

SEUBE

NLP – 2019

FFyL

última clase

Recorte del curso

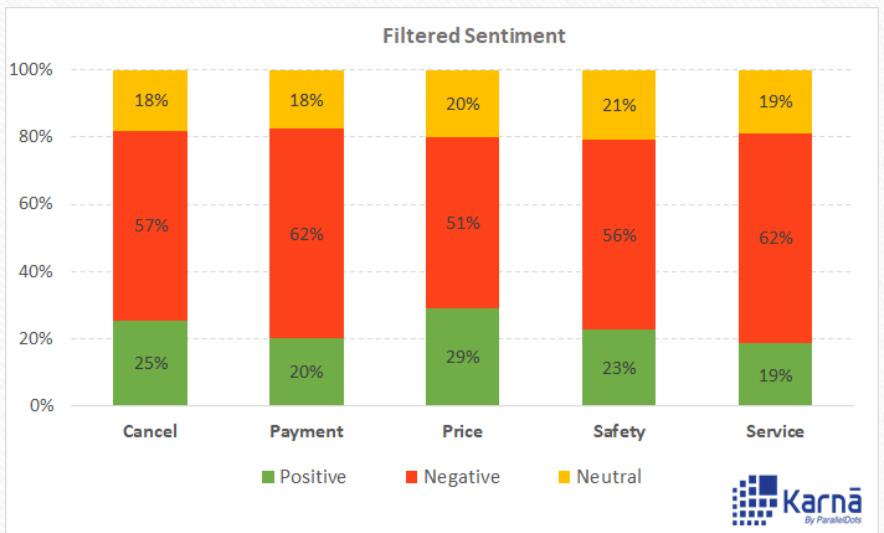
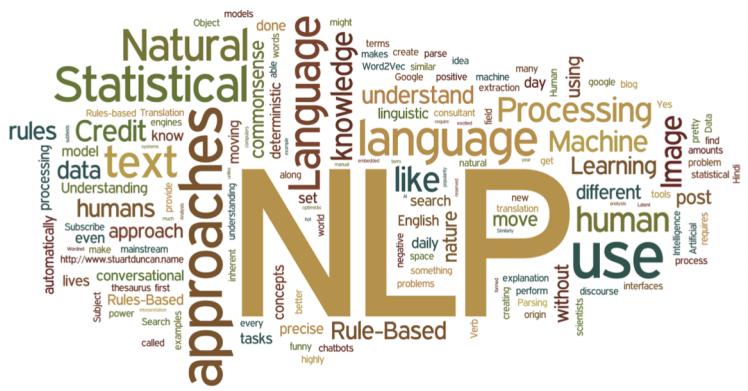
- Presentacion del grupo. Estado de la cuestión.
Presentación de los alumnos
- Teoría gramatical y computacional.
- Jupyter Notebook (todas las clases)
- 4 clases de Python básico:
 - Teoría: Sintaxis básica, loops, condicionales, listas, regex
 - Práctica: taller Python, taller Regex
- Intro a NLTK, módulos, librerías. Stemming.
- Lemmatización con POSTagging. Modelos.
- Parsers + Gramáticas
- Corpus, tf-Idf
- Spacy, Pipeline, NER

Recorte del curso

- ✓ Orientado a alumnos formados en “ciencias blandas”
- ✓ Introducción desde cero a conceptos computacionales
- ✓ Aplicación de herramientas con una intuición básica del trasfondo
- ✓ Recorrido por los aportes lingüísticos en análisis y productos de NLP
- ✓ Primer acercamiento al abordaje de los problemas de la lingüística desde una perspectiva computacional

Recorte del curso

- ✗ Profundización en los algoritmos de fondo:
 - ✗ Clustering
 - ✗ Clasificadores
 - ✗ Estadística
- ✗ Realizar un análisis o producto de principio a fin
 - ✗ Herramientas de visualización
 - ✗ Programación avanzada
- ✗ Topic Modeling



