Literaturverzeichnis

[1] German stoppwords. http://solariz.de/download-7, April 2010.

[2] Burton H. Bloom. Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM*, 13(7):422–426, 1970.

[3] Robert S. Boyer and Strother Moore. A fast string searching algorithm. *Communications of the ACM*, 20(10), Oktober 1977.

[4] Andrei Broder and Michael Mitzenmacher. Network applications of bloom filters: A survey. *Internet Mathematics*, 1(4):485–509, 2005.

[5] Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach, Mike Burrows, Tushar Chandra, Andrew Fikes, and Robert E. Gruber. Bigtable: A distributed storage system for structured data. *7th Conference on Usenix Symposium on Operating Systems Design and Implementation*, 9, 2006.

[6] Richard Cole and Ramesh Hariharan. Tighter upper bounds on the exact complexity of string matching. *SIAM J. Comput.*, 26(3):803–856, 1997.

[7] Jeffrey Dean and Sanjay Ghemawat. Mapreduce: Simplified data processing on large clusters. In *OSDI,Sixth Symposium on Operating System Design and Implementation*, pages 137–150, 2004.

[8] M. L. Fredman, R. Sedgewick, D. D. Sleator, and R. E. Tarjan. The pairing heap: a new form of self-adjusting heap. *Algorithmica*, 1(1):111–129, 1986.

[9] Michael Fredman and Robert Tarjan. Fibonacci heaps and their uses in improved network optimization algorithms. *Journal of the ACM*, 34(3):596–615, 1987.

[10] C.A.R. Hoare. Quicksort. *Computer Journal*, 5(1):10–15, 1962.

[11] Richard M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. New York: Plenum, 1972.

[12] Donald E. Knuth. *The Art of Computer Programming. Vol. 3: Sorting and Searching*. Addison-Wesley, second edition, 1998.

[13] The Lucene Webpages. lucene.apache.org.

[14] Fredrik Lundh. Python hash algorithms. http://effbot.org/zone/python-hash.htm, 2002.

[15] Rob Pike, Sean Dorward, Robert Griesemer, and Sean Quinlan. Interpreting the data: Parallel analysis with sawzall. *Scientific Programming*, 13(4):277–298, 2005.

[16] William Pugh. Skip lists: a probabilistic alternative to balanced trees. *Communications of the ACM*, 33(6), June 1990.

[17] Gaston H. Gonnet Ricardo A. Baeza-Yates. A new approach to text searching. *Communications of the ACM*, 35(10):74–82, Oktober 1992.

[18] Jean Vuillemin. A data structure for manipulating priority queues. *Communications of the ACM*, 21:309–314, 1978.

[19] John Zelle. *Python Programming: An Introduction to Computer Science*. Franklin Beedle & Associates, Dezember 2003.

Index