



Evaluation Checklist:

Enterprise AI Search & Knowledge Management

A practical, RFX-aligned guide for evaluating
platforms in this category

Pre-RFX Supplement | 2026

✓ How To Use This Checklist

This checklist helps you run structured market research and vendor evaluation without writing a full, formal RFX. It delivers many of the same decision-support inputs procurement teams typically build during RFX intake: clear requirements, evidence requests, evaluation criteria, and a documented decision trail.

If you later choose to run a formal RFX, you can use the outputs from this checklist as a plug-in for your requirements attachment, vendor demo script, and scoring rubric.

Step 1.

Start with outcomes and constraints. Capture mission outcomes, risk boundaries, and required timelines before comparing features.

Step 2.

Run a structured vendor screen. Share this checklist in advance and ask vendors to respond in writing (short, direct answers).

Step 3.

Hold at least one focused screening call per vendor (we recommend about 60 minutes). Use the “Minimum bar” items in each section to eliminate poor fits early. If needed, schedule a second deep-dive session for the short list.

Step 4.

Require proof for the requirements that matter most. For high-risk or complex requirements, ask for a demo, architecture evidence, or a reference that confirms the claim. Use judgment on lower-impact items.

Step 5.

Score what matters. Weight governance, security, and implementation realities more heavily than “nice to have” capabilities.

Step 6.

Translate into procurement language. Use the “Evidence to request” lines to build an RFX addendum or requirements attachment.

What “done” looks like: You can explain, in plain language, why the selected platform is the best value and/or lowest-risk way to deliver trusted enterprise search and answers across your environment, within your timeline and compliance boundaries.

1. Public Sector Readiness

Security & Compliance Framework

- ☐ Platform demonstrates pathway to FedRAMP, GovRAMP, or equivalent authorization, or already holds relevant certifications
- ☐ Demonstrates encryption at rest and in transit with clear data residency controls
- ☐ Provides hybrid or private cloud deployment options where data never leaves your environment
- ☐ Can articulate how it handles sensitive data, FOIA requests, and public records retention requirements

Procurement Access

- ☐ Available through GSA Schedule, NASPO, state cooperative contracts, or other relevant vehicles
- ☐ Can provide contract numbers, NAICS codes, and ordering instructions
- ☐ Has successfully completed public sector procurements before

Checklist

Proven Client Base

Why it matters: Public sector references signal fit in environments with complex governance, security, and change constraints, and can reduce execution risk. Private sector references can be a signal of product maturity and speed-to-value, and can help confirm a vendor can win on outcomes, not just navigate compliance.

- ☐ Provides references from federal, state, or local agencies

- ☐ Can describe relevant private sector experience that demonstrates speed-to-value and measurable outcomes (not just "we pivoted to government after struggling commercially")

- ☐ Can name proven use cases that map to your environment. Same-vertical examples are a strong signal, but results from other verticals can often still validate fit for your needs.

- ☐ References include organizations with comparable size, complexity, and regulatory environment

Red Flags

- Vague claims about "government-ready" without specific compliance certifications
- Cannot explain how permissions and access controls align with your existing governance policies
- No experience navigating appropriations cycles, budget constraints, or oversight requirements

2. Technical Architecture

Integration Philosophy

- ☐ Works with your existing technology stack without requiring migration or replacement
- ☐ Connects natively to the apps you already use (M365, Google Workspace, Salesforce, ServiceNow, Jira, etc.)
- ☐ Explains how it integrates across cloud, on-premises, and legacy systems
- ☐ Provides extensive native connectors (50+) across major enterprise systems or a clear path to custom integrations

Permissions & Governance

- ☐ Enforces real-time permissions from source systems (respects existing ACLs)
- ☐ Never shows users content they cannot access in the original application
- ☐ Explains how it handles document-level security and more granular access controls where applicable
- ☐ Does not require data replication or movement to function

AI Architecture & Transparency

- ☐ Uses Retrieval Augmented Generation (RAG) or similar approach that grounds AI in your data, not the public internet
- ☐ Explains how it prevents hallucinations and ensures answer accuracy
- ☐ Provides source citations and lineage for every AI-generated answer
- ☐ Offers access to multiple LLM models (not locked into a single vendor)

Checklist

Evidence To Request

- Architecture diagram showing how data flows (or doesn't flow) between systems
- Explanation of how permissions are enforced and refreshed
- Demo showing a user's access being revoked in a source system and immediately reflected in search results
- Technical documentation on RAG implementation and source citation

Red Flags

- Requires moving or replicating data in ways that create security, governance, or compliance concerns
- Cannot explain how it prevents data leakage or unauthorized access
- AI answers lack source citations or provenance
- Forces you into a single AI model with no flexibility

