

**Publishers/Service Providers which allow TDM**

Provider	Free?	Details	Further Information
Early English Books Online	Yes	<p>All restrictions will be lifted from EEBO-TCP Phase I, which consists of the first 25,000 texts transcribed and encoded by the TCP from 2000-2009.</p> <p>These 25,000 (plus a few hundred) texts will be freely available to anyone wishing to use them, and there will no longer be any restrictions on sharing these files.</p> <p>They will be licensed under the Creative Commons Public Domain Dedication (CC0 1.0 Universal), which will be indicated in the header of each text.</p>	<p><a href="http://www.textcreationpartnership.org/">http://www.textcreationpartnership.org/</a></p>
Gale	Yes	<p>Made content available from its Gale Primary Sources to academic researchers for data mining and text mining purposes.</p>	<p><a href="https://www.cengage.com/search/showresults.do?N=197+4294891079">https://www.cengage.com/search/showresults.do?N=197+4294891079</a>  <a href="http://assets.cengage.com/pdf/faq_DataMining.PDF">http://assets.cengage.com/pdf/faq_DataMining.PDF</a>  <a href="https://www.cengage.com/search/showresults.do?N=197+4294891079">https://www.cengage.com/search/showresults.do?N=197+4294891079</a></p>

		Currently offers the Gale Digital Learning Lab.	<a href="https://www.gale.com/uk/primary-sources/digital-scholar-lab">https://www.gale.com/uk/primary-sources/digital-scholar-lab</a>
JSTOR	Yes	Data for Research (DfR) provides datasets of content on JSTOR for use in research and teaching. Researchers may use DfR to define and submit their desired dataset to be automatically processed. Data available through the service includes metadata, n-grams, and word counts for most articles and book chapters, and for all research reports and pamphlets on JSTOR. Datasets are produced at no cost to researchers and may include data for up to 25,000 documents.	<a href="https://www.jstor.org/dfr/">https://www.jstor.org/dfr/</a>  Washington Lee University: <a href="https://digitalhumanities.wlu.edu/introduction-to-using-jstor-data-for-research/">https://digitalhumanities.wlu.edu/introduction-to-using-jstor-data-for-research/</a>
IEEE	Yes	IEEE permits non-commercial text and data mining of articles published open access with either the Open Access Publishing Agreement (OAPA) or the Creative Commons license (CC BY). No permission is required for non-commercial mining	

		of open access articles. Mining for commercial purposes or mining of non-open access content requires permission from IEEE.	
Lexis Nexis	Not officially	Does not 'officially' support or provide data/text mining options. However, since text files can be downloaded, TDM is possible. You can batch download up to 500 articles at a time in one text file (Source: University of Southern California)	<a href="https://libguides.usc.edu/contentmining/databases">https://libguides.usc.edu/contentmining/databases</a>
Oxford University Press	Yes	Oxford University Press recognizes the research benefit of Text and Data Mining (TDM) across a variety of research fields. As such, we are happy to accommodate TDM for non-commercial use. Researchers are not required to request permission for non-commercial text-mining.	
ProQuest	Yes	ProQuest does allow free text mining for the newspapers to which USC Libraries have purchased	

		<p>perpetual access licenses. Those newspapers are: Los Angeles Times (1881-1931), Los Angeles Sentinel (1934-2005), and New York Times (1851-1934). USC Libraries will have to request this data on your behalf.</p> <p>ProQuest offers two methods of data delivery:</p> <ol style="list-style-type: none"> <li>1. Online cloud-based delivery - a link to a sitemap is sent from which the documents can be downloaded.</li> <li>2. Hard drive can be mailed</li> </ol> <p>(Source: University of Southern California)</p>	
ScienceDirect/Scopus/Elsevier	Yes with an API	You can text mine all subscribed content so long as it is for non-commercial purposes through using Elsevier's Science Direct APIs. However, you must	<a href="https://dev.elsevier.com/tecdoc_text_mining.html">https://dev.elsevier.com/tecdoc_text_mining.html</a>

