

ОСНОВНІ ЗАСАДИ РЕФЕРУВАННЯ ДОКУМЕНТІВ ЯК СФЕРИ ПРАКТИЧНОЇ ДІЯЛЬНОСТІ

I.O. Коханова

Розглядається процес створення національної системи реферативної інформації в Україні. Аналізуються методологічні та практичні засади реферування документів та проблеми, які з ними пов'язані. Висвітлюються теоретичні положення, які допоможуть їх уникнути.

Ключові слова: реферування документів, реферативний аналіз, інформаційна цінність, інформаційна ємність, індикативний реферат, інформативний реферат, індикативні маркери, інформативні маркери, екстрагування, автоматичне реферування.

THE MAIN PRINCIPLES OF ABSTRACTING AS BRANCH PRACTISE ACTIVITY

I.O. Kokhanova

The article discusses process of creating national system of review information in Ukraine. Methodological and practical basis of this process and possible problems are analyzed. The theoretical basis, witch can help to avoid it are explored.

Keywords: abstracting, abstract analysis, informative value, informative capacity, indicative abstract, informative abstract, indicative marker, informative marker, extracting, automatic abstracting.

Зі здобуттям незалежності та входженням України до світового інформаційного простору, постала нагальна необхідність створення інформаційних ресурсів держави, вироблення і забезпечення суспільства інформаційними продуктами і послугами. І, насамкінець – створення якісно нової ланки практичної діяльності – системи вітчизняного реферування документів, яка б охоплювала всі галузі знань та не була дискретною.

Більшість вітчизняних фахівців, які займаються вивченням процесів становлення національної системи реферативної інформації, серед яких: В. Бірбраєр, М. Вівчарик,

Л. Костенко, Л. Кочубей, М. Сорока [1; 2; 3; 5] та інші, нині вбачають початковий етап формування концепції вітчизняної системи реферування та видання реферативної інформації; реферативні джерела з суспільних наук розглядають як основу системи інформаційного забезпечення державотворчого поступу України; намагаються сформулювати методологічні та практичні засади розробки та створення українських реферативних видань.

Вивчаючи сучасний стан національної системи реферування в Україні, ставимо за мету дослідити основні підвалини практичного реферування та проблеми, які можуть постати перед референтом у процесі опрацювання першоджерел. Слід зазначити, що, не зважаючи на широкий спектр методів реферування, цей процес залишається вкрай суб'єктивним. Звісно, референт не може у тексті реферату подавати власні думки та точки зору з приводу первинного документа, на відміну від анотатора. Але, методи перефразування та інтерпретації, застосовані у процесі згортання інформації, можуть досить сильно змінити контекст викладу змісту оригіналу документа.

Нині головним центром, який здійснює опрацювання наукових текстів та підготовку реферативної інформації з усіх галузей знань в межах країни, є Книжкова палата України.

Реферування документів є новим напрямом діяльності для цього закладу.

Для його виконання навесні 1997 р. було засновано окремий структурний підрозділ – відділ реферування інформації, діяльність якого спрямована на створення реферативних баз даних шляхом підготовки реферативних журналів і генерування реферативно-бібліографічної інформації та на дослідження процесів її підготовки.

Не зважаючи на вичерпну повноту отримання усіх друкованих джерел, які дозволять охопити та задовольнити всі інформаційні потреби вітчизняних науковців інформаційними ресурсами – працівникам Книжкової палати необхідно скооперувати свою діяльність із провідними реферативними агенціями світу, науковцями Академії наук України для окреслення професійних засад реферування як сфери практичної діяльності, створення відповідної нормативно-методичної бази з підготовки реферативних видань та широкого видового спектру рефератів до них (індикативних, інформативних, рефератів-екстрактів, рефератів-резюме, аналітичних оглядів тощо).

Кваліфікований референт чітко має усвідомлювати таке. Від викладу матеріалу у рефераті залежатиме його адекватне сприйняття користувачами. Матеріал у рефераті можна викладати в тій самій послідовності, як і в першоджерелі, а можна цю послідовність обумовлювати значенням різних аспектів змісту. Тому, доцільно аспекти ранжувати, тобто встановлювати послідовність за ступенем значущості. Визначаючи вид майбутнього реферату, враховувати категорію потенційних користувачів, для яких складається реферат, характер інформаційного

запиту (фактографічний чи документний); розуміти, чи буде реферат елементом видавничого оформлення документу або його надрукують у реферативному виданні тощо.

На етапі реферативного аналізу референту слід виявити основні аспекти змісту документа, усвідомити сутність основного об'єкта вивчення, його особливостей та властивостей, переваг та недоліків, мети та методів дослідження, отриманих результатів, висновків та сфери застосування. І, обов'язково, подати ті групи важливих та принципових відомостей про зміст першоджерела, які не можна проігнорувати.

