Leitfaden: DSGVO-konformer Einsatz von Voice-AI-Agenten (Inbound & Outbound)

1. Überblick: Warum Voice-AI datenschutzrelevant ist

Voice-AI-Agenten verarbeiten **gesprochene Sprache**, die als *personenbezogene Daten* gilt. Eine Sprachaufzeichnung kann Rückschlüsse auf **Identität**, **Geschlecht**, **Herkunft** oder sogar Gesundheitszustand einer Person zulassen [1]. Somit unterliegt bereits die Stimme selbst dem Datenschutz. Hinzu kommt, dass Inhalte, die Anrufer mitteilen (z. B. Name, Anliegen, Adress- oder Gesundheitsdaten), personenbezogene Informationen enthalten. Die **automatisierte Verarbeitung** dieser Spracheingaben durch KI (Spracherkennung, Interpretation, Antwortgenerierung) stellt eine *Datenverarbeitung* im Sinne der DSGVO dar. Entsprechend greifen sämtliche Datenschutzgrundsätze (Transparenz, Zweckbindung, Datenminimierung, Speicherbegrenzung etc.). Auch kann eine Stimme je nach Einsatz als **biometrisches Merkmal** gelten, wenn sie zur eindeutigen Identifizierung genutzt wird [2] – dann handelt es sich um eine *besondere Kategorie personenbezogener Daten* (Art. 9 DSGVO) mit nochmals strengerem Schutz. Kurz: **Voice-AI ist datenschutzrelevant**, weil sie persönliche Informationen (Audio und abgeleitete Daten) automatisiert verarbeitet und potenziell sensible Details preisgibt.

2. Typische Datenflüsse in Voice-AI-Agents

Ein Voice-AI-Agent durchläuft meist eine **Pipeline** aus mehreren KI-Komponenten. Typischer Ablauf bei einem Anruf:

- **Speech-to-Text (STT)**: Das gesprochene **Audio** des Nutzers wird aufgezeichnet und an einen Spracherkennungsdienst geschickt, der daraus **Text** transkribiert. Dieser Dienst kann eine Cloud-Komponente (z. B. von Microsoft, Google) oder eine lokale Engine sein. Dabei verlässt das Audio evtl. das Unternehmen, um im Rechenzentrum des STT-Dienstleisters verarbeitet zu werden (wichtig für die Datenschutzbewertung!).
- **Sprachverstehen & Logik (NLU/LLM)**: Der transkribierte Text wird von einer **KI-Modell** (z. B. einem *Large Language Model*) analysiert, um die **Intention** des Anrufers zu verstehen und die passende Antwort oder Aktion zu ermitteln. Dies kann intern passieren oder via externem KI-Dienst (etwa einer Dialog-KI in der Cloud). Hier werden evtl. auch Unternehmensdaten hinzugezogen (z. B. Kundendatenbank bei Kundenanfragen).
- **Antwortgenerierung & TTS**: Aus der ermittelten Antwort wird per **Text-to-Speech (TTS)** eine **synthetische Sprachausgabe** erzeugt, die dann dem Anrufer vorgespielt wird. Oft nutzt man einen TTS-Cloudservice für natürlich klingende Stimmen. Wiederum fließt dafür der Antworttext (der u. U. den Namen des Kunden oder andere persönliche Details enthält) an den TTS-Anbieter.
- **Orchestrierung**: Eine Vermittlungs-Software (“Agent Orchestrator”) koordiniert STT, KI und TTS, verwaltet den Gesprächsablauf und trifft eventuelle Entscheidungen (z. B. Weiterleitung an einen menschlichen Mitarbeiter, wenn der Bot nicht weiterweiß).
- **Logging & Monitoring**: Üblicherweise werden **Protokolle** geführt – etwa die Text-Transkripte des Gesprächs oder Meta-Daten (Zeit, Dauer, Telefonnummer). Mitunter werden auch Audioaufnahmen gespeichert. Diese Daten nutzt man zu **Qualitätssicherungs-** und **Trainingszwecken** (z. B. um das KI-Modell zu verbessern) oder zur Nachweisführung (welche

Auskunft wurde gegeben?). Wichtig: Solche Logs sind ebenfalls personenbezogen und bedürfen Schutz und klarer Zweckbindung.

Bei **Inbound-Agents** (der Kunde ruft an) fließen die Daten primär vom Kunden zum Unternehmen. Bei **Outbound-Agents** (die KI ruft aktiv Kunden an) initiiert das Unternehmen den Anruf, nutzt also zuvor erhobene Kontaktdaten; während des Gesprächs läuft die Pipeline analog. In beiden Fällen ist nachvollziehbar zu dokumentieren, *welche Daten in welcher Phase an wen übermittelt und verarbeitet werden*. Unternehmen sollten diesen Ablauf als **Datenfluss-Diagramm** skizzieren, um für jeden Verarbeitungsschritt Risiken und Zuständigkeiten bewerten zu können. Beispielsweise zeigt ein technisches Leitfadenbeispiel, dass ein produktiver Sprachagent mehrere Dienste umfasst und **GDPR-compliant infrastructure** erfordert – idealerweise ohne unnötige Drittanbieter-Abhängigkeiten, um Kontrolle über alle Datenflüsse zu behalten [3].

Fazit: Die Sprachpipeline involviert oft mehrere Akteure (eigene Systeme und externe KI-Dienste). Jeder dieser Schritte muss datenschutzkonform gestaltet werden, da überall personenbezogene Daten "durchfließen".

3. Prüfpunkte für Unternehmen (Compliance-Checkliste)

Bevor Voicebots live geschaltet werden, sollten Unternehmen – insbesondere Datenschutzbeauftragte und Entscheider – eine **gründliche Prüfung** durchführen. Wichtige Fragen und Handlungspunkte sind:

3.1 Welche Daten werden verarbeitet?

Machen Sie **transparent**, welche personenbezogenen Daten der Voice-Agent berührt. Typischerweise sind das: **Audioaufnahmen der Stimme**, daraus erzeugte **Textdaten** (Transkripte der Unterhaltung) sowie alle **Inhalte**, die der Anrufer preisgibt (Name, Kundennummer, Anliegen, ggf. sensible Angaben). Auch Meta-Daten wie **Rufnummer des Anrufers**, Anrufzeitpunkt, Dauer und Ergebnis (Erfolgsquote, Weiterleitung ja/nein) zählen dazu. Werden Gespräche **aufgezeichnet**, entsteht besonders schützenswertes Material (Audio = Originalbiometrie). Prüfen Sie außerdem, ob **besondere Kategorien** von Daten tangiert werden: z. B. Gesundheitsdaten, falls Kunden am Telefon Symptome schildern; oder **biometrische Daten**, falls eine Identitätsverifikation über Stimmabgleich erfolgt. In solchen Fällen greift Art. 9 DSGVO (Verbot mit Erlaubnisvorbehalt) – eine Verarbeitung ist nur mit *ausdrücklicher Einwilligung* oder einem engen Ausnahmetatbestand erlaubt. Oft versuchen Unternehmen, solche sensiblen Daten zu **vermeiden**. Dennoch: eine vollständige Risikoanalyse verlangt, alle *Datenkategorien* aufzuschlüsseln.

Tipp: Führen Sie ein **Verzeichnis der Verarbeitungstätigkeiten** für den Voicebot ein. Dokumentieren Sie Zweck, Datenarten, Kategorien betroffener Personen (z. B. Anrufer = Kunden, Interessenten) und Empfänger (z. B. KI-Dienstleister). Das hilft bei der weiteren Prüfung (und ist ohnehin nach Art. 30 DSGVO Pflicht).

3.2 Ist eine Datenschutz-Folgenabschätzung (DSFA) nötig?

Bei innovativen Technologien wie Voice-AI ist häufig eine **Datenschutz-Folgenabschätzung** (DSFA, Art. 35 DSGVO) erforderlich. Die DSFA bewertet, ob die geplante Verarbeitung ein *hohes Risiko* für die Rechte der Betroffenen mit sich bringt und wie man dieses Risiko mindert. Sprachassistenten **erfüllen sehr wahrscheinlich die DSFA-Kriterien** [4]: Es findet eine **systematische Überwachung** statt (Aufzeichnung von Gesprächen), oft in *großem Umfang*, und es werden möglicherweise neue technologische Ansätze mit hohem Risiko eingesetzt (z. B. biometrische Identifikation oder Profiling

durch Sprachanalyse). Der Europäische Datenschutzausschuss (EDPB) hat betont, dass Voice Assistant Services **in der Regel DSFA-pflichtig** sind [4].