		<p>register first to use these APIs.</p> <p>(Source: University of Southern California)</p> <p>Researchers at subscribing academic institutions can text mine subscribed full-text ScienceDirect content via the Elsevier APIs for non-commercial purposes. Researchers will also require a personal API key.</p>	<p><a href="https://www.elsevier.com/about/policies/text-and-data-mining">https://www.elsevier.com/about/policies/text-and-data-mining</a></p>
SpringerLink	Yes	<p>For subscribed journals and books, Springer Nature grants researchers text and data mining rights via their institutions, provided the purpose is non-commercial research.</p>	<p><a href="https://www.springernature.com/gp/researchers/text-and-data-mining">https://www.springernature.com/gp/researchers/text-and-data-mining</a></p>
Royal Society	Yes	<p>We support the stance that the right to read is the right to mine. We believe that the ability to use computers to extract information from scholarly material is one of many tools available to researchers, and we</p>	<p><a href="https://royalsociety.org/journals/ethics-policies/data-sharing-mining/">https://royalsociety.org/journals/ethics-policies/data-sharing-mining/</a></p>

		<p>support this activity on our journals.</p> <p>Members of subscribing institutions have our permission to mine journal content for either commercial or non-commercial purposes. We ask that you respect the copyright of the original papers, and where possible cite original works when you reuse them.</p>	
Cambridge Journals Online	Yes	<p>You may download, extract, store and index Content for the purposes of text and data mining ("TDM") for non-commercial research purposes only and may mount, load, integrate and analyse the results of TDM subject to the inclusion of a link to the underlying Content on our Site. Any copies of the Content stored locally by you for the purposes of TDM shall be deleted once such research project ends.</p>	<p><a href="https://www.cambridge.org/core/legal-notice/terms">https://www.cambridge.org/core/legal-notice/terms</a></p>

Hathi Trust	Yes	The HathiTrust Research Center (HTRC) facilitates non-profit and educational uses of the HathiTrust Digital Library by enabling computational analysis of public domain works and (on limited terms) in-copyright works from its collection.	<a href="https://www.hathitrust.org/one-of-worlds-largest-digital-libraries-opens-doors-to-text-mining-scholars">https://www.hathitrust.org/one-of-worlds-largest-digital-libraries-opens-doors-to-text-mining-scholars</a>
EBSCO	No		<a href="https://www.ebsco.com/blog-archives/article/ebsco-and-digital-humanities-data-and-text-mining">https://www.ebsco.com/blog-archives/article/ebsco-and-digital-humanities-data-and-text-mining</a>
Jisc Journal Archives	Not yet		<a href="https://scholarlycommunications.jiscinvolve.org/wp/2018/09/11/jisc-gateway-to-text-and-data-mining-status-update/">https://scholarlycommunications.jiscinvolve.org/wp/2018/09/11/jisc-gateway-to-text-and-data-mining-status-update/</a>
Library of Congress	Yes	See databases such as Chronicling America.	<a href="https://chroniclingamerica.loc.gov/about/">https://chroniclingamerica.loc.gov/about/</a> <a href="https://chroniclingamerica.loc.gov/about/api/">https://chroniclingamerica.loc.gov/about/api/</a>
SAGE Journals Online	Yes	SAGE allows text and data mining, but you must contact us to be authorized and set up correctly to avoid your IP from being blocked. You may extract or use information contained on SAGE Journals for educational purposes, including, but not limited to, text and data mining, extraction and manipulation of information for the purposes of illustration,	

		explanation, example, comment, criticism, teaching, research, or analysis.	
Web of Science (See Elsevier)	Yes	Requires an API.	<a href="https://www.elsevier.com/en-gb/about/open-science/research-data/text-and-data-mining">https://www.elsevier.com/en-gb/about/open-science/research-data/text-and-data-mining</a>