

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Отчет по лабораторной работе №2
по курсу «Распределенные системы»

Выполнил: студент 1-го курса магистратуры, гр. 21224

Гафиятуллин А.Р

Новосибирск, 2022

1. ОПРЕДЕЛЕНИЯ

RUSMARC — российский коммуникативный формат представления библиографических записей в машиночитаемой форме.

Zebra — свободный программный пакет для управления информацией.

Prefix Query Format — запрос в префиксном формате.

Contextual Query Language — формальный язык для представления запросов к системам информационного поиска, таких как поисковые системы, библиографические каталоги и музейные информационные коллекции.

2. СОКРАЩЕНИЯ

Prefix Query Format — **PQF**

Contextual Query Language — **CQL**

3. ОБЪЕКТ ИССЛЕДОВАНИЯ

Zebra — свободный программный пакет для управления информацией.

4. ЦЕЛИ РАБОТЫ

Создать репозиторий для поиска и извлечения информации на основе языков запросов PQF и CQL.

5. МЕТОДОЛОГИЯ/ПОРЯДОК ПРОВЕДЕНИЯ РАБОТ

5.1. Развернута виртуальная машина.

5.2. Установлен программный пакет **zoopark-ds-1.0.2**:

```
libuser@ds102:~/tmp/zoopark-ds-1.0.2 % make install
Making install in src
../config/install-sh -c -d '/home/libuser/local/bin'
/bin/sh ../libtool --mode=install /usr/bin/install -c zebra-client zebra-worker zebrash '/home/libuser/local/bin'
libtool: install: /usr/bin/install -c zebra-client /home/libuser/local/bin/zebra-client
libtool: install: /usr/bin/install -c zebra-worker /home/libuser/local/bin/zebra-worker
libtool: install: /usr/bin/install -c zebrash /home/libuser/local/bin/zebrash
Making install in include
Making install in cuba
../config/install-sh -c -d '/home/libuser/local/include/zoopark-ds'
/usr/bin/install -c -m 644 ctrl.h funupdate.h param.h version.h worker.h '/home/libuser/local/include/zoopark-ds'
Making install in etc
../config/install-sh -c -d '/home/libuser/local/etc/zoopark-ds'
/usr/bin/install -c -m 644 rc.conf.dist zebra-worker-start.sh zebra-worker-stop.sh yazgfs.xml pqf.properties '/home/libuser/local/etc/zoopark-ds'
Making install in profiles
Making install in dom
Making install in tab
../config/install-sh -c -d '/home/libuser/local/share/zoopark-ds/profiles/dom/tab'
/usr/bin/install -c -m 644 string.chr urx.chr dom.xml extract.xml input.xml store2dc.xml '/home/libuser/local/share/zoopark-ds/profiles/dom/tab'
../config/install-sh -c -d '/home/libuser/local/share/zoopark-ds/profiles/dom'
/usr/bin/install -c -m 644 zebra.cfg zebra.json '/home/libuser/local/share/zoopark-ds/profiles/dom'
Making install in tests
Making install in dom
libuser@ds102:~/tmp/zoopark-ds-1.0.2 % ls -
distfiles      local          tmp
libuser@ds102:~/tmp/zoopark-ds-1.0.2 %
```

5.3. Созданы репозиторий и база данных:

```
libuser@ds102:~/wdir % $HOME/local/bin/zebra-client -w . commit -r r1
20:39:42-26/03 [log] zebra_start 2.0.55 bd2bc9360225e695bbaba2c2d1cd6925c4eb23a5
20:39:42-26/03 [log] config ./r1/zebra.cfg
20:39:42-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-alvis.so
20:39:42-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-dom.so
20:39:42-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-marc.so
20:39:42-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-regex.so
20:39:42-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-xml.so
20:39:42-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-safari.so
20:39:42-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-text.so
20:39:42-26/03 [log] enabling shadow spec=./r1/shadow:10G
20:39:42-26/03 [log] cache_fname = ./r1/shadow/cache
20:39:42-26/03 [log] zebra_stop: 0.00 0.01 0.00
{"success": true}
libuser@ds102:~/wdir %
```

5.4. Выполнено индексирование записей из тестового набора данных:

```
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 59438
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 60062
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 60585
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 61109
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 62435
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 62949
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 63572
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 64201
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 64832
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 65801
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 66519
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 67355
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 68006
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 68692
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 69341
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 70030
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 70676
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 71302
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 71929
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 72556
20:40:46-26/03 [log] add dom.dom.xml ./simple-rusmarc-utf8 73198
20:40:46-26/03 [log] Iterations: isam/dict 3523/1137
20:40:46-26/03 [log] Dict: inserts/updates/deletions: 1136/1/0
20:40:46-26/03 [log] Records: 72 i/u/d 72/0/0
20:40:46-26/03 [log] enabling shadow spec=./r1/shadow:10G
20:40:46-26/03 [log] cache_fname = ./r1/shadow/cache
20:40:46-26/03 [log] zebra_stop: 0.14 0.03 0.06
{"success": true}
libuser@ds102:~/wdir % |
```

