Searching in AtoM

Exploring Advanced search tips and techniques in Access to Memory
Global search box

- Available throughout AtoM via the header bar
- Shows autocomplete results from:
  - Descriptions
  - Authority records
  - Archival institutions
  - Places
  - Subjects
- If you enter a search and hit return, it defaults to archival description results!
- Supports BOOLEAN expressions (must be in caps)
  - AND
  - OR
  - AND NOT
- Supports wildcards and other special search characters
- When multi-repository support is enabled, allows search to be scoped to descriptions from last repository visited
To improve the relevancy of search results, AtoM includes some basic weighting added to certain fields. When a search term matches the data in one of these weighted fields, it is boosted and returned higher in the list of matching results.

**Archival descriptions**
- **10x weight:** Title
- **6x weight:** Creator
- **5x weight:** Identifier, Subject access point, Scope and content
- **3x weight:** Name access point, Place access point

**Accessions**
- **10x weight:** Accession number, Donor name, Title, Scope and content
- **5x weight:** Location information, Processing notes, Immediate source of acquisition, Archival/custodial history
**Search overview**

**Default operator: AND**
- When multiple terms are entered in a search, AtoM will use AND as the default operator.
- This means that a search for **vancouver city hall** will be interpreted as:
  - Return records with Vancouver AND city AND hall appearing in any field.

**SEARCHING EXACT PHRASES**
- To search for a phrase in the exact order you enter the query, use quotations.
- A search for **vancouver city hall** will return any records that have vancouver AND city AND hall appearing anywhere.
- A search for *"vancouver city hall"* will return only records that have those terms together in that exact order.
**RESERVED CHARACTERS**

- AtoM uses Elasticsearch (ES) as its search index. In ES, certain characters are normally reserved to be used as operators in advanced searching.
- If you need to use any of the characters which function as operators in your query itself (and not as operators), then you should escape them with a leading backslash. For instance, to search for (1+1)=2, you would need to write your query as \
\((1+1)\)=2.\n
- The reserved characters are:
  - `+ - = && || > < ! ( ) { } [ ] ^ " ~ * ? : \ /`
- Failing to escape these special characters correctly could lead to a syntax error which prevents your query from running.

**STOPWORDS**

- Stopwords are very commonly occurring words in a particular language, that are often removed from search indexes during processing to produce better results and performance.
- English stopword list in AtoM:
  - a, an, and, are, as, at, be, but, by, for, if, in, into, is, it, no, not, of, on, or, such, that, the, their, then, there, these, they, this, to, was, will, with
The Advanced Search Panel

Narrow your results by:

- **Language**
  - Unique records: 123
  - English: 123
  - French: 3

- **Archival Institution**
  - All: 45
  - Art Gallery of Ontario: 45
  - Western Lutheran University Archives: 40
  - City of Greater Sudbury Archives: 18
  - Women's College Historical Association: 10
  - Burlington Historical Society: 6
  - Yukon Archives: 3
  - University of British Columbia Archives: 1

- **Creator**
  - King, A. Richard, 1919-2005: 1
  - Brew, Elaine: 1
  - Yukon Dog Musher's Association: 1
  - Mount, Alain: 1
  - Munro, W. A.: 1
  - Warren, Andy: 1
  - Mink, David: 1
  - Brownstone, S. Shelly: 1
  - Vefa, Florence: 1
  - Gallery 44 Centre for Contemporary Art: 1

- **Name**
- **Place**
- **Subject**
- **Genre**
- **Level of Description**

Showing 123 results

Archival description

Find results with:
- Search
- Add new criteria

Limit results to:
- Repository
- Top-level description

Filter results by:
- Level of description
- Digital object available
- Finding aid
- Copyright status
- General material designation

Filter by date range:
- Start
- End
- Overlapping
- Exact
Query builder

- Provides a graphical interface for building BOOLEAN queries
- Supports wildcards and other special search characters
- Search parameters can be limited to specific fields:
  - Any field
  - Title
  - Archival history
  - Scope and content
  - Extent and medium / Physical description
  - Subject / Place / Genre / Name access points
  - Identifier / Reference code
  - Digital object transcript text
  - Finding aid text or any field except finding aid text
  - Creator (coming in 2.5)
To prevent conflicting results, you can only use one of these two limiting options at a time.

