С.М. Новічонок¹, О.А. Усачова¹, О.В. Бабіч¹, О.О. Савченко¹, І.О. Солдатенко², П.В. Беляєв¹

¹Харківський національний університет Повітряних Сил ім. І. Кожедуба, Харків

²Управління регулювання державної авіації України, Київ

МЕТОДИКА СКОРОЧЕННЯ ЧАСОВИХ ВИТРАТ НА ЗАСВОЄННЯ ВИМОГ НОРМАТИВНО-ПРАВОВИХ АКТІВ

В статті запропоновано методику скорочення часових витрат на засвоєння вимог нормативних документів у предметній галузі. Основною метою досліджень є розробка такого способу автоматизації аналізу нормативних документів, який дозволить би на основі порівняння з певним еталоном визначити ранг окремого документу в множині документів. Запропоновано критерії повноти вихідного переліку документів, а також показники, які дозволяють попередньо, без перечитування, оцінити передбачуване застосування окремого документу. Наведена послідовність складання вихідного пулу достатнього обсягу, за допомогою сайту Верховної Ради України. Придатність розглянутої методики була підтверджена збігом рангів наданих розглянутим документам експертами з рангами, наданими програмним забезпеченням. Методика може бути запропонована для широкого кола напрямків діяльності.

Ключові слова: нормативно-правові акти, документ, програмне забезпечення, словосполучення, критерії, семантичний аналіз текстів.

Вступ

Постановка проблеми. Зважаючи на швидкі темпи розвитку суспільства та процеси глобалізації й інформатизації, новою реальністю сучасності стало те, що знання людини застарівають дуже швидко. Принцип “навчання впродовж життя” сьогодні діє в усіх цивілізованих країнах [1].

В таких умовах слід прагнути до того, щоб час на опанування нових знань був якомога меншим, однак не втрачав своєї якості.

За своєю суттю процес засвоєння або усвідомлення вимог нормативно-правових актів (НПА) та інших нормативних документів є процесом навчання. Час, який витрачається на цей процес можна розбити на три основних складових:

- підготовку до навчання;
- безпосередньо навчання;
- перевірку успішності навчання.

Традиційно, подібні дослідження присвячуються безпосередньо навчанню та способам перевірки набутих знань.

Отримані в результаті таких досліджень методики та форми навчання ґрунтуються на урахуванні особливостей сприйняття матеріалу тими, хто навчається [2–4]. Однак, складова підготовки до навчання набуває все більшої важливості через необхідність аналізу великої кількості нової інформації, для добору сучасного наповнення змісту навчання. Постійна повторюваність цього процесу підвищує

значущість цієї складової.

Слід також враховувати, що в Україні на теперішній час більшість фахівців, за різних причин, повинні підвищувати свій рівень підготовки шляхом самоосвіти.

До самоосвіти також слід віднести безперервне вивчення змін у відповідних НПА та інших нормативних документах. Однак більшість фахівців не мають необхідних навичок швидкого відбору інформації і не мають достатнього часу на ознайомлення з великою кількістю змінюваних документів традиційним способом.

Використання пошукових систем дозволяє швидко отримати певний перелік документів, але час для визначення їх придатності для зазначених цілей дуже великий. Легко показати, що для обсягу близько 250 тисяч слів (що відповідає приблизно 5-10 документам) необхідно біля 12 годин “чистого часу” тільки на прочитання та не менш 72 годин на усвідомлення матеріалу. Ще більше часу буде потрібно на складання докладного реферату за цим обсягом інформації.

Дослідження, проведені авторами, показали, що навіть у відносно вузькому напрямку (наприклад аеродромно-технічне забезпечення польотів) перелік необхідних документів може бути більше сотні. Тобто на традиційну підготовку матеріалу на основі такого обсягу інформації потрібно не менш як 90 робочих діб. Тільки підготовчий період підготовки матеріалу для підтримання юридичних знань профе-

сіоналом на актуальному рівні вимагає значних витрат часу (до чверті загального бюджету часу), що не є прийнятним.

Таким чином, можна стверджувати, що дослідження, спрямовані на пошук доступних широкому колу фахівців способів скорочення часових витрат на підготовчий етап до засвоєння вимог НПА та інших нормативних документів, є актуальними.

Аналіз останніх досліджень і публікацій. Підготовка до навчання взагалі і до вивчення НПА, зокрема, пов'язана з пошуком, накопиченням та аналізом інформації.

Звичайно робота починається з бібліографії. Щоб максимально швидко знайти бібліографічну інформацію й скласти перелік літератури, необхідно знати коло джерел за своєю галуззю знань, і загального характеру, звідки можна отримати необхідні відомості про видання.

Достатньо широкий огляд видів інформації та способів її пошуку наведено, наприклад в [5–6]. В [5] відмічається, що, по-перше, в загальному випадку далеко не вся потенційно корисна інформація (особливо ретроспективна) існує в електронному вигляді, і поки немає підстав очікувати швидкого перетворення всієї інформації на паперових носіях в електронну форму.

По-друге, багато представників гуманітарних і суспільних наук не погоджуються публікувати свої роботи в електронному вигляді, вважаючи подібний вид публікації найменш престижним.

