

Список використаних джерел

1. *Еременко Т. В.* Информатизация вузовских библиотек в России и США: сравнительный анализ : моногр. / Т. В. Еременко. – М. : Пашков дом, 2003. – 297 с.
2. *Кислюк Л.* Інтегральна оцінка рівня автоматизації бібліотек ВНЗ в Україні / Л. Кислюк // Вісн. Харк. держ. акад. Культури : зб. наук. пр. – Х., 2009. – Вип. 26. – С. 124–132.
3. *Колесникова Т. О.* Етапи інформатизації бібліотек вищої школи України і США : порівняльна характеристика / Т. О. Колесникова // Бібліотекознавство. Документознавство. Інформологія. – 2010. – № 4. – С. 51–59.
4. *Коряковцева Н. А.* Роль бібліотеки в формуванні корпоративної ідентичності вуза / Н. А. Коряковцева // Науч. и техн. б-ки. – 2009. – № 7. – С. 5–10.
5. *Кремень В.* Освіта і наука України: шляхи модернізації (факти, роздуми, перспективи) / В. Кремень. – К. : Грамота, 2003. – 216 с.
6. *Кузьмінська О. Г.* Інституційний репозитарій як складова електронної бібліотеки сучасного ВНЗ [Електронний ресурс] / О. Г. Кузьмінська. – Режим доступу: <http://www.ukrdeti.com/firstforum/m19.html>. – Назва з екрана.
7. *Панарина О. С.* База даних «Книгообеспеченность» – інструмент інформаційного забезпечення навчального процесу і система управління формуванням фонду вузовської бібліотеки / О. С. Панарина, В. І. Сосипатрова // Науч. и техн. б-ки. – 2010. – № 7. – С. 10–17.
8. Примірне положення про бібліотеку вищого навчального закладу III–IV рівнів акредитації [Електронний ресурс]. – Режим доступу : http://lib-gw.univ.kiev.ua/ukr/for_lib/nakazmou.html. – Назва з екрана.
9. Про національну програму інформатизації : Закон України від 4 лютого 1998 року N 74/98-ВР // Відом. Верхов. Ради України. – 1998. – № 27/28. – Ст. 181.
10. *Редькина Н.* Эффективность библиотечных сайтов / Н. Редькина // Науч. и техн. б-ки. – 2010. – № 3. – С. 31–34.
11. *Цветкова А. Л.* Электронная библиотека на основе информационной среды образовательного учреждения / А. Л. Цветкова // Науч. и техн. б-ки. – 2010. – № 5. – С. 67–75.
12. *Швецова-Водка Г.* Бібліотека як документальна комунікаційна система / Г. Швецова-Водка // Бібл. вісн. – 2011. – № 6. – С. 3–7.
13. *Шрайберг Я. Л.* Мировые тенденции развития библиотечно-информационной сферы, их отражение на пространстве СНГ / Я. Л. Шрайберг // Науч. и техн. б-ки. – 2005. – № 1. – С. 25–47.

УДК 519.765(048.3)

Артур Романюк,

мол. наук. співроб. НБУВ

СУЧАСНИЙ СТАН І ТЕНДЕНЦІЇ РОЗВИТКУ СИСТЕМ АВТОМАТИЧНОГО РЕФЕРУВАННЯ

У статті висвітлено історію походження поняття «реферат», подано його характеристику та визначення. Подано детальну інформацію про реферування статей, книг і розширені реферати, визначено їх характеристики. Увагу приділено особливостям реферування звітної науково-технічної та дослідно-конструкторської документації. Акцентовано на реферуванні описів винаходів, подано перелік аспектів змісту, які необхідно відобразити в рефераті: об'єкт, його призначення та галузь використання, суть, мета, технічне завдання, його технічні й економічні переваги. З'ясовано, що основною проблемою автоматичного реферування є створення моделі розуміння текстової інформації. Досліджено модель лінійно-вагових коефіцієнтів. На основі результатів дослідження встановлено, що для автоматичної побудови реферату існують такі методи: статистичний (словоформа, слововживання, анотація та словесний реферат тексту), позиційний (метод заголовку та локалізації) та логічно-семантичний.

