

The Turnitin difference:

Our web crawler

Turnitin's plagiarism checking service outperforms other plagiarism checkers in large part because our proprietary web crawler has found and indexed more than 91 billion web pages that are highly likely to be used by students and researchers in their writing. Our crawler has been built from the start to be horizontally scalable, smart about re-crawling to keep the content updated, and flexible enough to provide comprehensive results with every similarity check.

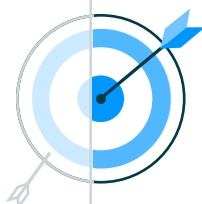
Average search engine

vs

Turnitin web crawler

Generic results

Most search engines only find content to index by following links from page to page. With this unguided and meandering methodology, many crawlers get stuck in crawler traps, capture irrelevant content, or completely miss high-value content that might not be linked to properly.



Targeted content

Turnitin's search methodology targets sites that are key to support students and researchers. We have partnerships with top open access repositories such as CORE to guide us directly to the best content. Every day, more than 102 million pages are re-crawled for updates and 22 million new pages are added! Need something specific added? Just let us know!

Inconsistent quality

Service providers may choose to sacrifice the match quality in order to reduce the number of searches required to keep a low cost to the customer.



Consistent quality

We are able to match the entirety of a submission against the entirety of our content database every single time.

Transient web content

Search engines will only return the most recent copy of a web page. If content was copied from a page which has since changed or is no longer live then a match will not be found.



Archived web content

Turnitin retains historical copies of all web pages it has crawled.

Hidden costs

Search engines costs are based on the number of searches required per document. Services using search engines work by extracting multiple phrases of text from a submitted document and running each of them through the search engine, thus making costs unpredictable.



Fixed cost

Because we have the content indexed within our database there is no external cost to be paid in order to perform similarity matches. Our web crawler has indexed over 91 billion web pages to date.

Learn more about our content and search our databases at turnitin.com/about/content