Unternehmen sollten daher **vor Inbetriebnahme** des Voice-Agenten eine DSFA durchführen. Aspekte der DSFA beinhalten u. a.: Beschreibung des geplanten Systems und Ablaufs, Notwendigkeit und Verhältnismäßigkeit der Verarbeitung, identifizierte Risiken (z. B. unbefugter Zugriff auf Sprachdaten, Fehlentscheidungen der KI zu Ungunsten des Nutzers, Diskriminierung) und geplante Schutzmaßnahmen (Verschlüsselung, Zugriffsbeschränkung, menschliche Review-Prozesse etc.). **Dokumentieren** Sie diese DSFA schriftlich. Sollte die DSFA ein **Rest-Risiko** identifizieren, das als hoch eingestuft wird und nicht ausreichend gemindert werden kann, **muss die Aufsichtsbehörde konsultiert werden** (Art. 36 DSGVO) bevor das System live geht.

Praxis-Tipp: Nutzen Sie Checklisten der Aufsichtsbehörden (z. B. das Kurzpapier Nr. 5 der DSK zur DSFA oder Online-Tools) als Hilfestellung. Eine gründliche DSFA dient nicht nur der Compliance, sondern auch als interner Lernprozess, um Schwachstellen des Projekts früh zu erkennen.

3.3 Wann sind Einwilligungen erforderlich?

Die **Rechtsgrundlage** der Verarbeitung muss für alle Verarbeitungsschritte geklärt sein. Oft werden *mehrere* Rechtsgrundlagen parallel relevant:

- **Vertragserfüllung (Art. 6 Abs. 1 lit. b DSGVO)**: Bei Inbound-Anrufen, die der Kunde *selbst initiiert*, um z. B. eine Dienstleistung in Anspruch zu nehmen (Kundenservice, Support), kann die Verarbeitung der Gesprächsdaten oft als *erforderlich zur Vertragserfüllung* oder Durchführung vorvertraglicher Maßnahmen angesehen werden. Beispiel: Ein Bankkunde ruft an, um den Kontostand zu erfragen – die automatisierte Verarbeitung seiner Anfrage durch den Voicebot ist Teil der Vertragserfüllung im Rahmen des Bankservices. Hier wäre *keine zusätzliche Einwilligung* nötig, sofern die Daten nur zur Bearbeitung dieser Anfrage genutzt werden. **Achtung:** Sobald der Voicebot mehr tut als nur das vom Kunden Gewollte (z. B. Gespräch mitschneiden zu Qualitätszwecken, oder die Daten später zu Trainingszwecken wiederverwenden), reicht Vertragserfüllung alleine nicht mehr aus.
- **Berechtigtes Interesse (Art. 6 Abs. 1 lit. f DSGVO)**: Unternehmen können argumentieren, dass der Einsatz eines Voice-AI-Agenten im *berechtigten Unternehmensinteresse* liegt (Effizienzsteigerung, 24/7 Erreichbarkeit, Verbesserung der Servicequalität) – insbesondere, wenn er für *nicht sensible* Abläufe genutzt wird und der Kunde eine entsprechende Betreuung erwarten kann. Voraussetzung ist aber immer eine **Interessenabwägung**: Die Interessen des Unternehmens dürfen die *Grundrechte der Anrufer* nicht überwiegen. Transparenz und Widerspruchsmöglichkeit sind hier wichtig (siehe Nutzeraufklärung unten). **Nicht geeignet** ist berechtigtes Interesse für Szenarien, die einen **Eingriff in die Privatsphäre** darstellen, den der Betroffene nicht erwarten würde.
- **Einwilligung (Art. 6 Abs. 1 lit. a DSGVO)**: In mehreren Fällen kommt man an einer expliziten *Einwilligung* der Betroffenen nicht vorbei. **Outbound-Werbeanrufe** durch einen KI-Agenten sind ein klassischer Fall: Ruft das Unternehmen aktiv Kunden zu Werbezwecken an, ist *vorherige ausdrückliche Einwilligung* nötig – andernfalls ist es rechtswidrige Telefonwerbung [5], was erhebliche Bußgelder nach sich ziehen kann. Daher: Für *jedes Marketing-Telefonat* (ob Mensch oder Maschine ruft, spielt keine Rolle) **vorab opt-in einholen**. Auch bei Umfragen oder automatisierten Zufriedenheitsabfragen per Sprachbot sollte eine Einwilligung vorliegen, sofern der Anruf nicht ausdrücklich vom Kunden initiiert oder gesetzlich erlaubt ist.
- **Einwilligung für Aufzeichnung**: Soll der Inhalt des Gesprächs **mitgeschnitten** werden (Audioaufnahme), braucht es in Deutschland im Regelfall die Zustimmung des Anrufers *zu Beginn des Gesprächs*. Andernfalls verstößt die Aufnahme gegen das Persönlichkeitsrecht/

Strafrecht (§201 StGB, “Verletzung der Vertraulichkeit des Wortes”). Viele Unternehmen lösen das durch einen Hinweis ala *"Dieses Gespräch wird zu Qualitätszwecken aufgezeichnet. Falls Sie nicht einverstanden sind, sagen Sie bitte Bescheid."* – technisch muss dann der Bot auf Opt-out reagieren (z. B. Aufnahme stoppen oder menschlichen Agenten hinzuschalten).

- **Einwilligung bei sensiblen Daten (Art. 9 Abs. 2 lit. a DSGVO)**: Falls der Voicebot z. B. **Gesundheitsdaten** erhebt (etwa Symptomabfrage bei einem medizinischen Hotline-Bot) oder **biometrische Identifikation** vornimmt (Stimm-ID als Login), *bedarf es ausdrücklich der Einwilligung* hierfür, außer es greift ein anderer Ausnahmegrund des Art. 9 Abs. 2. Diese Einwilligung muss *informiert und ausdrücklich* sein (also aktiv zugestimmt, z. B. per Sprache “Ja”). Praktisch könnte man z. B. sagen: *"Um Sie anhand Ihrer Stimme zu erkennen, benötigen wir Ihr Einverständnis."*
- **ePrivacy-Aspekte**: Neben der DSGVO gilt in Telefonie-Kontexten auch das **Telekommunikations- und Telemedien-Datenschutzgesetz (TTDSG)** bzw. ePrivacy-Richtlinie. So ist z. B. für das **Auslesen oder Speichern von Informationen im Endgerät** eine Einwilligung nach §25 TTDSG erforderlich. Bei einem reinen Telefongespräch ist das weniger relevant, aber falls der Voicebot etwa eine Smartphone-App nutzt und dort Audio puffert, könnte es greifen. Auch **Rufnummernanzeige-Unterdrückung** ist geregelt (Outbound-Anrufe dürfen die Nummer nicht fälschen; für Werbeanrufe gilt Präsentationspflicht).

Fazit: Unternehmen müssen genau festzurren, auf welcher Rechtsgrundlage welcher Teil der Verarbeitung erfolgt. **Einwilligungen** sollten *gezielt* dort eingeholt werden, wo keine andere Rechtsgrundlage tragfähig ist – insbesondere Outbound-Calls zu Marketing/Sales, Aufzeichnungen und besondere Daten. Wichtig: Einwilligungen müssen den DSGVO-Anforderungen genügen (freiwillig, informiert, spezifisch, dokumentiert und jederzeit widerrufbar). Und: Falls keine Einwilligung vorliegt, sind die Daten *entsprechend eingeschränkt* zu verarbeiten (z. B. Gespräch nicht aufzeichnen, Bot nur inbound verwenden oder im Zweifel auf das Feature verzichten).

3.4 Auftragsverarbeitung oder gemeinsame Verantwortlichkeit?

Beim Einsatz externer Voice-AI-Plattformen stellt sich die Frage der **datenschutzrechtlichen Rolle** der Beteiligten: Ist der Anbieter ein **Auftragsverarbeiter** im Sinne von Art. 28 DSGVO, oder liegt evtl. eine **gemeinsame Verantwortlichkeit** (Art. 26 DSGVO) zwischen Unternehmen und Anbieter vor?

