Privacy-by-Design
Toolkit v0.1

YOUTH-CENTRED GUIDELINES FOR SMART VOICE ASSISTANTS AND VOICE-ENABLED SMART DEVICES

Prepared by:
Dr. Ajay Shrestha

csci-viu.github.io/privyouth-smart
Vancouver Island University, BC, Canada

Prepared for: Educators, policymakers/regulators, and developers

Table of Contents

1. Overview

1.1 Purpose

This toolkit translates evidence from youth survey studies, youth focus groups, and a Privacy-by-Design technical audit of mainstream smart voice ecosystems into practical, youth-centered guidance. It supports three goals:

1) Reduce the highest-priority youth privacy risks in voice-enabled smart devices.

2) Increase youth ability to understand, navigate, and control privacy settings.

3) Provide implementable design and governance recommendations that can be verified.

1.2 Intended audiences

- **Educators** (K–12, post-secondary, libraries, community programs): building youth privacy literacy and practical device-management skills.
- **Policymakers / regulators / public-sector practitioners**: translating youth evidence into requirements, guidance, and accountability expectations.
- **Developers / product teams** (UX, privacy engineering, data governance, trust & safety): implementing privacy-by-design features and defaults that work for youth.

1.3 Scope boundaries (what v0.1 covers / does not cover)

Covers

- Voice-enabled assistants and companion-app/account settings used to manage them.
- Youth privacy concerns supported by the uploaded evidence base, including: ambient listening uncertainty, retention/deletion opacity, cross-platform inference concerns, policy overload, hidden controls, low privacy navigation efficacy, consent/defaults, access to recordings, safeguards, and accountability.

Does not cover (v0.1)

- Comprehensive legal interpretation beyond the PIPEDA-oriented checklist included in the audit template.

- Full accessibility testing (e.g., disability accommodations), cross-cultural adaptations, or non-Canadian regulatory mapping.
- Detailed technical security architecture; only user-facing safeguard expectations supported by the audit and survey items.
- Child-specific parental consent regimes (v0.1 is grounded in youth evidence; age range varies by study as noted below).

Assumption (minimal): Where the evidence discusses "youth" without a single fixed age boundary, v0.1 adopts the age range used in the survey-based studies (16–24) as the working definition for toolkit guidance and examples.

Traceability: Survey items and constructs are drawn from "Privacy by Voice" [1] and the age- and gender-focused multigroup papers [2], [3]; qualitative themes and codebook categories are drawn from "Convenience vs. Control" [4]; audit observations and PIPEDA/heuristic criteria are drawn from "Privacy by Design Audit" [5]; persona-style segmentation is informed by the cluster-analysis paper [6]; negotiation framing is informed by [7]; program framing is consistent with PI's proposed activities.

2. How to use this toolkit

1) Start with Tier 1: implement Tier-1 guidelines first. These address the highest-severity and most frequently evidenced youth concerns and are feasible to mitigate.

2) Select the module for your role: educator, developer, or policymaker checklists translate the same priorities into role-specific actions.

3) Use the verification steps: each guideline includes observable indicators and simple "prove it" checks.

4) Document changes: use the change-log structure (Appendix C) to track what was implemented and why.

3. Evidence consolidation and prioritization

3.1 Evidence sources included (Referenced files)

- Quantitative modeling: (PLS-SEM model linking risk, benefit, transparency/trust, self-efficacy, and protective behavior) [1].
- Multigroup analyses: **Age-Differentiated Pathway**s…[2] and **Gender-Based Heterogeneity**…[3] (differences by age group and gender).
- Qualitative study: **Convenience vs. Control**… (youth focus groups; themes + codebook snapshot) [4].
- Technical audit: **Privacy by Design Audit** (setup review, heuristic evaluation, PIPEDA checklist, UX testing tasks across Google Home, Amazon Alexa, and Siri) [5].
- Persona/segmentation: **Privacy Profiles**… (cluster analysis; "Concerned Skeptics" vs "Confident Optimists") [6].
- Negotiation framing: **Negotiating Privacy**… (indices capturing risk–benefit and control–acceptance tensions) [7], [8].
- Project framing: **PI's Proposed Activities** (program context and intended stakeholder-facing outputs) [9].

3.2 Prioritization criteria

Each concern is scored using four criteria (1–5), summed to a 4–20 score:

- **Severity (S):** potential impact on youth privacy and autonomy.
- **Frequency proxy (F):** presence across multiple evidence types and prominence in youth themes/items.

Gap note: item-level prevalence statistics are not consistently available across all uploaded papers; v0.1 uses triangulation across sources as a transparent proxy.

- **Feasibility of mitigation (Fe):** realistic implementability in product design, governance, and education within 6–12 months.
- **Downstream harm potential (H):** likelihood of secondary harms (surveillance, profiling, exposure of sensitive contexts, loss of trust).

Tier thresholds:
Tier 1: 18–20
Tier 2: 15–17
Tier 3: 12–14
Tier 4: ≤11

3.3 Evidence matrix (youth privacy concerns in smart voice ecosystems)

The matrix shown in the Evidence matrix table (v0.1) consolidates **C**oncerns and assigns **T**iers. (Evidence citations are file-based; see traceability notes.)

Evidence matrix table (v0.1)

ID	Concern	Evidence (files)	Criteria (S/F/Fe/H)	Score	Tier
C1	Always-on listening and unintended recording (wake-word uncertainty); need clear recording status + rapid mute/disable options	Convenience vs. Control (Qualitative); Survey items; Audit	5/5/4/5	19	Tier 1
C2	Retention, deletion, and voice-history lifecycle opacity (including indefinite storage); need time-boxed defaults + deletion receipts + ability to delete/view recordings	Convenience vs. Control; Survey items; Audit	5/5/4/5	19	Tier 1
C5	Usability/discoverability barriers in privacy settings; low navigation efficacy blocks protection (especially across devices)	Convenience vs. Control; Survey items; Audit	4/5/5/4	18	Tier 1
C3	Third-party sharing and cross-app/cross-platform inference (voice-to-ads linkage); need explicit disclosure + opt-in + compartmentalization	Convenience vs. Control; Audit; Survey items	5/4/3/5	17	Tier 2
C4	Policy overload and low transparency; youth need plain-language, searchable, task-based explanations of data collection and processing ('data nutrition label' concept)	Convenience vs. Control; Survey items; Audit	4/5/4/4	17	Tier 2