5.5. Задача search с использованием языка PQF:


```
libuser@ds102:~/wdir % $HOME/local/bin/zebra-client -w . scan -r r1 -d db1 -t pqf -q '@1=4 Билл'
20:45:03-26/03 [log] Command SCAN:
20:45:03-26/03 [log] repo='r1'
20:45:03-26/03 [log] databasename='db1'
20:45:03-26/03 [log] type='pqf'
20:45:03-26/03 [log] term='@1=4 Билл'
20:45:03-26/03 [log] position=-1
20:45:03-26/03 [log] number=-1
20:45:03-26/03 [log] wdir='.'
20:45:03-26/03 [log] pdir='(null)'
20:45:03-26/03 [log] cql2pqf='(null)'
20:45:03-26/03 [log] zebra_start 2.0.55 bd2bc9360225e695bbaba2c2d1cd6925c4eb23a5
20:45:03-26/03 [log] config ./r1/zebra.cfg
20:45:03-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-alvis.so
20:45:03-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-dom.so
20:45:03-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-marc.so
20:45:03-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-regx.so
20:45:03-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-xml.so
20:45:03-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-safari.so
20:45:03-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-text.so
20:45:03-26/03 [log] dom.xml dom filter: loading config file ./r1/tab/dom.xml
20:45:03-26/03 [log] zebra_stop: 0.01 0.00 0.01
{"success": true, "data": {"terms": [{"numberOfRecords": 2, "value": "Билл", "displayTerm": "Билл"}, {"numberOfRecords": 1, "value": "болезни", "displayTerm": "болезни"}, {"numberOfRecords": 1, "value": "больных", "displayTerm": "больных"}, {"numberOfRecords": 1, "value": "буквы", "displayTerm": "буквы"}, {"numberOfRecords": 1, "value": "бухгалтерский", "displayTerm": "бухгалтерский"}, {"numberOfRecords": 1, "value": "бухгалтерскому", "displayTerm": "бухгалтерскому"}, {"numberOfRecords": 7, "value": "в", "displayTerm": "в"}, {"numberOfRecords": 1, "value": "вдохновения", "displayTerm": "вдохновения"}, {"numberOfRecords": 1, "value": "век", "displayTerm": "век"}, {"numberOfRecords": 1, "value": "великий", "displayTerm": "великий"}]}}
```

5.8. Задача scan с использованием языка CQL:

```
libuser@ds102:~/wdir % $HOME/local/bin/zebra-client -w . -c ./pqf.properties scan -r r1 -d db1 -t cql -q "dc.title=Билл"
20:45:39-26/03 [log] Command SCAN:
20:45:39-26/03 [log] repo='r1'
20:45:39-26/03 [log] databasename='db1'
20:45:39-26/03 [log] type='cql'
20:45:39-26/03 [log] term='dc.title=Билл'
20:45:39-26/03 [log] position=-1
20:45:39-26/03 [log] number=-1
20:45:39-26/03 [log] wdir='.'
20:45:39-26/03 [log] pdir='(null)'
20:45:39-26/03 [log] cql2pqf='./pqf.properties'
20:45:39-26/03 [log] CQL transform result: ' @attr 6=1 @attr 7=3 @attr 4=1 @attr 3=3 @attr 6=1 @attr 5=100 @attr 1=4 "Билл" '
20:45:39-26/03 [log] zebra_start 2.0.55 bd2bc9360225e695bbaba2c2d1cd6925c4eb23a5
20:45:39-26/03 [log] config ./r1/zebra.cfg
20:45:39-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-alvis.so
20:45:39-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-dom.so
20:45:39-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-marc.so
20:45:39-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-regx.so
20:45:39-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-grs-xml.so
20:45:39-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-safari.so
20:45:39-26/03 [log] Loaded filter module /usr/local/lib/idzebra-2.0/modules/mod-text.so
20:45:39-26/03 [log] dom.xml dom filter: loading config file ./r1/tab/dom.xml
20:45:39-26/03 [log] zebra_stop: 0.01 0.01 0.00
{"success": true, "data": {"terms": [{"numberOfRecords": 2, "value": "Билл", "displayTerm": "Билл"}, {"numberOfRecords": 1, "value": "болезни", "displayTerm": "болезни"}, {"numberOfRecords": 1, "value": "больных", "displayTerm": "больных"}, {"numberOfRecords": 1, "value": "буквы", "displayTerm": "буквы"}, {"numberOfRecords": 1, "value": "бухгалтерский", "displayTerm": "бухгалтерский"}, {"numberOfRecords": 1, "value": "бухгалтерскому", "displayTerm": "бухгалтерскому"}, {"numberOfRecords": 7, "value": "в", "displayTerm": "в"}, {"numberOfRecords": 1, "value": "вдохновения", "displayTerm": "вдохновения"}, {"numberOfRecords": 1, "value": "век", "displayTerm": "век"}, {"numberOfRecords": 1, "value": "великий", "displayTerm": "великий"}]}}
```

5.9. Удаления репозитория:

```
libuser@ds102:~/wdir % $HOME/local/bin/zebra-client -w . destroy -r r1
20:46:40-26/03 [log] Command DESTROY:
20:46:40-26/03 [log] repo='r1'
20:46:40-26/03 [log] wdir='.'
20:46:40-26/03 [log] pdir='(null)'
20:46:40-26/03 [log] cql2pqf='(null)'
{"success": true}
libuser@ds102:~/wdir % |
```

6. ОПИСАНИЕ РЕЗУЛЬТАТА

На основе программного продукта zoopark-ds-1.0.2 был создан репозиторий для поиска и извлечения информации из тестового набора

данных с использованием языков запросов PQF и CQL для задач scan и search.

7. ССЫЛКИ

Prefix Query Format (PQF) — URL:

<https://software.indexdata.com/yaz/doc/tools.html#PQF> (Дата обращения: 26.03.2022)

CQL — URL: <https://www.loc.gov/standards/sru/cql/index.html> (Дата обращения: 26.03.2022)

Формат RUSMARC — URL: <http://rusmarc.ru/rusmarc/format.html> (Дата обращения: 26.03.2022)

Zebra - User's Guide and Reference — URL:

<https://software.indexdata.com/zebra/doc/index.html> (Дата обращения: 26.03.2022)