat för varje operatör och att behandlingen av personuppgifter är så förutsägbar som möjligt för den registrerade.

### Att utveckla talförståelse som ett berättigat intresse

Personuppgifterna som samlas in i kampanjen kommer att användas för utveckling och beforskning av applikationer och tjänster som förstår och producerar tal, samt för språkforskning. Att donera prat och lämna ut personlig information är helt frivilligt.

Det berättigade intresset i Donera prat-kampanjen är utveckling och beforskning av applikationer och tjänster som förstår och producerar tal, såväl som högskoleundervisning relaterad till dessa syften.

Det berättigade intresset är betydande eftersom det leder till bättre applikationer och tjänster som gynnar både enskilda aktörer och samhället som helhet. Företag som utvecklar artificiell intelligens får nytta av det i sin verksamhet. Andra företag och offentliga aktörer kommer att kunna utveckla bättre tjänster. Vanliga medborgare får bättre tjänster, och samhällsekonomin gynnas av ökad ekonomisk aktivitet. I slutändan kan därför donatorn, det vill säga den registrerade, också dra nytta av resultatet av projektet och behandlingen av sina personuppgifter, även om fördelen är indirekt och inte kan garanteras för varje enskild donator.

I Donera prat-kampanjen är ett berättigat intresse i praktiken den lämpligaste grunden för att behandla utomståendes personuppgifter, så länge det på god grund uppväger de registrerades rättigheter och friheter. För att fastställa detta bedöms riskerna för den registrerade inom ramen för detta jämviktstest. Om det berättigade intresset inte kan vägas mot riskerna för den registrerade ska behandlingen av personuppgifter omfattas av något av de andra behandlingsgrunderna som anges i dataskyddsförordningen.

## Andra rättsliga behandlingsgrunder

### Den registrerades samtycke

Samtycke betyder en medveten, genuint frivillig och entydig viljeyttring för ett på förhand fastställt, uttryckligt och lagligt syfte. Den registrerade kan därför inte ge sitt samtycke till att hans uppgifter används för alla möjliga syften och på alla möjliga sätt. Kravet på samtyckets ändamålsenlighet är t.ex. den registeransvariges skyldighet att kunna styrka förekomsten av samtycke. Därför rekommenderas en skriftlig blankett. Donera prat-kampanjens behandlingsgrund är mycket nära samtycke till behandling, eftersom uppgifterna erhålls direkt från en person som ger dem frivilligt.

Samtycke kan dock inte betraktas som en lämplig allmän grund för behandling av personuppgifter i Donera prat-kampanjen av följande skäl:

- Att ge samtycke till hela processen i Donera prat-kampanjen är svårt. Om nätverket av aktörer är omfattande, kan det vara svårt att formulera samtycket så att det på ett giltigt sätt omfattar behandlingen av personuppgifter i hela nätverket.
- Samtycke kan återkallas när som helst. Detta är en betydande begränsning av samtyckets användbarhet, givet behandlingens art och omfattning (lagring av information i talarkivet och utlämnande till ett stort antal aktörer).
- Om personen är omyndig borde man bedöma, om han eller hon själv kan lämna giltigt samtycke eller om hans eller hennes föräldrar eller annan vårdnadshavare ska lämna samtycke för hans eller hennes räkning. Enligt den finländska dataskyddslagen kan den som har fyllt 13 år själv samtycka till behandlingen av sin personliga information i informationssamhällets tjänster, det vill säga tjänster som tillhandahålls som elektroniska fjärrtjänster på mottagarens personliga begäran. Samtycke av en förälder eller annan vårdnadshavare krävs av ett barn som är under 13 år och för annat än informationssamhällets tjänster. Det bör noteras att det på andra ställen i lagstiftningen finns andra bestämmelser om vid vilken ålder ett barn giltigt kan ingå avtal eller vara juridiskt ansvarigt för sina handlingar.

I framtiden är det möjligt att samtycke från taldonatorer måste sökas separat om det i vissa fall också är nödvändigt att samla och bearbeta personuppgifter såsom hälsouppgifter, som inte kan hanteras utifrån ett berättigat intresse. I sådana fall är syftet att undvika ovannämnda problem genom att reducera hanteringen baserad på samtycke till ett minimum.

### Genomförande av ett avtal

Det finns inget avtal mellan donatorn och de personuppgiftsansvariga i Donera prat-kampanjen som skulle kräva behandling av personuppgifter.