Ключові слова: реферат, реферування, системи автоматичного «розуміння» тексту, комп'ютерна лінгвістика, модель лінійно-вагових коефіцієнтів.

На сьогодні, в умовах великої кількості наукової інформації, актуальним є питання створення систем автоматичного реферування. З кінця ХХ ст. науковці активно проводили дослідження з проблем комп'ютерного «розуміння» тексту. Нині існує багато спеціальних комп'ютерних програм, які автоматично створюють реферат, але єдиної, універсальної системи автоматичного реферування, яка б могла науково обґрунтовано, логіко-семантично скласти реферати з будь-якого виду тексту, так і немає. Тому це питання залишається актуальним і сьогодні.

Метою роботи є науковий аналіз становлення та розвитку систем автоматичного реферування тексту та виявлення основних проблем комп'ютерного «розуміння» тексту.

Історія походження поняття «реферат» досить неоднозначна. Багато хто вважає, що це слово бере свій початок від латинського «teferere», що

дослівно перекладається як «сповіщувати, доповідати», інші ж догримуються думки, що воно пішло від латинського терміна «abstractus», тобто «виводити заключення». На сьогодні ж існує багато визначень терміна «реферат». Найбільш загальним є таке: «Реферат – це коротке викладення змісту наукового документа». Але таке формулювання не достатньо розкриває особливості реферату як виду наукової продукції. Досить вдало дав визначення реферату В. І. Соловйов: «семантически адекватное, ограниченное малым объемом и вместе с тем возможно полное изложение основного содержания первичного документа, отличающееся постоянством структуры и предназначенное для выполнения разнообразных информационных функций при использовании его читателями различных категорий». Підсумовуючи викладене, можна зробити висновок – це поняття старовинне і процесом реферування займаються не одне століття.

У світі перші спеціалізовані служби з реферування наукової літератури почали з'являтися наприкінці XVIII – першій половині XIX ст. Цьому сприяли значні досягнення в фізико-математичних, хімічних і технічних науках. В Росії ж основоположником реферативної справи можна вважати М. В. Ломоносова. 1758 р. він запропонував програму з випуску щотижневого реферативного журналу, в якій передбачалося друкувати скорочені перекази творів з їх оцінкою для швидкого ознайомлення вчених з новими книгами. Видавництво першого реферативного журналу (РЖ) в Радянському Союзі відбулося в 1920-х роках та було покладено на плечі організації, яку назвали Бюро іноземної науки та техніки (БІНТ). Створення даної установи та вихід першого РЖ СРСР дало поштовх для утворення і інших подібних видань в різних галузях науки (РЖ «Сообщения о научно-технических работах в республике» (1920–1932 рр.), РЖ «Индексы», серія журналів «За овладение техникой» (ЗОТ), РЖ «Новости иностранной металлургии» (1935–1938 рр.), «Реферативный биологический журнал» (1933–1937 рр.), «Физико-математический реферативный журнал» (1938–1941 рр.), «Центральный медицинский [реферативный] журнал» (1928–1941 рр.) та інші).

Реферування – це стисле викладення тексту. Мета реферату – найбільш стисло передати зміст оригіналу та виділити найважливіше або щось нове, що міститься в матеріалі. Також за необхідністю створюється референтний коментарій, який відображує оцінку реферату.

Нині існує така класифікація рефератів:

1) за характером викладення: реферат-конспект (загальні тези з оригіналу в узагальненому вигляді), реферат-резюме (більший рівень узагальнення, ніж в рефераті-конспекті);

2) за оформленням: письмовий, усний;

3) за охопленням джерел: монографічний (одне джерело), зведений (декілька джерел), оглядовий (тема або напрям у вигляді оглядів), вибірковий (окремі розділи, глави).