100 Most Frequent Words in Books for Beginning Readers

- | | | | |
|----------|------------|-----------|-----------|
| 1. the | 26. he | 51. be | 76. cat |
| 2. a | 27. out | 52. now | 77. them |
| 3. and | 28. that | 53. when | 78. tree |
| 4. to | 29. one | 54. there | 79. where |
| 5. I | 30. big | 55. into | 80. away |
| 6. in | 31. go | 56. day | 81. time |
| 7. is | 32. was | 57. look | 82. as |
| 8. on | 33. like | 58. eat | 83. water |
| 9. you | 34. what | 59. make | 84. home |
| 10. it | 35. not | 60. his | 85. made |
| 11. of | 36. do | 61. here | 86. long |
| 12. said | 37. then | 62. your | 87. has |
| 13. can | 38. this | 63. an | 88. help |
| 14. for | 39. no | 64. back | 89. good |
| 15. my | 40. too | 65. mom | 90. going |
| 16. but | 41. she | 66. dog | 91. by |
| 17. all | 42. went | 67. very | 92. how |
| 18. we | 43. see | 68. did | 93. house |
| 19. are | 44. will | 69. her | 94. dad |
| 20. up | 45. so | 70. from | 95. or |
| 21. at | 46. some | 71. had | 96. two |
| 22. with | 47. down | 72. got | 97. red |
| 23. me | 48. little | 73. put | 98. am |
| 24. they | 49. come | 74. came | 99. over |
| 25. have | 50. get | 75. just | 100. saw |

Text in

Brazil ranks number 5 in the list of countries by population.

The term "Ibu Negara" (Lady/Mother of the State) is used for wife of the President of Indonesia.

Game of Thrones is an adaptation of A Song of Ice and Fire, George R. R. Martin's series of fantasy novels. It ranks fourth among the IMDB Top Rated TV Shows

Data out

THE COUNTRIES WITH THE LARGEST POPULATION

China	1	1,388,232,693
India	2	1,342,512,706
United States	3	326,474,013
Indonesia	4	263,510,146
Brasil	5	174,315,386

THE COUNTRY'S FIRST LADIES

Brigitte Macron

- Spouse: Emmanuel Macron, President of France (2017 -)

Melania Trump

- Spouse: Donald J. Trump, U.S. President (2017 -)

Iriana Widodo

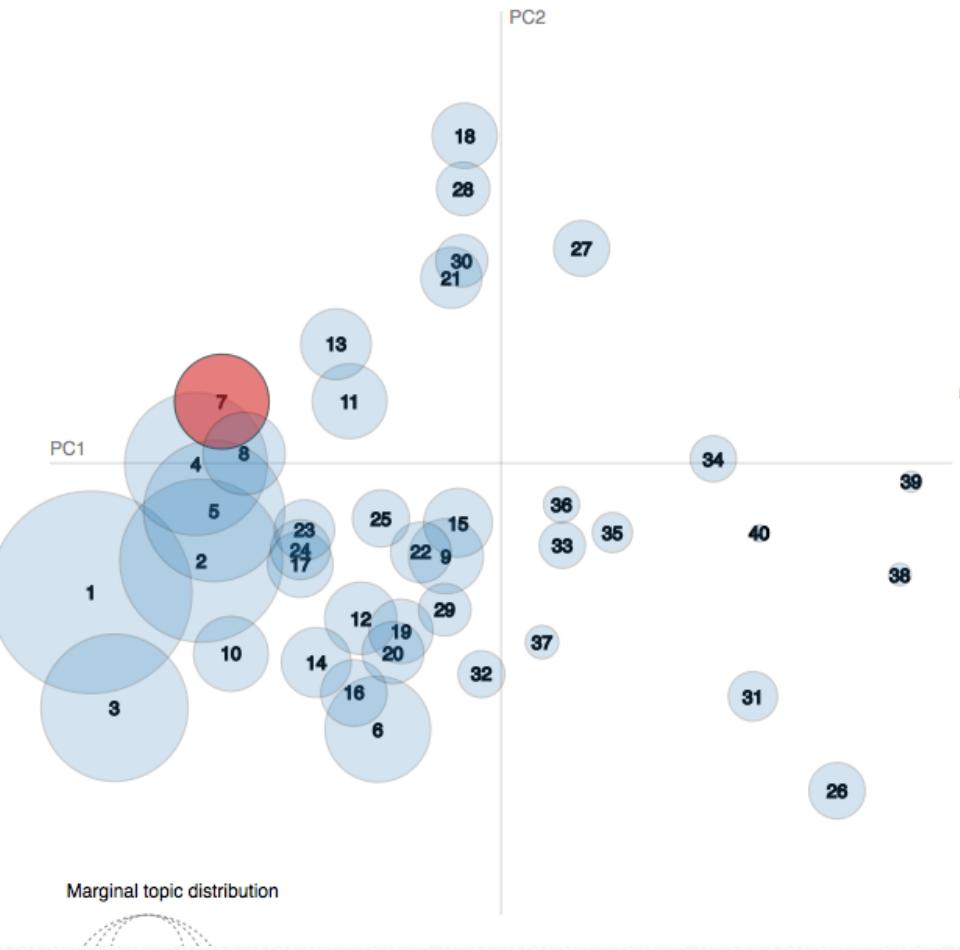
- Spouse: Joko Widodo, President of Indonesia (2014 -)