Однак, НПА державного рівня повинні бути оприлюднені в Інтернет та розміщені в спеціальній інформаційній системі [7]. Інформація щодо електронних джерел правової інформації та способів її пошуку наведено в [8]. Отже застосування сучасних засобів автоматизації для пошуку правової інформації є можливим.

Найбільшого досвіду автоматичного аналізу текстової інформації набуто у галузі суспільних наук. Застосування комп'ютерів в правовій сфері не є новиною. Існує поняття – “правова інформатика”. Широкий огляд програмних засобів роботи з правовою інформацією, в широкому сенсі, наведено в [9]. Виходячи з наведеної в [8] та [9] інформації можна сказати, що система пошуку НПА набула достатнього розвитку в Україні для пошуку документа за реквізитами чи за ключовими словами.

Проведені дослідження показали, що можливостей автоматичного добору документів, за певним напрямком діяльності, поки що не пропонується такими системами. Тобто відбір інформації за придатністю, та визначення достатності її обсягу користувачу слід проводити самостійно. Загальне представлення, щодо методів аналізу документів з метою встановлення придатності їх для використання можна отримати з роботи [10].

Методам аналізу НПА як джерел соціально-правових відносин, виникаючих через їх застосування, присвячена робота [11].

Взагалі, аналіз текстової інформації за певними ознаками прийнято називати контент-аналізом [12], який на теперішній час є одним з основних методів досліджень в галузі суспільних наук [13]. Узагальнюючи наведений в згаданих джерелах матеріал, можна сказати, що автоматизований пошук НПА доступний, але потрібні методики ефективного використання засобів такого пошуку.

В основу розробки такої методики можуть бути покладені методи частотного аналізу текстів.

Мета статті – розробити методику скорочення часових витрат на засвоєння вимог нормативних документів, яка полегшує фахівцю знаходження достатньо повного переліку НПА (скласти необхідну бібліографію) в конкретній предметній галузі, визначення першочерговості їх вивчення. Для досягнення поставленої мети необхідно визначити набір показників, які давали б представлення про відносний корисний обсяг та широту інформації, що міститься в документах, які розглядаються, без необхідності безпосереднього традиційного читання цих документів.

Виклад основного матеріалу

При подальшому викладенні матеріалу у якості основного Інтернет-джерела НПА будемо розглядати тільки інформаційну систему Верховної Ради України (далі – ВР).

У випадку, коли документи розміщені в Інтернеті, але не підтримуються спеціальною інформаційною системою, знайти їх можна за допомогою, гіпертекстових посилань, використання пошукових машин, або спеціальних програм – спайдерів, які в автоматичному режимі переглядають web-сторінки, відшуковуючи потрібну інформацію [5].

Якщо в зазначених засобах знайти необхідні НПА не вдається, прийдеться їх шукати традиційними способами, розгляд яких виходить за рамки статті.

Проте, навіть в цьому випадку можна порекомендувати знайдені НПА конвертувати в цифрову форму і організовувати у вигляді електронної бібліотеки. Для створення такої бібліотеки існує багато засобів, зокрема менеджер електронної бібліотеки Calibre [14].

Отже первинний пошук документів пропонуємо виконувати за допомогою розділу “Законодавство” офіційного порталу Верховної Ради України [15]. На порталі можна знайти необхідні НПА та, скориставшись вкладкою “Картка документу” (рис. 1), можна отримати інформацію про наявні на сайті документи, пов'язані з тим, що розглядаються.

Після того, як перелік необхідних документів

буде складений і документи отримані (завантажені), настає етап їх вивчення або аналізу.

Тут слід звернути увагу на дві вимоги:

– гарантування того, що буде отримано достат-

ньо повний перелік НПА для обраної галузі знань;

– скорочення часу на підготовку до вивчення саме того матеріалу, який належить до необхідної галузі знань.

Про затвердження Правил пошуково-рятувального забезпечення польотів державної авіації України

Рис. 1. Виклик картки документу
Джерело: розроблено авторами за даними [7].

Для оцінки ступеня виконання першої з цих вимог слід визначити критерій повноти переліку НПА та розробити методику складання достатньо повного переліку НПА, який би задовольняв визначеному критерію.

Очевидно, що при виконанні другої вимоги слід прагнути підготувати до вивчення саме такий обсяг матеріалу який, як очікується, може бути опановано в наявний час. Найпростішим рішенням буде скорочення обсягу матеріалу для вивчення. При цьому скорочення обсягу матеріалу повинне якомога менше вплинути на його корисність. Інакше кажучи, треба розташувати наявні джерела інформації в порядку корисності і починати вивчення з найбільш корисних. Корисність матеріалу в документі, може бути оцінена зокрема за широтою охоплення питань необхідної діяльності та обсягом шуканої інформації.

Таким чином, методику скорочення часових витрат на засвоєння вимог НПА можна розподілити на дві окремі методики: методику складання достатньо повного переліку НПА та методики аналізу обраних текстів за широтою охопленої інформації та її обсягом. Для першої з них необхідно також розробити критерій достатності переліку НПА.