**By repository**
- Drop-down menu
- Will not display if "multiple repositories" setting is disabled
- Will return only matching results linked to the selected repository

**By top-level desc.**
- Autocomplete field
- Begin typing to see matching results
- Will return only matching results within the selected archival unit
**Level of description**
- Limit results to a particular LOD
- Populated via the Levels of description taxonomy
- When new terms are added to the taxonomy, they become available here

**Digital object available**
- Limit results to those with (or without) digital objects
- Can be used in conjunction with the Media types facet filter for specific types of digital object (audio, video, text, etc.)

**Finding aid**
- Limit results to those with or without finding aid
- Can limit to those with a particular type of finding aid (generated from description or manually uploaded)
Filters

- Limit results to those with a particular copyright status
- Depends on presence of PREMIS rights statements
- Can be hidden!

- Limit results to those with a particular General material designation (aka material type)
- Depends on use of Canadian RAD description template
- Can be hidden!

- Limit results to just top-level descriptions
- On by default when arriving on descriptions browse page
- Off by default when entering query into global search box

Go to Admin > Settings > Default page elements to hide these filters if not in use
Date range searching

- Uses the internal description Start and End dates for matching
- Will NOT search against the Display date (what the public sees on the description view page)
- Expects dates to be in YYYY-MM-DD format (ISO 8601)
  - Will add -01-01 to YYYY start dates and -12-31 to YYYY end date parameters
  - Will not accept YYYY-MM date parameters
- Can enter text directly into field, or use the graphical date-picker widget
Two date range search options:

- Exact
- Overlapping

**Exact** means the start and end dates of descriptions returned must fall entirely within the date range entered.

**Overlapping** means that any description whose start or end dates touch or overlap the target date range will be returned.
Query range: 1950 start date

- 1940-01-01 - 1949-12-31
- 1950 - 1958
- 1952 - 1968
- 1966 - 1970-12-31
- 1945 - 1955
- 1940-01-01 - 1980-12-31
- 1955 start (no end date)
- 1969 - 1980
- 1971 - 1981
- 1955 end (no start date)
- 1945 end (no start date)
- 1971 start (no end date)
Query range: 1950 start date

- 1940-01-01 - 1949-12-31
- 1950
- 1950 - 1958
- 1952 - 1968
- 1966 - 1970-12-31
- 1940-01-01 - 1980-12-31
- 1955 start (no end date)
- 1955 end (no start date)
- 1959 - 1980
- 1971 - 1981
- 1945 end (no start date)
- 1971 start (no end date)
Boolean operators

• Use **AND** to combine search terms to narrow your search:
  • e.g. search `vancouver AND "city hall"` to return only descriptions that contain both search terms

• Use **OR** to combine search terms to broaden your search:
  • e.g. search `vancouver OR "city hall"` to return descriptions that contain either search term

• Use **AND NOT** to exclude terms:
  • e.g. search `vancouver AND NOT "city hall"` to return descriptions that include vancouver but do not contain the phrase “city hall”

AND, OR, AND NOT

• Can be used in the global search box and the advanced search panel
• Can be combined with other filters and facets to further refine results
### Boolean operators

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>Term after + <em>must</em> be in the result. Example: <strong>+tea cricket</strong> requires that results that must contain the term tea in them, and may have the term cricket.</td>
</tr>
<tr>
<td>-</td>
<td>Term after - <em>must not</em> be in the result. Example: <strong>-tea cricket</strong> requires that results that must not contain the term tea in them, and may have the term cricket.</td>
</tr>
<tr>
<td>?</td>
<td>Single character wildcard. Example: <strong>p?per</strong> will find paper and piper, but not pepper.</td>
</tr>
<tr>
<td>*</td>
<td>Multiple character wildcard. Example: *<em>galax</em> will find galaxy and galaxies, but not galactic.</td>
</tr>
<tr>
<td>~</td>
<td>Fuzzy search. Will return results with words similar to the term. Example: <strong>fjord~</strong> will find fjord, fjords, ford, form, fonds, etc.</td>
</tr>
<tr>
<td>&amp;&amp;</td>
<td>Boolean operator. Can be used in place of AND. Will cause an error if combined with spelled-out operators. Example: <strong>Arthur &amp;&amp; Ford AND Zaphod</strong> will fail; <strong>Arthur &amp;&amp; Ford &amp;&amp; Zaphod</strong> will succeed.</td>
</tr>
<tr>
<td>!</td>
<td>Boolean operator. Can be used in place of NOT. Will cause an error if combined with spelled-out operators.</td>
</tr>
</tbody>
</table>
# Boolean operators