Grundsätzlich gilt: **Verarbeitet ein Dienstleister personenbezogene Daten strikt nach Weisung des Unternehmens zu dessen definierten Zweck**, spricht vieles für eine *Auftragsverarbeitung*. Der Voice-AI-Anbieter hat dann *keine eigenständigen Zwecke* mit den Daten, sondern stellt lediglich die Infrastruktur/Software bereit. Beispiel: Das Unternehmen nutzt einen Cloud-Sprachbot-Service, um Kundentelefonate abzuwickeln, und der Anbieter verarbeitet alle Anrufe *ausschließlich* zur Bereitstellung dieser Dienstleistung – **nicht** etwa, um eigene Sprachmodelle zu trainieren oder die Daten anderweitig zu verwenden. Hier liegt klassisch eine Auftragsverarbeitung vor [6]. Konsequenz: Es *muss* ein **Auftragsverarbeitungsvertrag (AVV)** abgeschlossen werden, und der Verantwortliche (das Unternehmen) bleibt für die DSGVO-Compliance federführend verantwortlich [7].

Eine **gemeinsame Verantwortlichkeit** hingegen kommt in Betracht, **wenn Anbieter und Unternehmen gemeinsam die Zwecke und Mittel bestimmen** [8]. Das ist z. B. denkbar, wenn *beide Parteien die Daten für eigene Interessen nutzen*: Etwa wenn der KI-Anbieter die Mitschnitte aller Kunden nutzt, um sein generelles Spracherkennungsmodell zu verbessern (d.h. er verfolgt den Zweck “KI-Training” neben dem Zweck des Unternehmens “Kundenservice”). Oder wenn ein Voicebot von zwei Firmen gemeinsam betrieben wird, die Datenpools zusammenwerfen, um einen **gemeinsamen KI-Service** zu trainieren [9]. In solchen Fällen müssen die Beteiligten eine **Art. 26 DSGVO Vereinbarung** schließen, die festlegt, wer welche Pflichten übernimmt (z. B. wer informiert Betroffene, wer

beantwortet Auskunftsanfragen usw.). *Wichtig:* Gemeinsame Verantwortlichkeit ist juristisch komplex – Unternehmen sollten versuchen, diese Konstellation zu vermeiden, außer es ist klar von Vorteil. Oft deklarieren Anbieter vertraglich, dass sie *nur Auftragsverarbeiter* sind, um die Rollen eindeutig zu halten.

Praxis-Hinweis: Prüfen Sie die **Vertragsdokumente** des Voice-AI-Anbieters. Seriöse Anbieter bieten einen AV-Vertrag an. Lesen Sie im Kleingedruckten, ob der Anbieter sich zusätzliche Rechte an den Daten einräumt (Stichwort "*Daten für Forschungszwecke nutzen*"). Falls ja, hinterfragen Sie, ob dies zulässig im Rahmen Auftragsverarbeitung abgedeckt werden kann (meist nur, wenn im Vertrag als *auftragsgemäßer Zweck* definiert und transparent gegenüber Betroffenen gemacht) – oder ob Sie dem widersprechen/nicht zustimmen können. Im Zweifel: Lieber vertraglich klarstellen, dass keine gemeinsame Verantwortung gewollt ist, und auf *strikte Zweckbindung* pochen.

Zuletzt: Unabhängig von der rechtlichen Konstruktion muss dem *Anrufer gegenüber* klargemacht werden, wer verantwortlicher **Ansprechpartner** für den Datenschutz ist. In der Regel wird das Ihr Unternehmen sein (da der Anrufer Ihre Dienstleistung nutzt), selbst wenn Sie einen Dienstleister einschalten. Machen Sie also nicht den Fehler zu glauben, man könne Verantwortung “auf den Anbieter abwälzen” – **das Unternehmen steht in der Pflicht**, die DSGVO-Prinzipien einzuhalten, egal wie leistungsfähig oder autonom der KI-Assistent arbeitet.

4. Auswahl eines DSGVO-konformen Anbieters

Die Wahl des richtigen Voice-AI-Anbieters entscheidet maßgeblich darüber, wie leicht Sie die DSGVO-Anforderungen umsetzen können. Achten Sie bei Evaluierung und Vertragsgestaltung auf folgende Punkte:

✔ **Auftragsverarbeitungsvertrag (AVV):** Stellen Sie sicher, dass der Anbieter bereit ist, einen **AVV nach Art. 28 DSGVO** zu schließen. In diesem Vertrag muss stehen, **zu welchem Zweck** und **in welchem Umfang** der Dienstleister die Daten **in Ihrem Auftrag** verarbeitet, und dass er **keine eigenen Verwendungsrechte** daran hat [10] [11]. Prüfen Sie den AVV-Entwurf auf Vollständigkeit: Er sollte u. a. **Gegenstand und Dauer** der Verarbeitung, **Art der personenbezogenen Daten** (z. B. Audiomitschnitte, Kundendaten) und **Kategorien der Betroffenen** nennen [12]. Ebenfalls geregelt sein müssen die **Pflichten des Anbieters**: Verarbeitung nur auf dokumentierte Weisung, Vertraulichkeit der Mitarbeitenden, ausreichende Sicherheit (**TOM**, s.u.), Unterstützung bei Betroffenenrechten, Meldung von Datenschutzvorfällen, **Löschung** bzw. Rückgabe der Daten nach Auftragsende [13] [14]. Wichtig ist auch eine Klausel, die den Anbieter nur mit *Zustimmung des Auftraggebers* **Unterauftragsverarbeiter** einsetzen lässt [15] – so behalten Sie Kontrolle über evtl. weitere Parteien (z. B. wenn der Voicebot-Anbieter wiederum Cloud-Services einbindet). **Tipp:** Achten Sie darauf, dass **alle Subdienstleister namentlich benannt** sind oder zumindest in einer stets aktuellen Liste verfügbar gemacht werden, der Sie zustimmen müssen. So wissen Sie, ob z. B. **Drittländer** im Spiel sind.

✔ **Datenübermittlung in Drittländer (SCC):** Viele Voice-AI-Lösungen greifen auf KI-Services aus den USA zurück (z. B. Speech-to-Text von Google, OpenAI-API etc.). Sobald **personenbezogene Daten außerhalb der EU** verarbeitet werden (oder von einem Anbieter, der dem Zugriff z. B. US-Behörden unterliegt), greift das DSGVO-Regime für **Datenübermittlung in Drittländer**. Das heißt, der Anbieter muss entweder in einem Land mit Angemessenheitsbeschluss sitzen (z. Z. hat die USA [mit dem "Data Privacy Framework"] hier nur teilweise einen, Vorsicht ist geboten), **oder** es müssen **Standardvertragsklauseln (Standard Contractual Clauses, SCC)** abgeschlossen werden sowie ggf. **zusätzliche Schutzmaßnahmen** implementiert werden. Ein Auftragsverarbeiter aus den USA wird im AVV i. d. R. die EU-Standardvertragsklauseln anhängen (oft als **Anhang/Exhibit** im Vertrag) [16] [17]. Prüfen Sie, dass dies vorhanden ist und unterschrieben wird. Zusätzlich verlangen viele

Aufsichtsbehörden eine **Transfer-Folgenabschätzung** (Transfer Impact Assessment), in der der Anbieter darlegt, welche *Risiken* durch Zugriffe z. B. nach US-Recht (Stichwort CLOUD Act) bestehen und welche *zusätzlichen Maßnahmen* (z. B. Verschlüsselung der Daten in der Cloud, Anonymisierung) ergriffen werden, um ein Schutzniveau wie in der EU zu gewährleisten. Als Unternehmen sollten Sie sich vergewissern, dass der Anbieter die Daten **nicht ungeschützt Behörden preisgeben** muss. (Hinweis: US-Anbieter unterliegen dem CLOUD Act, der sie verpflichten kann, Daten offenzulegen, selbst wenn die Server in der EU stehen [18]. Deshalb sind vertragliche und technische Schutzmaßnahmen so wichtig.) Wenn möglich, bevorzugen Sie **EU-Anbieter oder EU-Hosting**, um das Drittland-Thema zu umgehen. Falls US-Dienste unverzichtbar sind, holen Sie rechtlichen Rat ein und dokumentieren Sie die getroffenen Vorkehrungen.