Також існують вимоги, які висуваються до реферату:

1) лаконічність (відсутність вступних конструкцій, які теоретично можливі в усному рефераті, відсутність неясних формулювань, складних речень);

2) за наявності маловідомих імен в дужках робиться їх написання мовою оригіналу;

3) можливе скорочення термінів, що багато повторюються, для цього після використання таких термінів в дужках робиться його скорочення і далі використовується лише скорочення. Формули повинні відповідати оригіналу, але якщо в оригіналі існують позначення, що не прийняті в мові перекладу, їх необхідно замінити прийнятими;

4) референт повинен підписати реферат.

Відповідно до загальноприйнятої структури реферат містить:

1) бібліографічний опис (автор тексту та назва наводяться мовою оригіналу, також вказується кількість сторінок, рік і місце видання);

2) головна ідея, яка є відповіддю на задану тему (референт може сформулювати її самостійно, якщо формулювання від автора не досить чітко);

3) викладення змісту, яке враховує узагальнення матеріалу.

4) висновки – відповіді автора на питання, що розглядалося в статті, – є логічним розвитком головної ідеї. Залежно від внутрішньої структури статті вони можуть бути всередині або навіть на початку статті;

5) реферативний коментар, який може містити:

загальні зауваження юридичного й ідеологічного порядку;

зауваження з історії питання (зв'язок з минулим, теперішнім);

фактичне уточнення та роз'яснення;

довідку про автора та джерела (якщо нема бібліографічних даних);

посилання на інші джерела або матеріали з даного питання.

Незважаючи на загальні правила складання реферату та професійні навички референта, нині існують і більш широкі рекомендації з

опрацювання джерел з різної тематики. Розглянемо найважливіші аспекти опрацювання різних видів літератури.

Реферування статей

Наукові статті становлять нині більшість реферованих документів. На них складають пристатейні реферати, короткі та розширені реферати для реферативних журналів, розширені реферати для видань експрес-інформації тощо.

Обираючи статті для реферування та в процесі складання рефератів, ураховують їх найважливіші змістові та структурні ознаки, відмінності. Під цими ознаками розуміють чітку тематичну спрямованість, глибокий аналіз і ґрунтовне наукове оброблення матеріалу, широту теоретичних і практичних узагальнень, аргументованість висновків і пропозицій, оперативність, актуальність, інформаційну новизну, достатній читацький інтерес. З погляду структурних особливостей у тексті статей чітко виділяють вступну, основну частину та висновки.

Пристатейний реферат здебільшого складає автор статті і, як правило, такий реферат є індикативним за глибиною викладу. У ньому вміщують узагальнену характеристику змісту статті, викладають найсуттєвіші результати. Проте стислий реферат може бути й інформативним. Вибір виду реферату залежить від цінності змісту статті, її читацької адреси, доступності, а також від особливостей видання, в якому опубліковано статтю.

Реферування книг

На книги складають як стислі, так і розширені реферати. Стислі реферати вміщують у реферативних виданнях, у вихідних відомостях книг, розширені – у реферативних та інших виданнях.

Складаючи стислий реферат, мають на увазі, що зміст книги відрізняється багатотемністю, багатоаспектністю у висвітленні проблеми, бо в ній, як правило, подано фундаментальне дослідження.

Підготовку прикнижкового реферату найчастіше покладають на автора. Такі реферати, доповнені необхідними бібліографічними даними, також можна використовувати в реферативних журналах, у періодичних виданнях і виданнях, що продовжуються, у неперіодичних збірниках, де вони виконують сигнальну, інформативну, індикативну, адресну та пошукову функції.

Розширеному реферату більшою мірою властива інформативна функція. Такі реферати, як правило, складають на іноземні книги, що є в країні в обмеженій кількості примірників, або на вітчизняні, якщо це зумовлено необхідністю. Розширений реферат найчастіше складають на основі поаспектного аналізу змісту книги, що забезпечує його компактність і змістовність.

Реферування звітної науково-технічної та дослідно-конструкторської документації

Звітна науково-технічна та дослідно-конструкторська документація містить відомості про найновіші дослідження, тому вона становить особливий інтерес для користувачів належить до неопублікованих документів, кількість її примірників невелика.

