

Full text search HOWTO



This feature is available since midPoint 3.6.

Background information

MidPoint contains a simple full text search feature. Selected items can be indexed and then queries can be put against their content. More specifically,

1. Any textual properties of midPoint objects can be designated as "full text search indexed" in system configuration. These can be either standard properties (e.g. user's given and family names) or extension properties.
2. When updating these objects, midPoint extracts content of these properties, normalizes it, and stores in a separate `m_object_text_info` table. Normalization consists of applying default PolyString normalizer that removes diacritics marks and converts all characters to lowercase. Also, during normalization, duplicate words are eliminated, in order to optimize space and processing speed.
3. The normalized indexed content is available for querying. A user can specify a word (or a part of it) and midPoint will look up all objects that contain that word (or its part) in the indexed content.

Current limitations:

1. Because conventional database mechanisms (standard indices) are used, this feature is applicable to small or medium sized installations; we estimate the threshold to be few thousands or tens of thousands records. (In the future we expect to apply specialized approaches to full-text search to provide wider applicability.)
2. Because of the implementation chosen, it is not possible to search for a phrase. Since 3.6.1 if multiple words are entered, midPoint looks for objects that contain all of the specified words (possibly in different parts of the object). In "plain" 3.6, the result of multiple words searches is undefined and can be falsely negative.

Setting it up

By default, the full-text search feature is disabled. How to enable it:

1. Provide a list of items to be indexed. This is done in system configuration object, for example by including the following configuration:

Sample configuration for indexing objects' text content

```
<fullTextSearch>
  <indexed> <!-- no "objectType" element means this applies to all objects -->
    <item>name</item>
    <item>description</item>
  </indexed>
  <indexed>
    <objectType>UserType</objectType>
    <item>fullName</item>
    <item>givenName</item>
    <item>familyName</item>
    <item>additionalName</item>
    <item>nickName</item>
    <item>emailAddress</item>
    <item>telephoneNumber</item>
    <item>honorificPrefix</item>
    <item>honorificSuffix</item>
    <item>title</item>
    <item>employeeNumber</item>
    <item>employeeType</item>
    <item>costCenter</item>
    <item>organization</item>
    <item>organizationalUnit</item>
    <item>locality</item>
  </indexed>
  <indexed>
    <objectType>AbstractRoleType</objectType> <!-- use AbstractRoleType for self-service with Role
catalog ! -->
    <item>displayName</item>
    <item>identifier</item>
  </indexed>
  <indexed>
    <objectType>RoleType</objectType> <!-- there can be more "objectType" elements for a single entry
(not shown here) -->
    <item>roleType</item>
  </indexed>
  <indexed>
    <objectType>OrgType</objectType>
    <item>orgType</item>
    <item>costCenter</item>
    <item>locality</item>
    <item>mailDomain</item>
  </indexed>
  <indexed>
    <objectType>ServiceType</objectType>
    <item>serviceType</item>
    <item>locality</item>
  </indexed>
</fullTextSearch>
```

2. Re-index existing objects. This can be done by *About -> Reindex repository objects* action.

midPoint ≡ About this system Configuration About this system administrator

SELF SERVICE

- Home
- Profile
- Credentials
- Request a role

ADMINISTRATION

- Dashboard
- Users
- Org. structure
- Roles
- Services
- Resources
- Work items
- Certification
- Server tasks
- Reports
- Configuration
 - Bulk actions
 - Import object
 - Repository objects
 - System
 - Notifications
 - Logging
 - Profiling
 - Admin GUI
 - Shadows details
 - Internals configuration
 - Query playground

Basic

Version	3.6-SNAPSHOT
Git describe	S[describe]
Build at	Pi, 24 feb 2017 16:40:55 +0000

Repository

Implementation name	SQL
Implementation description	Implementation that stores data in generic relational (SQL) databases. It is using ORM (hibernate) on top of JDBC to access the database.
Is embedded	true
Driver short name	org.h2.Driver
Driver version	1.3
Repository URL	jdbc:h2tcp://127.0.0.1:5437/midpoint;DB_CLOSE_ON_EXIT=FALSE;LOCK_MODE=1;LOCK_TIMEOUT=10000;MAX_LENGTH_INPLACE_LOB=10240

Additional details

dataSource	
hibernateDialect	org.hibernate.dialect.H2Dialect
hibernateHbm2ddl	update
transactionIsolation	TRANSACTION_SERIALIZABLE (read from connection)

Provisioning

Connid framework version	1.4.2.18
--------------------------	----------

System properties

file.separator	\
java.class.path	C:\midpoint\apache-tomcat-8.0.30\bin\bootstrap.jar;C:\midpoint\apache-tomcat-8.0.30\bin\tomcat-juli.jar
java.home	C:\Java\jdk1.8.0_102\jre
java.vendor	Oracle Corporation
java.vendor.url	http://java.oracle.com/
java.version	1.8.0_102
line.separator	
os.arch	amd64
os.name	Windows 10
os.version	10.0
path.separator	;
user.dir	C:\midpoint\apache-tomcat-8.0.30\bin
user.home	C:\Users\Pavol
user.name	Pavol

