

KEYWORD EXTRACTION BY EDIA

What is Keyword Extraction?

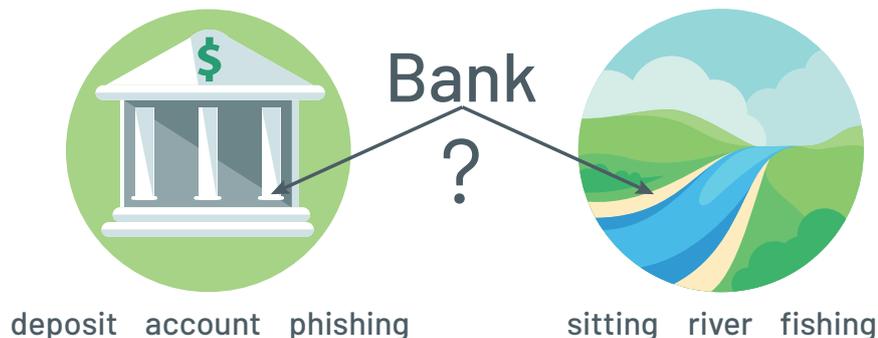
Keyword extraction is a form of metadata tagging, specifically finding the phrases in a text that give information about the text. Phrases can consist of multiple words like 'chemical composition' or a single word, like 'reaction'. Words like 'reaction' can mean multiple things in different contexts, so to make it more useful we also disambiguate the phrase to the keyword 'chemical reaction'. Specifically we link the keyword to a corresponding Wikipedia page.

Why is Keyword Extraction important? Who is it for? And what are the benefits?

Keywords are useful in several ways. Keywords ensure that someone can understand the high level meaning of the content without checking it in great detail. In organizations where a large volume of content is produced and processed, keywords can help content managers with content searchability as well as gaining strategic insight into their content library.

How does the EDIA Keyword Extractor work?

EDIA's Keyword Extractor scans through a text to find phrases that can be keywords. Based on the context of the surrounding text it determines whether the phrase is one of the keywords we associate with that phrase. For example, the word 'bank' can refer to a financial institution or a geographical feature alongside a river. Words like 'deposit', 'account', and 'phishing' can indicate the former, while 'river', 'sitting', and 'fishing' are more likely to occur in the context of the latter.



How is the EDIA Extractor made?

The EDIA Keyword Extractor uses Wikipedia to learn what different keywords and their associated phrases are. It also uses Wikipedia to learn about the contexts in which the phrases are used.

How can the EDIA Keyword Extractor be better than humans?

Our algorithm is better than humans. How does that happen? Let's consider an example. Radiologists assess your health by analysing MRI scans. In some cases, machine learning has shown to be better at this task. That is because machine learning is trained by the knowledge of many experts, not just one. Machine learning is trained using a huge dataset, one that is impossible to digest by humans. Machine learning doesn't need a coffee break and doesn't have

a bad night's sleep. As you can see, there are many similarities between this example to our technology, as it is trained by experts on a large, validated dataset and outperforms humans in keyword extractions.

With which software is the EDIA Keyword Extractor compatible?

The EDIA Keyword Extractor is available in the following content management systems (CMS) and editors: Microsoft Word, Google Docs, Alfresco, PublishOne, FontoXML, EDIA Papyrus, and others. In principle, EDIA Keyword Extractor can be integrated into any system that an organization uses.

How can you activate the EDIA Keyword Extractor in your workflow?

The EDIA Keyword Extractor is available in a form of API that you can easily activate in one of the software solutions mentioned above. All you need to get started is a valid license key which can be requested through the EDIA sales team.

How many costs can I save by using EDIA's Keyword Extractor?

On average it takes a human expert 5 minutes to label one piece of content. The hourly labour cost to tag manually 12 content items on average is € 48 (€ 4 per content item).

Our commercial proposition varies per volume of the content that you want to classify with EDIA Keyword Extractor. Typically we charge 50 cents per 1 API call (= 1 content item of 200 words or 1,000 characters). The more articles need to be classified, the lower the price per content item.



For example, to classify 1,000 content items would cost € 40,000 and more than 10 workdays using manual tagging (excluding hours required to find the grading experts). By using the EDIA automated Keyword Extractor it would take approximately 10 minutes for € 500, resulting in a **90% saving** for your organization, as well as significantly increased efficiency and time savings.

Contact information of EDIA sales team

Walter Montenarie

M +31 (0)6 2270 3526

E walter@edia.nl