✔ **Technische und organisatorische Maßnahmen (TOM):** Nach Art. 32 DSGVO muss der Anbieter **ausreichende Sicherheit** der Verarbeitung garantieren. Im AVV oder den Anhängen sollten **TOMs** aufgeführt sein – typische Maßnahmen sind: **Verschlüsselung** der Audio-Datenübertragung (z. B. TLS für Sprachstreaming), **Verschlüsselung von gespeicherten Mitschnitten** oder Transkripten, **Zugriffsbeschränkungen** (Role-Based Access, Need-to-know Prinzip für Support-Mitarbeiter), **Pseudonymisierung** (z. B. Trennung von Kundennummer und Inhaltsdaten in Logs), **Logging und Monitoring** von Zugriffsversuchen, regelmäßige **Backups** und getestete Wiederherstellungsverfahren, und Konzepte für **Datenschutz durch Voreinstellungen** (Privacy by Default) sowie **Datenschutz in der Technikgestaltung** (Privacy by Design). Der Dienstleister muss mindestens den Stand der Technik berücksichtigen. *Prüfen Sie kritisch:* Sind z. B. Aufzeichnungen nur abrufbar über gesicherte Interfaces? Wie stellt der Anbieter sicher, dass unbeabsichtigte Aufnahmen (z. B. wenn der Bot jemanden im Hintergrund mitschneidet) gelöscht werden? Hat der Anbieter Prozesse, um Daten bei Ihnen auf Wunsch zu löschen oder herauszugeben? – Ein guter Anhaltspunkt: Nach Art. 28 Abs. 1 DSGVO dürfen Sie **nur mit solchen Auftragsverarbeitern zusammenarbeiten, die durch geeignete TOM ein angemessenes Datenschutzniveau garantieren** [19]. Fordern Sie also gerne Nachweise ein, z. B. **Zertifikate** (ISO 27001, SOC 2) oder detaillierte Sicherheitskonzepte.

✔ **Serverstandort vs. Datenzugriff:** Viele Anbieter werben mit "Serverstandort in Deutschland/EU". Das ist positiv, jedoch zu kurz gedacht, wenn das Unternehmen hinter dem Service aus einem Drittland stammt. Entscheidend ist nicht nur, *wo* die Server physisch stehen, sondern **wer darauf zugreifen kann**. Ein US-Unternehmen mit EU-Rechenzentrum unterliegt immer noch US-Gesetzen (siehe CLOUD Act oben) [18]. Umgekehrt kann ein EU-Anbieter, der Subunternehmer in den USA nutzt, ebenfalls ein Risiko darstellen. Achten Sie deshalb auf die **juristische Kontrolle**: Ist der Anbieter in Europa rechtlich selbständig? Werden Daten wirklich nur in der EU verarbeitet oder fließen sie für Support, Wartung, AI-Auswertung doch wieder in die USA? Dieses Thema ist z. B. relevant im Vergleich *“VAPI vs. Aurili”*: Wenn *Voice API*-Plattformen global agieren (mit evtl. US-Muttergesellschaft), müssen strenge vertragliche Vorkehrungen getroffen werden. Ein Anbieter wie *Aurili* hingegen, der explizit DSGVO-Konformität betont, könnte z. B. ein rein europäisches Hosting mit voller Datenhoheit in der EU bieten. Allerdings auch hier: Prüfen Sie, ob möglicherweise externe KI-Module angebunden werden. **Bottom Line:** Lassen Sie sich *schriftlich zusichern*, wo und durch wen die Daten verarbeitet werden. Wenn EU-only versprochen wird, sollte das im Vertrag stehen (“Datenverarbeitung ausschließlich innerhalb EU/EWR durch Unterauftragnehmer X, Y, Z…” plus Klausel, dass jede Abweichung zustimmungspflichtig ist).

✔ **Reputation und Transparenz:** Wählen Sie einen Anbieter, der **offenlegt**, wie sein System funktioniert. Dazu gehören **Dokumentationen** zur Datenverarbeitung, eine leicht auffindbare **Datenschutzerklärung** des Anbieters, ggf. Whitepapers zur DSGVO-Compliance. Prüfen Sie auch Referenzen: Haben andere (vergleichbare) Unternehmen diesen Dienst schon erfolgreich datenschutzkonform im Einsatz? Wurden evtl. schon **Audits** durchgeführt? Ein vertrauenswürdiger Anbieter wird Ihnen dabei helfen, alle nötigen Infos zusammenzutragen – sei es für Ihre DSFA, für Rückfragen der Aufsicht oder für die Nutzerinformation.

5. Umgang mit externen (insb. US-) Plattformen

Manche der besten Spracherkennungssysteme und Sprach-KIs stammen von **großen Tech-Plattformen** (Google, Amazon AWS, Microsoft Azure, OpenAI etc.). Diese zu nutzen kann in puncto *Funktionalität* Vorteile bringen, wirft aber in puncto *Datenschutz* besondere Fragen auf. Grundsätzlich gilt: **US-Anbieter dürfen für DSGVO-zweckkonforme Verarbeitung eingesetzt werden**, *wenn* entsprechende **Zusatzmaßnahmen** ergriffen werden. Hier ein Leitfaden zum Vorgehen:

➡ **Bedingungen für den zulässigen Einsatz von US-Services:** Zunächst muss – wie in Abschnitt 4 erwähnt – ein gültiger **Rechtsmechanismus** für den Datentransfer vorliegen (SCC oder EU-US Data Privacy Framework, sofern anwendbar). Zusätzlich sollten Unternehmen **Schutzmaßnahmen** implementieren: Dazu gehört z. B. **Verschlüsselung** der Inhalte *vor* der Übertragung, soweit praktikabel. In einem Voicebot-Kontext ist client-seitige Verschlüsselung schwierig (der Dienst muss ja den Klartext hören, um ihn zu verstehen). Alternativ kann man technische Maßnahmen ergreifen wie **Datensparsamkeit** (nicht mehr personenbezogene Info mitschicken als nötig – z. B. Audio ohne zugehörige Kundendaten, sofern möglich), oder das **Opt-Out von Datenpersistenz**: Viele Cloud-STT-Dienste bieten an, dass sie die Audio-Daten *nicht für eigene Zwecke speichern*, oft gegen Aufpreis [20] . So kann man verhindern, dass z. B. Google die Sprachdaten ins Training aufnimmt. Eine weitere Maßnahme: **Pseudonymisierung** – statt Kundennamen könnte der Bot z. B. nur sagen *"Guten Tag, Kunde 12345"* zum STT schicken, und die Zuordnung Name↔Nummer erfolgt erst intern. All das reduziert das Risiko bei US-Plattformnutzung.

➡ **Vertragsgestaltung und Kontrolle:** Wenn Sie einen US-Dienst direkt nutzen (z. B. eigenständig die Azure Cognitive Services für STT/TTS einbinden), dann schließen Sie selbst mit Microsoft einen AVV und SCC. Achten Sie genau auf die **Einstellungen**: Viele Cloud-Angebote haben Konfigurationsmöglichkeiten für Datenspeicherung (z. B. "Cognitive Services: Data Logging off"). Nutzen Sie möglichst **Rechenzentrumsregionen in der EU** bei der Buchung. Dokumentieren Sie die Konfiguration als Nachweis. Führen Sie idealerweise eine **Transferfolgenabschätzung** durch, in der Sie bewerten, ob bspw. im konkreten Einsatzfall ein Zugriff durch US-Behörden wahrscheinlich ist oder ob die Art der Daten (z. B. belanglose Hotline-Dialoge) kein ernsthaftes Risiko darstellen – und notieren Sie, welche technischen Maßnahmen (siehe oben) ergriffen wurden.

