# Kashi Legal-Max Hardening Memo

How far the product can go while remaining legally and politically defensible

Document type: internal strategy and governance memo
Prepared for: Kashi project use
Date: 21 April 2026

**Bottom line**

Kashi can go further than the current deck implies, but only by becoming more disciplined about purpose, monitoring honesty, worker-facing safeguards, meeting-type validity, and post-signal procedure. The legally aggressive path is not more AI bravado. It is stronger constraint architecture, stronger notice, stronger consultation, and stronger decision-use walls.

This memo is for product, governance, rollout, and deck revision. It is not formal legal advice.

# 1. Executive summary

- Japan-first deployment can support recorded-meeting analytics if the employer specifies concrete purposes, writes the monitoring logic into internal rules, gives workers clear notice, consults majority representatives or unions where needed, and keeps access, retention, and downstream use tightly bounded.
- EU-facing deployment is materially stricter. Workplace emotion recognition is prohibited, and AI used to monitor or evaluate workers' behaviour or performance sits in the high-risk employment zone. If Kashi is used on the employer side in the EU, assume high-governance discipline from the start.
- The current Kashi deck and memo stack are strategically aligned, but not yet internally clean. The main unresolved issues are the transcript/content contradiction, the over-absolute 'not monitoring' line, the tension between CEO-facing rhetoric and worker-protective architecture, the incomplete validity model, and an underbuilt procedural spine.
- The practical fix is a five-part hardening pass: choose a stable data doctrine, tell the truth about monitoring, split worker-facing and institutional lanes, formalize validity and abstention, and turn governance from prose into system behaviour.

### Jurisdictional ceiling at a glance

| Jurisdiction | What Kashi can push | Hard legal ceiling | Recommended posture |
|---|---|---|---|
| **Japan** | Structural meeting analytics; worker self-view; aggregate institutional views; review-support workflows | No covert or undefined purpose; no casual reuse outside stated purpose; consultation and anti-retaliation expectations matter | Japan-first core market with strong internal rules, notice, worker-facing safeguards, and access discipline |
| **EU** | Very narrow structural review-support posture with strong human oversight and worker/rep information | No workplace emotion recognition; employment behaviour-evaluation can trigger high-risk AI obligations | Either stay extremely narrow and descriptive or build for high-risk discipline from day one |
| **Hong Kong / Singapore** | Bounded monitoring can be workable if policy, notice, and necessity are clear | Weakness shows up when monitoring is broader than needed or policies are vague | Treat as viable after Japan if the control pack is already mature |
| **China** | Behaviour-analysis tools | PIPL is strict on behaviour analysis, transparency, and individual rights | Separate regime; no improvised rollout |

# 2. Gap-to-fix mapping

These are the five current fault lines that must be closed if Kashi wants to move closer to the legal edge without falling into obvious surveillance, validity, or fairness traps.

| Gap | Why it matters | What to change now | Why this is the legally stronger move |
|---|---|---|---|
| **Transcript/content contradiction** | The deck says 'patterns, not content' and 'never transcribe for analysis', but some detectors imply transcript interpretation. | Choose structural-only MVP or explicitly declare a constrained hybrid with a structural-first core. | Cleaner purpose limitation, easier worker notice, and fewer overclaim problems. |
| **'Not monitoring' is too absolute** | The product still processes workplace interaction data over time, so a flat | Replace with 'not a general surveillance or performance- | Honest, proportionate framing is more defensible than |

| | | | |
|---|---|---|---|
| | denial looks evasive. | monitoring system' and explain the limits. | rhetorical denial. |
| **CEO rhetoric vs worker-protective logic** | The buyer is not the same as the protected person; current language makes the product sound more top-down than it is. | Separate public wedge, worker lane, and sponsor story. Keep CEO logic as buyer logic, not product identity. | Aligns the commercial story with privacy, labour, and anti-retaliation constraints. |
| **Validity model is incomplete** | Self-baseline alone is not enough; meeting purpose, role, dyad, and data quality all change interpretation. | Add meeting-type normalisation, evidence grades, abstention rules, and input-quality gating. | Weak validity becomes a legal and procedural fairness problem once workers are involved. |
| **Procedure is underbuilt** | The memos are strong on what Kashi detects and refuses, but weaker on contest rights, access doctrine, and downstream control. | Write the procedural spine: challenges, reason-coded access, review authority, anti-retaliation, and misuse sanctions. | Law and trust both care more about procedure than founders like to admit. |

## 3. Doctrine rewrite for the project

The legal-max version of Kashi should be governed by a small number of hard doctrines. These are not rhetorical flourishes. They are design constraints.

