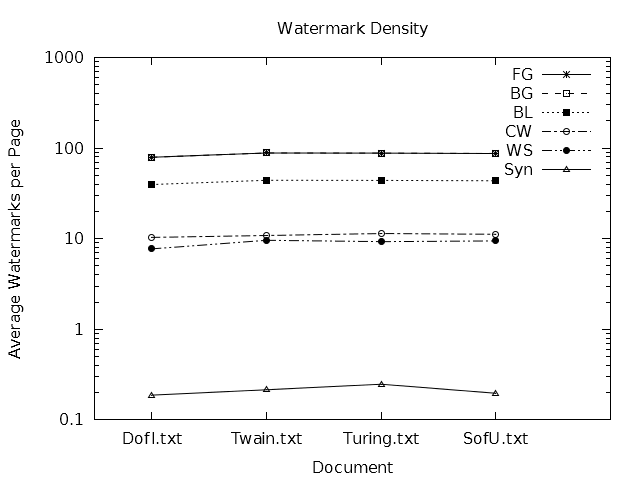
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **File** | **Title** | **Author** | **Date** | **Genre** | **Chars** | **Words** | **Pages\*** | **Source** |
| DofI.txt | Declaration of Independence | “Founding Fathers” | 1776 | Political | 8148 | 1339 | 3 | [Link](http://www.archives.gov/exhibits/charters/declaration_transcript.html) |
| Twain.txt | The McWilliamses and the Burglar Alarm | Mark Twain | 1882 | Literary | 14704 | 2630 | 5 | [Link](https://americanliterature.com/author/mark-twain/short-story/the-mcwilliamses-and-the-burglar-alarm) |
| Turing.txt | Computing Machinery and Intelligence | Alan Turing | 1950 | Academic | 67245 | 11757 | 19 | [Link](http://www.loebner.net/Prizef/TuringArticle.html) |
| SofU.txt | State of the Union Address+ | Barack Obama | 2016 | Spoken | 34243 | 6093 | 10 | [Link](https://www.whitehouse.gov/the-press-office/2016/01/12/remarks-president-barack-obama-%E2%80%93-prepared-delivery-state-union-address) |

\*Using 12-point Arial font, 8.5 x 11 inch paper, 1-inch margins all around, default leading and kerning, no indents before or blank lines between paragraphs. Subtract one for the number of full pages. Counts are before watermarking.

+Applause and laughter removed from transcript as well all HTML errors. Smart quotes replaced



|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Document | Chars | Words | Pages | FG | BG | BL | CW | WS | Syn |
| DofI.txt | 8157 | 1339 | 3 | 79.0625 | 79.0625 | 39.53125 | 10.3203125 | 7.7109375 | 0.1875 |
| Twain.txt | 14724 | 2630 | 5 | 88.125 | 88.125 | 44.0625 | 10.83203125 | 9.58984375 | 0.21484375 |
| Turing.txt | 67433 | 11757 | 20 | 87.7845394736842 | 87.7845394736842 | 43.8922697368421 | 11.380756578947368 | 9.264802631578947 | 0.24753289473684212 |
| SofU.txt | 34302 | 6093 | 11 | 86.959375 | 86.959375 | 43.4796875 | 11.1625 | 9.4359375 | 0.196875 |

These numbers of characters and pages are counted after watermarking and are affected by the synonym watermarking scheme.