➡ **Hybride Modelle (EU-only vs. gemischt):** Einige Anbieter – wie z. B. der erwähnte Aurili – erlauben einen **hybriden Ansatz**. Das bedeutet, Sie können konfigurieren, ob der Voicebot **ausschließlich mit EU-Komponenten** arbeitet oder ob er bestimmte externe KI-Dienste einbindet, um die Erkennungsrate oder Fähigkeiten zu verbessern. Beispiel: Ein Voicebot läuft auf europäischer Infrastruktur, nutzt aber optional für komplexe Fälle die API eines US-LLM (Large Language Model) zur Sprache-Analyse. In so einem "gemischten" Modus müssen Sie sehr genau steuern, **wann** Daten in die externe Cloud gehen. Möglicherweise kann man dies von Fall zu Fall abhängig machen (z. B. nur wenn der Nutzer eingewilligt hat oder nur für nicht-sensitive Anfragen). Als Unternehmen sollten Sie solche Optionen mit dem Anbieter besprechen. **Idealerweise** gibt es einen **"EU-only"-Schalter**, mit dem Sie strikt alle Daten innerhalb EU halten können – dann haben Sie datenschutzrechtlich weniger Kopfzerbrechen (dafür evtl. etwas weniger KI-"Intelligenz" oder höhere Kosten). Wenn Sie sich für **gemischte Verarbeitung** entscheiden, brauchen Sie ein besonders gutes Konzept zur **Nutzeraufklärung** (siehe nächster Punkt) und müssen alle oben genannten Transferbedingungen einhalten. Halten Sie auch vertraglich fest, welche Drittanbieter konkret zum Einsatz kommen dürfen.

➡ **Beispiel "VAPI mit SCC":** Angenommen, Sie nutzen eine *Voice API* Plattform aus den USA (hier als Platzhalter "VAPI"), weil diese hervorragende mehrsprachige Spracherkennung bietet. Sie würden dann: a) einen AVV/SCC mit dem VAPI-Anbieter abschließen, b) sicherstellen, dass VAPI alle

Unterauftragsverarbeiter offenlegt, c) ggf. dem Endnutzer **transparent mitteilen**, dass seine Spracheingaben von einem Dienstleister in den USA verarbeitet werden (inkl. Hinweis auf Standardklauseln), d) soweit möglich Einstellungen wählen, die die Datenverarbeitung beschränken (z. B. keine Speicherung nach Transkription) und e) regelmäßig prüfen/auditieren, ob VAPI die Zusagen einhält. Die Nutzung wäre dann *zulässig*, solange die DSGVO-Prinzipien eingehalten werden. **Im Falle eines Falles** (z. B. bei einer Beschwerde) müssen Sie darlegen können, dass die Drittlandsverarbeitung rechtmäßig abgesichert war.

Zusammengefasst: Externe Plattformen können genutzt werden, erfordern aber **besondere Sorgfalt**. Viele Unternehmen fahren zweigleisig: Sie bevorzugen EU-Lösungen, wo möglich, und dort, wo (noch) US-Dienste im Spiel sind, sorgen sie mit Verträgen, Technik und Information für maximalen Schutz. Bleiben Sie zudem auf dem Laufenden zu rechtlichen Entwicklungen (Stichwort Schrems-II-Nachfolger, neue Abkommen), denn was heute als Schutzmaßnahme reicht, könnte morgen angepasst werden müssen.

6. Privacy Policy & Nutzeraufklärung

Eine **klare und vollständige Information** der Nutzer ist essenziell – sowohl rechtlich (Transparenzpflicht nach Art. 13 DSGVO) als auch für die **Akzeptanz** Ihres Voicebots. Nutzer (Kunden, Anrufer) sollten verstehen, *was mit ihren Daten passiert*, wenn sie mit dem Sprachassistenten interagieren. Folgende Punkte sollten in **Datenschutzhinweisen** und ggf. Ansagen berücksichtigt werden:

Wesentliche Inhalte für die Datenschutzerklärung:

- **Verantwortlicher:** Nennen Sie Ihr Unternehmen mit Kontaktadresse und den Datenschutzbeauftragten (falls vorhanden) als verantwortliche Stelle.
- **Datenkategorien:** Erklären Sie, *welche Daten* der Voice-AI-Agent verarbeitet. Z. B.: *"Wenn Sie unsere Hotline anrufen, verarbeitet unser Sprachassistenzsystem Ihre Sprachdaten. Konkret werden Ihre gesprochenen Worte aufgezeichnet und in Text umgewandelt. Gegebenenfalls werden dabei Ihr Name, Kundennummer oder andere von Ihnen genannte personenbezogene Informationen erfasst."* Auch Meta-Daten erwähnen: *"Zudem protokollieren wir Datum und Uhrzeit des Anrufs, Ihre Rufnummer sowie das Anliegen (in Stichworten)."*
- **Zwecke der Verarbeitung:** Legen Sie dar, *wozu* diese Daten verwendet werden. Etwa: *"zur automatischen Bearbeitung Ihrer Anfrage und Beantwortung in Form einer computergenerierten Sprachansage"*. Mögliche Zwecke sind z. B. **Kundenservice**, **Terminvereinbarung**, **Bearbeitung von Aufträgen** oder **Weiterleitung von Informationen**. Wenn Sie die Daten auch zu Trainingszwecken nutzen, muss auch *dieser Zweck* genannt werden (z. B. *"Fortlaufende Verbesserung unseres Sprachsystems (Qualitätssicherung)"*). Wichtig: *Keine Zweck-Offenlassung!* (also nicht pauschal “wir nutzen die Daten für KI” ohne nähere Erklärung).
- **Rechtsgrundlage:** Geben Sie für jeden Zweck die **Rechtsgrundlage** an (siehe Punkt 3.3 oben). Z. B.: *"Die Verarbeitung Ihrer Sprachdaten erfolgt auf Grundlage von Art. 6 Abs. 1 lit. b DSGVO (Vertragserfüllung), da sie erforderlich ist, um Ihre Support-Anfrage zu beantworten."* Bei Werbeanrufen würde stehen: *"... auf Grundlage Ihrer Einwilligung nach Art. 6 Abs. 1 lit. a DSGVO."* Falls besondere Daten verarbeitet werden mit Einwilligung, auch Art. 9 Abs. 2 lit. a erwähnen.
- **Einsatz von Dienstleistern:** Geben Sie an, ob und welche **Auftragsverarbeiter** oder Partner eingebunden sind. Beispiel: *"Hierfür setzen wir den KI-Telefonassistenz-Dienst* XYZ *ein, der in unserem Auftrag die Spracheingaben technisch verarbeitet."* Falls Daten an Drittstaaten gehen, **muss** ein entsprechender Hinweis hinein, inkl. Verweis auf die *Sicherungen* (z. B. *"Datenübermittlung in die USA auf Grundlage von EU-Standardvertragsklauseln"*). Transparenz

erfordert, dass der Nutzer erfährt, dass z. B. seine Stimme von einem bestimmten Cloud-Service transkribiert wird. Viele Unternehmen verlinken hier auf die Datenschutzerklärung des Anbieters oder fassen das Wesentliche zusammen.

- **Speicherdauer:** Nutzer sollten erfahren, wie lange die Daten gespeichert werden. Z. B.: *"Die Audio-Aufzeichnungen Ihres Anrufs werden nicht dauerhaft gespeichert, sondern nach Ende des Telefonats unmittelbar gelöscht."* (Falls das System ohne Aufzeichnung auskommt). Oder: *"Transkripte der Gespräche bewahren wir für max. 30 Tage zu Beweiszwecken auf und löschen sie dann automatisch."* Wenn Trainingsdaten länger vorgehalten werden, angeben, wie lange bzw. nach welchen Kriterien (z. B. *"Wir überprüfen alle 6 Monate, ob eine weitere Aufbewahrung der Mitschnitte zu Schulungszwecken erforderlich ist, und löschen ansonsten die personenbezogenen Teile."*).
- **Rechte der Betroffenen:** Selbstverständlich müssen die **Betroffenenrechte** aufgelistet werden: Auskunft, Berichtigung, Löschung, Einschränkung, Widerspruch, Datenübertragbarkeit, Widerruf von Einwilligungen, Beschwerderecht bei der Aufsichtsbehörde. Wichtig in diesem Kontext: Weisen Sie auf das **Widerspruchsrecht nach Art. 21 DSGVO** hin, insbesondere wenn Sie sich auf berechtigtes Interesse stützen (*"Sie können der Verarbeitung Ihrer Daten durch den Sprachassistenten jederzeit widersprechen, z. B. indem Sie dies im Gespräch angeben oder uns nachträglich kontaktieren."*). Bei Einwilligung: Hinweis, dass *Widerruf jederzeit mit Wirkung für die Zukunft* möglich ist, ohne Angabe von Gründen.
- **Automatisierte Entscheidung/Bewertung:** Falls der Voice-AI-Agent Entscheidungen trifft, die *rechtliche Wirkung* oder *ähnlich erhebliche Beeinträchtigung* für den Nutzer haben (Art. 22 DSGVO), müssten Sie das besonders erwähnen und Rechte einräumen. In der Regel wird ein Sprachassistent aber kaum alleine solch gravierende Entscheidungen treffen (er verweigert höchstens eine Auskunft mangels Verifikation o. ä.). Zur Sicherheit könnte man klarstellen: *"Es findet keine automatisierte Entscheidungsfindung im Sinne des Art. 22 DSGVO statt."* (sofern zutreffend).
- **Kontakt & sonstiges:** Verweisen Sie darauf, wie der Nutzer weitere Infos bekommen kann oder seine Rechte ausübt (z. B. *"Bei Fragen zur Verarbeitung durch unseren KI-Assistenten können Sie sich an datenschutz@firma.de wenden."*).

