

|             | Documents | Terms | Postings | Occurrences |
|-------------|-----------|-------|----------|-------------|
| TREC GOV2   |           |       |          |             |
| Text        | 25 M      | 35 M  | 5.5 G    | 23 G        |
| Title       | 25 M      | 1.1 M | 135 M    | 150 M       |
| Web .uk     |           |       |          |             |
| Text        | 130 M     | 99 M  | 21 G     | 62 G        |
| Title       | 130 M     | 3.2 M | 609 M    | 691 M       |
| Mimir index |           |       |          |             |
| Token       | 1 M       | 49    | 27 M     | 1.2 G       |
| Tweets      |           |       |          |             |
| Text        | 13 M      | 2.3 M | 147 M    | 156 M       |