

TRUSTED TOOLS TO OPTIMIZE
YOUR INFORMATION LANDSCAPE

MUSE
KNOWLEDGE

HARVEST

A powerful content aggregator that lets you unlock information scattered across silos and stored in varied formats, Muse Knowledge Harvest automates your entire extraction, transformation, normalization and content delivery process.

Its intelligent data harvesting capabilities ensure you capture your content in the context you need and keep your information current. Whether gathering information from mainstream services or secure private platforms, you can reliably and precisely power your data collection and delivery with Muse Knowledge Harvest.





IMPROVE INFORMATION OUTCOMES

When you need to connect your users to important data locked in legacy systems or spread across proprietary systems.

When you need to deliver new content streams to address your organization's expanding requirements.

When you need to raise your existing content's quality to better meet your organizational objectives.

Muse Knowledge Harvest can resolve the challenges of a fractured content and information landscape.

With content harvesting processes that are tailored for your specific situational relevance, you can employ Muse Knowledge Harvest to help your organization:

- Improve the quality of existing content
- Build custom datasets from diverse feeds
- Migrate legacy data into modern systems
- Streamline real-time web content collection
- Create new data products and services

CUSTOMIZE CONTENT COLLECTION

The Muse Knowledge Harvest process is uniquely configured for each individual project. By leveraging key components of the powerful Muse Knowledge Software Platform, Muse Knowledge Harvest can reliably and cost-effectively deliver optimal content harvesting results. In close consultation with your organization, the Muse Knowledge team of experts will create a custom Muse Knowledge Harvest plan to:

- **Identify** the content sources you need to harvest: from the surface web, deep web, databases, applications and archives to third-party sources, social media, polling feeds and more.
- **Determine** what information you need extracted from each harvest source: from general term filters to specific single item validators, detailed contiguous extraction is defined for every harvest source to meet your objectives.
- **Tailor** the harvesting processes to extract content when and how you need: from extracting search-based, date-based, and feed-based content, to timed or 'on-demand' extraction run at granular intervals down to the minute.
- **Configure** the content delivery stream directly into your databases, engines and dashboards: from consistent delivery formats for each individual record, to optional enhanced 'virtual records', content is delivered by both file and feed, in full result sets, in deltas or both.

DEPEND ON DATA DELIVERY

Whether to support a one-time content migration project or a continual content curation service, there are many reasons why your organization needs a trustworthy content harvesting solution. Muse Knowledge Harvest is a powerful suite of content aggregation tools built to help you optimize your information landscape and increase your return-on-content investment.

To learn how Muse Knowledge Harvest can reliably and cost-effectively augment your organization's ability to access, use and interact with information, contact the Muse Knowledge team today.



RETURNING THE RIGHT DATA

- Aggregated from any number of sources
- Consistent, normalized records
- Gathered by feed or query
- Enhanced, 'virtual records'
- Detailed contiguous extraction
- Delivered by file or feed
- Deliver both full results or deltas

Content Consumers



LOCAL FILES
RDBMS



CONTENT
PROVIDERS



FEED
REPOSITORY

Muse Knowledge Harvesting

IMPORT

DIRECT
SEARCH
IMPORT

CONNECTORS

VARIOUS PROTOCOLS:
Atom, HTTP/XML,
JSON, NCIP, OAI-PMH,
RSS1.0, RSS2.0, SIP2,
SQL, SRU, SRW, Telnet,
Z39.50

FILTERS

DE-DUPE, ENRICH,
RANK, RECORD
ENHANCEMENT

VALIDATORS
Normalization,
Entity Extraction

WRITERS

DELTA WRITERS
FULL RECORD
WRITERS

Content

DELTAS REPOSITORY

FULL RECORDS REPOSITORY

FOR MORE INFORMATION

visit our website at
museknowledge.com and for any
inquiries, contact our office.

MUSE KNOWLEDGE HEADQUARTERS

Calea Bucuresti, Bl. 27B,
Sc. 1, Ap. 10
Craiova 200675, România

phone +40 251 413 496
fax +40 251 418 935
email info@museknowledge.com
email support@museknowledge.com

MUSE KNOWLEDGE EMEA

21 El Khalifa El Mamounst.
Roxy Tower, Heliopolis, 11341,
Cairo, Egypt

phone +202 241 87 349
phone +202 229 01 994
email info@museknowledge.com
email sales@museknowledge.com

MUSE KNOWLEDGE NORTH AMERICA

340 Madison Avenue, 19th Floor
New York, NY 10173 USA

phone +1 212 220-9250
email info@museknowledge.com