Neben der formalen Datenschutzerklärung (etwa auf Ihrer Website oder als PDF zum Download) sollten Sie auch überlegen, **wie Sie den Anrufer direkt informieren**, denn bei telefonischen Diensten ist die unmittelbare Transparenz tricky: Der Nutzer nimmt ggf. gar nicht wahr, dass ein KI-System beteiligt ist, und er liest während des Anrufs keine Web-Erklärung. **Best Practices** sind hier:

- **Initiale Ansage:** Lassen Sie den Voicebot am Anfang des Gesprächs *klarstellen*, dass der Anruf gerade von einem automatisierten System bearbeitet wird. Z. B.: *"Hallo, ich bin der virtuelle Assistent von Firma X und helfe Ihnen heute weiter."* – So weiß der Anrufer, es spricht keine menschliche Person. Das ist zwar nicht explizit gesetzlich vorgeschrieben, aber entspricht dem Transparenzgebot und verhindert Irreführung.
- **Kurzinfo Datenschutz:** Eine kompakte Info für den Anrufer zu Beginn kann lauten: *"Hinweis: Dieses Gespräch wird von einer KI automatisiert verarbeitet. Ihre Angaben werden gemäß unserer Datenschutzerklärung verarbeitet."* – Das ist natürlich heikel, man will den Anrufer nicht verschrecken. Eine Möglichkeit ist, **sehr kurze Hinweise** zu geben (*"Dieser Anruf wird aufgezeichnet, wenn Sie nicht einverstanden sind, sagen Sie 'Stop'."*). Oder der Bot fragt: *"Darf ich Ihre Anfrage automatisiert verarbeiten?"* – was in etwa einer Einwilligung gleichkäme. In der Praxis versuchen viele, die notwendigen Hinweise *kurz und prägnant* in die Begrüßung zu packen, und verweisen dann auf ausführliche Infos online oder per SMS.
- **Mehrstufige Aufklärung:** Eine Idee ist, **nach dem Gespräch** (oder währenddessen) zusätzliche Infos bereitzustellen. Z. B. könnte man per SMS oder E-Mail einen Link zur Datenschutzerklärung

schicken – *"Details zum Datenschutz unseres Sprachsystems finden Sie hier: …"* – gerade bei Outbound könnte das vorab erfolgen.

Im Folgenden zwei **Beispieltexte** (vereinfacht), wie eine Datenschutzhinweis-Sektion für Voice-AI aussehen kann:

Beispiel (Inbound, Kundenhotline):

> **Datenschutzinformation zum KI-Telefonassistenten:** Wenn Sie unsere Hotline anrufen, kann Ihr Anruf durch einen digitalen Telefonassistenten bearbeitet werden. Dabei wird Ihre **Stimme** aufgezeichnet und in Text umgewandelt, um Ihr Anliegen zu verstehen. Der Assistent gibt Ihnen anschließend per computergenerierter **Sprachausgabe** die passende Antwort. **Zweck** dieser Verarbeitung ist die schnelle und zuverlässige Bearbeitung Ihrer Anfrage (Kundenservice). **Verarbeitet werden:** Ihre **Sprachdaten** (Inhalt des Gesprächs) sowie Meta-Daten des Anrufs (Zeit, Dauer, Rufnummer). **Rechtsgrundlage:** Art. 6 Abs. 1 lit. b DSGVO (Vertrag/Anfrage) – Sie haben durch den Anruf unsere Dienstleistung in Anspruch genommen. **Dienstleister:** Zur Bereitstellung des Telefonassistenten setzen wir die Firma XYZ (Auftragsverarbeiter) ein. Gespräche können dabei zur Spracherkennung temporär an Server von XYZ in Deutschland übertragen werden. Ein angemessener Datenschutz ist vertraglich gewährleistet. **Speicherung:** Das Gespräch wird **nicht dauerhaft aufgezeichnet**; es erfolgt lediglich eine automatische Echtzeit-Verarbeitung. Lediglich anonymisierte Gesprächsprotokolle (ohne Personenbezug) behalten wir für maximal 30 Tage zur Verbesserung des Systems. **Ihre Rechte:** Sie können jederzeit Auskunft über die Sie betreffenden Daten erhalten sowie Berichtigung oder Löschung verlangen. Weitere Informationen finden Sie in unserer ausführlichen Datenschutzerklärung auf unserer Website.

Beispiel (Outbound, Termin-Erinnerung):

> **Hinweis zum automatisierten Anruf:** Sie werden von unserem **automatischen Telefonassistenten** kontaktiert, um Sie an Ihren bevorstehenden Termin zu erinnern. Hierbei verwendet das System Ihren **Namen** und den Termin aus unseren Kundendaten, um Ihnen persönlich die Details zu nennen. **Zweck und Rechtsgrundlage:** Die Verarbeitung erfolgt zur Erfüllung unseres Vertrags/Services (Art. 6 Abs.1 lit. b DSGVO), da wir Sie wie vereinbart rechtzeitig informieren. Alternativ stützen wir uns auf Ihre **Einwilligung** (Art. 6 Abs.1 lit. a DSGVO), sofern gesetzlich erforderlich, welche Sie uns im Rahmen der Terminvereinbarung gegeben haben. **Ablauf der Datenverarbeitung:** Eine computergenerierte Stimme ruft Ihre gespeicherte Telefonnummer an und spielt die Erinnerung ab. Falls Sie antworten oder Fragen haben, wird Ihre Antwort von der KI verarbeitet, um ggf. gleich darauf zu reagieren (z. B. Terminbestätigung oder Umbuchung). **Dienstleistereinsatz:** Der Anruf erfolgt über den Dienst ABC, der in unserem Auftrag die KI-Technologie bereitstellt. Dabei können technisch bedingt Daten wie Ihre Telefonnummer an ABC übermittelt werden. Mit ABC besteht ein DSGVO-konformer Vertrag. **Keine Werbung:** Dieser Anruf dient ausschließlich Informationszwecken zu Ihrem Termin, nicht der Werbung. **Ihre Wahlrechte:** Wenn Sie solche automatisierten Anrufe zukünftig nicht erhalten möchten, können Sie uns dies jederzeit mitteilen – wir werden das respektieren (Widerruf Ihrer Einwilligung/ Widerspruch). Unsere vollständigen Datenschutzinformationen finden Sie unter …

Die obigen Beispiele müssten natürlich an den konkreten Fall angepasst werden, sind aber ein Anhaltspunkt.

Wichtig: Halten Sie die Sprache **verständlich und nüchtern**, vermeiden Sie zu juristische Formulierungen im gesprochenen Hinweis (am Telefon versteht keiner "Standardvertragsklauseln"). In der schriftlichen Erklärung dürfen Sie ruhig präzise sein – diese richtet sich ggf. auch an Datenschutzexperten, die prüfen wollen, ob alles genannt ist. Überfrachten Sie aber weder die Ansage noch den Text mit unnötigen Details. Das Ziel ist, dass ein *normaler Anrufer* grob versteht: *Wer verarbeitet meine Daten? Was passiert damit? Warum?* – dann sind Sie auf einem guten Weg.