C7	Consent and default settings (pre-selected options, history on, ad personalization); need privacy-protective defaults and meaningful consent prompts	Audit; Survey; Negotiation indices paper	5/4/4/4	17	Tier 2
C6	Device-conditional privacy self-efficacy and inconsistent control locations (phone vs smart speaker); need a unified privacy hub and consistent patterns	Convenience vs. Control; Survey; Audit	4/4/4/3	15	Tier 2
C8	Individual access and granular control over voice recordings and derived data (view, export, delete specific items)	Audit; Survey; Qualitative	4/4/3/4	15	Tier 2
C10	Security safeguards and unauthorized access risk (voice data access by unauthorized parties); need clear safeguards and user-facing security cues	Survey; Audit; Qualitative	4/3/4/4	15	Tier 2
C12	Permission and scope management (deny mic/location; uninstall; just-in-time prompts); need proportional permissions and clear scope controls	Qualitative; Survey; Audit	3/4/5/3	15	Tier 2
C14	Youth-facing privacy education and protective routines (skills to navigate settings, clear histories, assess trade-offs)	Survey; Qualitative; OPC proposal	3/4/5/3	15	Tier 2
C11	Household and bystander privacy in shared spaces (others	Qualitative; OPC proposal; Audit	4/3/3/4	14	Tier 3

	recorded; shared accounts); need household modes and contextual consent				
C9	Accountability and responsive redress (clear privacy officer/contact; easy complaint path; timely effect of changes)	Audit; PIPEDA checklist in audit; Qualitative	3/3/4/3	13	Tier 3
C13	Perceived fairness and bias of recommendations (algorithmic fairness expectations)	Survey; Modeling papers	3/2/2/3	10	Tier 4

Tier 1

- C1 Always-on listening & unintended recording uncertainty; clear status + rapid disable.
- C2 Retention/deletion opacity and indefinite storage; time-boxed defaults + deletion receipts + granular deletion.
- C5 Hidden controls and low navigation efficacy; privacy controls must be easy to find and use.

Tier 2

- C3 Third-party sharing and cross-platform inference (voice-to-ads linkage); disclose, bound, and control.
- C4 Policy overload and low transparency; plain-language, searchable explanations.
- C7 Consent and risky defaults (pre-selected history/ad personalization); privacy-protective defaults.
- C6 Device-conditional efficacy; unified privacy hub and consistent patterns across devices.
- C8 Granular access to recordings/derived data; view/export/delete specific items.
- C10 Safeguards and unauthorized access risk; user-facing security cues.
- C12 Permission and scope management; proportional permissions and just-in-time prompts.
- C14 Youth privacy skills and routines; strengthen self-efficacy and behavior.

Tier 3

- C11 Household/bystander privacy in shared spaces; contextual modes and shared-device controls.
- C9 Accountability and redress; clear contact and predictable change effectiveness.

Tier 4

- C13 Perceived fairness/bias of recommendations; not a primary driver of youth privacy behavior in the current toolkit scope but retained as an emerging concern.

Traceability: The concern list and wording align with the qualitative codebook categories (e.g., "ambient always-on listening," "retention unknowns," "policy overload," "hidden controls," "low navigation efficacy," "permission refusal," "physical mitigations") and the survey items (PPR1–PPR4, ATT1–ATT4, PSE1–PSE4, PPB1–PPB4) presented in the multigroup papers and the PLS-SEM paper. Audit findings inform device-specific feasibility and verification steps.

3.4 Key findings extracted from the uploaded evidence base (condensed)

Survey-based modeling (Privacy by Voice) [1]

- Privacy-protective behavior is positively associated with **perceived privacy risk** and privacy **self-efficacy** (reported direct effects: PPR → PPB $\beta = 0.343$, $p < 0.001$; PSE → PPB $\beta = 0.373$, $p < 0.001$).
- **Algorithmic transparency & trust** influences behavior primarily **through self-efficacy** (ATT → PSE $\beta = 0.434$, $p < 0.001$; ATT → PPB is reported as non-significant).
- Benefits show a tension-consistent role: benefits are positively associated with self-efficacy (PPBf → PSE $\beta = 0.121$, $p < 0.1$) while the reported direct effect on protective behaviour is negative (PPBf → PPB $\beta = -0.130$, $p < 0.1$).

Age differences (Age-Differentiated Pathways) [2]

- Construct means differ by age group: **perceived risk is the highest construct overall** and **transparency/trust is the lowest**, with reported mean differences across most constructs.

- Multigroup analysis reports one statistically significant pathway difference: **ATT → PSE is stronger for older youth (19–24: β = 0.567) than younger youth (16–18: β = 0.356), p = 0.024**, suggesting transparency improvements may translate into control confidence more strongly for older youth.

Gender differences (Gender-Based Heterogeneity) [3]

- Pathways to protective behavior differ by gender in the reported multigroup results:
- **Risk → behavior** is stronger for males (PPR → PPB: males β = 0.424; females β = 0.233; p ≈ 0.062).
- **Transparency/trust → self-efficacy → behavior** is stronger for females (ATT → PSE → PPB: females β = 0.229; males β = 0.132; p ≈ 0.091).
- Measurement results indicate meaningful group differences for self-efficacy (PSE), supporting usability- and skills-focused interventions.

Youth privacy profiles (Cluster analysis of risk, benefits, trust, and behavior) [6]

- Two distinct profiles are identified:
    - **Concerned Skeptics:** higher privacy risk, lower benefits, lower transparency/trust, lower self-efficacy, and moderate protective behavior.
    - **Confident Optimists:** lower risk, higher benefits, higher transparency/trust, higher self-efficacy, and comparable protective behavior.
- Cluster membership differs by age (older youth are more represented among Concerned Skeptics), suggesting targeted interventions.

