ARCADIA

Security & Data Privacy Overview

Prepared for academic library and institutional review

This document is intended for library directors, systems librarians, IT security teams, and institutional procurement officers evaluating Arcadia for deployment. It addresses the most common security and data privacy questions in plain language, without technical jargon where possible.

1. What Arcadia Does — and Doesn't Do

Arcadia is an AI-powered research assistant widget that libraries embed on their existing websites. When a student or faculty member asks a research question, Arcadia searches the library's catalogue via a read-only API connection and uses an AI model to synthesise a research briefing from the results.

It is important to be clear about what Arcadia does not do:

– Arcadia does not store, log, or retain researcher search queries.

– Arcadia does not collect any personally identifiable information about users.

– Arcadia does not modify, write to, or alter any library catalogue data.

– Arcadia does not require users to create accounts or provide credentials.

– Arcadia does not scrape, copy, or republish catalogue content.

---

2. Data Flow: Where Information Goes

Understanding the data flow is the most important part of any security evaluation. Here is exactly what happens when a researcher uses Arcadia:

Step 1 — The researcher types a question into the widget on the library's website. This question is sent over HTTPS to Arcadia's backend server.

Step 2 — Arcadia's server translates the question into optimised search queries and sends them to the library's Ex Libris Primo API using a read-only institutional API key. This is identical to what happens when a user searches OneSearch directly.

Step 3 — Primo returns catalogue records (titles, authors, abstracts, metadata) to Arcadia's server. No full text of licensed content is transferred — only the same metadata that is publicly visible in the catalogue.

Step 4 — Arcadia sends the catalogue records and the original research question to Anthropic's Claude API. Claude synthesises a research briefing and returns it. The question and catalogue records are not retained by Anthropic beyond the duration of the API call under their standard data processing terms.

Step 5 — The briefing is returned to the researcher's browser and displayed in the widget. Nothing is stored on Arcadia's servers.

---

3. What Data Arcadia Accesses

Arcadia accesses only data that is already publicly available through your library's catalogue search interface. Specifically:

– Bibliographic metadata: titles, authors, publication dates, subjects, abstracts

– Record identifiers used to construct deep links back to full catalogue records

– No licensed full-text content is accessed or transmitted

– No patron data, circulation records, or account information is accessed

– No internal administrative or acquisitions data is accessed

The read-only API key provided by the institution limits Arcadia's access to search queries only. It cannot be used to modify records, access patron accounts, or retrieve any data beyond public catalogue search results.

---

4. Security Architecture

4.1 What Could Go Wrong, and Why the Risk Is Low

The most common institutional concern is: could Arcadia be hacked, and what would an attacker gain? The honest answer is: very little. Arcadia's architecture is deliberately minimal:

– No database: Arcadia stores nothing persistently. There is no database of user queries, researcher profiles, or catalogue data to breach.

– No user accounts: Researchers do not log in to Arcadia. There are no credentials to steal.

– Read-only catalogue access: The worst case for a compromised API key is that an attacker could run searches against your catalogue — the same thing any anonymous user can do through OneSearch.

– No financial data: Arcadia processes no payments and holds no financial information.

4.2 Rate Limiting and Abuse Prevention

The /query endpoint is rate-limited to prevent abuse. Excessive automated requests are rejected before they reach either the Primo API or the Anthropic API. This protects both your institutional API key and Arcadia's operating costs from malicious use.

4.3 Transport Security

All communications between the researcher's browser, Arcadia's server, your Primo API, and Anthropic's API are encrypted using HTTPS/TLS. No data is transmitted in plain text at any point in the pipeline.

---

5. Third-Party Services

5.1 Anthropic (Claude API)

Arcadia uses Anthropic's Claude API to synthesise research briefings. Relevant data handling details:

– Anthropic is SOC 2 Type II certified.

– API inputs and outputs are not used to train Anthropic's models under their standard API terms.

– Data sent via the API is not retained beyond the duration of the request under standard API usage.

– Anthropic's full data processing agreement is available at anthropic.com/legal/privacy.

For institutions with strict data residency requirements — particularly EU institutions subject to GDPR — it is worth noting that Anthropic processes data in the United States. Arcadia can discuss data processing agreements and model deployment options for institutions where this is a constraint.

5.2 Ex Libris Primo API

Arcadia integrates with Ex Libris Primo via the standard v1 REST API, which is the same API used by Ex Libris's own developer ecosystem and documented publicly at developers.exlibrisgroup.com. The integration uses a read-only API key generated by the institution's own library systems administrator, who retains full control over the key and can revoke

it at any time.

5.3 Cloud Hosting

Arcadia's backend is hosted on Railway, a cloud infrastructure provider. Servers are located in the United States (US West region). Railway maintains SOC 2 compliance. For institutions requiring on-premise or regional hosting, this is a discussion Arcadia is willing to have.

---

6. The AI Accuracy Question

A concern specific to AI-powered tools in academic settings is the risk of hallucination — the AI generating plausible-sounding but false information. This is a legitimate concern that Arcadia takes seriously.

Arcadia addresses this through a strict system prompt that instructs the model to:

– Synthesise only from the catalogue records returned by the Primo search — never from the model's general training knowledge.

– Cite every claim with the source title and author.

– Explicitly acknowledge when the source set is limited or does not fully address the research question.

– Never invent sources, authors, or publication details not present in the search results.

In practice this means the briefing Arcadia generates is grounded in real catalogue records that are linked back to your OneSearch interface. Researchers can verify every source independently with one click.

---

7. Security Summary

Security Area	Status	Notes
No persistent data storage	**Addressed**	Nothing retained after request
No user PII collected	**Addressed**	No accounts, no tracking
Read-only catalogue access	**Addressed**	Cannot modify any library data
Transport encryption (HTTPS)	**Addressed**	All connections TLS encrypted

Security Area	Status	Notes
Third-party AI data handling	**Addressed**	Anthropic SOC 2, no training use
Rate limiting / abuse prevention	**Addressed**	Implemented on /query endpoint
AI hallucination risk	**Addressed**	Grounded synthesis, cited sources only
Data residency (US processing)	**Minimal**	Discuss if EU/GDPR constraint applies
On-premise deployment option	**On request**	Available for enterprise agreements

8. For Your IT Security Team

If your institution requires a formal security review, the following information is typically requested:

Penetration testing: Not yet conducted. Available on request for institutions proceeding to contract.

Data Processing Agreement: Available on request.

Subprocessor list: Anthropic (AI synthesis), Ex Libris / Clarivate (catalogue API), Railway (hosting).

Incident response: Arcadia commits to notifying affected institutions within 72 hours of any confirmed security incident.

API key management: Institutions retain full control of their Primo API key and can revoke access at any time without notifying Arcadia.