Методика складання достатнього переліку нормативно-правових актів

Дослідження, що наводяться у статті, були частиною досліджень щодо організації аеродромно-технічного забезпечення (далі – АТЗ) польотів Державної авіації України, тому матеріал статті ілюстровано прикладами НПА АТЗ польотів. Докладне вивчення НПА АТЗ польотів показало, що в них не завжди, а точніше, досить рідко посилаються на документи, що уточнюють зміст основного матеріалу. Таким чином, виходячи зі змісту документів, однозначно визначити, чи є наданий перелік достатньо повним, неможна.

Запропонуємо наступний критерій достатності

переліку НПА. Перелік НПА може вважатись достатнім, якщо змістовна частина НПА в повному обсязі забезпечує однозначне тлумачення технології (у широкому розумінні) здійснення професійної діяльності фахівця. Звідси витікає розуміння того, що на окремій посаді такий перелік може мати відмінності, через те, що різними посадовцями здійснюються різні функції. Таким чином, отримати повний перелік НПА можна тільки тоді, коли досконало вивчений та задокументований технологічний процес. Однак, опис процесу міститься в НПА та інших документах. Тут виникає протиріччя: складання повного переліку документів вимагає наявності повного опису технології, складання повного опису технології вимагає повного переліку НПА.

Запропонуємо методику складання достатньо повного переліку НПА.

1. Складання первинного переліку документів:
 - на основі певного первинного переліку НПА (що є в наявності у професіонала);
 - на основі виявлених взаємозв'язків, що надають юридичні портали (в нашому випадку інформаційна система Верховної Ради України);
 - на основі знань експертів (бесіди з фахівцями, вивчення відповідних матеріалів поза межами законодавчої бази України).
2. На основі змісту первинного переліку документів та досвіду експертів створюється загальний опис діяльності, як технологічного процесу, для чого із документів відбирається наступний матеріал:
 - об'єкти;
 - суб'єкти;
 - функції;
 - перелік документів, які визначають кожну функцію та учасників процесу.

3. Уточнити перелік, складений на першому етапі шляхом додавання до переліку НПА, які описують функції, об'єкти, суб'єкти, що додатково були виявлені при аналізі діяльності, як технологічного процесу.

4. У разі виявлення застарілих НПА або відсутності документів, що відповідають певній функції технологічного процесу, можна звернутись до керівництва, за фактом виявленого недоліку або з пропозиціями щодо його усунення.

Множину НПА, що входять до первинного переліку позначимо R . НПА множини R мають певні зв'язки між собою. НПА з множини R сумісно з наявними зв'язками можна представити у вигляді графа. При чому, вершини графа r_i будуть мати сенс окремих документів, індекс i номера документа в множині R , а ребра x_j будуть показувати зв'язки, де j номер ребра в множині ребер X .

Дуга призначається тільки в тому випадку, як-

що зв'язок між НПА прямо вказаний в тексті хоча б одного з них. Звичайно, в певному виді діяльності є відомим "основний" НПА, тобто документ, що визначає основні визначення, призначення та базовий зміст діяльності. Принаймні, на підставі досвіду, зазвичай відокремлюють такий документ від решти.

Позначимо такий документ r_0 , він буде основною вершиною графу.

В нашому випадку, в якості НПА r_0 було використано наказ Міністерства оборони України від 24.12.2015 року № 761 (далі Наказ № 761) [16]. Користуючись сайтом ВР, знайдемо документи, що з ним пов'язані.

Такі документи наведені в табл. 1.

Таблиця 1

Перелік документів пов'язаних
із наказом №761

Вершина графа	Назва документу
r_1	Повітряний кодекс України. Закон України від 19.05.2011 № 3393-VI. – Первинний акт (https://zakon.rada.gov.ua/laws/show/3393-17)
r_2	Наказ України Міноборони України від 05.01.2015 № 2. Про затвердження Правил виконання польотів державної авіації (https://zakon.rada.gov.ua/laws/show/z0082-15)
r_3	Наказ Міноборони України від 22.01.2014 № 47. Про затвердження Інструкції з організації контролю якості стиснутих та зріджених газів у державній авіації України https://zakon.rada.gov.ua/laws/show/z0234-14
r_4	Наказ Міноборони України від 01.07.2013 № 441. Про затвердження Інструкції з експлуатації аеродромів державної авіації України https://zakon.rada.gov.ua/laws/show/z1229-13
r_5	Постанова Кабінету Міністрів України від 10.10.2001 № 1306 Про Правила дорожнього руху https://zakon.rada.gov.ua/laws/show/1306-2001-п

Джерело: розроблено авторами.

Результати пошуку на даному етапі подано у вигляді графу на рис. 2.

Рис. 2. Результат первинного пошуку НПА з АТЗ польотів на сайті ВР України
Джерело: розроблено авторами.

Просуваючись далі від документа до документа, вдалося отримати граф із 134 елементів (рис. 3). Тобто перелік необхідних для вивчення НПА склав 134 документа. Незалежно від пошуку на сайті ВР, на підставі досвіду до переліку було додано шість документів. Для спрощення рисунку у випадках, коли існує зв'язок одного документа до багатьох замість окремих ребер зображено одне ребро більш жирною лінією. Також вважалося, що коли мова йде про конкретну техніку (чи технологію), слід засто-

совувати відповідне технічне керівництво, стандарт, тощо.