### Doctrine 1 — constrained visibility by default

Institutional visibility is deliberately weaker, slower, narrower, and more procedural than the technology could technically permit.

### Doctrine 2 — structural review support, not truth claims

Kashi estimates repeated interaction asymmetry under uncertainty, within comparable meeting contexts, for review support. It does not detect harm, intent, illegality, or emotional state.

### Doctrine 3 — no affect, no psychologising, no open-ended employer content surveillance

Emotion, stress, tone, and affect inference remain off-limits. Employer-facing outputs do not become a general searchable archive of what workers said.

### Doctrine 4 — every consequential representation is challengeable

Transcript errors, speaker attribution, context windows, thresholds, summaries, and review decisions must all be open to dispute and correction.

### Doctrine 5 — concern formation is private by default

Opening a private pattern page, creating a vault, drafting a concern, or marking confounds must not create employer-visible signals before explicit sharing or a separately governed threshold event.

### Doctrine 6 — hard wall against employment decision use

No performance, promotion, discipline, compensation, or managerial ranking use. Product, policy, contract, and interface all say the same thing.

# 4. Required product architecture changes

## 4.1 Data and detector doctrine

- Default to a structural-first core: speaking share, overlap, interruption, latency, directional concentration, chilling-delta, and comparable wrappers.
- If transcript-semantic features are kept, label them as a separate, constrained module rather than pretending the whole system is metadata-only.
- Do not let employer-facing UX become open-ended content search, message summarisation, or 'what was said' browsing.

## 4.2 Two-lane visibility architecture

- Worker lane: self-view, confounds, private explainer, private evidence staging, optional user-driven sharing.
- Institutional lane: aggregate views, threshold-triggered review objects, exceptional drill-down only, audited access, no casual raw transcript browsing.
- Manager Mirror is not the product's identity. It is a bounded lane inside a broader accountability system.

## 4.3 Validity stack

- Baseline stack should include self-history, within-meeting peers, meeting-type baseline, role baseline, and dyad baseline.
- Unsupported or low-confidence meeting types should fall back to observational-only mode and not generate review-worthy events by default.
- Every signal should carry an evidence grade, uncertainty posture, and explicit abstention logic when evidence is weak or confounded.
- Input-quality gating must account for transcript confidence, diarisation confidence, overlap quality, multilingual complexity, and relevant confounds.

## 4.4 Example evidence-grade ladder

| Grade | When it should appear | System behaviour |
|---|---|---|
| **Insufficient evidence** | Too little comparable exposure; poor input quality; strong confounds | Show descriptive metrics only; no review-worthy event by default |
| **Weak pattern** | Some directional signal, but sparse or unstable | Show cautionary preview; no institutional escalation without additional corroboration |
| **Emerging pattern** | Repeated, comparable, and interpretable signal; still contestable | Allow worker-facing explanation and bounded review support |
| **Stable pattern** | Repeated, comparable, and stronger evidence with acceptable | Permit review-worthy event construction and human review under documented |

| | | |
|---|---|---|
| | input quality | procedure |

## 5. Procedural spine Kashi must add

This is the main gap between the current deck and a legally serious deployment posture. The system needs an operating model, not only detector logic.

### 5.1 Challenge and correction state machine

- Allow disputes over transcript accuracy, speaker attribution, context windows, summaries, thresholds, and dispositions.
- Log suppression, correction, override, and reviewer rationale.
- Treat challengeability as part of product function, not customer support overflow.

### 5.2 Access doctrine

- Aggregate-first by default; named raw-context access is exceptional.
- Require reason codes, approval path, and access logging for any deeper review.
- Give affected individuals access to their own access history where legally and operationally appropriate.

### 5.3 Anti-retaliation controls

- Private awareness, concern formation, draft creation, and vault activity remain employer-invisible unless explicitly shared or threshold-governed.
- Use minimum-group thresholds, batching, and suppression rules to avoid small-team inference leakage.
- Escalation should share the minimum necessary event object, not a wide context dump by default.

### 5.4 Decision-use controls

- No export into appraisal, performance, compensation, promotion, or discipline workflows.
- No named subordinate telemetry to line managers.
- No league tables, ranking, or 'manager quality score' features.

### 5.5 Post-deployment quality regime

- Include appeals, overrides, incident response, misuse reporting, rollback, and decommission triggers.
- Treat suspiciously clean metrics as potentially adaptive rather than automatically exculpatory.
- Assume absence of signal is not proof of absence of harm.