JVM properties

```
-Dmidpoint.home=C:\midpoint\home\scratch
-Dcom.sun.management.jmxremote=
-Dcom.sun.management.jmxremote.port=1099
-Dcom.sun.management.jmxremote.ssl=false
-Dcom.sun.management.jmxremote.authenticate=false
-Djava.rmi.server.hostname=127.0.0.1
-Djava.util.logging.config.file=C:\Users\Pavol\IntelliJidea2016.3\system\tomcat\Unnamed_midpoint_3\conf\logging.properties
-Djava.util.logging.manager=org.apache.juli.ClassLoaderLogManager
-Djava.endorsed.dirs=C:\midpoint\apache-tomcat-8.0.30\endorsed
-Dcatalina.base=C:\Users\Pavol\IntelliJidea2016.3\system\tomcat\Unnamed_midpoint_3
-Dcatalina.home=C:\midpoint\apache-tomcat-8.0.30
-Djava.io.tmpdir=C:\midpoint\apache-tomcat-8.0.30\temp
```

Repository self test | Check and repair org closure consistency | **Reindex repository objects** | Provisioning self test

Non-subscription.

The `m_object_text_info` table content will then look like this:

```
OWNER_OID | SYSTEM_TEXT
00000000-0000-0000-0000-000000000001 | systemconfiguration
00000000-0000-0000-0000-000000000002 | administrator midpoint
00000000-0000-0000-0000-000000000003 | default password policy
00000000-0000-0000-0000-000000000004 | role that gives user full authorization in midpoint system superuser
00000000-0000-0000-0000-000000000005 | a system task that repeatedly checks for outdated objects eg closed tasks old audit records and removes them Cleanup
00000000-0000-0000-0000-000000000006 | a system task that maintains effective status of objects based on their validity time constraints scanner
00000000-0000-0000-0000-000000000007 | a system task that executes triggers attached to objects trigger scanner
00000000-0000-0000-0000-000000000008 | role authorizing end users to log in change their passwords and review assigned accounts system user
00000000-0000-0000-0000-000000000009 | report made from audit records logs
00000000-0000-0000-0000-00000000000a | role authorizing users to make approval decisions on work items system approver
00000000-0000-0000-0000-00000000000b | role authorizing users to make decisions on certification cases system reviewer
00000000-0000-0000-0000-00000000000c | role authorizing users to delegate their own privileges any other user delegator
00000000-0000-0000-0000-000000000010 | reconciliation report for selected resource
00000000-0000-0000-0000-000000000011 | users listed in midpoint
00000000-0000-0000-0000-000000000012 | default security policy
00000000-0000-0000-0000-000000000013 | all certification definitions with basic information on related campaigns report
00000000-0000-0000-0000-000000000014 | all certification campaigns along with their state report
00000000-0000-0000-0000-000000000015 | cases within a given certification campaign report
00000000-0000-0000-0000-000000000016 | decisions of individual reviewers for a certification campaign as whole or given stage report
00000000-0000-0000-0000-000000000020 | languages
00000000-0000-0000-0000-000000000021 | locales
00000000-0000-0000-0000-000000000022 | timezones
00000000-0000-0000-0000-000000000023 | lifecycle states
0c36e274-9475-4879-b440-d0dbf809ae3a | reindex repository objects
6420a8a1-9fc0-45a6-a1a9-0da4f45b24a37 | defaultnode
66103c7a-5a62-4bab-b7f7-f5090b47863a | icf orgidentityconnectorsdatabasetabledatabasetableconnector v1420
8288190b-8437-48b8-92f1-7a5844fcd00b | icf comevolveumpolygonconnectorldapedirectorydirectoryldapconnector v144snapshot
a1ef0f91-91c2-4b55-8e7c-cb5301ff36f0 | icf comevolveumpolygonconnectorldapedadidapconnector v144snapshot
eaf25ab5-db39-40c7-bb0e-d7df0c198804 | icf comevolveumpolygonconnectorldapdapconnector v144snapshot
f040fd80-3851-42b5-90d4-d13d90cf931b | icf comevolveumpolygonconnectorldapdapconnector v1420
```

Using it

If the full text search feature is enabled in the system configuration, it can be used from the GUI.

- SELF SERVICE
 - Home
 - Profile
 - Credentials
 - Request a role
- ADMINISTRATION
 - Dashboard
 - Users
 - Org. structure
 - Roles
 - Services
 - Resources
 - Work items
 - Certification
 - Server tasks

All roles view

Request for me

Full text search



Basic

Advanced



Approver



Details

Add to cart

Delegator



Details

Add to cart

End user



Details

Add to cart

Reviewer



Details

Add to cart

Superuser



Details

Add to cart