Negotiation framing (Risk-benefit and control-acceptance) [7]

- Youth privacy decision-making is framed as negotiation: **risk–benefit tension** and **control–acceptance tension** indices are reported as meaningfully associated with protective action.
- Frequent users tend to show more benefit-dominant and acceptance-leaning profiles, implying the need for protective defaults and low-friction controls.

Qualitative themes (Convenience vs. Control) [4]

- High-salience concerns include **ambient listening anxiety**, **retention/deletion uncertainty**, **policy fatigue**, **hidden controls**, and **cross-platform inference concerns**.
- Practical barriers include low navigation efficacy and device-conditional confidence; youth often rely on **permission refusal** and **physical mitigations** when software controls are hard to use.

Technical audit (Privacy by Design Audit) [5]

- Audit evidence shows measurable trade-offs between usability and compliance. Heuristic evaluation (7 criteria, 0-2 scale) scored Google Home 14/14 and Alexa and Siri 13/14; feedback was weaker for Alexa and Siri due to delayed or missing confirmation, and at least one critical Siri control was located outside Siri's main configuration menu.
- PIPEDA compliance scoring (10 principles, 0-2 scale) scored Siri highest (18/20), followed by Google Home (16/20) and Alexa (15/20). Youth UX testing found disabling voice history to be the most difficult task across devices (average time approx. 95-140 seconds; Siri highest), and identified a delayed effect on Alexa, where disabling voice recording history could take up to 36 hours to take effect;
- The summary also notes indefinite retention on Alexa unless users configure auto-delete or manually delete data. The audit's UX task list provides a practical verification set aligned with youth capability barriers (disable recording history; locate third-party sharing; delete voice commands; change a setting and verify).

UX task averages (ease 1-5; steps; time seconds): disable voice history (Google 3.5, 5 steps, 95s; Alexa 4.3, 5 steps, 103s; Siri 2.5, 5 steps, 140s); find data sharing info (Google 4.3, 3 steps, 80s; Alexa 4.5, 3 steps, 41s; Siri 4.3, 3 steps, 100s); delete commands (Google 4.3, 5 steps, 95s; Alexa 4.8, 4 steps, 38s; Siri 4.7, 4 steps, 100s); verify change (Google 4.3, 5 steps, 50s; Alexa 4.7, 5 steps, 29s; Siri 4.5, 3 steps, 60s).

Traceability: All bullets above are paraphrased from the uploaded papers and the audit's "Findings Summary," results tables (path coefficients / group comparisons), and qualitative theme sections.

4. Validation against referenced frameworks and guidance (alignment check)

4.1 Referenced frameworks/guidance available in the referenced materials

- **PIPEDA-oriented privacy principles checklist** (audit template): accountability; identifying purposes; consent; limiting collection; limiting use/disclosure; accuracy; safeguards; openness; individual access; challenging compliance.
- **Privacy-control usability heuristics** (audit template): discoverability; comprehensibility; control; minimization; feedback; reversibility; consistency.
- **Evidence-driven behavioral framework** (survey papers): perceived privacy risk (PPR), perceived privacy benefits (PPBf), algorithmic transparency & trust (ATT), privacy self-efficacy (PSE), privacy-protective behavior (PPB).
- **Persona/segmentation lens** (cluster analysis): differing youth profiles with distinct trust/self-efficacy and usage patterns.

4.2 Alignment summary (normative coverage)

Coverage classification reflects whether the referenced frameworks directly address the concern in principle (not whether platforms currently comply).

- **Clearly addressed:** C2, C7, C8, C9, C10
- **Partially addressed:** C1, C3, C4, C5, C6, C11, C12
- **Not addressed / weakly addressed:** C13, C14 (education is outside PIPEDA; fairness is only indirectly captured via ATT4)

Alignment mapping table (v0.1)

ID	Mapped framework elements (audit)	Coverage
C1	PIPEDA: Consent, Identifying purposes, Limiting collection, Openness	Heuristics: Control, Feedback, Consistency
C2	PIPEDA: Limiting use/retention, Individual access, Openness, Consent	Heuristics: Control, Feedback, Reversibility
C5	PIPEDA: Openness, Consent	Heuristics: Discoverability, Comprehensibility, Feedback, Reversibility

C3	PIPEDA: Identifying purposes, Consent, Limiting use/disclosure, Openness	Heuristics: Control, Comprehensibility, Discoverability
C4	PIPEDA: Openness, Identifying purposes, Consent	Heuristics: Comprehensibility, Discoverability
C7	PIPEDA: Consent, Identifying purposes, Limiting collection	Heuristics: Control, Comprehensibility
C6	PIPEDA: Openness, Consent	Heuristics: Consistency, Discoverability
C8	PIPEDA: Individual access, Accuracy, Openness	Heuristics: Control, Discoverability
C10	PIPEDA: Safeguards, Accountability, Openness	Heuristics: Comprehensibility, Control
C12	PIPEDA: Limiting collection, Consent, Openness	Heuristics: Control, Discoverability
C14	PIPEDA: — (education/skills outside PIPEDA checklist)	Heuristics: —
C11	PIPEDA: Consent, Identifying purposes, Limiting collection	Heuristics: Control, Comprehensibility
C9	PIPEDA: Accountability, Challenging compliance, Openness	Heuristics: Feedback, Consistency
C13	PIPEDA: — (only indirectly via ATT fairness item)	Heuristics: —

"Note: '—' indicates no direct corresponding principle/heuristic applies to this concern."

Traceability: PIPEDA checklist and usability heuristic criteria are taken from the audit template. Behavioral constructs and items are taken from the survey papers. Persona lens is taken from the cluster-analysis paper.

5. Gap analysis

5.1 Gap types used

- **Transparency gap:** unclear data flows, purposes, processing, or disclosures.
- **Control gap:** missing, weak, delayed, or non-granular controls.
- **Comprehension/usability gap:** controls exist but are hard to find/use or written beyond youth comprehension.

- **Defaults gap:** privacy-invasive options enabled by default or consent not meaningful.
- **Third-party sharing/inference gap:** unclear or uncontrollable sharing, profiling, or cross-platform inference.
- **Retention/deletion gap:** unclear or ineffective lifecycle management, indefinite retention, no deletion proof.
- **Accountability/redress gap:** unclear contact paths, complaint routes, or verification of changes.
- **Education/skills gap:** youth lack practical routines and self-efficacy to act on controls.