### Rättslig förpliktelse, skydd av vitala intressen, offentlig makt

Den rättsliga grunden för behandlingen av personuppgifter kan vara efterlevnad av en rättslig förpliktelse, skydd av vitala intressen eller utövande av offentlig makt. Det finns ingen rättslig förpliktelse att samla in personlig information i samband med Donera prat-kampanjen, det är inte

fråga om att skydda någons vitala intressen och utövandet av offentlig makt kräver inte behandling av personuppgifter som samlats in i Donera prat-kampanjen, så dessa behandlingsgrunder är inte tillämpliga.

### Allmänt intresse

Utförande av en uppgift i allmänhetens intresse är en berättigad grund för behandling av personuppgifter, till exempel inom vetenskaplig forskning. I den mån informationen som samlats in genom Donera prat-kampanjen kan komma att användas i framtida vetenskaplig forskning kan det motiveras av att en uppgift utförs i allmänhetens intresse.

## 2. Uppfylls de grundläggande kraven?

För att anses vara berättigad måste ett intresse först uppfylla följande krav:

- Intresset måste vara berättigat (i enlighet med tillämplig EU-lagstiftning eller nationell lagstiftning).
- Intresset ska tydligt anges så att dess avvägning i förhållande till den registrerades intressen och rättigheter kan bedömas.
- Intresset ska representera ett verkligt och omedelbart behov. Fördelen kan inte vara spekulativ.

Materialet som genereras i Donera prat-kampanjen innehåller fyra typer av personlig information:

Information relaterad till tal och talinnehåll:

1. En talares röst kan vara personlig information om den kan kopplas till talaren.
2. Pratet kan innehålla information som kan kopplas till en person.

Information relaterad till talinsamling och inspelning:

3. All information som talaren lämnat om sig själv under inspelningen, såsom talarens namn, bostadsort, dialekt och ålder.
4. Den tekniska information som samlas in om användningen av applikationen, såsom informationen relaterad till användarens terminal, tidpunkten för lagring och identifieraren (UUID) som skapades under installationen av applikationen.

**Intresset är lagligt och tydligt angivet**

Utveckling och forskning av applikationer och tjänster som förstår och producerar tal, samt språkforskning och högre utbildning i anslutning till dessa ändamål, är lagligt i Finland och EU, så intresset är lagligt.

Intresset är tydligt angivet och definierat (utveckling och forskning av talförståelse och produktion samt högre utbildning relaterad till dessa ändamål).

Intresset basera sig på ett verkligt och omedelbart behov

Intresset representerar ett verkligt och omedelbart behov och är inte spekulativt. Utvecklingen av applikationer för artificiell intelligens som förstår och producerar de språk som talas i Finland är för närvarande avsevärt begränsad av bristen på taldata. Det finns till exempel mycket mera material tillgängligt på engelska för utveckling och forskning av artificiell intelligens, varför det finns många fler artificiella intelligensapplikationer som förstår och producerar engelska och är betydligt bättre än artificiellintelligensapplikationer som förstår och producerar finlandssvenska.

Det finns klart färre tjänster baserade på artificiell intelligens på finlandssvenska på marknaden, t.ex. jämfört med andra EU-språk. Därför behövs mer data på de språk som talas i Finland så att Finland inte ytterligare blir efter i utvecklingen av artificiell intelligens och i utvecklingen av talbaserade tjänster och applikationer.

### **3. Är behandlingen av personuppgifter nödvändig för att uppnå det berättigade intresset?**

Förståelse och produktion av tal kan inte utvecklas utan bearbetning av autentiskt talmaterial. Syftet är att behandla personuppgifter så lite som möjligt i Donera prat-kampanjen och i förhållande till det resulterande materialet. Det är dock omöjligt att undvika att de insamlade uppgifterna innehåller personuppgifter. Tal i sig är personlig information om rösten går att känna igen. Dessutom kan personer, när de donerar prat, dela information som kan kopplas direkt eller indirekt till personen, såsom talaren själv eller en tredje part. Dessutom har deltagarna instruerats att inte dela personlig eller känslig information om sig själva eller andra.

Kampanjen samlar också in viss bakgrundsinformation av deltagarna själva, till exempel information om deras ålder och dialektområde. Denna information behöver bearbetas eftersom den påverkar utvecklingen och beforskningen av applikationer och tjänster som förstår och producerar tal, samt språkforskningen.

