



Nexis® Data as a Service **Data Delivery Guide**

Table of Contents

INTRODUCTION & API OVERVIEW

NEWS MONITORING APIs

- Overview
- Metabase Search & Retrieve
- Metabase Firehose
- Custom Feed
- Metabase Archive
- Discoverydesk

WEB SERVICES API

- Overview
- Content Types

BULK DELIVERY

- Overview

SECTION 1: INTRODUCTION & API OVERVIEW

Get the exact data you need—delivered precisely how you need it—with Nexis® Data as a Service (DaaS)

Easily integrate into your analytics or applications using flexible APIs and platform options designed for specific data use cases—from trend analysis, event-based triggers, or financial modeling to Natural Language Processing (NLP) development or machine learning.

What APIs and data environments are available with Nexis DaaS?

- **Search and Retrieve APIs** offer access to a wide variety of datasets with search and retrieve functionality, as well as post-search filters to further refine results.
- **Ongoing Monitoring Options** are used to receive updates and alerts throughout the day on developments relevant to you and your use case.
- **Bulk Delivery Options** deliver news, legal and regulatory content reaching back more than 30 years. These deliver large volumes of content for historical analysis, predictive modeling and more.
- **Data Lab** leverages the best-in-class Jupyter Notebook environment, allowing users to search, refine and analyze our unmatched enriched news data—all in one place.

Click on the links below to learn more about Nexis DaaS APIs and delivery platforms.

NEWS MONITORING APIS

- Overview
- Metabase Search & Retrieve
- Metabase Firehose
- Custom Feed
- Metabase Archive
- Discoverydesk

WEB SERVICES API

- Overview
- Content Types

BULK DELIVERY

- Overview

SECTION 2: NEWS MONITORING APIS

The Nexis® Metabase API supports around-the-clock coverage of online news and social commentary from sources worldwide. With unparalleled breadth of content, Metabase powers the business applications of leading companies across multiple markets, including media monitoring firms, financial and business information services, as well as search engines, professional networking sites, and corporate intranets.

Using proprietary data analysis and extraction technology, the Metabase API provides customers with a stream of news and social commentary, delivered in XML, JSON, RSS or ATOM. Leverage the Metabase API for any database or application to search, index and analyze as well as monitor millions of articles for mentions of and commentary on companies, customers, competitors, products, trends, announcements or any other topic pertinent to business intelligence.

Each article is analyzed, categorized and tagged based on more than 100 data elements, such as:

- Sentiment
- Industry
- Company
- Subject
- Language
- Source location
- Media type (news, blogs, social commentary, etc.)
- Source ranking

Global coverage, fabrication and deep metadata enrichments enable customers to monitor, manage and leverage the global swell of online news, opinion, commentary, and buzz.

Specifically developed for NEWS, the Metabase API offers access to an average of 3-5 million news articles and social media posts per day, spanning 235+ countries and territories in 90+ languages.

Delivering relevant content through XML, JSON, RSS or ATOM data feeds, Metabase provides a stream of open web and licensed content enabling standard, flexible integration into any database or application.

Users can monitor coverage and commentary on companies, customers, products, trends, events and pertinent topics, or analyze and report on media trends and business information by entities, locations, vertical markets, stock symbols, MozRank, news source category and more.

Choose the content type needed, in any combination:

Print—Global news and industry publications direct from the publisher in the format in which their print publication was distributed.

Publicly Available—Global news and industry sites curated by the publisher for a web audience.

Social Commentary—Thought-leader blogs, industry forums and other social commentary

Data delivery:

- Content delivery in XML, JSON RSS or ATOM format
- Typical Metadata included:
 - Entity Extraction
 - Reach Metrics
 - Source Metadata
 - Sentiment
 - Duplicate tag to filter out duplicates
 - LexisNexis® SmartIndexing

Metabase API Options

Search – Search and Retrieve API for fast, comprehensive access across all media types

- Search the up to 10 years of Print or the past 100 days of publicly available content and social commentary.
- Up to 200 articles are returned for a search, with the ability to make subsequent calls every 10 seconds.
- Articles are typically available about 2 minutes after initially coming into our system.
- Access to Metabase Search is through an HTTP request like this: `http://metabase.moreover.com/api/v10/searchArticles?key=(your_API_key)&query=(your_query)` This call will return the most recent article(s) matching your query. Metabase Search queries support Boolean syntax and search operators.

