

УДК 004.91

**Артюхова Д.І., Мазурець О.В.**

*Хмельницький національний університет, Україна*

**ВИКОРИСТАННЯ СИНТАКСИЧНИХ ШАБЛОНІВ ПРИ  
АВТОМАТИЗОВАНОМУ ПОШУКУ КЛЮЧОВИХ ТЕРМІНІВ У  
НАВЧАЛЬНИХ МАТЕРІАЛАХ**

**Artiukhova D.I., Mazurets A.V.**

**USE OF SYNTAXIC TEMPLATES AT AUTOMATED SEARCH OF  
KEY TERMS IN EDUCATIONAL MATERIALS**

В інформаційній технології автоматизованого створення тестів до навчальних матеріалів курсу [1] на етапі пошуку ключових термінів до елементів навчальних матеріалів [2] передбачено необов'язкову можливість коригування одержаних множин ключових термінів, що може підвищити якість вихідних даних, шляхом фільтрування за синтаксичними шаблонами.

За результатами проведеного аналізу вибірки з понад 1300 елементів навчальних матеріалів із визначеними укладачем (автором) репрезентативними множинами ключових термінів, встановлено, що всі елементи множин ключових термінів авторів відповідають наступним закономірностям:

- Кількість слів у терміні  $n=1..6$ .
- Якщо термін є словом ( $n=1$ ), то воно входить до множини іменників  $M_I$ .
- Якщо термін є словосполученням ( $n>1$ ), то до його складу входять елементи множини  $M_M$ . До складу множини  $M_M$  входять множини семантично значущих елементів (іменників  $M_N$ , числівників

$M_{Num}$  і прикметників  $M_A$ ) та семантично зв'язуючих елементів (сполучників  $M_S$  і прийменників  $M_P$ ).

– Якщо  $n > 1$ , то до складу словосполучення входить принаймні один елемент із множини іменників  $M_N$ .

– Якщо  $n > 1$ , то першим ( $k=1$ ) та останнім ( $k=n$ ) словом є елементи множини семантично значущих елементів  $M_N \cup M_{Num} \cup M_A$ .

– Якщо  $n > 1$ , то між елементами словосполучення відсутні розділові знаки (окрім дефісу та апострофу, які є частинами слова).

При цьому, якщо термін є словосполученням, то для слів з його складу визначено наступні обмеження:

– До множини сполучників  $M_S$  належать тільки сполучники, які належать до виду простих, що морфологічною будовою не пов'язані з іншими частинами мови. Їх множина є визначеною:  $M_S = \{i, й, а, чи, та\}$ .

– До множини прийменників  $M_P$  належать тільки прийменники, які належать до виду первинних, тобто які мають праслов'янське походження. Їх множина є визначеною:  $M_P = \{без, у, в, від, для, по, через, при, над, під, до, з, із, як, за\}$ .

– До множини числівників  $M_{Num}$  належать тільки порядкові числівники (наприклад: «*дев'яності*», «*перша*», «*двадцять*»  $\in M_{Num}$ ). Також до цієї множини можна віднести групи цифр розмірності не більше 4 (наприклад: «*2019*», «*1053*», «*256*»  $\in M_{Num}$ ).

Таким чином, після автоматизованого визначення множини ключових термінів навчального матеріалу, з неї видаляються ті терміни, що не відповідають наступній умові:

$$\begin{aligned}
M_T = \{ \langle x_1, x_2, x_3, x_4, x_5, x_6 \rangle \mid & x_1 \in M_N \cup M_{Num} \cup M_A, \\
x_2 \in M_M, x_3 \in M_M, x_4 \in M_M, x_5 \in M_M, & \\
x_6 \in M_N \cup M_{Num} \cup M_A, \langle x_1, x_2, x_3, x_4, x_5, x_6 \rangle \cap M_N \neq \emptyset \}. & \quad (1)
\end{aligned}$$

Одержані в результаті елементи множини ключових термінів можуть бути відсортовані за номінальним значенням оцінки важливості, в їх кількість – обмежена згідно показника щільності слів.

За результатами проведеного аналізу елементів навчальних матеріалів було отримано закономірності, що дозволили побудувати інформаційний шаблон ключових термінів у цифрових навчальних матеріалах. Після автоматизованого визначення множини ключових термінів навчального матеріалу, з неї видаляються терміни, що не відповідають інформаційному шаблону. Наведена фільтрація за допомогою шаблону одержаних елементів множини ключових термінів дозволяє підвищити якість формування множин ключових термінів і якість вирішення ряду похідних задач.

#### Список літератури

1. Бармак О. В., Мазурець О. В., Кліменко В. І. Інформаційна технологія автоматизованого формування тестових завдань / О. В. Бармак, О. В. Мазурець, В. І. Кліменко // Науковий журнал «Вісник Хмельницького національного університету» серія: Технічні науки. Хмельницький, 2017, №5. – С.93-103.
2. Крак Ю. В., Бармак О. В., Мазурець О. В. Практична реалізація інформаційної технології автоматизованого визначення множини семантичних термінів в контенті навчальних матеріалів / Ю. В. Крак, О. В. Бармак, О. В. Мазурець // Науковий журнал «Проблеми програмування». Київ, 2018, №2-3. – С.245-254.