На підставі змісту доданих документів встановити зв'язки з іншими документами на пряму (за текстом документів) не вдалося. Таким чином, вершини $r_{135} - r_{|R|}$ є висячими.

З певною точністю можна стверджувати, що цей перелік є достатнім, але його обсяг завеликий і може виявитися навіть надлишковим для конкретного фахівця. Таким чином, треба відібрати з нього найбільш необхідні документи. Відбір документів можна здійснювати за наступними показниками – глибина та широта висвітлених питань. За цими показниками можна складати рейтинги. На основі отриманих рейтингів та інформації про загальний обсяг документів можна зробити обмежені переліки документів для першочергового вивчення звичайним способом. Здійснення такого відбору традиційним способом займає багато часу з причин, зазначених вище. Таким чином, необхідно розробити методику, що дозволить виконувати такий відбір автоматично.

Розробка методики аналізу обраних текстів за широтою охопленої інформації та її обсягом

При зменшенні обсягу матеріалу слід бути впевненим, що не втрачаються його важливі складові.

На теперішній час множина НПА R може складатись з підмножини R' текстових та R'' графічних. Таким чином, $R = R' \cup R''$. До множини R'' будемо відносити суто графічні документи. Якщо документ має пояснювальну текстову частину, тоді

пропонується відносити його до множини R' . В цьому випадку $|R''| \ll |R'|$.

Швидкість сприйняття графічної інформації значно більша ніж текстова. На основі графічної інформації фахівець легше та скоріше усвідомить ступінь її корисності.

Отже визначити рейтинг документу з множини R'' можна без застосування засобів автоматизації. Текст сприймається значно повільніше, отже в подальшому будемо розглядати множину R' .

Рис. 3. Аналіз зв'язків НПА, що мають відношення до АТЗ польотів
Джерело: розроблено авторами за даними [7].

Відомим способом пошуку необхідної інформації у текстовому документі є пошук за ключовими словами. Існує велика кількість пошукових систем [17–19]. Найбільш потужні системи мають можливість шукати не тільки конкретні слова, але і словосполучення в різних комбінаціях. Одне з передових місць в цьому питанні займає пошукова система Google [20]. Однак пошук інформації сам по собі не допоможе скоротити час на опанування документу. Для того, щоб шукати інформацію треба знати, що саме треба шукати, а при створенні первинної бібліографії користувач може мати лише певне представлення про зміст наявної в документах інформації.

В першу чергу, за допомогою машинного аналізу треба з'ясувати, чи містить документ інформацію щодо шуканої предметної галузі і якщо так, то в якому обсязі.

Критерієм наявності необхідної інформації

може бути частота появи в документі слів або словосполучень визначеної предметної галузі. На наш погляд словосполучення є більш інформативними. Наприклад, якщо текст містить слова “літак” та “обслуговування” це зовсім не означає, що мова йде про обслуговування літака.

Позначимо множину словосполучень, яка міститься в досліджуваному R_i -му документі A . Множину словосполучень, яка необхідна для описання технологічних процесів у необхідному виді діяльності B . Позначимо множину корисних для цього словосполучень, що потрапила до досліджуваного документу C . Тоді $C = A \cap B$. З множини B , а отже і з множини C можна відокремити унікальні словосполучення B' та C' . Тоді чим більше потужність $|C|$ тим більший обсяг інформації в цьому документі належить до необхідної діяльності.

Чим більш потужність $|C'|$, тим ширше охоплення необхідних питань в документі.

Для можливості розподілу досліджуваних документів за характером матеріалу розділимо множину словосполучень B на підмножини.

Слід зазначити, що, теоретично, вихідну множину B можна розділяти на будь-яку кількість підмножин, але у зв'язку з тим, що людина здатна утримувати в короткостроковій пам'яті одночасно 7 ± 2 елементи (число Інґве-Міллера [21]). Очевидно, що без додаткового інструментарію (передбачається, що експерт буде діяти саме так) порівнювати між собою експерт зможе таку кількість елементів. Отже, нормальна кількість градацій від 5 до 9.

Розподілимо B' на 5 підмножин:

B_1 – безпосередньо містить словосполучення, які прямо вказують на необхідну діяльність;

B_2 – словосполучення мають безпосереднє відношення скоріше до шуканого виду діяльності, ніж до інших видів;

B_3 – словосполучення мають значне відношення до процесів у необхідному виді діяльності, але і для інших процесів;

B_4 – мають значне відношення до процесів, які підтримують процеси необхідної діяльності;

B_5 – решта словосполучень, що мають певне відношення до необхідного виду діяльності.

Оскільки $C \in B$, а $C' \in C$ то ці підмножини також розбіються на аналогічні підмножини:

$$C = C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_5,$$

$$C' = C'_1 \cup C'_2 \cup C'_3 \cup C'_4 \cup C'_5.$$

Частота появи в досліджуваному тексті корисних словосполучень визначатиметься за формулою:

$$v = \frac{|C|}{|A|}, \quad (1)$$

або за підмножинами

$$v_i = \frac{|C_i|}{|A|}, \quad (2)$$

де $i \in \{1, 2, 3, 4, 5\}$ – індекс підмножин множин B, B', C, C' .