Традиційно фахівці виокремлюють 3 групи відомостей, які завжди слід відбивати у тексті майбутнього наукового реферату:

1) нові ідеї, гіпотези, експериментальні дані, нові методики, опис нових оригінальних конструкцій механізмів, інформацію про відкриття та дослідження нових явищ або процесів тощо;

2) фактичні дані, що не є принципово новими – відомості про застосування традиційних методів та технологій, фактичні показники, формули, цифрові дані, будь-які обчислення та відсоткові показники, які будуть зрозумілими користувачу-науковцю та можуть істотно скоротити текст реферату;

3) відомості, які використовують як ілюстративний матеріал, це може бути певна аргументація автора первинного документа.

Також слід усвідомлювати, що вивчення будь-якого наукового тексту потребує уміння визначати інформаційну цінність його окремих елементів, а саме, термінів як основних структуроутворюючих та змістовних компонентів. Не лише автоматичне реферування, анотування, інформаційний пошук, але й переклад (наукового технічного) тексту різною мірою об'єднується з його семантичною компресією (змістовним стисненням). У зв'язку з цим виникає низка питань, головним серед яких є те – як адекватно визначити інформаційну цінність тієї чи іншої одиниці тексту.

Інформаційну цінність терміна слід розуміти як його здатність підвищувати інформаційний потенціал (тезаурус) реципієнта [6].

Інформаційна цінність будь-якого терміна залежить від потреб користувача, його знань, кваліфікації, попереднього досвіду тощо. Взагалі немає сенсу говорити про інформаційну цінність терміна безвідносно до конкретної ситуації.

Розрізняють системну та дискурсивну цінність терміна.

Системна цінність – це цінність терміна як елемента термінологічної системи певної галузі знань;

Дискурсивна цінність – цінність терміна як елемента наукового тексту.

Як системна так і дискурсивна цінності мають ієрархічний характер (цінність у межах системи, підсистеми, в межах тексту, підтексту).

Крім того, в обох випадках додатково виокремлюється низка аспектів, які відображають галузь (підгалузь) знання, по відношенню до якої визначається цінність терміна та конкретне питання, яке цікавить реципієнта-користувача.

У зв'язку з цим можна виокремити: предметно-орієнтовану та проблемно-орієнтовану цінність терміна.

Серед семантичних характеристик слова як елемента лексичної системи важливу роль відіграють дві – семантична складність та семантична продуктивність.

Семантична складність розуміється як роль слова у формуванні певної логіко-семантичної системи, а саме – термінології.

Терміни з великою когнітивною (пізнавальною) значимістю співвідносяться з базовими поняттями відповідної галузі знань. Більшість з яких має широку сферу функціонування. Ці терміни відомі не лише фахівцям, але й широкому загалу.

Терміни з низькою когнітивною значимістю відображають вузькоспеціальні поняття високого ступеня диференціації (більшість із них не мають усталених російських або українських еквівалентів).

Терміни з великою інформативною ємністю у більшості випадків відповідають складним спеціальним поняттям. У спеціальних довідниках зафіксована лише частина таких лексичних одиниць. Серед відсутніх – багато нових термінів: мовних та авторських неологізмів, що відбивають поняття, які нещодавно виникли.

Терміни з малою інформаційною ємністю частіше за все передають поняття досить широкої сфери функціонування.

Оскільки в мові терміни з великою когнітивною значимістю відображають певні категоріальні поняття відповідної галузі знань, вони відіграють важливе значення у формуванні терміносистеми [4].

Значення термінів у науковому тексті відрізняється від їхнього значення у терміносистемі. У тексті терміни з великою когнітивною значимістю несуть досить низьке змістовне навантаження. Вони не стільки передають нову інформацію, скільки відносять її до певних базових категорій відповідної галузі науки, тобто виконують роль своєрідних *категоріальних або індикативних маркерів*.

У зв'язному тексті їхня головна функція – індикативна: вони є показниками тих понять, до яких відноситься текст. Саме в цьому їхня цінність для розуміння наукового тексту, а також для різних видів його інтелектуальної обробки, пов'язаних зі стисненням інформації та вилученням нових знань.

Терміни з великою інформаційною ємністю означають складні спеціальні поняття малої уживаності. В науковому тексті саме вони набувають особливого статусу, оскільки саме вони передають нову інформацію – саме ту, за ради якої і створюється текст. Вони відіграють роль *інформативних маркерів* [7].

Лінійно-композиційна структура наукового тексту визначається міцністю семантичних зв'язків між лінійними фрагментами тексту. Вона знаходить своє відображення в членуванні тексту на відносно самостійні у змістовному значенні лінійні відрізки –вступ, основну частину тексту (з можливим подальшим поділом на підрозділи, в яких описуються завдання, методи дослідження, проведені експерименти, обговорюються їхні результати тощо) та заключну частину.

Когнітивна (пізнавальна) структура наукового тексту визначається розподіленням у тексті елементів, що несуть нові знання, відбивають та пояснюють нові поняття.

У центральних розділах тексту (в основній частині) найчастіше зустрічаються терміни з великою, відносно вступу та заключної частини, інформаційною значимістю. Це свідчить про те, що у вступі та заключній частині перевага надається термінам з більш загальним значенням, а саме, *індикативним маркерам*, а у центральних розділах – інформаційно насиченим термінам, тобто *інформаційним маркерам* [6].