## 6. Pilot and rollout gate checklist

Kashi rollout should be treated as a governance-program launch, not a normal software enablement plan. A pilot should not begin until the gate pack below exists.

| Gate | Required artifact | Why it exists | Minimum status |
|---|---|---|---|
| **Scope** | In-scope / out-of-scope meeting classes; external attendee | Prevents quiet scope creep | Approved |

| | | | |
|---|---|---|---|
| | rules; transcription policy | | |
| **Notice** | Worker notice, meeting notice, plain-language FAQ, role-based explanation | Monitoring legitimacy depends on bounded notice | Approved |
| **Representation** | Union / works-council / majority-rep consultation packet and responses | Worker involvement is not cosmetic | Completed or documented |
| **Policy** | Purpose statement, retention schedule, no-go uses, anti-retaliation rules | Turns ethical intent into internal rules | Approved |
| **Challenge** | Transcript, diarisation, summary, threshold, and review dispute path | Meaningful oversight requires recourse | Live before pilot |
| **Access** | RBAC matrix, reason-coded drill-down, audit-log rules | Prevents casual browsing and after-the-fact improvisation | Live before pilot |
| **Misuse** | Sanction path for unauthorised reuse or repurposing | Use walls fail unless breaches have consequences | Approved |
| **Exit** | Pause, rollback, and decommission triggers | A serious system needs a stop condition | Approved |

## 7. Immediate work plan

The shortest viable path is a staged hardening sequence rather than one giant rewrite.

| Phase | What to do | Concrete outputs |
|---|---|---|
| **P0** | Fix copy and doctrine | Replace weak claims; decide structural-only vs constrained hybrid; add two-lane architecture slide; rewrite monitoring language; publish evidence-grade and abstention posture |
| **P1** | Build procedural and validity controls | Meeting-type normalisation spec; challenge flow; anti-retaliation state model; access doctrine; Manager Mirror guardrails; no-go use wall |

| | | |
|---|---|---|
| **P2** | Harden rollout and buyer readiness | Pilot gate pack; worker/rep consultation materials; Security & Assurance Pack; subprocessor map; retention and deletion memo; incident response summary |

## Appendix A. High-priority wording replacements

| Current risky line | Safer replacement |
|---|---|
| 'Not an employee-monitoring tool.' | 'Not a general surveillance or performance-monitoring system. A restricted meeting-governance system with strict procedural and visibility limits.' |
| 'Never transcribe for analysis.' | 'Kashi ingests transcript-linked meeting records, but employer-facing detection is restricted to structural interaction signals rather than open-ended semantic content classification.' |
| 'The pattern is the harm.' | 'The pattern may constitute evidence consistent with uneven conversational treatment over time.' |
| 'Kashi is the CEO's instrument for seeing the bill before it arrives.' | 'Economic sponsor logic: Kashi shortens the time that hidden people-risk remains invisible and unmanaged. Public product identity: restricted meeting-governance infrastructure.' |

## Appendix B. Selected official legal and governance anchors

This appendix lists the main official or near-primary materials used to shape the memo. It is intentionally selective and practical.

- Japan Personal Information Protection Commission (PPC) — APPI Q&A on employee monitoring and purpose specification — https://www.ppc.go.jp/personalinfo/faq/APPI_QA/
- Japan PPC — purpose specification Q&A — https://www.ppc.go.jp/all_faq_index/faq1-q2-1/
- MHLW — workplace harassment / employer measures — https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/koyou_roudou/koyoukintou/seisaku06/index.html
- MHLW — startup labour Q&A on work rules and majority representative opinion — https://www.startup-roudou.mhlw.go.jp/qa/zigyonushi/syuugyoukisoku/q3.html
- European Commission — AI Act policy / FAQ / navigating the Act — https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
- EU AI Act text (EUR-Lex) — https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689
- EDPB — consent guidance under GDPR — https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-consent-under-regulation-2016679_en

- Hong Kong PCPD — Monitoring and Personal Data Privacy at Work — https://www.pcpd.org.hk/english/data_privacy_law/code_of_practices/files/Monitoring_and_Personal_Data_Privacy_At_Work_revision_Eng_20151103.pdf
- Singapore PDPC — data protection obligations — https://www.pdpc.gov.sg/overview-of-pdpa/the-legislation/personal-data-protection-act/data-protection-obligations
- China PIPL official English text — https://en.spp.gov.cn/2021-12/29/c_948419.htm

## Appendix C. One-line project synthesis

**Legal-max version: narrower purpose, honest monitoring language, stronger worker-private states, stronger validity and abstention, and a real procedural backbone.**