- Also known as: "Ibu Negara" (Lady/Mother of the State)

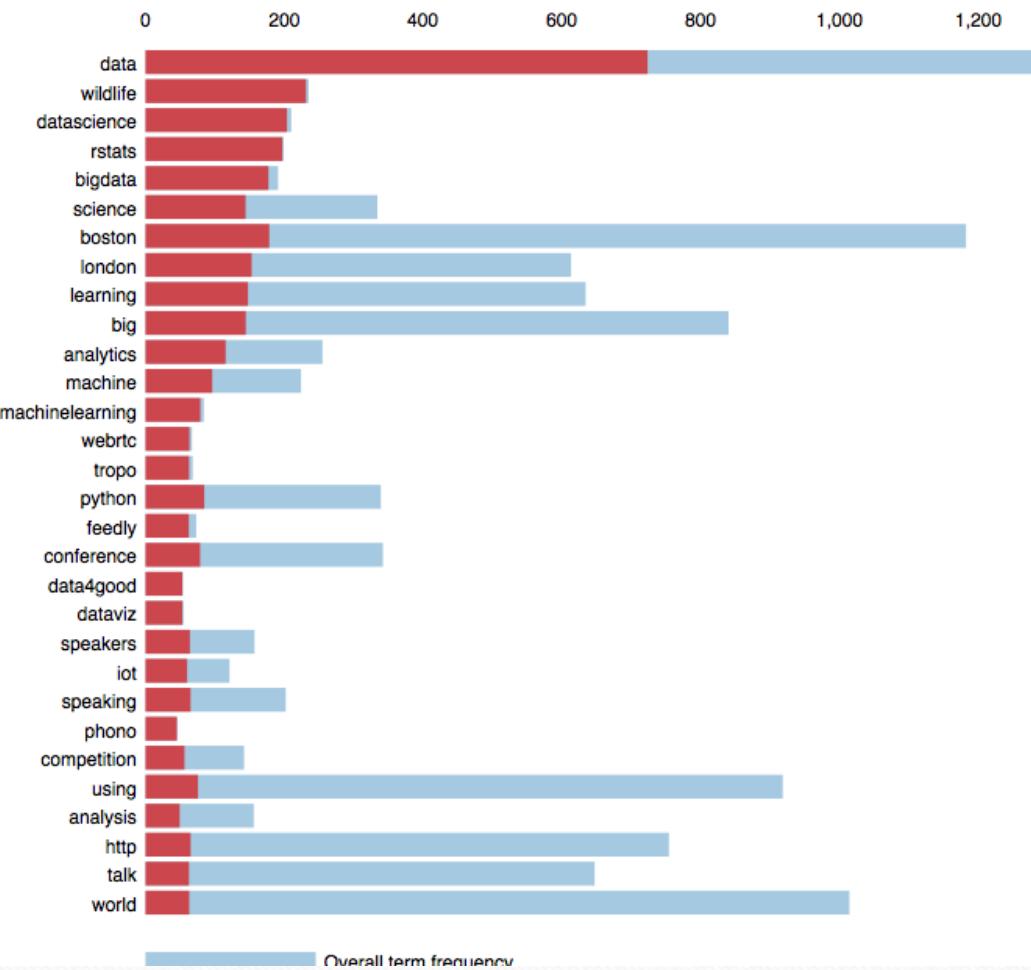
IMDB TOP RATED TV SHOWS

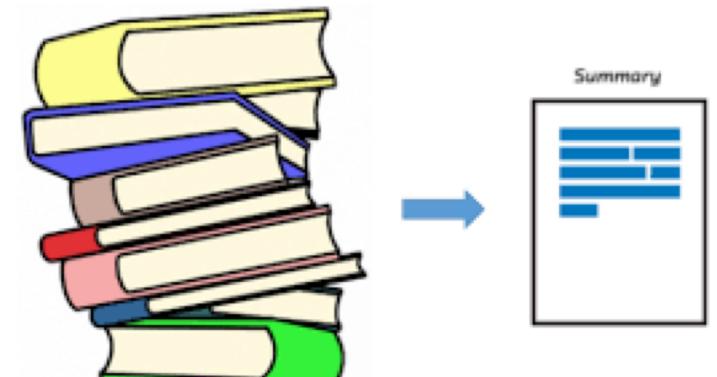
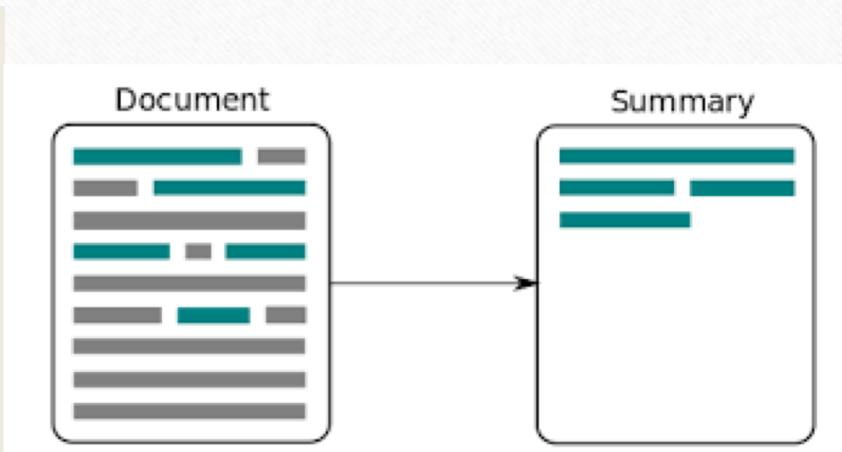