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot; &quot;</td>
<td>Term enclosed in quotes must appear exactly as provided. Example: &quot;towel&quot; will find towel, but not towels. Used for phrase searching as well – a search for city hall will return any records that contain city AND hall anywhere, while a search for &quot;city hall&quot; will only return records where the terms appear together in the order of the query.</td>
</tr>
<tr>
<td>^</td>
<td>Boost relevance. Multiplies the relevance of the preceding term by the number following the symbol, affecting the sorting of the search results. Example: paranoid android(^5) gives results containing the term &quot;android&quot; 5x the relevance as results containing only the word &quot;paranoid&quot;, and will sort them closer to the start of the search results.</td>
</tr>
<tr>
<td>\</td>
<td>Escapes the immediately following character, so that it is treated as text, rather than as a special character. For example, to search for &quot;(1+1):2&quot;, use the following: ((1+1))</td>
</tr>
<tr>
<td>( )</td>
<td>Used to group search clauses. This can be useful if you want to control the precedence of boolean operators for a query, e.g. (coffee NOT tea)OR cream will return different results than coffee NOT(tea OR cream). Without grouping, by default in Elasticsearch, NOT takes precedence over AND, which takes precedence over OR.</td>
</tr>
</tbody>
</table>
Expert searching with Elasticsearch
What is Elasticsearch?

Elasticsearch is an open-source, broadly-distributable, readily-scalable, enterprise-grade search engine. Accessible through an extensive and elaborate API, Elasticsearch can power extremely fast searches that support your data discovery applications.

– John Vanderzyden, Qbox blog - https://qbox.io/blog/what-is-elasticsearch

- Search index
- Use to support browsing and searching in AtoM
- Version used in AtoM 2.4: 1.7
- Version used in AtoM 2.5: 5.5

https://www.elastic.co/products/elasticsearch
Expert searching – using ES fields

- Fields are organized in the index hierarchically
  - Nested by entity type, culture, field type, etc.
- Example: AtoM’s alternative identifiers consist of 2 fields - a label and an identifier value. In the ES index, these are expressed as:
  - `alternativeIdentifiers.identifier`
  - `alternativeIdentifiers.label`
- Field names are written in camelCase
- Nesting is represented by a period between elements

Example format for searching:
- `field.name:searchterm`
- `field.name:"search term"`

Examples
- Search for records with an alternative identifier value of “123”:
  - `alternativeIdentifiers.identifier:123`
- Search for records with an alternative identifier label of “Legacy call number”:
  - `alternativeIdentifiers.label:"Legacy call number"`
Expert searching – using ES fields

Because AtoM is a multilingual application, some translatable fields are organized together under the common prefix of “i18n”

- i18n is a common abbreviation in multilingual application development, short for internationalization.
- When searching these fields, the culture of your search must be entered as part of the field name, because nested within each i18n element of the index there will be a different term for each language.

Examples

- To search your English data for “linear feet” in the Extent and medium field, you would enter the search term as follows:
  - `i18n.en.extentAndMedium: “linear feet”`
- To search your French data for “albums de coupures” in the Scope and content field, you would enter the search term as follows:
  - `i18n.fr.scopeAndContent: “albums de coupures”`

List of languages supported in AtoM, and their corresponding language codes:

AtoM ES fields: archival descriptions

- alternativeIdentifiers.identifier
- alternativeIdentifiers.label
- createdAt
- digitalObject.filename
- digitalObject.mediaTypeId
- digitalObject.thumbnailPath
- findingAid.status
- findingAid.transcript
- hasDigitalObject
- i18n.%LANG%.accessConditions
- i18n.%LANG%.accruals
- i18n.%LANG%.acquisition
- i18n.%LANG%.alternateTitle
- i18n.%LANG%.appraisal
- i18n.%LANG%.archivalHistory
- i18n.%LANG%.arrangement

Replace the placeholder `%LANG%` with the 2-letter culture code of the language in which you want to search.
AtoM ES fields: archival descriptions