Величини v можна використовувати для порівняння між собою документів різного обсягу щодо відносного вмісту в досліджуваному документі матеріалу, який відповідає необхідній діяльності. Більші значення v_i будуть свідчити про більший відносний вміст в тексті словосполучень відповідної підгрупи. При вірному підборі словосполучень до множини B' можна вважати, що більші значення v_i відповідатимуть документу, який має директивний характер саме для необхідної діяльності, більші зна-

чення v_2 свідчать про документ з технологічних особливостей. Більші значення v_3 свідчать про наявність загальної інформації, яка в той же час важлива для необхідної діяльності. Більші значення v_4 свідчать про те, що документ скоріш описує суміжні з необхідною діяльністю процеси. Більші значення v_5 відповідатимуть документу, що визначає загальні процеси на фоні яких відбувається необхідна діяльність. В загальному випадку $|B'_1| \neq |B'_2| \neq |B'_3| \neq |B'_4| \neq |B'_5|$, тому замість потужностей значення $|C'_i|$ для порівнянь документів між собою слід також використовувати частоти. Однак в цьому випадку можна використати різні способи підрахунку частот. Величина

$$v'_i = \frac{|C'_i|}{|A|}, \quad (3)$$

буде мати сенс аналогічний величинам v_i , а величина

$$v''_i = \frac{|C'_i|}{|B'_i|}, \quad (4)$$

буде показувати вміст унікальних корисних словосполучень відносно можливого. Більші значення v''_i відповідають більшому тематичному охопленню. Тобто, коли $v''_i = 1$, це означає, що всі можливі слова i -ї підмножини множини B' потрапили до множини A . А це означає, що досліджуваний документ стосується усіх можливих питань галузі знань.

Таким чином, величини v, v_i, v'_i, v''_i є частковими показниками досліджуваного тексту на приналежність до необхідного виду діяльності за різними ознаками. Запропонований спосіб є частковим випадком реалізації частотного аналізу текстів, який широко використовується у лінгвістиці, а саме в контент-аналізі [12–13]. Статистичний показник терміна документа доцільно використовувати із попередньою обробкою [22]. Це обумовлено необхідністю врахування значущості терміна. Існує два підходи. Один базується на заздалегідь підготовленій оцінці важливості термінів, а інший на загальних мовних законах, які надають можливість визначити важливість терміну на підставі лише частоти.

До часткових показників v, v_i, v'_i може бути застосований перший підхід, а до v''_i другий. Це обумовлено тим, що важливість (рейтинг) групи термінів відносно всього тексту визначений потраплянням терміну до підмножин B_i множини B , а внутрішній розподіл за важливістю між елементами кожної з підмножин B_i не передбачається. За першим підходом, при наявності заздалегідь визначених n підмножин B_i , тобто при наявності n града-

цій значимості, ваговий коефіцієнт для кожної градації можна визначити за формулою

$$\lambda_{1i} = \frac{n - R_i + 1}{\sum_{i=1}^n R_i}, \quad (5)$$

де R_i – ранг підмножини B_i у певному виді діяльності;

i – індекс рангу підмножини.

При визначенні вагового коефіцієнту за другим підходом (на основі частоти) слід враховувати те, що значення частоти повторення терміна, який частіше інших був вжитий у документі, абсолютно не впливає на його значимість. Більшого значення буде мати ранг частоти, який дозволяє зрівнювати значимість самих термінів, що зустрічаються, будь-яких текстів і одночасно розподіляє значимості термінів усередині одного тексту. При цьому терміни з однаковою частотою повторення, маючи однаковий ранг частоти, будуть однаково значимі для передачі змісту.

В [23] для врахування частоти повторення при визначенні вагового коефіцієнта використано наступну формулу

$$\lambda_{2i} = 1 - \log_{\max}(r) r_i, \quad (6)$$

де r_i – ранг частоти терміна (ранг частоти найбільш розповсюдженого слова дорівнює одиниці);

i – індекс рангу частоти.

Формула (6) відповідає закону Зіпфа [24], який стверджує, що слова з найбільшими рангами частоти будуть мати найменший вплив на сенс тексту. Значення λ_2 для слів з максимальним рангом частоти буде дорівнювати нулю.

Відомо, що сума вагових коефіцієнтів повинна дорівнювати одиниці, тобто

$$\sum_{i=1}^n \lambda_i = 1. \quad (7)$$

Ця умова не завжди виконується для формули (6). Тому пропонується використання наступної формули

$$\lambda_{2i} = \frac{(1 - \log_{\max}(r) r_i)}{\sum_{i=1}^n (1 - \log_{\max}(r) r_i)}. \quad (8)$$

Таким чином, часткові показники будуть визначатись за формулами:

$$w_i = \lambda_{2i} \cdot v_i; \quad (9)$$

$$w'_i = \lambda_{2i} \cdot v'_i; \quad (10)$$

$$w''_i = \lambda_{2i} \cdot v''. \quad (11)$$

На підставі (9) та (10) можуть бути отримані формули узагальнених показників

$$W_i = \sum_{i=1}^n w_i; \quad (12)$$

$$W' = \sum_{i=1}^n w'_i; \quad (13)$$

$$W'' = \sum_{i=1}^n w''_i. \quad (14)$$

Для можливості застосування запропонованої методики отримання показників наявності необхідної інформації в досліджуваному тексті необхідно визначити зміст множини B і розподілити її за підмножинами B_i .

