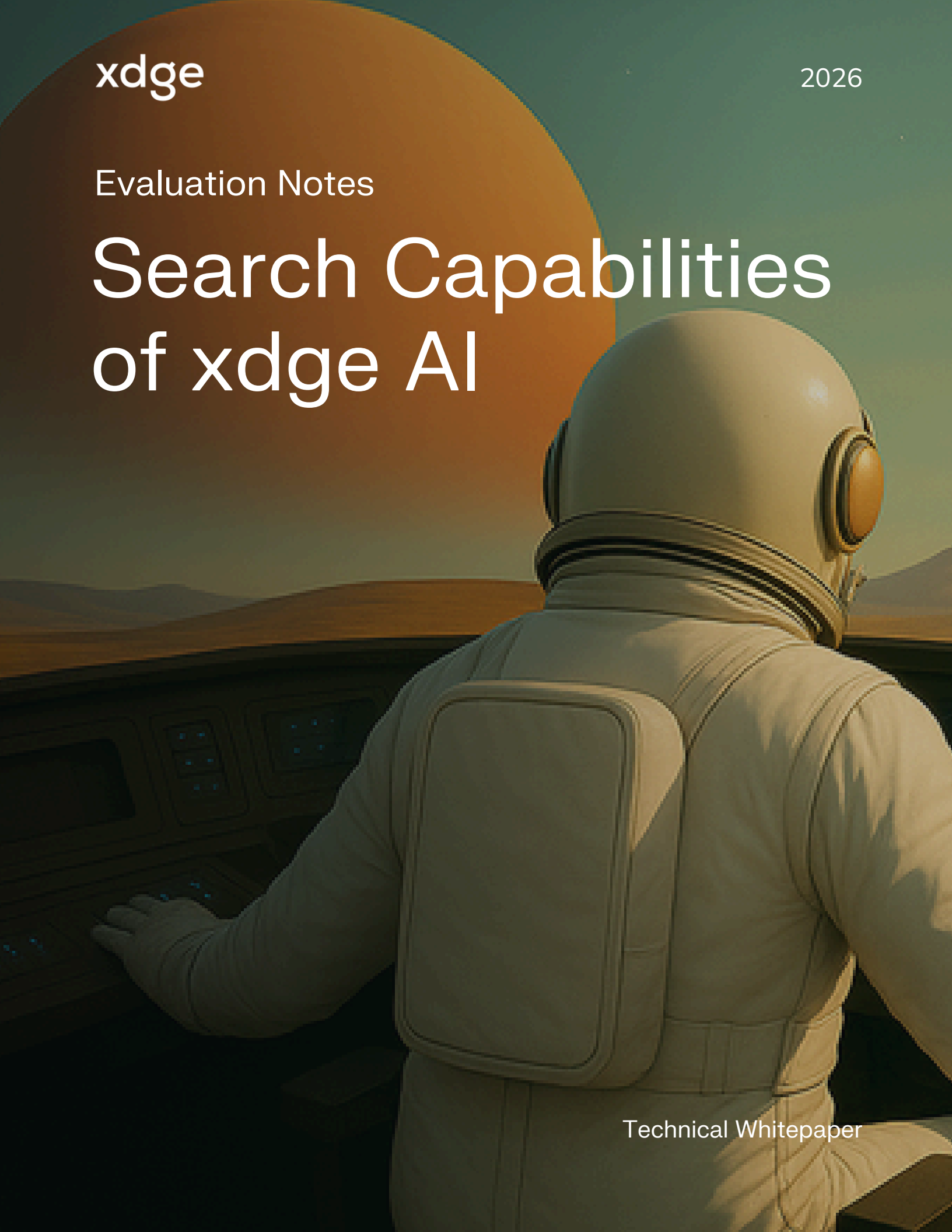
xdge

2026

Evaluation Notes

Search Capabilities of xdge AI



Technical Whitepaper

Evaluating Search Capabilities of xdge AI

Executive Summary

This document is aimed at technical due diligence teams evaluating xdge AI's search capabilities. It explains how to test the platform's search capabilities, including the product design choices made and the rationale behind them.

Wait ~3 hours before testing newly uploaded content

Many enterprise applications take time for newly created documents to become searchable in their own UIs and APIs. In addition, xdge performs staggered periodic crawling across connectors (typically 30 minutes to ~3 hours, configurable). The staggering allows crawlers to be distributed over time. Because of these combined delays, testing search results immediately after uploading content can produce misleading results. For accurate evaluation, we recommend waiting approximately 3 hours after uploading new content before testing search behavior. This is a safe upper limit, and most of the time, content is searchable much sooner.