- i18n.%LANG%.edition
- i18n.%LANG%.extentAndMedium
- i18n.%LANG%.findingAids
- i18n.%LANG%.institutionResponsibleIdentifier
- i18n.%LANG%.extentAndMedium
- i18n.%LANG%.findingAids
- i18n.%LANG%.institutionResponsibleIdentifier
- i18n.%LANG%.locationOfCopies
- i18n.%LANG%.locationOfOriginals
- i18n.%LANG%.physicalCharacteristics
- i18n.%LANG%.relatedUnitsOfDescription
- i18n.%LANG%.reproductionConditions
- i18n.%LANG%.revisionHistory
- i18n.%LANG%.rules
- i18n.%LANG%.scopeAndContent
- i18n.%LANG%.sources

Returns all occurrences of **scrapbook** (i.e. scrapbooks, scrapbooking, etc.) in scope and content in English descriptions.
AtoM ES fields: archival descriptions

- i18n.%LANG%.sources
- i18n.%LANG%.title
- identifier
- levelOfDescriptionId
- materialTypeId
- publicationStatusId
- referenceCode
- referenceCodeWithoutCountryAndRepo
- slug
- sourceCulture
- transcript
- updatedAt
- generalNotes.i18n.%LANG%.content
- alphaNumericNotes.i18n.%LANG%.content
- conservationNotes.i18n.%LANG%.content
- physicalDescriptionNotes.i18n.%LANG%.content
AtoM ES fields: archival descriptions

Linked entities

- creators.i18n.%LANG%.authorizedFormOfName
- creators.otherNames.i18n.%LANG%.name
- creators.parallelNames.i18n.%LANG%.name
- creators.standardizedNames.i18n.%LANG%.name
- genres.i18n.%LANG%.name
- names.i18n.%LANG%.authorizedFormOfName
- names.otherNames.i18n.%LANG%.name
- names.parallelNames.i18n.%LANG%.name
- names.standardizedNames.i18n.%LANG%.name
- places.i18n.%LANG%.name
- repository.i18n.%LANG%.authorizedFormOfName
- subjects.i18n.%LANG%.name

Return all English descriptions linked to a subject term “Arts and culture”
Searching for empty or populated fields

**EMPTY / MISSING FIELDS**

- Syntax:
  - `!_exists_:field.name`

**Examples**

Search for English descriptions with no scope and content:
- `!_exists_: i18n.en.scopeAndContent`

Search for French descriptions with no extent or medium statement:
- `!_exists_: i18n.fr.extentAndMedium`

**POPULATED FIELDS**

- Syntax:
  - `_exists_:field.name`

**Examples**

Search for English descriptions with a subject access point:
- `_exists_:subjects.i18n.en.name`

Search for Spanish descriptions with a statement on Sources in the Description control area:
- `_exists_:i18n.es.sources`
Publication status and modification dates

**PUBLICATION STATUS**

- Field name:
  - `publicationStatusId`

Use **159** for draft; **160** for published descriptions

**Examples**

Return all draft descriptions:
- `publicationStatusId:159`

Return all published descriptions:
- `publicationStatusId:160`

**CREATION AND MODIFICATION**

- Field names:
  - `createdAt`
  - `updatedAt`

**Examples**

Return all records last modified between Jan 1 2010 and Dec 31 2012:
- `updatedAt:[2010-01-01 TO 2012-12-31]`

Return all records created after January 1, 2012:
- `createdAt:[2012-01-01 TO *]`
See the docs for more tips

Search
- https://www.accesstomemory.org/docs/latest/user-manual/access-content/advanced-search/

Advanced search
- https://www.accesstomemory.org/docs/latest/user-manual/access-content/advanced-search/

Browse
- https://www.accesstomemory.org/docs/latest/user-manual/access-content/browse/
Documentation links for ES fields

**Descriptions**

**Authority records**
- [https://www.accesstomemory.org/docs/latest/user-manual/access-content/search-atom/#elasticsearch-authority-record-fields](https://www.accesstomemory.org/docs/latest/user-manual/access-content/search-atom/#elasticsearch-authority-record-fields)

**Accessions**
- [https://www.accesstomemory.org/docs/latest/user-manual/access-content/search-atom/#elasticsearch-accession-record-fields](https://www.accesstomemory.org/docs/latest/user-manual/access-content/search-atom/#elasticsearch-accession-record-fields)

**Archival institutions**
- [https://www.accesstomemory.org/docs/latest/user-manual/access-content/search-atom/#elasticsearch-archival-institution-record-fields](https://www.accesstomemory.org/docs/latest/user-manual/access-content/search-atom/#elasticsearch-archival-institution-record-fields)

**Terms**
Questions?

info@artefactual.com