Створення текстів рефератів у складі звітної науково-технічної документації з науково-дослідних і дослідно-конструкторських робіт доцільно здійснювати з використанням анкетного методу реферування. Основну увагу слід приділяти викладенню головних результатів роботи, характеристики виробу або його складників, опису істотних особливостей, відомостей про наукове значення та технічну й економічну ефективність, рекомендаціям для використання результатів. Текст реферату рекомендують подавати в абзацній формі, виділяючи в окремих абзацах різні аспекти змісту. Разом з тим, текст має бути єдиним цілим, в якому всі змістові аспекти узгоджені та підводять до головного – результатів роботи. У кінці реферату вказують кількість і характер ілюстрацій, таблиць, назв у списку літератури.

Реферування описів винаходів

Відомості в описах винаходів завжди однотипні й обов'язково належать до нових технічних рішень, тексти описів винаходів мають стандартну, суворо регламентовану форму надання інформації, їм притаманні типові усталені мовні звороти. Завдяки цьому визначено перелік аспектів змісту опису винаходу, які необхідно відобразити в рефераті: об'єкт (предмет) винаходу, його призначення та галузь використання, суть винаходу і мета, якої досягають за його допомогою, технічне завдання, яке вирішують за допомогою винаходу, його технічні та економічні переваги. Цей перелік універсальний, його можна використовувати для аналізу описів винаходів незалежно від країни патентування.

На сьогоднішній день глобальна комп'ютеризація й автоматизація торкнулася і процесу реферування. Так, однією з основних задач комп'ютерної лінгвістики є побудова систем автоматичного реферування наукових, технічних, політичних, ділових текстів. Поштовхом для створення даних продуктів було створення повнотекстових баз даних. Розроблення системи автоматичного реферування є можливим на основі експериментально-семантичного дослідження міжфразової структури тексту з наступною реалізацією інформаційних систем, що володіють спроможністю гіперсинтаксичного «розуміння» текстів.

За декілька останніх десятиліть було зроблено багато кроків у дослідженні проблем комп'ютерного розуміння тексту. Початковими етапами розвитку системи автоматичного реферування було просте зменшення тексту. Так, в книзі Н. Дарчук зазначено, що найпростішим прийомом стиснення тексту є автоматичне здобуття тексту тих речень, які містять одне або більше ключових слів чи словосполучень, що є свого роду «піками» у розподілі смислової інформації тексту. Ці речення виведені комп'ютером на екран у порядку їх слідування, утворюють машинний реферат документа. Основним недоліком даного методу є те, що речення для реферату формуються без урахування існуючих між ними смислових зв'язків. Наступним етапом розвитку автореферування та створення фундаменту для сьогоднішніх систем, була розробка В. Берзона та М. Блехмана алгоритму згортання, який ураховує надфразні єдності у тексті.

Надфразовий зв'язок виявляється в експліцитному й імпліцитному смислового насиченні одних речень тексту іншими реченнями. Якщо має місце експліцитне насичення, то воно виявляється через формальні валентні зв'язки, які встановлюються між реченнями за схемою $S_1 \rightarrow S_2$ або $S_1 \leftarrow S_2$.

Окрім цього, багато систем працює на основі тезауруса, тобто на основі попередньо укладених словників ключових (термінологічних) слів і словосполучень, списків антиознак (загальноновживаних слів), списків релятивів, фреймів. Сукупність цих і подібних їм даних є відтворювальною інженерно-лінгвістичною моделлю, яка має бути введена в комп'ютер, що розглядається як його навчання [5, с. 48].

Для створення автоматичного реферату більшість систем використовують модель лінійно-вагових коефіцієнтів. Основою цієї моделі є призначення вагових коефіцієнтів для кожного блоку тексту, які характеризуються такими параметрами: розміщення блоку в оригіналі,

частота появи в тексті, частота використання ключових термінів та показники статистичної вагомості. Сума даних показників разом з індивідуальними налаштуваннями системи визначають загальну вагу блоку тексту:

$$\text{Weight}(U) = \text{Location}(U) + \text{CuePhrase}(U) + \text{StatTerm}(U) + \text{AddTerm}(U)$$

Ваговий коефіцієнт розміщення (Location) визначає розміщення даного фрагменту в тексті.

