

ІНТЕГРАЦІЯ СИСТЕМ ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ ДОКУМЕНТІВ У ВИДАВНИЧО-ПОЛІГРАФІЧНІ ПРОЦЕСИ

Мороз Р. Б.

Національний університет «Львівська політехніка», м. Львів

Для автоматизованого опрацювання неструктурованих та напівструктурованих даних у видавничо-поліграфічних інформаційних потоках на сьогодні доцільно застосовувати інтелектуальну обробку документів (Intelligent Document Processing, IDP). На відміну від традиційних систем оптичного розпізнавання символів (OCR), IDP поєднує технології машинного навчання, обробки природної мови (NLP), глибинного навчання та комп'ютерного зору. Це забезпечує здатність системи розуміти контекст, витягати релевантну інформацію, класифікувати документи та формувати структуровані дані для подальшої обробки в інформаційних системах. Ключовими функціональними компонентами IDP-рішень є автоматичне виявлення типу документа, аналіз його структури, витягування сутностей (таких як імена, дати, адреси, фінансові показники), побудова семантичних зв'язків між фрагментами тексту, а також перевірка та валідація отриманих даних. Системи IDP також передбачають механізми інтерактивного навчання моделей за участі експертів, що забезпечує адаптацію до нових форматів та стилістичних особливостей документів.

У контексті цифрової трансформації документообігу в науковій, адміністративній, правовій і видавничій сферах застосування IDP дозволяє істотно зменшити навантаження на персонал, підвищити точність обробки контенту та забезпечити інтеграцію отриманих даних у системи управління знаннями. Використання IDP стає особливо актуальним для організацій, які працюють з великими масивами текстової інформації, що потребує високої точності аналізу та оперативності прийняття рішень. Інтеграція систем інтелектуальної обробки документів у видавничі процеси сприяє підвищенню якості передредакційної підготовки матеріалів, поліпшенню взаємодії між редакторами, авторами та рецензентами, а також прискоренню прийняття рішень щодо прийняття чи доопрацювання поданих матеріалів. IDP забезпечує автоматичне вилучення релевантних даних з рукописів, таких як ідентифікатори авторів, ключові слова, бібліографічні посилання, конфлікти інтересів тощо, що істотно оптимізує попередню перевірку та індексацію матеріалів.

Таким чином, застосування IDP у науковому видавництві відкриває можливості для автоматизованої перевірки відповідності поданих матеріалів вимогам журналів або конференцій, а також для аналізу структурних і змістових характеристик рукописів за допомогою машинного класифікування та тематичного моделювання. Це дозволяє видавничим онлайн-платформам не лише підвищити ефективність обробки поданого контенту, але й сформувати підґрунтя для інтелектуальної аналітики, що слугує базою для стратегічного розвитку редакційної політики та цифрових сервісів наукової комунікації. Подальші дослідження будуть зосереджені на виборі адекватних онлайн-платформ.