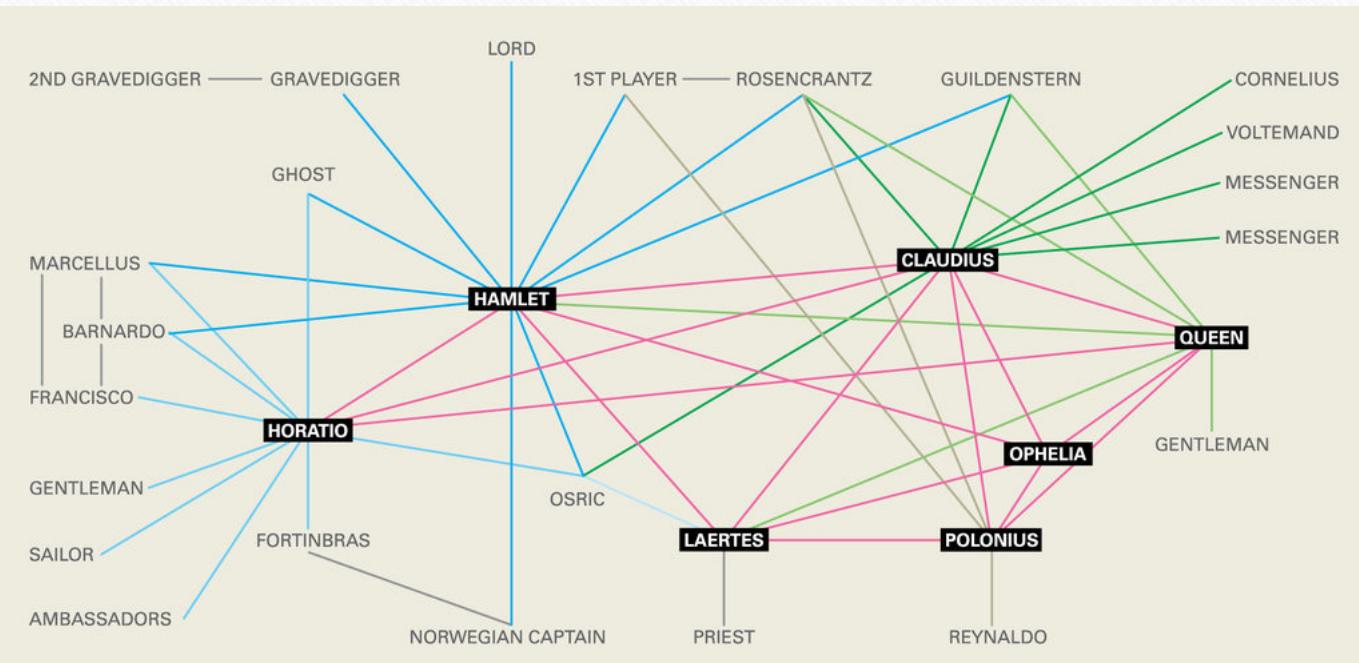
- 1 Planet Earth II (2016) 9.6.
- 2 Band of Brothers (2001) 9.5.
- 3 Planet Earth (2006) 9.5.
- 4 Game of Thrones (2011) 9.4.
- 5 Breaking Bad (2008) 9.4.

