



Procesamiento de Datos Masivos

## Material complementario Módulo 2

- Documentación de índices invertidos de Elastic:  
<https://www.elastic.co/blog/found-indexing-for-beginners-part3>.
- TF-IDF en Scikit Learn: [https://scikit-learn.org/stable/modules/generated/sklearn.feature\\_extraction.text.TfidfTransformer.html#sklearn.feature\\_extraction.text.TfidfTransformer](https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfTransformer.html#sklearn.feature_extraction.text.TfidfTransformer).
- Búsqueda con índices en BigQuery:  
<https://cloud.google.com/bigquery/docs/search>.
- Función SEARCH en BigQuery:  
[https://cloud.google.com/bigquery/docs/reference/standard-sql/search\\_functions](https://cloud.google.com/bigquery/docs/reference/standard-sql/search_functions).
- Charla Ted de Tim Berners-Lee sobre el futuro de la Web:  
[https://www.ted.com/talks/tim\\_berniers\\_lee\\_the\\_next\\_web?language=en](https://www.ted.com/talks/tim_berniers_lee_the_next_web?language=en).
- Historia de La Web: <https://web30.web.cern.ch/web-history.html>.
- Tutorial de HTML: <https://www.w3schools.com/html/default.asp>.
- Documentación de BeautifulSoup:  
<https://www.crummy.com/software/BeautifulSoup/bs4/doc/>.