5.2 Gap Register (v0.1)

Tier 1 gaps

- **C1 (Comprehension/Control):** youth perceive "always listening" risk; platforms need clearer status indicators and more direct controls.
- **C2 (Retention/Deletion):** retention timelines and deletion outcomes remain uncertain; audit flags indefinite storage and limited recording management.
- **C5 (Usability):** hidden controls and navigation barriers suppress protective actions.

Tier 2 gaps

- **C3 (Third-party sharing/inference):** youth suspect voice-to-ads linkage; disclosures and control boundaries are not sufficiently visible or actionable.
- **C4 (Transparency/Comprehension):** policy fatigue undermines consent; youth request plain-language "data label" style summaries and a unified privacy hub.
- **C7 (Defaults/Consent):** pre-selected settings and history/ad personalization weaken meaningful consent.
- **C6 (Usability/Consistency):** control locations differ by device; confidence varies by device.
- **C8 (Control/Access):** inability to view/delete individual recordings creates persistent uncertainty.
- **C10 (Safeguards/Transparency):** safeguards may exist but are not explained in youth-legible ways.
- **C12 (Control/Scope):** youth rely on denial/uninstall/physical mitigation; platforms should support proportional permissions.

- **C14 (Education/skills):** youth need repeatable routines (review, limit, delete) and understanding of trade-offs.

Gap Register table (v0.1)

ID	Tier	Gap type	Gap statement
C1	Tier 1	Comprehension / control (contextual consent)	Listening/recording cues and stop controls are not consistently clear or immediate for youth; perceived 'always listening' persists.
C2	Tier 1	Retention / deletion (lifecycle proof)	Retention periods and deletion outcomes are not always visible or provable; audit notes indefinite storage and limited per-record management in some ecosystems.
C5	Tier 1	Comprehension / usability (hidden controls)	Privacy settings can be hard to locate and operate; navigation barriers suppress protective actions.
C3	Tier 2	Third-party sharing / inference transparency	Cross-service personalization and third-party sharing boundaries are not sufficiently visible; youth infer voice-to-ads linkage.
C4	Tier 2	Transparency / comprehension (policy overload)	Long policies and fragmented explanations undermine consent; youth request short, searchable summaries.
C7	Tier 2	Defaults / consent (meaningful opt-in/out)	Some data uses are enabled by default or pre-selected; youth may accept defaults without informed choice.
C6	Tier 2	Usability / consistency (fragmented control locations)	Controls vary by device and account layer; youth confidence varies by device and platform.
C8	Tier 2	Control / access (granular history management)	Users may be unable to view/delete specific recordings or understand derived data handling.
C10	Tier 2	Safeguards transparency (user-facing security cues)	Safeguards may exist but are not communicated clearly; youth remain concerned about unauthorized access.
C12	Tier 2	Scope control (proportional permissions)	Permissions are not always proportional or easy to manage; youth resort to refusal/uninstall or physical mitigations.

C14	Tier 2	Education / skills (capability building)	Practical privacy routines and skills need structured teaching and reinforcement beyond device UI.
C11	Tier 3	Contextual consent (shared spaces)	Shared-space recording and bystander privacy are not strongly operationalized in controls and guidance.
C9	Tier 3	Accountability / redress (contact + change effectiveness)	Privacy contacts/complaints and the effectiveness/timeliness of setting changes can be unclear.
C13	Tier 4	Algorithmic fairness (emerging)	Fairness/bias expectations are measured but not translated into privacy controls in the current evidence base.

Traceability: Gap statements synthesize findings from the qualitative themes (*A1/A2/C1/C2/D1/D2/E1/E2*), the audit weakness notes (e.g., retention and deletion limitations; privacy officer; effect delays), and the survey constructs linking trust/self-efficacy to protective behavior.

Note: *A1–E2* refer to qualitative codebook/theme identifiers used in the Convenience vs. Control study [4].

- *A1* (PPR): Ambient Listening and Uncertain Retention Raise Baseline Risk
- *A2* (PPR): Suspected Cross-App Inferences Amplify Surveillance Concerns
- *C1 (*ATT): Policy Overload and Hidden Controls Undermine Transparency (different from the concern ID used in the toolkit *C1*)
- *C2* (ATT): Retention/Deletion Opacity Depresses Trust (different from the concern ID used in the toolkit *C2*)
- *D1* (PSE): Low Navigation Efficacy Blocks Protective Action
- *D2* (PSE): Efficacy Is Device-Conditional; Youth Ask for Brief Scaffolds
- *E1* (PPB): Permission and Scope Management Are Primary Mitigations
- *E2* (PPB): Physical and Situational Strategies Supplement Software Controls

6. Tiered guidelines (privacy-by-design recommendations)

Tier 1 Guidelines (implement first)

G1. Make listening/recording states explicit and controllable (addresses C1; supports C12/C11)

Guideline statement (youth-centered):
Youth should be able to tell (at a glance) when a device is listening or recording, and stop it immediately without hunting through menus.

Rationale (evidence):
Youth report ambient "always-on" anxiety and uncertainty about when collection occurs (qualitative themes). Survey evidence includes concern about recording without awareness/consent and the need to prevent recording when undesired. Audit activities include setup defaults and control mechanisms.

Who should act
Developers: product UX, device firmware teams, companion-app teams.
Policymakers/regulators: guidance expectations for visible states and meaningful controls.
Educators: teach youth to locate and test listening indicators and mute controls.

What to implement

- **Design actions (developers)**
    - Persistent, unambiguous **listening/recording indicator** (device + app) with plain-language text ("Listening for wake word," "Recording," "Mic off").
    - **One-step control** to mute/disable mic and to pause voice capture for a time window (e.g., "Pause for 1 hour").
    - **Just-in-time prompts** when enabling voice features: why mic access is required; what is stored; how to turn it off later.
    - **Household mode**: a visible "shared space" toggle that increases indicator visibility and reduces background capture.
- **Governance actions (policymakers)**
    - Require vendors to document listening/recording states and how indicators map to actual capture states.