Intertopic Distance Map (via multidimensional scaling)

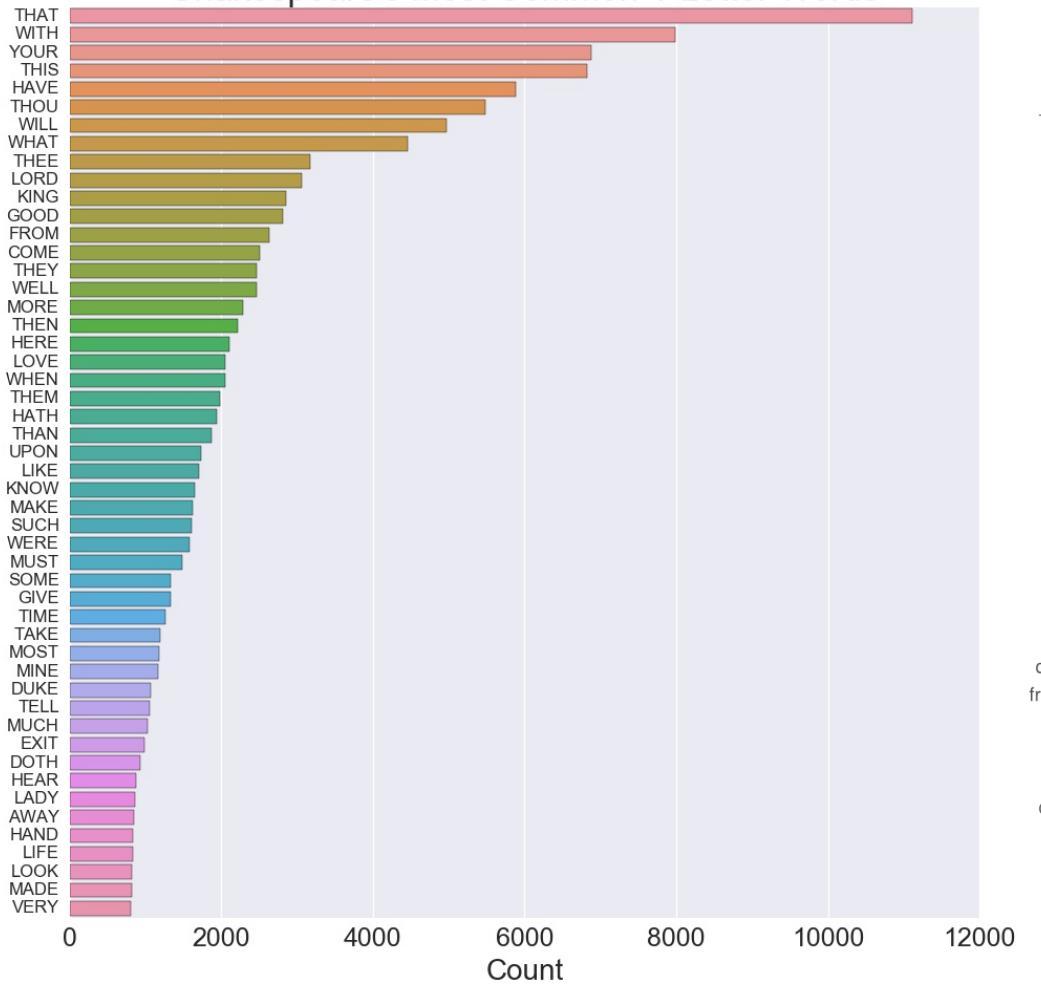


Top-30 Most Relevant Terms for Topic 7 (3.5% of tokens)

