
JSON MDQ Server Documentation

Release 1.1.1

Leif Johansson

Nov 18, 2020

Contents:

1	Installing and Running thiss-mdq	3
2	API	5
2.1	MDQ	5
2.2	Search	5
2.3	WebFinger	5
2.4	Monitoring	6
3	Release Notes	7
3.1	Version 1.1.4	7
3.2	Version 1.2.0	7
3.3	Version 1.2.1	7
3.4	Version 1.2.2	7
3.5	Version 1.2.3	7
3.6	Version 1.2.4	7
3.7	Version 1.2.5	8
3.8	Version 1.2.6	8
3.9	Version 1.3.0	8
3.10	Version 1.3.1	8
3.11	Version 1.3.2	8
4	Indices and tables	9

thiss-mdq is a minimal implementation of the metadata query protocol (MDQ) that only supports JSON data. Metadata in discojson format is loaded from a JSON-file referenced by \$METADATA. The file is watched for updates and is reloaded automatically. A server is started on \$HOST (0.0.0.0) and \$PORT (3000). Run it using “npm start” or the supplied Docker container.

Installing and Running thiss-mdq

The recommended way to install thiss-mdq is via docker using the supplied Dockerfile:

```
# docker build -t thiss-mdq:latest
```

The docker-image can be run thus:

```
# docker run -d -p 3000:3000 -v /some/where/metadata.json:/etc/metadata.json thiss-  
↪mdq:latest
```

Verify that the container has started by cURL:ing:

```
# curl -s http://localhost:3000/
```

The result should include information about the version of thiss-mdq and the number of entities in the metadata set.

There are two backend options: the default lunr-based backend and the redis backend. Setting the environment variable INDEXER to 'redis' turns on the latter.

The thiss-mdq server implements the metadata query protocol (MDQ) along with a few extensions.

2.1 MDQ

The following endpoints are specified by MDQ (using GET requests)

- `/entities/` - returns all JSON metadata in a single array.
- `/entities/{sha1}<sha1 hash of entityID>` - returns a single entity in JSON format

The MDQ protocol specifies standard content negotiation but currently the Accept header is ignored and `application/json` is always returned. This may change in the future and clients should always set the Accept header to include `application/json`. An optional “.json” efile extension in the URL is ignored.

In addition to these endpoints thiss-mdq supports a couple of extensions:

- Search
- WebFinger
- Monitoring

2.2 Search

Issue a GET request to `/entities/?q=<search>`. The result is a (possibly empty) JSON list of entities matching the search query.

2.3 WebFinger

Issue a GET request to `/.well-known/webfinger`. The result is a list of all URIs (all the entities) included in the indexed metadata. This can be used together with curl/wget to mirror the `_static_` contents of the mdq instance. It

is of course not possible to mirror all possible searches this way.

2.4 Monitoring

Issue a GET or HEAD request to `/status`. The result is a 200 if the metadata index is alive an well and 500 otherwise.

3.1 Version 1.1.4

- escape special characters in plain text queries

3.2 Version 1.2.0

- pluggable indexing and redis-support from Hannah Sebuliba

3.3 Version 1.2.1

- bugfixes

3.4 Version 1.2.2

- updates to the search algorithm to favor more exact searches over less

3.5 Version 1.2.3

- Integrate stopwords module instead of locally maintained stopwords

3.6 Version 1.2.4

- Enforce AND-logic for combining multiple search terms

3.7 Version 1.2.5

- Never treat the first token as a stopword

3.8 Version 1.2.6

- Reduce index memory footprint
- Use cluster controller for expressjs

3.9 Version 1.3.0

- Discontinue support for automatic restarts - rely on external restart
- Change tokenizer to split tokens on whitespace only
- Improved i18n-support
- Provide information about metadata mtime in monitor json

3.10 Version 1.3.1

- Updated and extended status json information: mtime and ctime for the metadata file is reported separately

3.11 Version 1.3.2

- Avoid including user-supplied data in header values

CHAPTER 4

Indices and tables

- `genindex`
- `search`