



ПЛАТФОРМА ІЗ ЗАХИСТУ ДІТЕЙ
У ПІВДЕННО-СХІДНІЙ ЄВРОПІ

Що можуть розповісти дані Twitter про міграційну кризу в Європі

У рамках спільного проекту УВКБ ООН і групою ООН за ініціативою «Глобальний пульс» (UN Global Pulse), який вивчає настрої навколо примусового переміщення, використовувалися дані соціальних мереж для збору реакції на поточну ситуацію в режимі реального часу і кращого розуміння переваг великих даних. Первинний аналіз був спрямований на вивчення поведінки біженців відносно один одного і по відношенню до постачальників послуг у Європі. Проте через підвищену увагу засобів масової інформації і негативні настрої у зв'язку з терористичними нападами на континенті дослідники вирішили змістити акцент і проаналізувати стосунки між вимушеними мігрантами та транзитними і приймаючими громадами. Дослідники сподівалися використати отримані результати для вдосконалення оперативних послуг по усій Європі.

Соціальні мережі надають унікальну можливість для аналізу таких ситуацій, як міграційна криза в Європі, при якій переміщуються великі громади, часто без їхнього відстеження, що перешкоджає традиційним методам збору даних (наприклад, опитуванням). Дані, зібрані з таких платформ, як Twitter або Facebook, дозволяють у режимі реального часу отримати уявлення про події від різних груп користувачів з різних місць. У спільному проекті використовувалися дані, зібрані в мережі Twitter, при аналізі різних блоків твитів, таких як хештеги або ключові слова, для створення класифікації. Далі система шукала повідомлення, що містять ідентифіковані слова або пропозиції, які потім вручну кодувалися дослідниками. Ця так звана «навчальна база даних» служила основою, на якій алгоритми класифікували нові пости, додаючи їх у набір даних. Відповіді розподілялися за трьома категоріями: від ксенофобських до нейтральних і, нарешті, до неактуальних.

Дослідники відзначили труднощі в розрізненні акцентів і сленгів при створенні запитів для системи. Враховуючи, що дані містили тільки витяги, було важко ідентифікувати людину, що стоїть за твитом. Крім того, як виявилось, існує деяка неузгодженість відносно мов, наприклад, система не була налаштована на фарсі, оскільки була отримана тільки невелика вибірка. Тому, не дивлячись на те, що дослідження і дозволило по-новому поглянути на аналіз великих даних стосовно гуманітарних питань, проект одночасно виявив і деякі важливі обмеження і проблеми, пов'язані з таким аналізом. Важливо пам'ятати, що, хоча соціальні мережі мають велику методологічну перевагу у поєднанні з традиційними джерелами даних, вони, проте, можуть давати повну картину самі по собі.

Джерела:

UNHCR and UN Global Pulse (2017) Social Media and Forced Displacement: Big Data Analytics and Machine Learning

<https://www.unglobalpulse.org/news/new-paper-un-global-pulse-and-unhcr-explores-use-digital-data-insights-forced-displacement>

Підготовлено Фані Ковесді для «Платформи із захисту дітей в Південно-Східній Європі» у березні 2018.