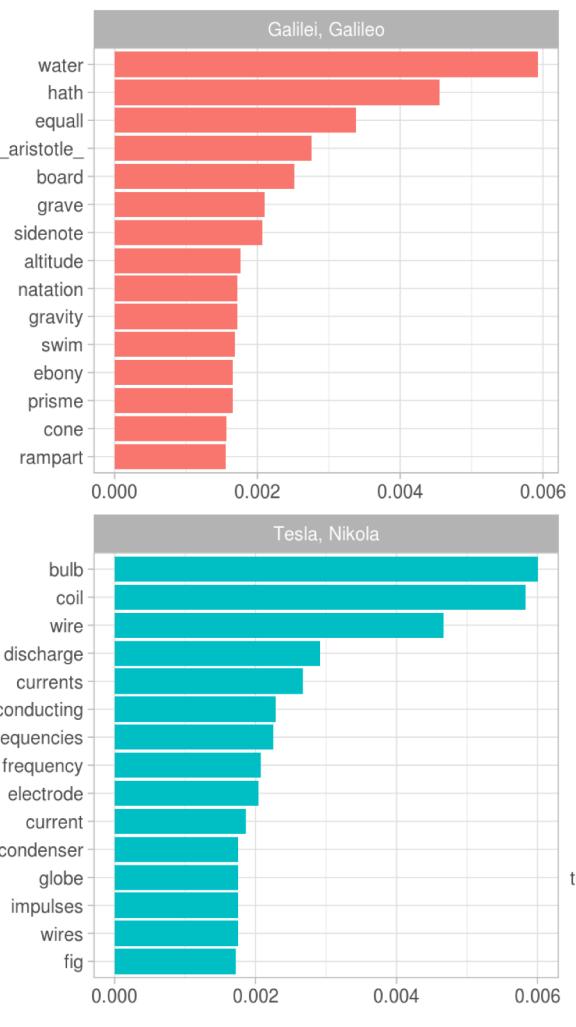




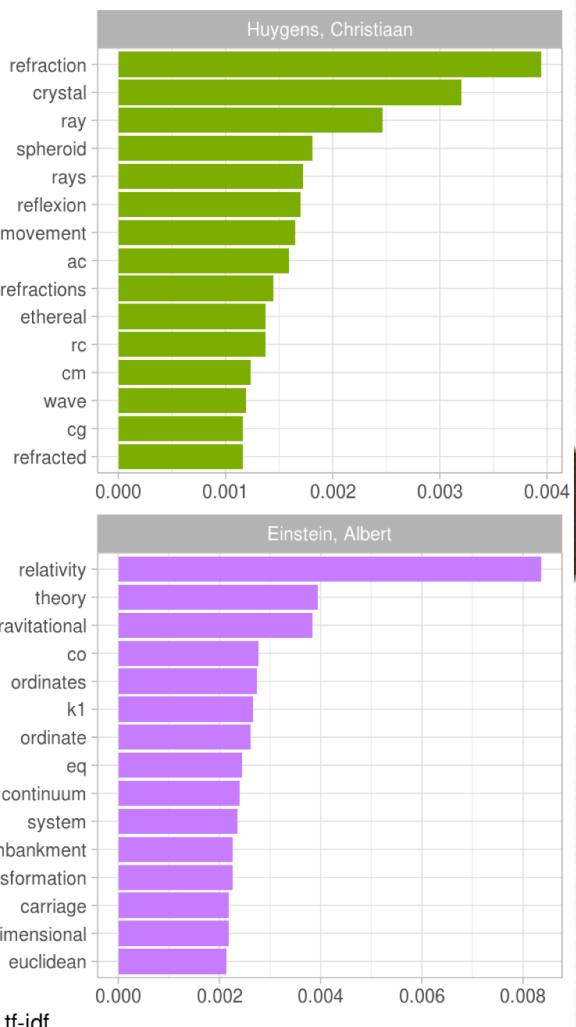
Shakespeare's Most Common 4-Letter Words



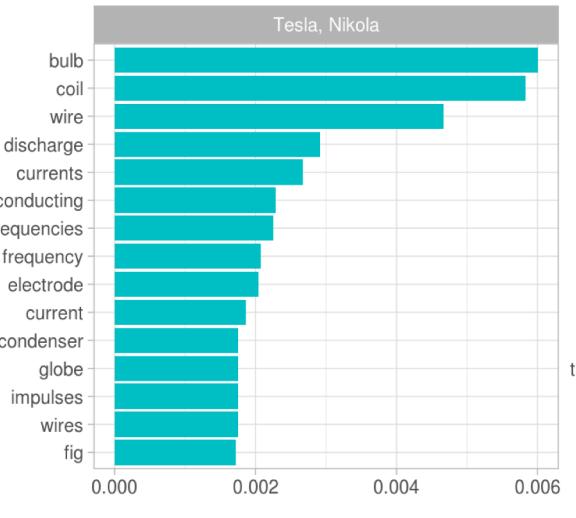
Galilei, Galileo



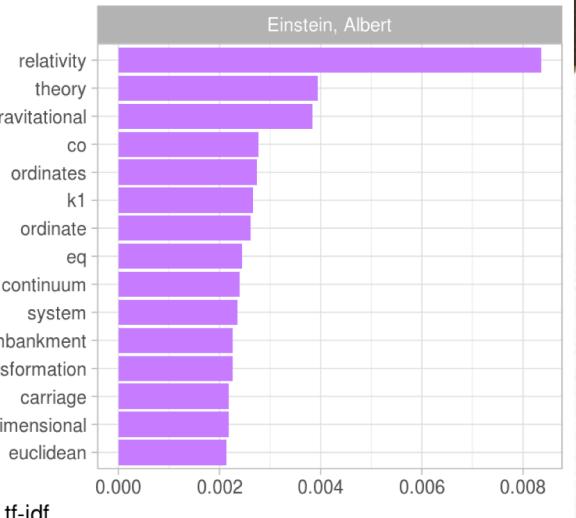
Huygens, Christiaan



Tesla, Nikola



Einstein, Albert



tf-idf



