

Group Number: IR24W-A2 41

Group members:

Name	Student Id	UCI Net Id
Heli Kothari	55969236	helik@uci.edu
Swami Sivananda Nadendla	62389275	nadendls@uci.edu
Hiya Kothari	52062518	hnkothar@uci.edu
Pratyush Patel	45854791	pratyujp@uci.edu

1. How many unique pages did you find?

Ans: 8342

2. What is the longest page in terms of the number of words?

Ans: <https://ics.uci.edu/~eppstein/bibs/eppstein>

Number of words: 167314

3. What are the 50 most common words in the entire set of pages crawled under these domains ?

Ans: most 50 common words:

1. 2019 = 38641
2. data = 28075
3. 2018 = 24365
4. news = 21724
5. 2021 = 20256
6. 2015 = 18432
7. research = 17557
8. 2016 = 17331
9. graduate = 16774

10. august = 15698
11. contact = 13896
12. science = 12893
13. bren = 11935
14. 2020 = 11861
15. study = 11670
16. july = 10837
17. model = 10726
18. informatics = 10527
19. software = 10396
20. statistics = 10387
21. information = 9422
22. event = 9304
23. december = 9290
24. project = 9100
25. resource = 8745
26. irvine = 8626
27. june = 8605
28. undergraduate = 8598
29. january = 8059
30. stats = 7781
31. november = 7683
32. privacy = 7639
33. engineering = 7093
34. report = 7077
35. february = 7035
36. may = 6957
37. spring = 6289
38. seminar = 5861
39. march = 5652
40. students = 5587
41. people = 4636
42. update = 4507
43. classes = 4397
44. school = 4100

- 45. department = 3669
- 46. media = 3040
- 47. major = 2688
- 48. events = 2298
- 49. computer = 2198
- 50. support = 1760

4. How many subdomains did you find in the ics.uci.edu domain?

Ans:

- 1. <http://acoi.ics.uci.edu> = 67
- 2. <http://admissions.ics.uci.edu> = 2
- 3. <http://aiclub.ics.uci.edu> = 12
- 4. <http://archive.ics.uci.edu> = 18
- 5. <http://asterix.ics.uci.edu> = 93
- 6. <http://cbcl.ics.uci.edu> = 7
- 7. <http://cert.ics.uci.edu> = 14
- 8. <http://cg.ics.uci.edu> = 11
- 9. <http://cml.ics.uci.edu> = 1
- 10. <http://code.ics.uci.edu> = 17
- 11. <http://courses.ics.uci.edu> = 23
- 12. <http://create.ics.uci.edu> = 15
- 13. <http://cwicsocal18.ics.uci.edu> = 7
- 14. <http://dgillen.ics.uci.edu> = 3
- 15. <http://duttgroup.ics.uci.edu> = 2
- 16. <http://dynamo.ics.uci.edu> = 3
- 17. <http://emj.ics.uci.edu> = 9
- 18. <http://evoke.ics.uci.edu> = 5
- 19. <http://fano.ics.uci.edu> = 2
- 20. <http://flamingo.ics.uci.edu> = 8
- 21. <http://fr.ics.uci.edu> = 3
- 22. <http://futurehealth.ics.uci.edu> = 1
- 23. <http://grape.ics.uci.edu> = 317
- 24. <http://graphics.ics.uci.edu> = 19

25. <http://graphmod.ics.uci.edu> = 1
26. <http://hack.ics.uci.edu> = 1
27. <http://hpi.ics.uci.edu> = 3
28. <http://hub.ics.uci.edu> = 17
29. <http://icde2023.ics.uci.edu> = 2
30. <http://ieee.ics.uci.edu> = 4
31. <http://igb.ics.uci.edu> = 11
32. <http://industryshowcase.ics.uci.edu> = 6
33. <http://informatics.ics.uci.edu> = 57
34. <http://isensorium.ics.uci.edu> = 12
35. <http://isg.ics.uci.edu> = 86
36. <http://keys.ics.uci.edu> = 2
37. <http://latis.ics.uci.edu> = 1
38. <http://mdogucu.ics.uci.edu> = 11
39. <http://mds.ics.uci.edu> = 28
40. <http://mhcid.ics.uci.edu> = 34
41. <http://motifmap.ics.uci.edu> = 1
42. <http://mswe.ics.uci.edu> = 4
43. <http://nalini.ics.uci.edu> = 2
44. <http://nlp.ics.uci.edu> = 12
45. <http://password.ics.uci.edu> = 1
46. <http://perennialpolycultures.ics.uci.edu> = 5
47. <http://psearch.ics.uci.edu> = 5
48. <http://radicle.ics.uci.edu> = 11
49. <http://rbcl.ics.uci.edu> = 3
50. <http://redmiles.ics.uci.edu> = 9
51. <http://scale.ics.uci.edu> = 6
52. <http://se.ics.uci.edu> = 1
53. <http://sherlock.ics.uci.edu> = 1
54. <http://sli.ics.uci.edu> = 3
55. <http://stairs.ics.uci.edu> = 15
56. <http://statconsulting.ics.uci.edu> = 3
57. <http://student-council.ics.uci.edu> = 34
58. <http://studentcouncil.ics.uci.edu> = 71
59. <http://support.ics.uci.edu> = 7

- 60. <http://tad.ics.uci.edu> = 2
- 61. <http://tippersweb.ics.uci.edu> = 1
- 62. <http://transformativeplay.ics.uci.edu> = 4
- 63. <http://tutors.ics.uci.edu> = 19
- 64. <http://ugradforms.ics.uci.edu> = 38
- 65. <http://unite.ics.uci.edu> = 2
- 66. <http://vision.ics.uci.edu> = 1
- 67. <http://wics.ics.uci.edu> = 213
- 68. <http://wiki.ics.uci.edu> = 16
- 69. <http://www.ics.uci.edu> = 3177
- 70. <http://wwwdb.ics.uci.edu> = 2
- 71. <http://xtune.ics.uci.edu> = 3