Firehose

- Ingest content going forward, with up to 500 articles returned with each call. Typically, calls should be scheduled 20 to 60 seconds apart, depending on the volume of content you are set to receive.
- Create scheduled HTTP calls to the Metabase servers to receive current data as it becomes available.
- Scheduled calls request a batch of articles from the polling API. Articles are delivered in an XML feed, with XML tags providing detailed information about each article.
- Article content encoding is in UTF-8. A single call will return up to 500 articles (maximum). You can instruct the call to only return new articles since the previous call.
- Articles are typically available about 2 minutes after initially coming into our system.

Custom Feed

- Use pre-set search queries to return matching content moving forward.
- Parse out based on search criteria/naming of each filter, eliminating issues of storage and indexing.
- Create filters as needed.

Authentication

The Metabase Filters API uses HTTP Basic authentication. Every request must contain an 'Authorization' header with the value 'Basic' followed by the value resulting from encoding in Base64 the username and password separated by ':' (colon). Example for username 'testUser' and password 'testPassword': Authorization: Basic dGVzdFVzZXI6dGVzdFBhc3N3b3Jk

Basic access authentication

In the context of an HTTP transaction, basic access authentication is a method for an HTTP user agent (e.g., a web browser) to provide a username and password when making a request.

In basic HTTP authentication, a filter has 4 properties:

1. filter id, identified by attribute "id"
2. filter name, identified by attribute "name"
3. filter query, identified by attribute "query"
4. active flag with possible values true/false, identified by attribute: "active"

Example:

```
{ "id": 1, "name": "query_us_gb", "query": "sourceCountryCode:us OR sourceCountryCode:gb", "active": true }
```

A filter must conform with the following:

- the filter name must not exceed 100 characters
- the query length must not exceed 10,000 characters
- the query must be syntactically valid (see details about filters syntax)

The name of the filter(s) that matched an article will be output in the <matchingFilters> tag in the Metabase API response. Example: <matchingFilters> <name>Video Games</name> <name>Nintendo</name></matchingFilters>

Metabase Archive Option:

- Extract up to 10 years archive currently of Print or 100 days of the Web Harvested content, from AWS or an FTP site based on the number of years you require for on premises hosting.
- Keep the archive for the life of the agreement. Required usage tracking is done by the URL that is in the XML. A maintenance key is also required.
- Ingest content going forward, with up to 500 articles returned with each call.
- We can cut this archive for you and you can keep adding to it.

Discoverydesk for Data Analysts & Engineers

Nexis Newsdesk™ can serve as a complementary discovery application that works well with our News API.

A great starting point when using our data, Newsdesk enables one-off searches across the same news content featured in the API. Useful for testing a search on an idea to see if it returns relevant results before embarking on a data analysis project.

Nexis Newsdesk lets you embark on discovery enabling searching and analysis from a single platform, running over the same news content available through the Metabase APIs.

- Easily use the guided search tool to quickly capture the power of our content. Review results, drilling down to the full text in a click or using built-in analytics without downloading any content.
- Embed code to an internal site and integrate interactive charts with visualizations that are live as the news changes.

SECTION 3: WEB SERVICES API (WSAPI)

Web Services API (WSAPI) provides access to enhanced functions such as search, retrieve and saved searches. The client application is responsible for providing all UI design and implementation, simplifying developers' interactions with the services.

- Same content types available in Lexis® Advance & Nexis products
- XML formatted data for ease of use
- Retrieve data on demand with no hosting
- Enhanced post filtering and LexisNexis® SmartIndexing taxonomy to reduce noise

The WSAPI uses an authentication protocol based on OAuth 2.0 Authentication Framework and exposes resources using REST-based data services based on Open Data Protocol (OData 4.0).

WSAPI Content Categories

- Company financial reports and filings
- Profiles of companies and company executives, annual reports and other SEC filings, bankruptcy reports, merger and acquisition transactions, earnings reports. The file covers all US public companies, more than 200,000 private US companies, and many thousands of other companies around the world.
- Biographies and directors' profiles
- Biographical references and directories.
- Sanctions and watchlists
- Politically Exposed Persons (PEPs) lists
- Global Full Text News Sources available in English and non-English Inaguage
- Legal sources

SECTION 4: BULK DELIVERY

Bulk API provides a backfill of historical content, allowing subscriptions to ATOM (XML) feeds of news and US legal content.

Integrate a large volume of on premise content from thousands of sources into your own system with this standard mechanism for bulk data delivery.

With more than 7,400 current news sources and a 30+ year archive of licensed content, the Bulk API is ideal for use in:

- Historical analysis
- Predictive analytics
- Training algorithms and other AI applications

Call **+31 (0) 20 485 3456** to speak with a data specialist, gain access to our Developers Portal and test your proof of concept.