Ongoing Operations & Governance

- | | |
|--------------------------|--|
| ☐ | Provides comprehensive audit trails and activity logging for compliance and oversight |
| <hr/> | |
| ☐ | Explains how compliance is maintained as source systems, users, and permissions change over time |
| <hr/> | |
| ☐ | Supports data retention policies and records management requirements |
| <hr/> | |
| ☐ | Can demonstrate how governance controls scale as the platform expands to more systems and users |
| <hr/> | |
| ☐ | Offers ongoing monitoring and alerting for security, usage, and compliance events |
| <hr/> | |

Evidence To Request

- Sample audit log showing user activity, data access, and system changes
- Documentation on how permissions sync frequency and change propagation work
- Explanation of retention policy controls and legal hold capabilities

3. Core Capabilities

What the Platform Should Demonstrate:

- **Unified Search:** Search across all connected systems from a single interface, with results ranked by relevance and personalized to the user
- **AI Assistant:** Generative AI that can answer questions, summarize documents, create content, and analyze information using your agency's data
- **Workflow Automation:** AI agents that can automate repetitive tasks, orchestrate multi-step processes, and respond to routine inquiries
- **Knowledge Discovery:** Surface institutional knowledge, tribal expertise, and relevant context based on what the user is working on

Concrete Deliverables to Look For:

- ☐ Can demonstrate search across at least 5 disparate systems in real-time
- ☐ Shows AI-generated answers with clear source citations and confidence indicators
- ☐ Provides examples of workflow automation or AI agents built for similar use cases
- ☐ Demonstrates personalization (results adapt based on user role, department, and activity)
- ☐ Shows how it surfaces institutional knowledge during onboarding or project transitions

Red Flags

- Only searches one ecosystem (e.g., Microsoft-only or Google-only)
- AI features feel like a bolt-on rather than core functionality
- Cannot show how it handles complex, multi-part questions
- Requires significant configuration before delivering basic search functionality

4. Implementation & Deployment

Speed to Deployment

- ☐ Can be operational in weeks, not months or quarters
- ☐ Offers phased rollout options (pilot with one department before agency-wide deployment)
- ☐ Provides clear onboarding timeline with milestones and deliverables
- ☐ Explains technical requirements and IT resource needs upfront

Integration & Configuration

- ☐ Demonstrates straightforward connector setup and configuration process
- ☐ Provides clear documentation and technical support during implementation
- ☐ Can work within annual budget constraints and appropriations timing
- ☐ Explains technical requirements and IT resource needs upfront

Checklist

Technical Support & Enablement

- ☐ Includes dedicated implementation support and technical guidance

- ☐ Provides ongoing optimization and tuning services

- ☐ Offers clear SLAs for uptime, support response, and issue resolution

- ☐ Has responsive technical support team familiar with government environments

Red Flags

- Implementation timelines extend beyond 90 days for basic functionality
- Requires extensive IT resources or dedicated team to manage
- Vague about technical requirements or dependencies
- No clear path for scaling from pilot to agency-wide deployment

5. User Adoption & Change Management

User Experience

- ☐ Offers search and AI features where employees already work (Slack, Teams, email, browser)
- ☐ Demonstrates intuitive interface that requires minimal training
- ☐ Shows that users can be productive within days, not weeks
- ☐ Provides personalized experience that adapts to individual roles and work patterns

Adoption Strategy & Support

- ☐ Includes change management resources and proven adoption strategies
- ☐ Provides training materials, documentation, and onboarding resources
- ☐ Offers adoption analytics and dashboards to track usage and identify barriers
- ☐ Shows measurable adoption metrics from similar organizations (target: 70%+ adoption within 6-12 months)
- ☐ Can articulate how they help organizations drive adoption beyond initial deployment

Checklist

Adoption Strategy & Support

- ☐ Provides usage analytics showing active vs. passive users

- ☐ Demonstrates measurable productivity improvements (time saved per employee, reduction in support tickets)

- ☐ Can show reduction in use of unauthorized tools or workarounds

- ☐ Tracks user satisfaction and engagement over time

Red Flags

- Adoption rates in case studies are below 50% after the first year
- No clear change management support or adoption resources
- Cannot explain how they measure or drive adoption
- Users describe the tool as "nice to have" rather than essential to their work

6. Cost & ROI

Pricing Transparency

- ☐ Provides clear per-user or per-employee pricing model

- ☐ Explains what is included vs. additional costs (connectors, support, AI usage)

- ☐ Can work within annual budget constraints and appropriations timing

- ☐ Offers flexible contract terms (annual, multi-year, or consumption-based)

Value Demonstration

- ☐ Provides ROI calculator or framework showing productivity gains

- ☐ Can quantify time savings per employee (benchmark: at least 2-5 hours per week)

- ☐ Shows reduction in "shadow AI" risk (employees using unauthorized tools like public ChatGPT)

- ☐ Demonstrates faster onboarding, reduced support burden, or improved decision-making

Checklist

Pilot & Proof-of-Concept Options

- ☐ Offers proof-of-concept or pilot options to demonstrate value before full commitment

- ☐ Explains what is included vs. additional costs (connectors, support, AI usage)

- ☐ Provides clear success criteria and measurement framework for pilots

- ☐ Can deliver measurable results within 30-60 days

Red Flags

- Pricing is opaque or changes significantly after initial discussions
- Cannot articulate ROI in terms that matter to budget officials
- No pilot option available; requires full enterprise commitment upfront
- Hidden costs emerge after contract signature (per-connector fees, overage charges, etc.)

