

Steps 1 & 2: Document listing te 5 topics and the 100 keywords (20 per topic)

=====

Software > OS > Linux

1. Fedora
2. flavor
3. ARM
4. Linux
5. servers
6. free operating system
7. OS
8. open source
9. Debian
10. Ubuntu
11. GNU
12. supervisor
13. daemon
14. x86_64
15. x86
16. kernel
17. compiler
18. compiling
19. distribution
20. nonprofit

=====

Software > Emulators > Nintendo

1. Emulator
2. computer
3. ROM
4. GBMac
5. advance
6. version
7. pokemon
8. gameboy
9. color
10. PPU
11. SNES
12. GNU
13. beta
14. free
15. N64
16. compatibility

17. systems
18. amd64
19. platform
20. GBE
21. virtual

=====

Software > Databases > MySQL

1. MySQL
2. cloud
3. service
4. IRC
5. detection
6. changelog
7. oracle
8. privacy
9. users
10. workbench
11. server
12. files
13. SQL
14. cluster
15. database
16. documentation
17. MyAccess
18. tables
19. query
20. data

=====

Software > Editors > Vim

1. Vim
2. Vi
3. IDE
4. syntax highlighting
5. UNIX
6. helptags
7. bash
8. editors
9. buffer
10. plugins
11. macros
12. commands

13. open source
14. power user
15. mode
16. text editor
17. GNU/Linux
18. programmer
19. shortcuts
20. vimrc

=====

Software > Licensing

1. free
2. software
3. open source
4. services
5. online
6. program
7. FreeBSD
8. official
9. public
10. mozilla
11. license
12. copyright
13. agreement
14. terms and conditions
15. internet
16. python
17. PHP
18. Language
19. privacy policy
20. Ruby

Step 3: Strategy used to find these keywords

In order to find these keywords, we used the demo version of IBM Watson's *Alchemy*. It allowed us to pinpoint keywords that were common across multiple links within the same subtopic. Doing so allowed us to get a set of keywords that are evenly distributed across the articles.