Ключові фрази (CuePhrase) – це лексичні або фразові конструкції («у цій статті», «згідно з результатами аналізу», «у підсумку»), які орієнтують на висновки з тексту.

Також на загальну вагу блоку тексту впливає показник статистичної важливості (StatTerm), який вираховується на основі даних, одержаних в результаті аналізу автоматичної індексації, при якій визначаються вагові коефіцієнти терміну.

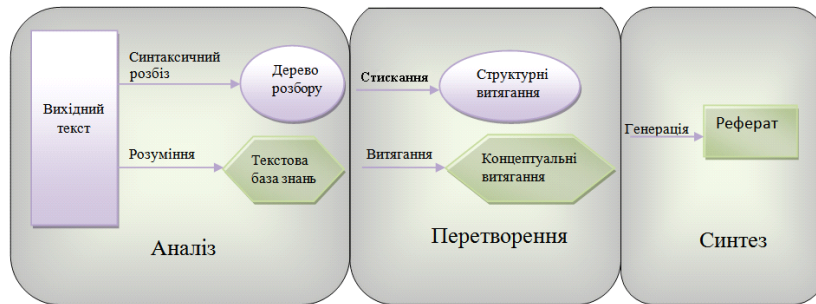
Окрім цього, на ваговий коефіцієнт розміщення впливає наявність додаткових термінів, їх розміщення (AddTerm) в блоці тексту.

Вагові коефіцієнти базуються на змінах статистичної важливості, наявності спеціальних термінів та розміщення речення в тексті. Вагові коефіцієнти речення, отримані на етапі аналізу, передаються відразу на вхід компоненту синтезу, де витягаються з найбільшими коефіцієнтами, визначеними за рівнем стиснення.

На етапі навчання вектор характеристик представляє кожне речення джерела. Класифікатор позначає кожен вектор відповідно до однорідності інформації в реченнях і рефераті. Потім направляється набір навчальних прикладів на алгоритм навчання, який запам'ятовує правила класифікації, для визначення – чи буде включене конкретне речення в реферат. На етапі тестування реферати відсутні. Замість цього класифікатор перетворює кожне речення тестового набору в вектор характеристик і перевіряє його відповідність правилам формування реферату.

Формування стислового викладу спирається на два базових підходи, які можна об'єднувати. Перший підхід (верхній на рисунку) на етапі аналізу характеризується створенням дерева розбору, яке на етапі перетворення піддається стисненню, шляхом відсікання або перегруповання гілок відповідно до структурних критеріїв, таких як дужки або підрядні речення. В результаті одержуємо структурні витягання, які мають значно простішу структуру. Другий метод (нижній) на етапі аналізу за рахунок текстової бази знань передбачає створення уявлення про всю інформацію, що

міститься в документі. На етапі перетворення шляхом об'єднання та узагальнення інформації відбувається процес стиснення. Результатом даних дій є формування концептуальних витягань. Етап синтезу передбачає генерацію результатів двох попередніх методів та створення єдиного більш правильного реферату.



Архітектура витягання фрагментів тексту
(реферування з опорою на знання)

Більшість сучасних методів автореферування базується на алгоритмах автоматичного екстрагування окремих речень, тематично та логічно зв'язаних між собою. Наукові праці Л. Н. Шавердної, А. Ф. Осики, В. П. Леонова та ін. надають вичерпний огляд наявних спроб автоматизації реферування та відображають загальну характеристику проблеми у вітчизняних і зарубіжних дослідженнях. Певний аналітичний огляд щодо теоретичних і прикладних досліджень лінгвістичних і структурних характеристик реферату й їх реалізацію в автоматизованому опрацюванні містить монографія В. І. Горькової та Є. А. Борохова. Основною проблемою автоматичного реферування є створення моделі розуміння текстової інформації. Побудова семантичного компонента, що забезпечує повноцінну роботу системи автореферування, має базуватися на вивченні та систематизації глибинних механізмів зв'язного тексту (В. І. Соловйов, А. І. Жолкова, Л. А. Гаврилов, Ш. Р. Ранганатан, А. А. Вейзе та ін.). Але задати єдиний алгоритм моделювання процесу розуміння текстової інформації неможливо. Кожен із розроблених алгоритмів висвітлює лише певні аспекти знань і в прикладному плані орієнтований на визначений клас завдань. Так, останнім часом для автоматичного аналізу текстових