Для цього пропонується спочатку обрати найбільш змістовний документ у предметній галузі, розбити його на окремі елементи і тим самим утворити множину A . Потім за допомогою експертного оцінювання, відокремити з множини A множину B одночасно розподіляючи її на підмножини B_i .

Запропонована методика була реалізована у спеціальному програмному забезпеченні.

За допомогою цього програмного забезпечення було проведено аналіз 134 документів, що були виявлені та отримані за допомогою інформаційної системи ВР. Так, запропонований метод дозволив автоматично, без прочитання документів відсортувати перелік таким чином, що 86% найбільш корисних документів опинились на початку переліку. Решта 14% опинилась в першій чверті загального рейтингу.

Щодо часових характеристик, то зміст документів з орієнтовним загальним обсягом 4700 сторінок був оброблений за 16 хвилин. В результаті чого і було отримано ранжований перелік документів. При цьому, для кожного документу був сформований окремий файл-звіт в якому зазначено:

– перелік слів, що стосуються АТЗ польотів, що увійшли до документу із зазначенням частоти та значимості кожного терміну;

– частотні характеристики, щодо відношення до АТЗ польотів та корисності документу в цьому сенсі, в цілому;

В нашому випадку підготовка множини вихідних документів склала близько 8 год. Тобто можна сказати що, при заздалегідь підготовленому еталонному файлі етап підготовки до вивчення НПА, що стосуються АТЗ польотів склав приблизно один робочий день. Зазначимо, що підготовка еталонного файлу, а саме рейтингування словосполучень експертами зайняло близько тижня. Термінологія виду діяльності змінюється достатньо повільно і в подальшому внесення коректив до еталонного переліку словосполучень буде займати незначний час.

Висновки

Таким чином в результаті проведених досліджень було розроблено методику скорочення часових витрат на засвоєння вимог НПА, що стосуються необхідної діяльності, яка в свою чергу складається з двох окремих методик – методики складання достатньо повного переліку документів для охоплення необхідної діяльності та методики аналізу обраних текстів за інформаційним охопленням та обсягом. Розроблена методика дозволяє отримати ранжований перелік документів, що визначають необхідну діяльність без безпосереднього ознайомлення із змістом цих документів. Методика аналізу обраних текстів за широтою охопленої інформації та її обсягом була реалізована в спеціальному програмному забезпеченні, що дозволило одночасно перевірити обидві складові методики скорочення часових витрат на засвоєння вимог нормативних документів в предметній галузі.

Автоматично (без прочитання документів) вдалося визначити 86% найбільш важливих документів. Часові показники для визначення найбільш важливих документів наступні:

- підготовка множини із 134 документів у предметній галузі склала близько 8 годин;
- автоматична обробка множини із 134, з орієнтовним обсягом 4700 сторінок, склала менше 20 хвилин.

Можна зробити висновок що, при заздалегідь підготовленому еталонному файлі етап підготовки до вивчення НПА загальним обсягом до 5000 сторінок складе приблизно один робочий день.

Розвиток запропонованих методик, щодо аналізу текстів може полягати в автоматизації роботи в пошукових системах (за допомогою так званих ботів) та застосуванні в еталонній множині не тільки словосполучень, а й цілих речень, або висловів.