7. On-Premise-Lösung als Sonderfall

Für bestimmte Branchen und Einsatzfälle kann es sinnvoll oder sogar erforderlich sein, den Voice-AI-Agenten **On-Premise** (also auf eigener Infrastruktur) zu betreiben, statt eine Cloud-Lösung zu nutzen.

Wann On-Premise erwägen? Vor allem **hochsensitive Bereiche** wie **Banken/Finanzwesen**, **Gesundheitswesen** (Krankenhäuser, Arztpraxen) oder auch öffentliche Stellen mit schutzwürdigen Informationen (Behörden, Polizei-Hotlines etc.) setzen oft auf maximale Kontrolle der Daten. Hier möchte man vermeiden, dass Kundendaten oder vertrauliche Inhalte überhaupt das eigene Netzwerk verlassen. Auch Unternehmen, die strengen **Geheimhaltungs- oder Compliance-Regeln** unterliegen (z. B. Anwaltskanzleien mit Berufsgeheimnis, Rüstungsindustrie), könnten On-Prem bevorzugen. In solchen Fällen minimiert eine Inhouse-Lösung die Abhängigkeit von externen Diensten und reduziert das Risiko von **Datenabflüssen** oder Zugriffen Fremder.

Vorteile: Alle Sprachdaten bleiben innerhalb der eigenen IT-Landschaft, unter Kontrolle der eigenen IT-Abteilung. Man kann Systeme im eigenen Rechenzentrum härten, Zugriffe fein steuern und hat keine Drittlandtransfers. Dadurch entfallen viele der oben diskutierten Probleme (SCCs, Cloud Act etc.) von vornherein. Zudem lässt sich On-Prem oft **besser in Legacy-Systeme** integrieren (z. B. direkt ans interne CRM anbinden) und man ist nicht von den AGB externer Provider abhängig.

Nachteile/Herausforderungen: On-Premise hat natürlich auch Haken. Die **Initialkosten** sind höher – man benötigt Hardware oder mindestens VMs/Kubernetes-Kapazitäten, eventuell spezielle Beschleuniger (GPUs für KI-Modelle). Die **Implementierung** kann komplex sein (siehe Beispiel aus der Technik: eine eigene Infrastruktur mit WebRTC-Server, STT/LLM/TTS-Services, Logging, Monitoring muss aufgesetzt werden [3]). Zudem ist man selbst verantwortlich für Updates, Model-Optimierungen etc., außer der Anbieter liefert ein Paket mit Wartung. Viele On-Prem-Lösungen hinken funktional etwas hinterher, da Cloud-Lösungen schneller mit den allerneuesten KI-Modellen arbeiten können.

Aurili On-Premise Beispiel: Der Anbieter Aurili (ein KI-Telefonassistent) bietet nach eigenen Aussagen eine **DSGVO-konforme** und **sofort einsatzbereite** Lösung. Es ist anzunehmen, dass Aurili für z. B. Bankkunden oder im Gesundheitsbereich auch einen **On-Premise-Betrieb** unterstützt. D.h. die Software (STT, Dialog-KI, TTS) kann auf Servern direkt beim Kunden laufen. So was funktioniert oft als Appliance oder Docker-Setup, geliefert vom Hersteller. Der Vorteil: Das Unternehmen behält die *Datenhoheit*. Gerade wenn Spracheingaben möglicherweise sensible Kontodaten oder Patientendaten enthalten, ist das für die Risikoabwägung Gold wert. Auch entfallen hier viele vertragliche Konstruktionen – der Anbieter fungiert primär als **Softwarelieferant**; sofern er keine Einblicke in den Live-Betrieb hat, fließen personenbezogene Daten nicht an ihn, was datenschutzrechtlich einfacher ist (ggf. gar kein AVV nötig, außer für Support mit Testdaten).

Wann sollte man wirklich On-Prem in Betracht ziehen? Wenn **Gesetzgebung oder Aufsicht** es praktisch verlangen. Z. B. Banken könnten von der BaFin Auflagen bekommen, bei kritischen Kundendaten keine US-Cloud zu verwenden. Oder ein Krankenhaus möchte sicherstellen, dass Patienteninformationen nur inhouse verarbeitet werden, um allen Datenschutzgesetzen (und Ethik)

gerecht zu werden. Auch wenn eine DSFA ergibt, dass das Risiko bei Cloud zu hoch wäre, könnte On-Prem eine risikominimierende Maßnahme sein.

Hybrid als Alternative: Manche Unternehmen wählen einen **hybriden Ansatz**: Die *nicht-sensiblen* Teile laufen in der Cloud, für *sensible Daten* schwenkt man auf On-Prem-Module um. Beispielsweise könnten Routinefragen vom Cloud-Bot beantwortet werden, aber sobald es um persönliche Gesundheitsdaten geht, übernimmt ein lokal installiertes Modul (oder ein Mensch). Solche Modelle sind komplexer, aber denkbar.

Fazit: On-Premise lohnt sich, wenn Datenschutz **oberste Priorität** hat und man bereit ist, dafür mehr Ressourcen in Technik und Betrieb zu stecken. Anbieter wie Aurili, die explizit mit EU-Compliance werben, zeigen, dass Marktbedarf hierfür besteht. Falls Sie On-Prem in Erwägung ziehen: Planen Sie genügend Zeit für Implementierung und Tests ein. Prüfen Sie auch, ob der Anbieter regelmäßige **Updates/Security-Patches** liefert – denn die Verantwortung für den laufenden Betrieb liegt dann bei Ihnen. Und vergessen Sie nicht: Auch ein On-Prem-System muss DSGVO-konform sein (Privacy by Design, Logs, Löschkonzepte etc. in der Software). Es ist kein Freifahrtschein, aber in gewissen Szenarien der sicherste Hafen für Daten.

8. Best Practices & häufige Fallstricke

Abschließend einige bewährte Praktiken und Hinweise, um typische Fehler beim Einsatz von Voice-AI-Agenten zu vermeiden:

- **Einwilligung bei Outbound-Calls einholen:** Einer der häufigsten Fehler ist, **automatisierte Anrufe ohne gültige Einwilligung** durchzuführen. In Deutschland sind Werbeanrufe ohne vorheriges Opt-in *illegal* [5] – das gilt auch, wenn ein KI-Bot anruft statt einer Person. Stellen Sie also sicher, dass Sie vorab die ausdrückliche Zustimmung des Kontakts haben, *bevor* ein Voice-Agent einen Kunden z.B. zu Marketingzwecken kontaktiert. Dokumentieren Sie diese Einwilligung (z. B. Häkchen im CRM mit Timestamp). Auch bei Bestandskunden darf man nicht einfach automatisiert anrufen, es sei denn, es handelt sich um erlaubte Service-Calls (z. B. Terminerrinnerung, siehe oben) – und selbst da ist Vorsicht geboten ohne Einwilligung, je nach Inhalt. **Praxis-Tipp:** Lassen Sie Outbound-Bots am Anfang fragen *"Haben Sie jetzt kurz Zeit? Darf ich fortfahren?"* – so kann der Angerufene zumindest den Anruf abblocken, falls er ihn nicht wünscht. Und natürlich: bei jedem Outbound-Call eine **Abmeldemöglichkeit** bieten (z. B. “Drücken Sie 2, um keine weiteren Anrufe zu erhalten”). So bewegen Sie sich auf der sicheren Seite und wahren die Nutzerrechte.

- **DSFA sauber durchführen und Nachweise aufbewahren:** Wie erwähnt, gehört eine **Datenschutz-Folgenabschätzung** in fast allen Fällen zum Voicebot-Projekt. Ein Stolperstein ist, diese zwar irgendwie zu machen, aber nicht ordentlich zu **dokumentieren** oder die festgelegten Maßnahmen dann nicht umzusetzen. Halten Sie Ihre DSFA **schriftlich fest** (ggf. als Bericht mit Datum, Verantwortlichen, Freigabe). Sollten die Aufsichtsbehörden später nachfragen, können Sie mit diesem Dokument zeigen, dass Sie die Risiken geprüft haben. Achten Sie besonders darauf, in der DSFA **Maßnahmen** abzuleiten (z. B. “Audio wird Ende-zu-Ende verschlüsselt”, “Benutzer wird zu Beginn informiert”) und diese dann auch tatsächlich einzuhalten. Ein häufiger Fallstrick ist, dass eine DSFA zwar Risiken erkennt (z. B. “US-Dienst, Risiko behördlicher Zugriff”), aber keine ausreichenden Maßnahmen dazu beschrieben oder umgesetzt werden – was im Audit negativ auffällt. Daher: Nutzen Sie die DSFA wirklich als *Lenkungsinstrument* und updaten Sie sie, falls sich etwas ändert (neuer Dienstleister, neue Funktion im Bot etc.). Und: Sollte eine

DSFA mal ergeben, dass das Risiko trotz allem zu hoch ist, **ziehen Sie die Reißleine**, bevor es ein Datenschutzvorfall wird.