даних усе частіше застосовуються методи Text Mining, які об'єднують технологічні й методологічні підходи до аналізу змісту, комп'ютерної лінгвістики та штучного інтелекту у вигляді потужних сучасних систем, що включають елементи автореферування [7, с. 6].

Для більш конкретного знайомства з системами автоматичного «розуміння» тексту, виявлення в них переваг і недоліків достатньо розглянути декілька їх прикладів.

Smart Search System (3S)

Проект 3S був запущений компанією InnerSpace наприкінці 1999 р. До теперішнього часу було створено систему, яка об'єднує високу релевантність результатів, тобто використовується великий семантичний розбір тексту. На відміну від простого пошуку за наявністю в тексті тих чи інших слів, алгоритм семантичного пошуку використовує асоціації між термінами, що виявлені під час структурування колекції документів. Можливість пошуку розширюється за рахунок автоматичного додавання у запит термінів, що пов'язані з запитом, які вже знаходяться в семантичній сітці цієї колекції документів.

Copernic Summarizer

Програма, що дозволяє продивлятися й аналізувати різні документи та створювати їх короткий огляд. Copernic Summarizer підтримує роботу із документами Word, веб-сторінками, PDF-файлами, поштовими повідомленнями та із файлами, що знаходяться в буфері обміну. Завдяки статистичним і лінгвістичним алгоритмам програма визначає ключову концепцію і витягає найбільш важливу для розуміння інформацію. Copernic Summarizer створює огляд документів, не відкриваючи їх, та зберігає одержані дані в різних форматах (.doc, .txt, .rtf, .pdf, .htm, .html). Програма може працювати з англійською та німецькою мовами.

Inxight Summarizer

На ринку знаходиться досить невелика кількість традиційних програм реферування, тобто таких, які виділяють найбільш вагомі речення із тексту, при цьому використовуються статистичні алгоритми або слова-підказки. Inxight Summarizer – одна із найбільш відомих систем реферування, що комерційно розповсюджуються. Програма була створена у Випробувальному центрі Ксероксу в Пало Альто.

Extractor

Програма створена Інститутом технологій національної випробувальної ради Канади, є модулем, що виділяє із представленого йому на вхід тексту найбільш інформаційні іменні групи. За замовчуванням кількість таких груп – 7 незалежно від довжини тексту. Extractor використовується в програмних продуктах фірм ThinkTank Technologies і Tetranet, а також в системі пошуку Журналу досліджень в області штучного інтелекту.

TextAnalyst

Програмний продукт створений в Московському науково-видавничому інноваційному центрі «МікроСистеми». TextAnalyst працює лише з російською мовою, виділяючи іменні групи, та на основі їх будує семантичну сітку – структуру взаємозалежностей між іменними групами.

Золотий ключик

Це програмна бібліотека, що працює за принципом фільтрації на базі тезауруса. Як вхідні дані програмі дається будь-який текст російською мовою, на стандартному виході програма формує анотацію цього тексту та список рубрик, до яких він відноситься. Як анотацію використовують речення з вхідного тексту, що найповніше відображає його тематику. Під час рубрикації тексту використовується фіксований список раніше визначених рубрик.

МЛ Анотатор

Програма складає зв'язаний реферат документа. Відносний розмір реферату («коефіцієнт стиснення») задається користувачем. Програма має два режими роботи: безпосередньо реферування та виділення ключових слів. Програма виділяє в тексті значущі слова, самостійні та залежні речення, визначає семантичну вагу речень і видаляє не важливі фрагменти. Відібрані речення за необхідності дещо перефразовуються. Використовуються спеціальні ймовірні моделі, машинна морфологія російської мови й інші інтелектуальні алгоритми.

