

Dillon Montag

Andrey Kitsen

Crawling the Bible Assignment

1) It took 161.814 seconds to crawl the domain.

2) There are 1256 unique URL's.

3) I found 6075 links.

4) The longest page is:

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/PS+119.html

5) 25 most common words:

NOTE: Only the stop words from the first list given by the link were eliminated. Other words like "will" that appeared as stop words in other lists on the linked website were not eliminated.

will : 10,285

lord : 7904

chapter : 4762

god : 4275

said : 3199

one : 2602

next : 2563

king : 2532

son : 2391

previous : 2379

man : 2256

people : 2253

israel : 1904

men : 1862

like : 1588

land : 1486

come : 1480

us : 1452

day : 1450

go : 1425

father : 1292

jesus : 1284

may : 1274

also : 1263

went : 1241

6) 25 most common Two-Grams:

See note for problem 5.

next chapter : 2376

previous chapter : 2376

chapter previous : 1124

o lord : 392

lord says : 319

lord said : 302

sovereign lord : 297

lord almighty : 264

lord will : 263

will come : 245

will make : 205

will bring : 197

will give : 196

hebrew manuscripts : 162

burnt offering : 161

will go : 158

chapter psalm : 150

will know : 145

jesus christ : 140

see also : 132

lord came : 129

o god : 119

will become : 119

jesus said : 118

will never : 117

7) Palindromes:

NOTE: Several palindromes tied for having a length of 9. The top 10 palindromes are listed alphabetically if the length tied.

1) asurejerusa :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/ZECH+2.html

2) ullammallu :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/EZRA+10.html

3) ahandnaha :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/JOSH+21.html

4) alliwilla :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/ISA+65.html

5) arofffora :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/ACTS+2.html

6) earsisrae :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/JUDG+11.html

7) emandname :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/JOSH+19.html

8) emansname :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/RUTH+1.html

9) erewewere :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/ACTS+20.html

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/ACTS+21.html

10) erofafore :

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/MAL+2.html

8) Dolcinator Venexus shows up on these 5 pages:

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/REV+22.html

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/EXOD+31.html

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/EXOD+35.html

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/JOHN+11.html

http://djp3.westmont.edu/classes/2015_09_CS150/tasks/crawl_bible/Bible/bible/PS+42.html