

*corpus*  
*(uncompressed)*

http://in.tum.de  
http://cwi.nl  
www.uni-jena.de  
www.wikipedia.org  
http://www.vldb.org  
...

*symbol table*

0	http://	7
1	www.	4
2	uni-jena	8
3	.de	3
4	.org	4
5	a	1
6	in.tum	6
7	cwi.nl	6
8	wikipedi	8
9	vldb	4
...		
255		

*symbol length*

*corpus*  
*(compressed)*

063  
07  
123  
1854  
0194  
...