Список літератури

1. Маркозова О. О. Навчання впродовж життя – необхідна передумова досягнення життєвого успіху людини. *Вісник Національного університету “Юридична академія України імені Ярослава Мудрого”*. 2016. № 2. С. 198–205. URL: http://nbuv.gov.ua/UJRN/vnuuaup_2016_2_19 (дата звернення: 03.05.2021).
2. Шляхи підвищення якості освіти. Київський університет імені Бориса Грінченка. *Педагогічний інститут* : веб-сайт. URL: https://kubg.edu.ua/images/stories/Departaments/yakist_ospity/pi_yakist_ospity.pdf (дата звернення: 03.05.2021).
3. Новіченок С. М., Тарасова В. В. Использование опорных сигналов при изучении технических дисциплин. *Збірник наукових праць “Теорія та методика навчання фундаментальних дисциплін у вищій технічній школі”*. 2003. С. 139–143. URL: <http://elibrary.kdpu.edu.ua/xmlui/handle/0564/651> (дата звернення: 03.05.2021).
4. Ключко В. І., Ключко Н. О. Формування методологічної компетентності студентів технічних університетів. *Збірник наукових праць “Теорія та методика навчання фундаментальних дисциплін у вищій школі”*. 2008. № 5. С. 151–158.
5. Косиченко О. О. Правові інформаційні ресурси Інтернет. Дніпро : ДДУВС, 2017. 92 с.
6. Тарасова В. В., Новіченок С. М., Ручка О. О. Науково-дослідна робота : метод. посіб. Харків : ХНУПС, 2020. 212 с.
7. Положення про веб-ресурси Верховної Ради України : Розпорядження Голови Верховної Ради України від 19 трав. 2015 р. № 699. URL: <https://zakon.rada.gov.ua/laws/show/699/15-%D1%80%D0%B3#Text> (дата звернення: 03.05.2021).
8. Косиченко О. О. Правові інформаційні ресурси інтернет. Дніпро : ДДУВС, 2017. 91 с.
9. Правовая информатика / Акопов Г. Л. та ін. Ростов-на-Дону : Ростовский юридический институт МВД России, 2006. 149 с. URL: <https://www.twirpx.com/file/243002/> (дата звернення: 30.04.2021).
10. Методы сбора информации в социологическом исследовании. *Омский университет*. веб-сайт : URL: <http://www.univer.omsk.su/omsk-old/socstuds/crawlinfo/index.html> (дата звернення: 30.04.2021).
11. Каргин К. В. Юридические документы : автореф. дис. на соискание уч. степени канд. юрид-х наук : 12.00.01. Нижний Новгород, 2005. 28 с.
12. Аверьянов Л. Я. Контент-анализ 2007 г. *Электронная библиотека научной литературы*. 2007. 286 с. URL: http://window.edu.ru/resource/277/54277/files/averjanov_kontent.pdf (дата звернення: 30.04.2021).
13. Контент-анализ. *Wikipedia*. веб-сайт : URL: <https://uk.wikipedia.org/wiki/Контент-анализ>.
14. Calibre e-book management. *Calibre*. веб-сайт : URL: <https://calibre-ebook.com/about> (дата звернення: 30.04.2021).
15. Верховна Рада України. *Законодавство України*. веб-сайт : URL: <https://zakon.rada.gov.ua/laws> (дата звернення: 30.04.2021).
16. Про затвердження Правил аеродромно-технічного забезпечення польотів повітряних суден державної авіації України : Наказ міністерства оборони України від 24 груд. 2015 р. № 761. URL: <https://ips.ligazakon.net/document/view/Re28260> (дата звернення: 30.04.2021).
17. Пошукова система. *Wikipedia*. веб-сайт : URL: https://uk.wikipedia.org/wiki/Пошукова_система (дата звернення: 30.04.2021).
18. Пошук вмісту за допомогою функції розширеного пошуку. *Сайт підтримки Microsoft office*. веб-сайт : URL: <https://support.office.com/uk-ua/article/Пошук-вмісту-за-допомогою-функції-розширеного-пошуку-2979bb4d-56e9-4ccd-8316-472eba777349> (дата звернення: 30.04.2021).
19. Всі документи бази даних “Законодавство України”. *Верховна Рада України*. веб-сайт : URL: <https://zakon.rada.gov.ua/laws/main/a#Find>.
20. Пошукові системи світу, статистика 2018. *Marketer*. веб-сайт : URL: <https://marketer.ua/ua/search-engine-stat-2018/> (дата звернення: 30.04.2021).

21. Магічне число сім, плюс-мінус два. *Wikipedia*. веб-сайт : URL: [https://uk.wikipedia.org/wiki/ Магічне_число_сім_плюс-мінус_два](https://uk.wikipedia.org/wiki/Магічне_число_сім_плюс-мінус_два).

22. Аюшеева Н. Н., Кушеева Т. Н. Способ расчета весовых коэффициентов вершин семантической сети научного текста. *Фундаментальные исследования*. 2012. № 6(3). С. 626–630. URL: <https://www.fundamental-research.ru/ru/article/view?id=30087> (дата звернення: 30.04.2021).

23. Орлов А. И. Высокие статистические технологии: Экспертные оценки. Москва : Экзамен, 2007. 372 с.

Надійшла до редколегії 20.05.2021

Схвалена до друку 13.07.2021

Відомості про авторів:

Новічонко Сергій Михайлович

кандидат технічних наук доцент
старший науковий співробітник
Харківського національного університету
Повітряних Сил ім. І. Кожедуба,
Харків, Україна
<https://orcid.org/0000-0003-1269-5797>

Усачова Ольга Анатоліївна

кандидат технічних наук старший науковий співробітник
начальник науково-дослідного відділу
Харківського національного університету
Повітряних Сил ім. І. Кожедуба,
Харків, Україна
<https://orcid.org/0000-0003-0864-5017>

Бабіч Олексій Вікторович

науковий співробітник
Харківського національного університету
Повітряних Сил ім. І. Кожедуба,
Харків, Україна
<https://orcid.org/0000-0003-4171-6008>

Савченко Ольга Олександрівна

кандидат філософських наук доцент
професор кафедри іноземних мов
Харківського національного університету
Повітряних Сил ім. І. Кожедуба,
Харків, Україна
<https://orcid.org/0000-0003-0085-7189>

Солдатенко Ігор Олегович

начальник відділу управління регулювання
держаної авіації України,
Київ, Україна
<https://orcid.org/0000-0001-5199-5499>