Use Recent Mode by default, and switch to Anytime Mode when needed

Recent Mode is used as the default for searches because it retrieves results from the indexed dataset created through periodic crawling, providing the fastest and most representative experience. Anytime Mode performs live API searches and is automatically used when indexed data is not yet available—such as during initial onboarding and after adding a new connector—before the first crawl runs. The user can also choose to use Anytime Mode when searching very old historical data that has not been indexed.

Technical Details Explaining the Architecture and Rationale

Background

When evaluating xdge AI, it helps to understand how its search behaves in steady-state operation and how that differs from certain transitional states, such as during a first-time crawl or immediately after adding a new connector. The sections below describe what you should expect while testing.

Steady-State Operation (Normal Production Behavior)

In a steady state, all connected applications have already completed their initial crawl and are maintained through periodic incremental crawling.

Crawling runs on a staggered schedule across applications, typically ranging between 30 minutes and 3 hours, though these intervals are fully configurable depending on operational preferences and scale.

In this mode, xdge provides two complementary retrieval paths described below.

Recent Mode

Recent Mode retrieves documents from the xdge index, which is populated through the periodic crawling process.

Because content is pre-indexed:

- Searches are very fast
- Results reflect content captured during the most recent crawl
- Crawling across applications happens on staggered intervals (30 minutes to ~3 hours by default)

This is the default mode used for most searches.

Since crawling happens periodically, there is naturally a window where very recently created items may not yet appear in the index.

Example scenario during testing:

- A document is uploaded to SharePoint or a ticket is created in Jira.
- Until the next crawl runs for that application, the item may not appear in Recent Mode.

Anytime Mode

Anytime Mode performs live API searches across connected applications.

While Anytime Mode can surface newly created objects if they are available through the application's API, the primary purpose of Anytime Mode is historical retrieval.

In many enterprise environments:

- Systems may contain 10–20 years of data
- Crawling all historical content continuously can become operationally expensive
- Older content often has diminishing retrieval frequency

Because of this, Anytime Mode provides a way to search very old historical content on demand through APIs without requiring xdge to maintain a fully indexed copy of that entire historical dataset.

Native Latency in Source Applications (UI and API)

During testing, it is important to recognize that many SaaS applications do not make newly uploaded content searchable immediately.

For example:

- Platforms like SharePoint often take time before a newly uploaded document becomes searchable in their own UI
- The same delay appears in their API search results

Because of this, there can be variability between:

- The application's native UI search
- API search (used by xdge Anytime Mode)
- Indexed search (used by xdge Recent Mode)

To avoid confusion during testing, we recommend waiting approximately 3 hours after uploading new content before evaluating search results. This is a safe upper limit, and most of the time, content is searchable much sooner.

This ensures that:

- The source application has made the content searchable through its own indexing
- Staggered crawling schedules across connectors have had time to run
- API-level propagation delays do not affect evaluation results.

First-Time Crawl (New Tenant Onboarding)

When a new tenant connects applications for the first time, xdge begins an initial crawl and indexing process.

During this phase:

- Some connectors may already be indexed
- Others may still be in the crawling process

For connectors that are not yet fully indexed, Recent Mode temporarily performs API-based searches, effectively behaving like Anytime Mode.

This means:

- Results may be slower during this phase
- But objects that are available through the application's APIs will still be retrievable

Once the initial crawl completes, Recent Mode returns to its normal behavior of serving results from the index.

Adding a New Connector to an Existing Account

A similar temporary state occurs when a user adds a new connector to an already-indexed tenant.

In this situation:

1. Existing connectors continue operating through indexed search in Recent Mode.
2. The newly added connector begins its first crawl.
3. During that crawl, Recent Mode may temporarily perform API searches for that connector only, similar to Anytime Mode.

This allows the new data source to be searchable sooner, even before its indexing completes. Note again that, even with the API-based search, the API itself may not surface newly uploaded content until the app's native indexing has occurred.

Once crawling finishes, that connector behaves like the others, and results are served from the index in Recent Mode.

What This Architecture Enables

Fast search in the default mode

Recent Mode provides high-speed indexed retrieval across all connectors that have completed their crawl cycles.

On-demand historical retrieval

Anytime Mode allows users to search very old enterprise data (often 10–20 years old) without requiring the system to maintain a fully indexed copy of that entire historical dataset.

Operational efficiency at scale

Not crawling decades of historical data continuously reduces infrastructure cost and indexing load while still allowing retrieval via Anytime Mode.

Fallback retrieval during transitional states

When crawling is not yet complete—such as during new tenant onboarding, adding a new connector, or before the next crawl runs after new content is created - Recent Mode may temporarily behave like Anytime Mode, using API-based retrieval so that available objects remain searchable.

Configurable crawling behavior

Crawling schedules are configurable and staggered across applications, typically ranging from 30 minutes to 3 hours, allowing organizations to balance freshness, scale, and operational cost.

xdge

Welcome to the future that works for you

xdge.ai