    - Require a “no dark patterns” expectation for voice capture controls (controls must be as easy to turn off as on).
- **Educational actions (educators)**
    - Classroom activity: “Find the state, test the mute.” Youth identify indicator cues, mute device, and confirm the app reflects the change.

How to verify (observable indicators)

- Indicator changes immediately when the mic is muted/unmuted.
- A “pause recording” option exists and is reachable in ≤2 steps.
- Companion app confirms a change was applied (feedback message).
- A youth tester can successfully disable voice capture without external help.

Priority tier & expected impact

- **Tier:** 1
- **Expected impact:** Reduces ambient surveillance anxiety; improves consent meaningfulness; increases youth agency and trust.

Traceability: Convenience vs. Control (A1; physical mitigations), survey items PPR2/PSE2, and audit control/feedback heuristics.

G2. Make retention, deletion, and history management time-bounded and provable (addresses C2; supports C8)

Guideline statement (youth-centered):
Youth should be able to see what was stored, delete it easily, and trust that deletion actually happened—by default and by design.

Rationale (evidence):
Retention/deletion opacity depresses trust in focus groups. Survey evidence includes concern about the duration stored and behavioral items about deleting voice history. Audit flags indefinite retention and limited ability to view/delete recordings for some ecosystems.

Who should act

- **Developers:** data governance, privacy engineering, UX.
- **Policymakers:** retention and deletion expectations; proof-of-deletion norms.

- **Educators:** teach simple routines for clearing histories and checking retention settings.

What to implement

- **Design actions (developers)**
    - Default **auto-delete** retention (e.g., 30–90 days) and easy adjustment (not buried).
    - A clear **voice activity timeline** (what was captured, when, and why) with per-item delete.
- **Deletion receipts** (confirmation plus "what remains" explanation—e.g., model training exclusions, backups, or derived data).
    - "Clear now" flows that complete in ≤60 seconds and ≤5 taps for typical users.
- **Governance actions (policymakers)**
    - Require published retention periods and "what deletion means" explanations.
    - Require a minimum deletion capability: per-item delete + account-wide delete + export.
- **Educational actions (educators)**
    - Teach the "3D routine": **Discover** (find history), **Delete** (remove), **Default** (set auto-delete).

How to verify

- Auto-delete exists and is ON by default for youth-facing contexts (or strongly recommended at setup).
- User can delete a single recording and confirm that it no longer appears.
- System provides a confirmation message and timestamp of deletion action.
- Vendor documentation clearly states retention and post-deletion handling.

Priority tier & expected impact

- **Tier:** 1
- **Expected impact:** Directly addresses highest-rated risk area; improves trust; reduces long-term exposure.

Traceability: Convenience vs. Control (C2), survey item PPR4 and PPB2, audit weaknesses (indefinite storage; per-record deletion limits).

G3. Make privacy controls discoverable, youth-legible, and task-based (addresses C5; supports C4/C6)

Guideline statement (youth-centered):

Youth should not need to search online or read long policies to protect themselves; privacy controls must be easy to find, easy to understand, and organized around tasks youth actually do.

Rationale (evidence):

Focus groups describe policy fatigue, hidden controls, and low navigation efficacy. Survey results show self-efficacy is a key pathway to privacy-protective behavior (and is shaped by transparency/trust). Audit includes discoverability and comprehensibility heuristics and UX testing tasks.

Who should act

- **Developers:** UX/content design, privacy UX writing.
- **Policymaker**s: accessibility and usability expectations for privacy controls.
- **Educators:** reinforce practical navigation skills and routines.

What to implement

- **Design actions (developers)**
  - A unified **Privacy Hub** with top tasks: “Stop recording,” “Delete history,” “Manage sharing,” “Export my data,” “Ad personalization,” “Household mode.”
  - Plain-language microcopy written for youth comprehension, with short summaries plus optional deeper detail.
  - Searchable settings and direct links from help prompts to the relevant setting screen.
  - Confirmations after changes (feedback) and easy undo (reversibility).
- **Governance actions (policymakers)**
  - Add usability expectations to privacy guidance (controls must be discoverable and understandable, not merely present).
  - Encourage standardized “privacy tasks” across platforms.
- **Educational actions (educators)**
  - Assign a “privacy scavenger hunt” using the Privacy Hub tasks and reflection questions.

How to verify

- A youth tester can complete core tasks without external help: disable recording history; find third-party sharing details; delete voice commands; confirm a setting change.
- Controls are reachable within a small number of steps (e.g., ≤3 from home in the companion app).
- Language is plain and examples are included.

Priority tier & expected impact

- **Tier:** 1
- **Expected impact:** Converts concern into action; reduces skill barriers; supports equitable protection across youth groups.

Traceability: Convenience vs. Control (C1/D1/D2), Privacy by Voice (PSE mediates ATT → PPB), audit UX tasks and heuristics.

Tier 2 Guidelines (implement next)

G4. Make data flows transparent with plain-language "data labels" and searchable explanations (addresses C4; supports C3/C10)

Guideline statement:

Youth deserve short, clear summaries of what data is collected, why, where it goes, and how to control it—without policy overload.

Rationale:

Focus groups report policy overload and distrust driven by opaque data flows. Survey items directly measure understanding of what is collected/stored and whether platforms are upfront about processing.

Who should act: Developers; policymakers; educators.

What to implement

- **Developers:** “data label” style summaries (what/why/where/how long/who shared); in-app links; examples of common scenarios.
- **Policymakers:** encourage standardized label formats across devices; require that labels be available in-app.
- **Educators:** teach youth to read labels and identify the “control points” (turn off, limit, delete).

How to verify

- The label exists in the app and is reachable from the Privacy Hub.
- A youth tester can answer: what is collected, how long stored, who shared with, and how to turn it off.

Tier & expected impact: Tier 2; improves informed consent and trust.

Traceability: Convenience vs. Control (policy overload), survey ATT items, audit openness findings.

G5. Bound third-party sharing, ads personalization, and cross-platform inference with explicit opt-in controls (addresses C3; supports C7)

Guideline statement:

If voice interactions can influence ads or content elsewhere, youth must be told clearly and be able to opt out easily.