7. Competitive Positioning

Ecosystem Coverage vs. Walled Gardens

- ☐ Works across multiple ecosystems (Microsoft, Google, Salesforce, Atlassian, etc.)
- ☐ Does not lock you into a single vendor's suite of tools
- ☐ Can search and surface answers from proprietary systems and custom applications
- ☐ Can search and surface answers from proprietary systems and custom applications

AI Safety & Control

- ☐ Explains how it prevents data exfiltration to public AI models
- ☐ Provides governance controls for AI usage (audit logs, policy enforcement, usage limits)
- ☐ Demonstrates how it handles sensitive or classified information
- ☐ Offers "safe AI" alternative to employees using unauthorized tools

Differentiation & Innovation

- ☐ Articulates clear advantages over competitors (search quality, AI accuracy, speed to value)
- ☐ Shows ongoing investment in R&D and platform evolution
- ☐ Provides roadmap visibility for upcoming features and capabilities
- ☐ Offers "safe AI" alternative to employees using unauthorized tools

Checklist

Questions to Ask:

- "How do you differ from Microsoft Copilot or Google's AI offerings?"
- "What happens if an employee tries to search for classified or sensitive data?"
- "How do you handle multi-cloud or hybrid environments?"
- "What if we add a new system or retire an old one...how quickly can you adapt?"

Red Flags

- Cannot clearly differentiate from free or bundled offerings
- Defensiveness when asked about competitors or limitations
- Roadmap is vague or nonexistent
- Requires ripping out existing tools to function

8. References & Proven Results

Evidence of Impact

- ☐ Provides specific metrics from past deployments (hours saved, support tickets reduced, onboarding time cut)

- ☐ Can explain ROI in terms that matter to oversight and budget officials

- ☐ Shows measurable improvement in knowledge retention during employee turnover or retirements

- ☐ Demonstrates reduction in "search abandonment" or frustrated users

References That Matter

- ☐ At least three references from public sector organizations

- ☐ References include both technical and executive stakeholders

- ☐ Can speak to results achieved, not just features deployed

- ☐ References are from organizations of comparable size and mission complexity

Questions to Ask:

- "What specific productivity improvements did you measure?"
- "How long did it take to see value after deployment?"
- "What percentage of your workforce actively uses the platform?"
- "Did it reduce reliance on unauthorized AI tools?"
- "What surprised you (positively or negatively)?"
- "Would you expand your deployment or recommend it to peer agencies?"

Red Flags

- Cannot provide references from public sector clients
- Results are described in activities ("deployed search") not outcomes ("saved 110 hours per employee per year")
- References are reluctant to endorse or cite implementation challenges
- Case studies lack quantifiable metrics or business impact

✓ How To Apply This Checklist

Step 1: Initial Screening

Use sections 1, 4, and 6 to quickly eliminate platforms without basic qualifications (security, procurement access, deployment feasibility, reasonable pricing).

Step 2: Technical Deep Dive

Use sections 2 and 3 to assess whether the architecture and capabilities match your needs (permissions enforcement, ongoing governance, ecosystem coverage, AI transparency).

Step 3: Adoption & Change Readiness

Use section 5 to evaluate whether the platform can achieve meaningful adoption in your organization, not just technical deployment.

Step 4: Competitive Differentiation

Use section 7 to understand how the platform compares to alternatives and whether it solves problems that free or bundled tools cannot.

Step 5: Risk & Value Validation

Apply section 8 to validate claims through references and measurable evidence from similar organizations.

Step 6: Comparative Scoring

Create a simple scorecard assigning points for each criterion that matters most to your situation. Weight heavily the factors that address your most critical pain points (e.g., if Shadow AI risk is your top concern, prioritize governance and permissions enforcement in sections 2 and 7).

Final Guidance

The best platforms will welcome detailed questions and provide concrete evidence. If a vendor resists technical scrutiny, cannot produce references, or deflects when asked about limitations and trade-offs, that tells you something important.

The right platform will be transparent about its architecture, honest about what it can and cannot do, and able to demonstrate measurable results in environments like yours.

The "Knowledge Paradox" Test:

Ask yourself: Does this platform help us solve the core problem? For many, that core problem is that their workforce has access to more data than ever but knows less than ever because that data is trapped in silos. If the answer is yes, and it meets the criteria above, you have found a solution worth piloting.



REDLEIF