

АНАЛІЗ ЕФЕКТИВНОГО ОПРАЦЮВАННЯ ВЕЛИКИХ ТЕКСТОВИХ ДАНИХ ЗАСОБАМИ ХМАРНИХ СЕРВІСІВ

UDC 004.91+811.1

Kh. Yurkevych¹, A. Lutskiv¹, N. Popovich²¹(Ternopil I.Pulyu National Technical University, Ukraine)²(Uzhgorod National University, Ukraine)**ANALYSIS OF EFFICIENT PROCESSING OF BIG TEXT DATA BY CLOUD SERVICES**

Для того щоб максимально ефективно опрацювати природну мову комп'ютерна система повинна використовувати ефективні алгоритми та відповідні методи, а їх реалізації у вигляді програмних бібліотек повинні забезпечувати можливість паралельного виконання й добре масштабуватись, відповідно до розмірів вхідних даних. Все більше набувають популярності serverless-технології – набір сервісів, для створення спеціалізованих комп'ютерних систем. Розглянемо найпопулярніші сервіси для опрацювання природної мови людини:

Таблиця 1 – Хмарні сервіси опрацювання великих текстових даних

Назва сервісу	Підтримувані мови	Токенізація	Лематизація	NER	POS-тегування	Емоційне забарвлення	Ключові слова
Google Cloud AutoML Natural Language	English, Chinese-Simplified, Spanish, Italian, Japanese, Russian, Ukrainian	+	+	+	+	Від 0 до 1, magnitude, magnitude	+, визначає ваговий коефіцієнт
Amazon Comprehend	English, Chinese-Simplified, Spanish, Italian, Japanese, Russian, Ukrainian+93	+	+	+	+	Від 0 до 1, positive, negative, neutral, i mixed	+, визначає ваговий коефіцієнт
Microsoft Azure Text Analytics	English, Chinese-Simplified, Spanish, Italian, Japanese, Russian	+	+	+	+	Від 0 до 1, positive, negative, neutral, i mixed	Порівнює з заданими та визначає автоматично
IBM Watson	English, Chinese-Simplified, Spanish, Italian, Japanese, Russian	+	+	+	+	True, false, тільки для English	+

Література

1. Google Cloud AutoML Natural Language [Електронний ресурс] Режим доступу: URL: <https://cloud.google.com/natural-language/automl/docs/>
2. Amazon Comprehend [Електронний ресурс] Режим доступу: URL: <https://aws.amazon.com/comprehend/>
3. Text Analytics [Електронний ресурс] Режим доступу: URL: <https://azure.microsoft.com/en-us/services/cognitive-services/text-analytics/>