Rationale:

Youth interpret voice-to-ads patterns as surveillance. Audit notes ad personalization defaults and links to ad ecosystems; survey items show transparency/trust predicts self-efficacy and behaviour.

Who should act: Developers; policymakers; educators.

What to implement

- **Developers:** a single “Voice-to-ads / cross-service personalization” control; default OFF for youth contexts; clear explanation; compartmentalize voice data away from ad targeting unless opted in.
- **Policymakers:** require disclosure and opt-in for cross-service personalization; require a simple opt-out.
- **Educators:** teach youth to locate and disable ad personalization and cross-service sharing.

How to verify

- Control exists and is not buried.
- Default state is protective (or explicitly presented at setup as a meaningful choice).
- Opt-out takes effect promptly and is confirmed.

Tier & impact: Tier 2; reduces surveillance perception and downstream profiling harms.

Traceability: Convenience vs. Control (A2), audit setup defaults and evidence sources referencing ad ecosystems, ATT/PSE linkage in survey papers.

G6. Make consent meaningful and defaults protective (addresses C7; supports C1/C2/C3)

Guideline statement:

Default settings should minimize collection and retention; consent should not be pre-selected or bundled in ways youth cannot understand.

Rationale:

Audit identifies pre-filled consent-related settings and history/ad personalization defaults. Qualitative findings show policy fatigue; survey shows risk concern is high.

Who should act: Developers; policymakers.

What to implement

- **Developers:** privacy-protective default profiles; granular consent; unbundled choices; “explain in one screen” summaries.
- **Policymakers:** require that opt-in/out is symmetric; prohibit pre-ticked boxes for non-essential uses in youth contexts.

How to verify

- Non-essential data uses are OFF by default or require explicit opt-in.
- Consent choices are granular and reversible.

Tier & impact: Tier 2; prevents privacy erosion through default-driven inertia.

Traceability: Audit consent notes; qualitative policy overload; survey PPR/ATT.

G7. Provide granular access to recordings and derived data (view, export, delete) (addresses C8; supports C2)

Guideline statement:

Youth should be able to see and manage individual recordings and key derived data, not only broad account controls.

Rationale:

Audit highlights missing per-record management in some ecosystems; qualitative data shows deletion uncertainty.

Who should act: Developers; policymakers.

What to implement

- Per-record view/delete;
- export capability;
- explanations of derived data and what remains after deletion.

How to verify

- Per-item deletion exists.

- Export is available and understandable.

Tier & impact: Tier 2; reduces retention uncertainty and increases control.

Traceability: Audit weaknesses; retention theme; PPB2.

G8. Strengthen safeguards and explain them in youth-legible terms (addresses C10)

Guideline statement:

Youth should be confident that voice data is protected; security measures must be communicated in clear, actionable ways.

Rationale:

Survey includes unauthorized access concern; audit indicates safeguards may not be clearly described.

Who should act: Developers; policymakers; educators.

What to implement

- Clear security cues: encryption in transit/at rest statements; account protection reminders; device access controls; “who can access recordings” transparency.

How to verify

- Security information appears in the Privacy Hub with plain-language explanations.
- Users can review account access and connected devices.

Tier & impact: Tier 2; reduces fear of unauthorized access.

Traceability: Survey PPR3; audit safeguards notes.

G9. Build youth privacy self-efficacy through repeatable routines (addresses C14; supports C5/C6)

Guideline statement:

Youth should learn simple, repeatable routines for privacy protection that match how they actually use SVAs.

Rationale:

Survey evidence shows self-efficacy is a major pathway to protective behavior; qualitative data show navigation barriers and policy fatigue.

Who should act: Educators; policymakers (supporting curriculum and public education).

What to implement

Teach and practice routines:

- **Locate:** find Privacy Hub and key tasks.
- **Limit:** disable unnecessary features; manage permissions.
- **Log/Review:** review permissions and activity history.
- **Delete:** clear histories; set auto-delete.
- **Verify:** check indicators and confirmations.

How to verify

- Youth can complete the audit-style tasks without assistance.
- Youth can explain what data is collected and how to reduce it.

Tier & impact: Tier 2; improves capability and reduces reliance on extreme mitigations.

Traceability: Privacy by Voice mediation results; qualitative low navigation efficacy; audit UX tasks.

7. Tier 3–4 concerns (monitor and address as capacity allows)

Tier 3 (C11, C9)

- **Shared-space and bystander privacy:** develop household modes and guidance for shared-device consent.
- **Accountability and redress:** standardize privacy contacts, complaint pathways, and predictable change effectiveness.

Tier 4 (C13)

- **Fairness/bias perceptions:** retain as an emerging area for future toolkit expansion.

Traceability: Household context and accountability arise from audit and qualitative implications; fairness is measured via ATT4 in survey instruments.

8. Practical checklists (role-based)

8.1 Educator checklist (classroom / workshop ready)

1) Teach the Listen–Limit–Delete–Verify routine (G1–G2–G9).
2) Run the 4-task "privacy lab" (from the audit UX tests):
    a. Disable voice recording history.
    b. Find whether data is shared with third parties.
    c. Delete stored voice commands.
    d. Change a privacy setting and verify the change.
3) Discuss policy fatigue and how to use "data labels" instead of long policies (G4).
4) Assign a reflection prompt: "What convenience do you gain, and what control do you give up?" (negotiation framing).
5) Encourage respectful shared-space norms (ask before using voice in sensitive contexts).

8.2 Developer checklist (product/UX/privacy engineering)

1) Provide unambiguous listening/recording indicators + one-step stop/pause controls (G1).

2) Default to time-bounded retention and provide deletion receipts (G2).
3) Build a task-based Privacy Hub with searchable settings and plain-language copy (G3–G4).
4) Make cross-service personalization explicitly opt-in and easy to disable (G5–G6).
5) Provide per-record view/export/delete and explain derived data handling (G7).
6) Provide user-facing safeguards info and access reviews (G8).
7) Validate with youth testers using the audit tasks; record completion time and confusion points.

