

# Personalization guardrails checklist

A lightweight, executive-ready checklist to ship personalization responsibly—without slowing teams down.

## USE THIS WHEN

Shipping personalization, recommendations, ranking, search, or any AI-driven content selection that changes what users see.

### 1) Value and user benefit

- The user problem is explicit and measurable (not “AI for AI”).
- The recommendation improves decision quality or reduces effort (clear user win).
- We can explain what “better” means (success metrics defined).
- There is a non-AI fallback path if personalization is off or fails.

### 2) Transparency and control

- Each recommendation can show a simple “Why this?” explanation.
- Users can dismiss / hide items and the system learns from it.
- Users can opt out of personalization (or choose a less-personalized mode).
- No dark patterns: recommendations do not manipulate or mislead.

### 3) Privacy and data boundaries

- We only use signals we have a right to use (consent + policy alignment).
- We avoid surprising inferences (e.g., health, finances, sensitive traits).
- Sensitive features are excluded or heavily constrained; rationale documented.
- Data retention and access are least-privilege; auditability exists.

---

#### **4) Safety, fairness, and compliance**

- We reviewed for bias risks and documented mitigations (segments that could be harmed).
- We tested for harmful / inappropriate content exposure and added filters where needed.
- We have a compliance review checkpoint (legal, risk, privacy) before launch.
- We have an incident plan (what happens if recs are wrong, harmful, or exploitative).

#### **5) Measurement and rollout discipline**

- We defined leading metrics (engagement, adoption) and lagging metrics (retention, revenue, contribution).
- We defined guardrail metrics (complaints, dismiss rate, churn, negative feedback).
- We have stop / rollback criteria (what would make us turn it off).
- We launch via controlled rollout (A/B or phased exposure), not full blast.
- We have monitoring dashboards for model drift and experience health.

#### **6) Operational readiness**

- We know who owns the system post-launch (product, DS, eng on-call).
- We have a retraining / refresh plan and change log for model updates.
- We can troubleshoot issues (logging, reproducibility, versioning).
- We documented constraints and intended behavior for future teams.