### **4. Åsidosätter intresset verkligen den registrerades intressen och rättigheter?**

Den registeransvariges eller tredje parts intresse

Det berättigade intresset är utveckling och beforskning av applikationer och tjänster som förstår och producerar tal, språkforskning och högre utbildning med anknytning till dessa syften. Det bör noteras att detta inte bara är en fråga för registerförarerna (Helsingfors universitet, Svenska litteratursällskapet i Finland och Yleisradio Oy), utan framför allt för tredje parter som företag, offentliga tjänsteleverantörer, forskare och konsumenter och medborgare som drar nytta av bättre tjänster, funktionell talteknik eller tillämpningar. Fördelen med att behandla personuppgifter är att de insamlade uppgifterna kan användas för att utveckla bättre tjänster och få mer forskningsinformation relaterad till språk. Om vi vill vara helt säkra på att materialet inte innehåller personuppgifter, måste vi välja att inte utföra kampanjen.

## Inverkan på den registrerade

Det bör noteras att personuppgifter i första hand kommer från en individ och tillhandahålls frivilligt av individen med en förståelse för vad de ska användas till. Informationen är i princip inte särskilt känslig, även om det inte går att vara helt säker på att talaren själv inte berättar något om till exempel sin hälsa eller uppgifter som rör en annan person.

Behandlingsåtgärderna har ingen omedelbar inverkan på den registrerade. Uppgifterna behandlas för utveckling och forskning av applikationer och tjänster som förstår och producerar tal, för språkforskning och för högre utbildning med anknytning till dessa ändamål. Materialet används till exempel för att utveckla algoritmer och för att lära ett träningsbart system för artificiell intelligens att förstå eller producerar tal. Vid genomförandet av bearbetningen analyseras stora mängder ljud, och det är inte ett informationsintresse att specifikt analysera en enskild ljudinspelning eller person.

Nedan finns beskrivningar av de viktigaste identifierade riskerna, som den registrerade kan ställas inför, särskilt i händelse av fel eller avsiktligt missbruk vid behandlingen av materialet. De identifierade riskerna beskrivs mer i detalj i konsekvensbedömningen för dataskydd.

### Identifierade risker och skyddsåtgärder

Steg	Riskhändelse	Skada för den registrerade - beskrivning	Sannolikhet	Skadans storlek	Åtgärd
Röstinspelning	Ett fel uppstår under inspelningen	Inspelningen kan inte kopplas till högtalaren eller så är den kopplad till fel person	liten	liten	Röstinspelningens funktion verifieras genom att testa applikationen före i bruktagandet
Röstinspelning	Ett fel uppstår under inspelningen	Inspelningen får felaktig tilläggsinformation, vilket minskar materialets användbarhet	liten	liten	Röstinspelningens funktion verifieras genom att testa applikationen före i bruktagandet
Röstinspelning	Inspelningen misslyckas	Talaren måste spela in pratet på nytt eller lyckas inte donera sitt prat	liten	liten	Röstinspelningens funktion verifieras genom att testa applikationen före i bruktagandet
Röstinspelning	Personen förstår inte att den donerar	Personinformationens användningsändamål förblir oklart	liten	liten	Syftet med donation klargörs så tydligt det går
Röstinspelning	Personen tror att de donerar och senare erbjuds data mot en avgift	Personen känner sig lurad	liten	liten	Syftet med donation klargörs så tydligt det går
Röstinspelning	Personen tror att uppgifterna bara kommer att gå till Yle	Användningen av persondata förblir oklar	liten	liten	Syftet med donation klargörs så tydligt det går
Röstinspelning	Personen förstår inte att talmaterialet kommer att föras vidare till andra parter	Behandlingens omfattning är oförutsedd	liten	liten	Användaren informeras om överföringen i dataskydds- och kampanjbeskrivningen.
Röstinspelning	Personen spelar in prat i misstag ("fickinspelning")	Inspelningen kan innehålla personuppgifter som donatorn inte har avsett att berätta.	liten	liten	Gränssnittet planeras så att möjligheten för oavsiktlig inspelning är liten.