8.3 Policymaker / regulator checklist

1) Translate Tier-1 controls into minimum expectations: visible states, rapid stop controls, default retention limits, deletion proof.
2) Require plain-language, in-app transparency summaries and standardized "privacy task" navigation.
3) Require explicit opt-in and simple opt-out for cross-service personalization and third-party sharing.
4) Require per-record access and deletion for voice histories.
5) Strengthen accountability expectations: accessible contact, complaint path, and verifiable change effectiveness.
6) Encourage adoption of audit-style usability testing as evidence for compliance readiness.

Traceability: Checklists are derived from Tier 1–2 guidelines and are anchored in audit UX tasks and the survey/qualitative findings.

9. Templates (copy-paste ready)

Template T1: "Privacy Hub" task list (for developers)

Top tasks (must be front-and-centre):

- Stop/pause listening and recording
- Delete voice history (single item / today / all time)
- Set auto-delete (30/60/90 days)

- Manage third-party sharing and cross-service personalization
- Export my data
- Review connected devices and account access
- Household/shared-space mode

Template T2: Classroom "SVA Privacy Lab" worksheet (for educators)

Student name/initials: ____ Device/ecosystem: ____

1) Find the listening indicator. What does it show? ____
2) Mute the mic. Does the indicator change immediately? ____
3) Find voice history. What is stored? ____
4) Delete one item. Is there a confirmation? ____
5) Find third-party sharing or ad personalization. Is it on/off? ____
6) Set auto-delete (if available). What period did you choose? ____
7) Reflection: What do you gain from the assistant? What data do you give up? ____

Template T3: Policymaker "minimum expectations" one-pager

Objective: protect youth autonomy and reduce high-severity risks in voice ecosystems.

Tier-1 minimum requirements:

- Visible listening/recording state + one-step stop/pause
- Default retention limits + per-record deletion + deletion receipts
- Task-based Privacy Hub with youth-legible transparency summaries

Verification: require vendor usability evidence using the 4 audit tasks.

Template T4: Roundtable feedback form (for stakeholder consultation)

For each guideline (G1–G9), rate:

- Clarity (1–5): __
- Feasibility (1–5): __
- Expected impact (1–5): __
- Missing considerations: __
- Implementation barriers: __
- Verification evidence you would accept: __

10. Example workflows/use cases (supported by uploaded files)

Workflow W1: "Disable voice recording history"

- Navigate to Privacy Hub / Activity settings.
- Turn off history or set auto-delete.
- Confirm in the app that the change is active.

Workflow W2: "Find third-party sharing / cross-service personalisation"

- Locate "sharing," "ads personalization," or "cross-service" settings.
- Record whether it is enabled by default.
- Opt out and verify it remains off after app restart.

Workflow W3: "Delete stored voice commands"

- Open voice history/activity.
- Delete one item; then delete all for today; then all-time (if available).
- Confirm deletion receipt and what remains.

Workflow W4: "Change a privacy setting and verify"

- Change a single setting (e.g., mic permissions, personalization).
- Check for immediate feedback and confirm on-device indicator updates if relevant.

Traceability: These workflows are directly adapted from the audit's user-experience testing task list.

11. Glossary (toolkit terms)

- **SVA (Smart Voice Assistant):** voice-enabled assistant ecosystem (device + cloud services + companion apps).
- **Wake word:** phrase used to activate a voice assistant (e.g., "Hey …").
- **Ambient listening:** perception that a device is always listening while awaiting wake word.

- **Voice history / activity:** stored logs or recordings of voice interactions.
- **Retention:** how long voice data is stored.
- **Deletion receipt:** confirmation that a deletion action occurred and what residual data remains.
- **Third-party sharing:** disclosure or transfer of data to entities beyond the primary service provider.
- **Cross-service personalization:** using data from one service to personalize another (e.g., ads or content).
- **Algorithmic transparency & trust (ATT):** youth understanding and trust in how the system collects/processes data and produces outputs.
- **Privacy self-efficacy (PSE):** youth confidence in accessing and adjusting privacy controls.
- **Privacy-protective behavior (PPB):** actions such as reviewing permissions, deleting history, refusing features, and using additional measures.
- **PIPEDA principles (audit checklist):** accountability, identifying purposes, consent, limiting collection, limiting use/disclosure, accuracy, safeguards, openness, individual access, challenging compliance.
- **Heuristic evaluation (audit):** structured usability review of privacy controls (discoverability, comprehensibility, control, feedback, reversibility, consistency).

Traceability: Terms reflect those used in the survey constructs and the audit template.

Appendix A. Evidence notes (brief)

- Survey instruments include risk (PPR), benefits (PPBf), transparency/trust (ATT), self-efficacy (PSE), and protective behaviour (PPB) with four items each (see age-differentiated paper's constructs-and-items table).
- Age differences: older youth show stronger ATT → PSE linkage; mean differences indicate risk is rated highest overall.
- Gender differences: structural pathways differ by gender; self-efficacy shows mean differences, supporting targeted skill-building.
- Qualitative themes highlight always-on anxiety, policy overload, hidden controls, retention uncertainty, and reliance on physical mitigations.
- Audit identifies strengths and weaknesses across Google Home, Alexa, and Siri, including retention and deletion limitations and accountability gaps.

Appendix B. Assumptions and gaps (v0.1)

1) Frequency proxy: v0.1 uses triangulation across qualitative themes, survey item inclusion, and audit tasks rather than item-level prevalence percentages (not consistently available across uploaded files).
2) Age boundary: v0.1 uses 16–24 as the working youth range, consistent with the survey studies.
3) Platform variability: audit findings are device- and ecosystem-specific; v0.1 generalizes to "voice ecosystems" but flags where limitations are platform-dependent.

Appendix C. Change Log and Version Control

Document: *Privacy-by-Design Toolkit*
Current release: v0.1 (Draft)
Maintainer: Dr. Ajay Shrestha
Purpose: Provide a transparent record of key changes, evidence inputs, and verification updates.