Беляєв Павло Васильович

науковий співробітник
Харківського національного університету
Повітряних Сил ім. І. Кожедуба,
Харків, Україна
<https://orcid.org/0000-0003-0650-6232>

Information about the authors:

Serhii Novichonok

Candidate of Technical Sciences Associate Professor
Senior Researcher Associate
of Ivan Kozhedub Kharkiv
National Air Force University,
Kharkiv, Ukraine
<https://orcid.org/0000-0003-1269-5797>

Olga Usachova

Candidate of Technical Sciences Senior Researcher
Chief of Scientific Research Department
of Ivan Kozhedub Kharkiv
National Air Force University
Kharkiv, Ukraine
<https://orcid.org/0000-0003-0864-5017>

Oleksii Babich

Research Associate
of Ivan Kozhedub Kharkiv
National Air Force University,
Kharkiv, Ukraine
<https://orcid.org/0000-0003-4171-6008>

Olga Savchenko

PhD Associate Professor
Professor of Foreign languages Department
of Ivan Kozhedub Kharkiv
National Air Force University
Kharkiv, Ukraine
<https://orcid.org/0000-0003-0085-7189>

Ihor Soldatenko

Head of the Regulatory Department
State Aviation of Ukraine,
Kyiv, Ukraine
<https://orcid.org/0000-0001-5199-5499>

Pavlo Bieliaiev

Research Associate
of Ivan Kozhedub Kharkiv
National Air Force University,
Kharkiv, Ukraine
<https://orcid.org/0000-0003-0650-6232>

МЕТОДИКА СОКРАЩЕНИЯ ВРЕМЕННЫХ ЗАТРАТ НА УСВОЕНИЕ ТРЕБОВАНИЙ НОРМАТИВНЫХ ДОКУМЕНТОВ

С.М. Новичонко, О.А. Усачова, А.В. Бабич, О.А. Савченко, И.О. Солдатенко, П.В. Беляев

В статье предложена методика сокращения временных затрат на усвоение требований нормативных документов в предметной области. Основной целью исследований является разработка способа автоматизации анализа нормативных документов, который позволит на основе сравнения с определенным эталоном определить ранг отдельного документа в пуле. Предложены критерии полноты исходных перечисленных документов, а также показатели, которые позволяют предварительно, без прочтения, оценить предполагаемую применимость отдельного документа. Предложена последовательность составления исходного пула достаточного объема, с помощью сайта Верховного Совета Украины. Приведен пример схемы взаимосвязей руководящих документов, которую удалось получить без непосредст-

венного ознайомлення с ними, на основе сведений, предоставленных поисковой системой сайта Верховного Совета Украины. Эффективность предложенной методики была подтверждена высоким уровнем совпадения рангов предоставленных экспертами к рассмотрению документов по рангам, предоставленным программным обеспечением. Методика может быть предложена для широкого круга отраслей.

Ключевые слова: нормативно-правовые акты, документ, программное обеспечение, словосочетание, критерии, семантический анализ текстов.

METHODOLOGY OF TIME REDUCING FOR MASTERING REGULATORY DOCUMENTS

S. Novichonok, O. Usachova, O. Babich, O. Savchenko, I. Soldatenko, P. Biliyaev

The article proposes a time reducing methodology for the understanding of regulatory documents requirements in the subject area. The article does not deal with aspects of the peculiarities of individual learners' material perception, however, it describes the time reducing approach for preliminary selection of necessary material from a defined provided pool. The main goal of the investigation is to develop an automatic analysis method of output documents, which will allow, based on comparison with a certain standard, to determine the rank of an individual document in the pool. Criteria of the original listed documents completeness, as well as indicators that allow preliminary, without reading, to assess the intended applicability of a separate document are proposed. As a result of the studies, the corresponding indicators and criteria were found. It has been determined that the list of guidance documents can be considered sufficiently complete if the content of these documents fully provides an unambiguous interpretation of the technology for the implementation of the specialist's professional activity. This technique application is possible in case when words are sufficient to convey the content, while it is assumed that the language of the studied documents will be Ukrainian or have a similar structure. The technique is not applicable to documents where formulas, pictograms or other graphic images are used to convey content. Since most of the current Ukrainian governing documents correspond to the described conditions, the proposed methodology fully meets the needs of specialists who are guided in their work by the regulatory framework of Ukraine. The sequence of sufficient volume of initial pool assembling is shown using the website of the Supreme Council of Ukraine. Diagram example of governing documents relationship, which was obtained without direct acquaintance with them, based on information provided by the search engine of the website of the Supreme Council of Ukraine, is given. The list of documents in the field of aircraft ground handling of the State Aviation of Ukraine was used to check the proposed methodology. The pool volume is more than 130 documents with a total volume of 5000 pages. Total automation time of pool ranking is less than 20 minutes. The effectiveness of the proposed methodology was confirmed by a high level of coincidence of the ranks provided by experts for the consideration of documents with the ranks provided by the software. The methods can be proposed for a wide range of industries.

Keywords: regulatory legal acts, document, software, collocation, criterion, semantic analysis of texts.