Röstinspelning	Bakgrundssystemets kapacitet är otillräcklig.	Pratet kan inte sparas.	liten	liten	Bakgrundssystemets kapacitet säkerställs.
Röstinspelning	Inspelning i hemlighet	Den rättmätige talarens rättigheter åsidosätts.	måttlig	medelstor	Talare uppmanas att inte donera andras tal. Sådana inspelningar raderas så snart de kommer till kännedom.
Röstinspelning	Tal innehåller hatretorik eller liknande innehåll riktat mot en person.	Tal är kränkande för en persons heder, integritet etc.	måttlig	medelstor	Talare vägleds och avtal och licenser inkluderar mekanismer för att ta bort sådant material.
Lagring av talmaterialet	Dataintrång	Pratet kan användas för ändamål som talaren inte har godkänt.	liten	medelstor	Tekniska och organisatoriska åtgärder syftar till att säkerställa en adekvat nivå på datasäkerheten.
Lagring av talmaterialet	Överlåtelse av materialet i strid med avtalet.	Pratet kan användas för ändamål som talaren inte har godkänt.	liten	medelstor	Avtal säkerställer att materialet inte lämnas ut eller används i strid med det godkända ändamålet.
Lagring av talmaterialet	Åsidosättande av ansvarsskyldighet	Talaren vet inte om hans rättigheter har kränkts.	liten	liten	Varje organisation som är involverad i behandlingen av personuppgifter tar hand om och är hela tiden medveten om sin ansvarsskyldighet.
Lagring av talmaterialet	För stora befogenheter att använda data.	Förstöring av data, användning av data för andra ändamål.	måttlig	liten	Organisatoriska åtgärder säkerställer att rätt personer har rätt tillgång till uppgifterna.
Lagring av talmaterialet	En stor mängd felaktigt innehåll	Felaktig information används mot donatorn.	liten	medelstor	Felaktig information korrigeras eller raderas så snart den kommer till kännedom.
Användning av talmaterialet	Prat produceras med talarens röst	Identitetsstöld - autentisering med talarens röst	liten	medelstor	Pseudonymisering och dataminimering minskar möjligheten att en röst kopplas till en specifik donator.
Användning av talmaterialet	Prat produceras med talarens röst	Identitetsstöld - det produceras tal som påstås höra till en specifik talare	måttlig	medelstor	Pseudonymisering och dataminimering minskar möjligheten att en röst kopplas till en specifik donator.
Användning av talmaterialet	Pratet innehåller konfidentiell information om talaren	Talarens konfidentiella information avslöjas för utomstående	liten	liten	Talare uppmanas att inte dela konfidentiell information. Konfidentiell information kommer att raderas så snart den kommer till kännedom.
Användning av talmaterialet	Pratet innehåller konfidentiell information om en annan person	En annan persons konfidentiella information avslöjas för utomstående	liten	medelstor	Talare uppmanas att inte dela konfidentiell information. Konfidentiell information kommer att raderas så snart den kommer till kännedom.
Användning av talmaterialet	Indirekt implikation, t.ex. mentalstört prat	Talarens konfidentiella information avslöjas för utomstående	liten	medelstor	Talare uppmanas att inte dela konfidentiell information. Konfidentiell information kommer att raderas så snart den kommer till kännedom.
Användning av talmaterialet	Material avlägsnas enbart från originalsamlingen men inte från kopiorna.	Talarens lagliga rättigheter förverkligas inte.	liten	medelstor	Tekniska åtgärder, avtal och licensvillkor säkerställer att raderingar och reparationer överförs till alla kopior av materialet som används.

Användning av talmaterialet	Talaren känns igen på grund av sällsynta ord.	Det donerade pratet kan användas mot donatorn.	liten	medelstor	Det är alltid möjligt för något i talet att avslöja talarens identitet. Fullständig anonymitet kan inte utlovas. Risken minskas genom att sådan information raderas snabbt när den kommer till kännedom.
Användning av talmaterialet	Talaren känns igen på grund av metadata.	Det donerade pratet kan användas mot donatorn.	liten	medelstor	Det är alltid möjligt för något i metadat att avslöja talarens identitet. Fullständig anonymitet kan inte utlovas. Risken minskas genom att sådan information raderas snabbt när den kommer till kännedom.
Användning av talmaterialet	En övertränade modell upprepar i sin helhet prat / talstil / meningsstruktur / exakta ord.	Det donerade pratet kan användas mot donatorn.	liten	liten	Delning av personlig information med modellen förbjuds med avtal. Informationen raderas när den kommer till kännedom.

### Fastställande av preliminär jämvikt

Baserat på ovanstående avvägning av de olika parternas intressen kan man säga att vikten av de registeransvarigas, och i synnerhet tredje parts, intressen i förhållande till den registrerades grundläggande rättigheter och friheter är betydande och det identifierade berättigade intresset är förenligt med den registrerades rättigheter och friheter.

## 5. Säkerställ ytterligare dataskyddsgarantier

De skyddsåtgärder som beaktas vid hanteringen beskrivs ovan i samband med riskkarakteriseringen i avsnitt 4.

## 6. Påvisa åtgärdernas laglighet och säkerställ öppenhet

Detta dokument beskriver jämviktstestet och hållas tillgängligt. Vi kommer att agera öppet och vid behov förklara för registrerade och myndigheter varför behandlingen av personuppgifter i detta fall ligger i den registeransvariges och tredje parts berättigade intresse. All dokumentation hålls uppdaterad och jämviktstestet upprepas vid behov.