Shadow AI in the Workplace

Kantara Initiative Discussion Group — Working Outline

Contributor: Deb Bucci

April 2026

1. Framing the Problem — Why This Matters Now

- AI use in the workplace is already happening — regardless of policy
- "Shadow AI" is a signal, not just a violation
- Containment strategies are failing in practice

2. What Is Shadow AI? (Clarifying the Term)

- Unsanctioned AI use (tools, agents, automation)
- Semi-sanctioned (known but not governed)
- Fully sanctioned but not understood or controlled
- Spectrum, not binary

3. Drivers of Shadow AI Adoption

- Productivity gaps vs. enterprise tooling
- Accessibility of external AI systems
- Latency of internal governance vs. speed of work
- Individual optimization vs. organizational control

4. Risk Landscape

- Data exposure / leakage
- Intellectual property risks
- Regulatory / compliance exposure
- Model reliability / hallucination
- Loss of auditability and traceability

5. Identity vs. Authority — The Core Gap

- **Identity answers:** who is acting
- **Authority answers:** what they are allowed to do
- Current systems bind authority at access time, not execution time
- AI agents amplify this gap:
  - Act across systems
  - Reuse credentials
  - Obscure attribution

6. Delegation and Agent Behavior

- Humans increasingly delegate tasks, not just access

- Agents may:
  - Act repeatedly
  - Act asynchronously
  - Act across contexts
- **Key issue:** Authority persists beyond intent

7. Failure Modes

- Agent continues acting after user context changes
- Shared credentials → indistinguishable actors
- No ability to selectively revoke in-flight actions
- "Confused deputy" scenarios
- Actions taken without a clearly bound principal

8. Why Current Controls Fall Short

- IAM, OAuth, SSO → access grants, not execution governance
- Logging ≠ control
- Policy defined statically, not evaluated dynamically
- Revocation is coarse and delayed

9. Toward a New Control Point

- Need for execution-time evaluation of authority
- Separate:
  - **Intent** (what was meant)
  - **State** (what is true now)
  - **Authority** (what is allowed right now)
- Introduce a control layer that:
  - Evaluates each action at the moment it occurs
  - Produces a visible decision (ALLOW / DENY + reason)

10. Governance Considerations

- How is intent expressed and updated?
- Who can override or intervene?
- How are decisions audited and explained?
- Cross-system and cross-organization implications

11. Enterprise Posture Shift

From:

- "Block unauthorized tools"

To:

- "Govern actions regardless of tool"

12. Discussion Prompts for the Group

- Where are you seeing Shadow AI today?
- What risks are real vs. theoretical?
- What control points exist today (if any)?
- How should delegation be modeled going forward?
- Is execution-time enforcement a missing layer?