- **Unterauftragsverarbeiter-Ketten klären:** Voice-AI-Lösungen sind oft ein Konglomerat aus Diensten. Ein häufiger Fallstrick ist die **Intransparenz über Subunternehmer**. Beispiel: Sie schließen einen Vertrag mit Anbieter A. Im Kleingedruckten steht, dass A Dienste von B, C, D in Anspruch nimmt (z. B. Cloudplattform, SMS-Service, STT-Engine von Google). Wenn diese Kette nicht sauber offengelegt ist, laufen Sie Gefahr, dass irgendwo Daten hingehen, von denen Sie nichts wussten – und die weder im AVV noch in Ihrer Datenschutzerklärung berücksichtigt sind. Fordern Sie daher vom Anbieter eine **Liste aller Subprozessoren**. Prüfen Sie, ob für jeden einzelnen wiederum Verträge/SCC bestehen. Wenn der Anbieter US-Subdienste nutzt, fragen Sie nach, welche *zusätzlichen Vorkehrungen* bestehen. **Im Zweifelsfall** müssen Sie diese Kette dem Nutzer gegenüber transparent machen. Der schlimmste Fall ist, Sie versprechen "alles bleibt in EU", aber der Anbieter lässt doch irgendwelche Daten in die USA wandern – das wäre ein Vertrauensbruch und ein DSGVO-Problem. Also: **Lieber offen kommunizieren** und ggf. nach EU-Alternativen für kritische Subdienste fragen.

- **Datensicherheit nicht vernachlässigen:** Ein praktischer Fallstrick ist, dass man sich auf die Funktion des Voicebots konzentriert und dabei **IT-Security**-Basics übersieht. Beispielsweise sollten Sprachaufzeichnungen auf dem Server **verschlüsselt gespeichert** werden, sonst könnten bei einem Hack alle Kundengespräche geleakt werden (GAU!). Oder: Zugang zu Bot-Logs über ein Web-Dashboard – stellen Sie sicher, dass **nur Berechtigte** (z. B. Admins, ausgewählte Data Scientists) Zugriff haben, und dass ein Role-Concept verhindert, dass jeder Support-Mitarbeiter alle Gespräche einsehen kann. Setzen Sie **Passwortschutz/MFA** für die Bot-Konsole ein. Schauen Sie auch auf **Netzwerkebene**: Der Bot, falls On-Prem, sollte in einem Segment laufen, wo er keinen unnötigen Zugang zu internen Datenbanken hat, außer was er benötigt. **Logging**: Protokollieren Sie Zugriffe auf die Daten, damit im Falle eines Incidents nachvollziehbar ist, wer was wann abgerufen hat. Diese Sicherheitsmaßnahmen gehören zu den TOM und sollten teils vom Anbieter kommen, teils von Ihrer internen IT umgesetzt werden.

- **Privacy by Design und Default umsetzen:** Bei der Einrichtung des Sprachassistenten stellen Sie sicher, dass **datenschutzfreundliche Voreinstellungen** gewählt sind. Beispiel: Deaktivieren Sie *standardmäßig* jede Analyse oder Speicherung, die nicht unbedingt gebraucht wird. Oft bieten KI-Dienste Optionen an ("Daten zur Qualitätsverbesserung verwenden: Ja/Nein") – wählen Sie **Nein**, außer Sie haben einen guten Grund und Einwilligungen. Sorgen Sie dafür, dass der Bot **nur das Minimum** an Daten abfragt: Fragt er z. B. routinemäßig nach dem Namen, obwohl er es für eine einfache Auskunft nicht bräuchte? Dann besser weglassen. Filter Sie **Hintergrundgeräusche oder Fremdstimmen** raus, soweit möglich – der EDPB empfiehlt, unnötige Aufnahmen gar nicht erst zu erfassen [21]. Implementieren Sie auch eine **automatische Löschung**: z. B. alle Gesprächsprotokolle älter als X Tage werden aus dem System entfernt, sofern nicht anders vorgeschrieben. Solche Mechanismen sollten idealerweise **schon bei Implementierung** konfiguriert werden, nicht erst nachträglich.

- **Schulung und interne Prozesse:** Nicht zuletzt: **Schulen Sie Mitarbeiter**, die mit dem Voicebot-Projekt zu tun haben, in den Datenschutzfragen. Das betrifft Entwickler (die verstehen müssen, warum z. B. bestimmte Daten nicht geloggt werden sollen) ebenso wie das Kundenservice-Team (das wissen muss, wie es reagiert, wenn ein Kunde sagt "Löschen Sie meine Daten" am Telefon – auch wenn er mit dem Bot gesprochen hat). Legen Sie intern fest, wer für die **Überwachung der KI** zuständig ist – z. B. regelmäßige Reviews, ob der Bot falsch reagiert oder unzulässige Daten speichert. Halten Sie Prozessbeschreibungen parat, wie Sie z. B. bei einer **Betroffenenanfrage** die Daten aus dem Voicebot-System exportieren/löschen können. Und planen Sie einen

Fallback: Was ist bei Systemausfall? (Dann sollte z. B. ein normaler Anrufbeantworter oder Mensch übernehmen, damit keine Anfragen verloren gehen – auch das ist letztlich Datenschutz, nämlich Datentreue und Verfügbarkeit).

Wenn Sie all diese Punkte beachten, sind Sie auf einem sehr guten Weg, Voice-AI-Agenten **rechtskonform und verantwortungsvoll** einzusetzen. Die DSGVO-konforme Umsetzung erfordert zwar Mühe und interdisziplinäre Abstimmung, aber sie zahlt sich aus: Ihre Kunden (und Aufsichtsbehörden) werden es Ihnen mit Vertrauen danken, und Ihr Unternehmen vermeidet empfindliche Strafen oder Reputationsschäden. **Voicebots haben enormes Potenzial – nutzen wir sie datenschutzgerecht!**

---

[1] [20] GDPR, CCPA and Voice Recognition Privacy - Picovoice
https://picovoice.ai/blog/gdpr-ccpa-voice-recognition-privacy/

[2] [4] [21] edpb.europa.eu
https://www.edpb.europa.eu/system/files/2021-07/edpb_guidelines_202102_on_vva_v2.0_adopted_en.pdf

[3] Running Real-Time AI Voice Assistants in Kubernetes | by Péter Megyesi | L7mp Technologies | Medium
https://medium.com/l7mp-technologies/running-reel-time-ai-voice-assistants-in-kubernetes-136662bd031f

[5] Bundesnetzagentur - Unerlaubte Telefonwerbung
https://www.bundesnetzagentur.de/DE/Vportal/TK/Aerger/Faelle/UEW/start.html

[6] KI & Datenschutz – Wie passt das zusammen? - News | TÜViT
https://www.tuvit.de/de/aktuelles/newsroom/news/news-detail/article/ki-datenschutz-wie-passt-das-zusammen/

[7] AVV Datenschutz - Der Auftragsverarbeitungsvertrag nach DSGVO
https://www.dataguard.de/blog/der-auftragsverarbeitungsvertrag-avv/

[8] [9] [19] Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz | Der Landesbeauftragte für den Datenschutz und die Informationsfreiheit Baden-Württemberg
https://www.baden-wuerttemberg.datenschutz.de/rechtsgrundlagen-datenschutz-ki/

[10] [11] [12] [13] [14] [15] Auftragsverarbeitungsvertrag (AVV) – Was steht drin?
https://cortina-consult.com/datenschutzberatung/auftragsverarbeitungsvertrag/

[16] [17] voiceaiwrapper- privacy policy
https://voiceaiwrapper.com/dpa

[18] U.S. CLOUD Act vs. GDPR | activeMind.legal
https://www.activemind.legal/guides/us-cloud-act/