Для автоматичної побудови реферату існують методи:

1. *Статичний*: в цьому методі ключове слово – це значуще слово-текст, яке з урахуванням синонімів зустрічалося в тексті найбільшу кількість разів. За цим методом виникає задача, яка дозволяє отримати: а) анотацію тексту у вигляді слів-реляторів з наступними за ними ключовими

словосполученнями тексту – ключовими іменниками з визначеннями, що розміщені перед ними та виражені прикметниками або дієприкметниками; б) словесний реферат тексту у вигляді послідовного ланцюжка ключових речень, що містять три чи більше ключових слова; в) слововживання – ланцюжок символів, пов'язаних між двома пропусками; г) словоформа – слововживання поза текстом. Декілька словоформ, що мають однакове лексичне значення, утворюють слово.

Алгоритм розв'язання поставленої задачі:

а) ПК за кожним абзацом складає алфавітно-частотний словник словоформ;

б) всі словники об'єднуються в єдиний роздільний алфавітно-частотний словник усього тексту;

в) система «очищує» словник, стискає його до словника потенційних ключових слів (видаляється службова та загальноприйнята лексика; об'єднуються граматичні форми одного й того ж слова; об'єднуються синоніми; видаляються слова, що зустрічаються лише в одному абзаці);

г) словник потенційних опорних слів ділиться на словник головних опорних слів і допоміжних;

в) будується анотація, що складається зі слів-реляторів з ключовими словосполученнями, що йдуть за ними та складаються з опорного слова та визначення.

2. *Позиційний*: основним критерієм цього методу є місце або позиція речення в тексті. Існує два види позиційного методу: а) метод заголовку (основний зміст тексту береться з заголовку) – складається словник ключових слів на основі заголовків або підзаголовків; б) метод локалізації (передбачає опрацювання текстів вузької тематики) – ідея в тому, що в таких текстах речення, що містять результати, посідають фіксоване місце.

3. *Логічно-семантичні методи*: досліджують структуру та семантику тексту. Ключові речення в них мають найбільше значення. На них впливають такі фактори: зв'язок з лівим і правим оточенням; наявність в реченні семантично-значимих слів; виділення тексту різними шрифтами та інше.

Кожний з даних методів має свої переваги та недоліки. Вони відносяться до напряму квазіреферування, основою якого є виділення з тексту найбільш інформаційних речень, що передають основну ідею документа.

Отже, на сьогодні з'явилися методи нового напряму, що виділяють з тексту найбільш інформаційні дані й утворюють з них новий текст. Таке реферування наближається до інтелектуального. Описані системи

працюють відразу з декількома джерелами, а також можуть працювати з відео. Але, незважаючи на це, єдиного, універсального програмного забезпечення, яке б могло комплексно аналізувати, систематизувати та реферувати різномірну інформацію так і нестворено. Кожна з існуючих систем орієнтована на більш вузький, спеціалізований матеріал.

