Team: Group 5

jiliangn, 81224742

ycao16, 17878238

sibow2, 50617929

**Part 1**

1. The number of documents of the corpus is 37497.

2. The number of unique tokens present in the index is 577,040.

3. The total size of our index on disk is 106,488 KB, which is equal to 103MB.

|  |  |  |
| --- | --- | --- |
| Total number of documents | Total number of unique tokens | Total size (in KB) of index on disk |
| 37497 | 577040 | 106,488KB |

**Part 2**

The first 10 URLs retrieved for each of the search queries (“Informatics”, “Mondego”, “Irvine”):

**1. Informatics**

1.

DocID: 71/197

URL:

www.ics.uci.edu/alumni/mentor/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/stayconnected/stayconnected/stayconnected/stayconnected/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/stayconnected/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/inductees.aspx.php

2.

DocID: 71/192

URL:

www.ics.uci.edu/~eppstein/pubs/a-smitchell.html

3.

DocID: 22/263

URL:

www.ics.uci.edu/about/bren/index.php/bren\_vision.php/about\_contact.php/bren\_vision.php

4.

DocID: 22/261

URL:

www.ics.uci.edu/alumni/mentor/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/stayconnected/stayconnected/index.php

5.

DocID: 55/45

URL:

www.ics.uci.edu/alumni/mentor/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/stayconnected/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/stayconnected/stayconnected/index.php

6.

DocID: 22/264

URL:

www.ics.uci.edu/about/bren/index.php/bren\_vision.php/index.php/../search/index.php/../../grad/index.php/ICS/ics/about/about\_facts

7.

DocID: 42/173

URL:

sdcl.ics.uci.edu/2014/04/won-2nd-place-in-2014-annual-autism-appjam

8.

DocID: 42/172

URL:

www.ics.uci.edu/faculty/profiles/view\_faculty.php?ucinetid=gtsudik

9.

DocID: 42/175

URL:

www.ics.uci.edu/alumni/mentor/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/stayconnected/stayconnected/stayconnected/stayconnected/stayconnected/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/stayconnected/hall\_of\_fame/hall\_of\_fame/stayconnected/index.php

10.

DocID: 42/174

URL:

wics.ics.uci.edu/category/news/info-session?afg41\_page\_id=2page%2F2%2F&afg44\_page\_id=3&afg42\_page\_id=2

**2. Mondego**

1.

DocID: 38/381

URL:

mondego.ics.uci.edu/projects/clonedetection/tool/source?C=D;O=A

2.

DocID: 30/151

URL:

mondego.ics.uci.edu/projects/clonedetection/tool/latest?C=D;O=A

3.

DocID: 10/183

URL:

www.ics.uci.edu/community/news/notes/notes\_2013.php

4.

DocID: 43/436

URL:

mondego.ics.uci.edu/projects/clonedetection/tool?C=S;O=A

5.

DocID: 49/191

URL:

mondego.ics.uci.edu/projects/clonedetection/tool/latest/features?C=N;O=A

6.

DocID: 27/69

URL:

mondego.ics.uci.edu/datasets/wikipedia-events?C=D;O=A

7.

DocID: 17/28

URL:

mondego.ics.uci.edu/projects/clonedetection/tool/demo?C=S;O=A

8.

DocID: 65/378

URL:

mondego.ics.uci.edu/projects/clonedetection/tool/latest?C=N;O=D

9.

DocID: 27/335

URL:

mondego.ics.uci.edu/datasets/wikipedia-events/files

10.

DocID: 43/323

URL:

mondego.ics.uci.edu/projects/clonedetection/tool?C=D;O=A

**3. Irvine**

1.

DocID: 71/196

URL:

www.ics.uci.edu/community/events/openhouse/demos.html

2.

DocID: 71/190

URL:

www.ics.uci.edu/~jesmaeln

3.

DocID: 71/192

URL:

www.ics.uci.edu/~eppstein/pubs/a-smitchell.html

4.

DocID: 71/198

URL:

fano.ics.uci.edu/cites/Document/Fully-dynamic-biconnectivity-and-transitive-closure.html

5.

DocID: 22/263

URL:

www.ics.uci.edu/about/bren/index.php/bren\_vision.php/about\_contact.php/bren\_vision.php

6.

DocID: 4/214

URL:

cml.ics.uci.edu/aiml/page/12

7.

DocID: 4/211

URL:

sli.ics.uci.edu/Site/UserCreation

8.

DocID: 4/213

URL:

fano.ics.uci.edu/ca/gliders/g1434.html

9.

DocID: 22/264

URL:

www.ics.uci.edu/about/bren/index.php/bren\_vision.php/index.php/../search/index.php/../../grad/index.php/ICS/ics/about/about\_facts

10.

DocID: 55/49

URL:

www.ics.uci.edu/~ejw/authoring/collection/dt/CollSpec022.txt