Change log summary table (release-level)

Version (date)	Status	Key changes (high level)	Evidence basis + verification updates
v0.1 (Feb 4, 2026)	Draft	Initial complete toolkit draft: evidence matrix + tiering; alignment check; gap register; Tier 1–2 guideline packages; role-based checklists; templates; glossary; traceability notes.	Evidence: survey papers (PLS-SEM; age/gender MGA), qualitative focus groups, Privacy-by-Design audit, profile/cluster paper, negotiation draft, OPC proposal context. Verification: audit-task checks (disable history, locate sharing, delete commands, verify setting changes).
v0.2 (TBD)	Planned	Incorporate stakeholder roundtable feedback; strengthen prioritization with item-level statistics (Need/Leverage) where available; expand audit scoring rubric; refine verification criteria.	Evidence additions: stakeholder roundtable outputs; any added item-level stats; additional scored audit runs; youth usability evidence (if conducted). Verification: add acceptance criteria and completion benchmarks.

References (Toolkit v0.1 evidence base)

[1] M. Campbell and A. K. Shrestha, "Privacy by voice: Modeling youth privacy-protective behavior in smart voice assistants," accepted in *2026 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, Tokyo: IEEE, Feb. 2026.

[2] Y. Bobkova, M. Campbell, and A. K. Shrestha, "Age-differentiated pathways to privacy protection in smart voice assistants: A multigroup PLS-SEM study of youth," accepted in *5th IEEE International Conference on AI in Cybersecurity (ICAIC)*, Houston, Texas, USA: IEEE, 2026.

[3] Y. Bobkova, M. Campbell, and A. K. Shrestha, "Gender-based heterogeneity in youth privacy-protective behavior for smart voice assistants: Evidence from multigroup PLS-SEM," submitted to the *39th Annual Canadian Conference on Electrical and Computer Engineering (CCECE 2026)*, Montreal, Canada: IEEE, 2026.

[4] M. Campbell, T. De Clark, M. Sheikho Al Jasem, S. Joshi, and A. K. Shrestha, "Convenience vs. control: A qualitative study of youth privacy with smart voice assistants," in *2026 IEEE 16th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, USA: IEEE, Jan. 2026. "In press" preprint at doi: arXiv:2601.04399.

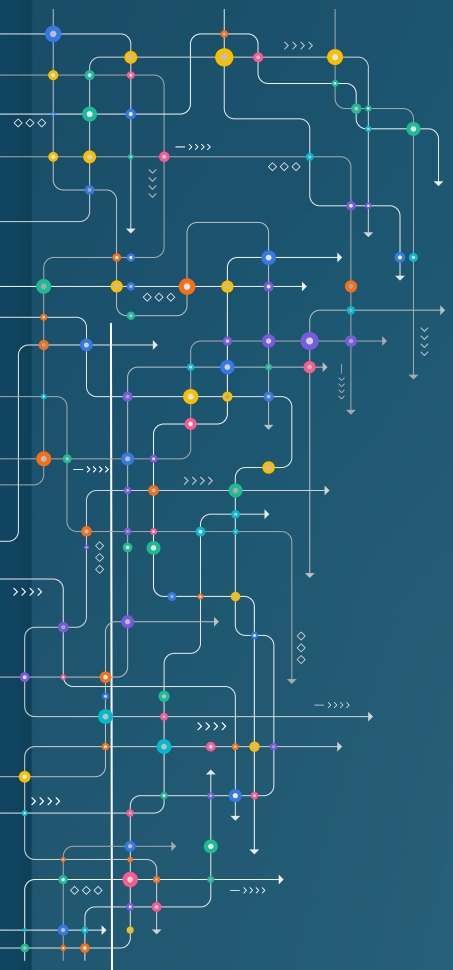
[5] T. De Clark, Y. Bobkova, and A. K. Shrestha, “Balancing usability and compliance in AI smart devices: A privacy-by-design audit of Google Home, Alexa, and Siri,” in *2026 IEEE 16th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, USA: IEEE, 2026. “In press” preprint at doi: arXiv:2601.04403.

[6] T. De Clark and A. K. Shrestha, “Privacy profiles of youth smart voice assistant users: A cluster analysis of risk, benefits, trust, and behavior,” Draft *Version*, 2026.

[7] M. Sheikho Al Jasem and A. K. Shrestha, “Negotiating privacy with smart voice assistants: Risk-benefit and control-acceptance,” Draft *Version*, 2026.

[8] M. Campbell, Y. Bobkova, and A. K. Shrestha, “From framework to practice: Youth negotiations of privacy with smart voice assistants through the PEA-AI lens,” Draft Journal Version, 2026.

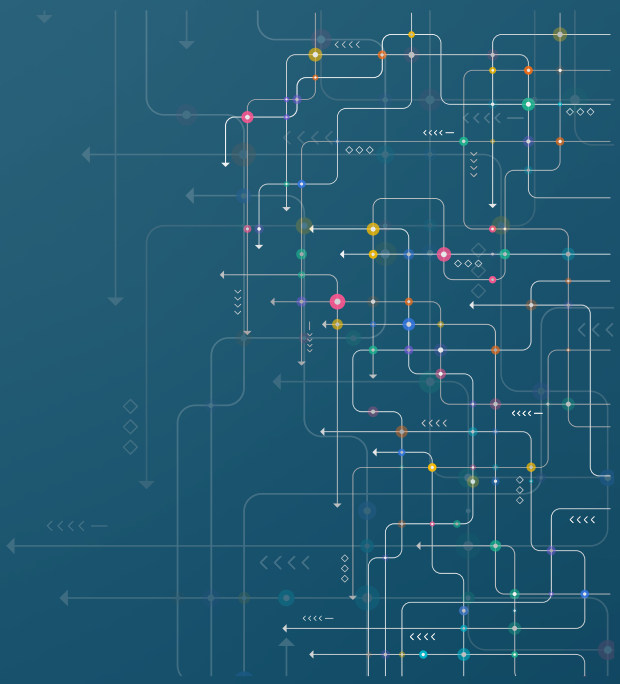
[9] Csci-viu, “Empowering young Canadians in the smart device era.” Accessed: Jan. 30, 2026. [Online]. Available: https://csci-viu.github.io/privyouth-smart/



CONTACT

For additional information,
guidance, or collaborations,
please reach out to:

Dr. **Ajay Shrestha**,
Principal Investigator
Vancouver Island University
ajay.shrestha@viu.ca



Privacy-Aware AI | Vancouver Island University