

Text & Data Mining

When we access text or data we don't only look at it, analyse it, take notes and set up new documents. Nowadays this happens digitally. Progress we have made in the fields of genetic research or astronomy, for example, would nowadays be unthinkable without text and data mining.



Text and data mining is not an activity limited to professionals, large companies and research organisations. Citizens, start-ups and journalists already heavily use mining techniques.

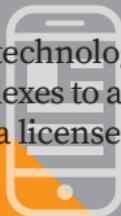


WIKIMEDIA

Who benefits from text & data mining?

Everyone

Everyday tasks rely on mining technologies. Our electronic devices build indexes to allow us to search files or documents. If we need a license to Ctrl+F, copyright goes too far.



Scientists

Researchers use information technology to obtain new insights and develop novel concepts from large collections of data and text. Practical examples include biomedical research of gene combinations, “distant reading” of thousands of literary works, or linguistic research.



Innovators and entrepreneurs

Ever heard of Treemetrics (forest management using earth observation), SentiOne (online monitoring data analysis), or iris.at (AI research assistant)? This is only a select number of European tech startups that rely on legacy being able to mine text and data.



Journalists

Journalists analyse large datasets to recognise dependencies or spot improper behaviour. The Panama Papers and Dieselgate are just the most prominent stories.



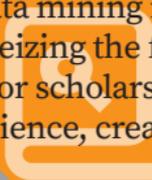
Citizens

Public institutions are providing data, yet the only way to assure that the public can really make sense of it is using technical tools. Anti-corruption kits are one powerful example.



Libraries

Unclear rules about text and data mining impede libraries and their users from seizing the full potential of the available information – for scholarship, research, improvements in health and science, creativity, and social inclusion.



The right to read is the right to analyse!

Licence Information

smartphone article by Royyan Wijaya from the Noun Project, Data Scientist by Becris from the Noun Project, Businessman by Chameleon Design from the Noun Project, journalist by Aldric Rodríguez from the Noun Project, public opinion by Symbolon from the Noun Project, Book by Ralf Schmitzer from the Noun Project