Список використаних джерел

1. *Варенко В. М.* Референтна справа : навч. посіб. для студ. вищ. навч. закл. / В. М. Варенко. – К. : Кондор, 2009. – 212 с. – Бібліогр. : с. 210–211.
2. *Дарчук Н. П.* Комп'ютерна лінгвістика (автоматичне опрацювання тексту) : підручник / Н. П. Дарчук. – К. : Видавн.-полігр. центр «Київ. ун-т», 2008. – 351 с.
3. *Задорожний Э. М.* Работа секретаря-референта в иностранных и совместных фирмах : практ. пособие секретарю, предпринимателю и менеджеру / Э. М. Задорожний, С. Э. Задорожний. – К. : [б. и.], 1994. – 176 с.
4. *Зайченко Н. Я.* Електронний реферативний журнал: яким йому бути? / Н. Я. Зайченко // *Наук. пр. НБУВ.* – К., 2002. – Вип. 8. – С. 379–383.
5. *Кондратюкова Л. К.* Аннотирование и реферирование английской научно-технической литературы : учеб. пособие для студ. техн. вузов / Л. К. Кондратюкова. – Омск : Изд-во ОмГТУ, 2001. – 184 с.
6. *Лазаренко О. В.* Моделирование процесса узагальнення в системі автоматичного реферування: монографія / О. В. Лазаренко, А. А. Яковенко. – Х. : Вид-во НУА, 2007. – 124 с.
7. *Панченко Д. І.* Моделирование семантических зв'язків «Текст-Реферат» у системах автоматичного реферування : автореф. дис. ... канд. філол. наук : 10.02.21 / Д. І. Панченко. – К., 2012. – 20 с.
8. *Севбо И. П.* Структура связного текста и автоматизация реферирования / И. П. Севбо. – М. : Наука, 1969. – 135 с.
9. *Сорока М. Б.* Національна система реферування української наукової літератури: монографія / М. П. Сорока. – К., 2002. – 210 с. – Бібліогр. : с. 196–207.
10. *Шаховська, Н.* Інформаційна система реферування множини документів, поданих у різних форматах, базована на онтології / Н. Шаховська, В. Литвин, В. Крайовський // *Вісн. Нац. ун-ту «Львів. Політехніка».* – 2010. – № 672 : Комп'ют. науки та інформ. технології. – С. 63–71.
11. *Шубкіна О.* Використання конкурентної ймовірнісної мережі у задачах обробки текстової інформації / О. Шубкіна, І. Пліс, С. Бодяньський // *Вісн. Нац. ун-ту «Львів. Політехніка».* – 2011. – № 710 : Комп'ют. науки та інформ. технології. – С. 219 – 223. – Бібліогр. : 11 назв. – укр.
12. *Шубкіна О. В.* Модели семантического аннотирования текстовых документов с использованием искусственных нейронных сетей специального вида / О. В. Шубкіна // *Системи оброб. інформації.* – 2011. – Вип. 2. – С. 221–225.

УДК: 002.1.(048.3):004.65

Оксана Сандул,

мол. наук. співроб. НБУВ

ЗАСТОСУВАННЯ ІННОВАЦІЙНИХ ТЕХНОЛОГІЙ У ПРОЦЕСІ РЕФЕРУВАННЯ ЛІТЕРАТУРИ З МЕТОЮ СТВОРЕННЯ ЯКІСНИХ ІНФОРМАЦІЙНИХ РЕСУРСІВ

Проаналізовано можливості використання сучасних інформаційних технологій у процесі реферування наукової літератури. Розкрито значення й особливості розвитку зарубіжних і вітчизняних реферативних журналів і баз даних. Подано інформацію про технічні засоби та програмні продукти для здійснення процесів сканування, розпізнавання, перекладу та редагування тексту. Увагу приділено досвіду та перспективним напрямкам застосування сучасних інноваційних технологій в роботі Відділу реферування наукової інформації НБУВ.

Ключові слова: реферування, реферативні журнали та бази даних, сканування та розпізнавання тексту, автоматизоване реферування.

Метою дослідження є аналіз сучасних технічних засобів і програмних продуктів у сфері створення реферативних ресурсів, а також можливості їх застосування в поточній роботі відділу реферування наукової інформації НБУВ.

Стрімкий розвиток світового інформаційного простору сприяє розширенню сфери впливу бібліотек на розвиток суспільства. Завдяки впровадженню передових бібліотечно-інформаційних технологій, комп'ютеризації та автоматизації бібліотечних процесів наукові бібліотеки перетворюються з книгозбірень на сучасні інформаційні центри.

Серед інформаційних ресурсів чільне місце посідає реферативна інформація, яка є загальноприйнятною формою обміну науковими досягненнями та спілкування вчених в усьому світі, допомагає зорієнтуватися у потоці наукової літератури, одержати найновіші відомості щодо наукових досліджень з певних галузей знання. Одним з найбільш популярних бібліографічних джерел є реферативні журнали, в структурі яких виділяють такі складові: бібліографічний опис, реферативний текст, посилання, примітки. В зазначених журналах наводяться реферати, які є аналітико-синтетичними переробками тексту, що несуть головну