Це можна проілюструвати наступним чином. Візьмемо дві опрацьовані наукові статті, на які складено реферати-екстракти. Вони є вкрай лаконічними та інформаційно насиченими, тому що складені, переважним чином, з маркерованих речень з означеними видами маркерів.

Григорьев В. А. Знак и понятие / В. А. Григорьев // Науч. техн. информация. – 2004. – №3. – С. 7-21.

В статье говорится, что в неформальной интеллектуальной системе знак, понятие и чувство неразрывно связаны единым процессом и потому возникают одновременно. Достоинством работы является разработанный поход, позволяющий решить ряд фундаментальных проблем: выполнить градацию понятия по уровню формальности, выявить общность понятия и неформальной интеллектуальной системы.

Брудный И. М. Политика идентичности и посткоммунистический выбор России // Полис. – 2002. – №1. – С. 87-105.

Автором ставится задача доказать, что национальная идентичность – критический фактор для понимания современных политических проблем России. В статье актуальным представляются посткоммунистические реформы, которые позволили бы сосредоточить в своих руках политическую и экономическую власть, но недостаток в том, что эти реформы делают быт россиян еще труднее, чем в конце правления коммунистов.

Суцність этой статъи заключається в том, что в России была фактически узаконена нелиберальная концепция ее государственности, а сулимая либералами «жизнь как в Европе» превратилась в отдаленную мечту.

Стосовно термінів наукових статей, то, як правило, їхні тексти складаються з фрагментів, які умовно можна поділити на *дескриптивні* та *конструктивні*. До перших належать: опис пристроїв та конструкцій; викладення сутності експерименту; його обговорення тощо. До другої – безпосередньо концептуальна частина наукового тексту – викладення гіпотези, наукових методів, винаходів, запропонованих пристроїв та систем, що фактично і є комунікативним завданням автора.

«Нові» терміни, як правило, концентруються в центральних фрагментах відповідних структур тексту: в центральних розділах статті їх більше, ніж у вступі чи заключній частині; в центральних абзацах розділу їх більше, ніж у перших та останніх абзацах; в центральних реченнях абзацу вони зтрапляються частіше, ніж у першому чи останньому [9].

Слід ще раз підкреслити, що характерною особливістю методів реферування є те, що вони спрямовані на виявлення та відбиття у вторинному документі основних положень змісту первинного документу. Якщо в процесі традиційного реферування це відбувається, в більшості випадків, інтуїтивно, то в межах, наприклад, автоматичного реферування під основним змістом зазвичай розуміють речення, які виділяють та екстрагують на основі зафіксованих у словниках лексико-граматичних ознак [10].

Таким чином, наголошується, що методика автоматичного реферування передбачає три основні етапи:

1. складання словника;
2. перегляд тексту оригіналу документа та виявлення речень, які містять лексико-граматичні одиниці, вказані у словнику;
3. екстрагування виявлених речень та складання тексту реферату.

Залежно від структури словника можна назвати два підходи до автоматизованого реферування: реферування на основі словників нетематичної лексики та реферування на основі словників тематичної лексики. Прикладом першого підходу є методика маркерного реферування. Такий вид реферування здійснюється на основі словника маркерів, індикаторів та конекторів. Під маркерами розуміються слова або словосполучення, які точно та адекватно відбивають аспекти змісту першоджерела (аспект цільової установки, запропонованого варіанта вирішення проблеми, недоліками або перевагами роботи, отриманими результатами тощо).

Індикатори, на відміну від маркерів, не пов'язані з певним аспектом змісту, але вони вказують на речення, яким автор надає особливого значення (слід підкреслити, необхідно зазначити).

Конектори слугують для виділення речень, пов'язаних із маркерованими.

Екстрагування здійснюється з використанням словника маркерів та індикаторів та передбачає такі етапи: перегляд тексту першоджерела та виписування речень, у яких є маркери. Речення з індикаторами референт виписує лише у тому випадку, якщо, на його думку, одних маркерованих речень недостатньо.

Маркерне реферування застосовують лише для документів великого обсягу (наукові статті, монографії), в яких можна виділити певні аспекти. Перевагою такого виду реферування є те, що воно може здійснюватися безвідносно до предметної сфери первинного документа.

Прикладом складання рефератів на основі словників тематичної лексики може бути реферування, яке базується на ідеї, що найбільш значуща інформація в першоджерелі міститься в окремих реченнях, які мають більшу функціональну вагу, ніж інші. Функціональна вага речень розглядається як кількість зв'язків між даним реченням та іншими реченнями даного тексту. Кожен зв'язок – це повторення іменника, дієслова, прислівника, прикметника, який належить до даної предметної сфери. У процесі реферування на основі словника тематичної лексики підраховується функціональна вага речень у первинному тексті і, речення з більшою функціональною вагою включаються в текст реферату. Недоліком функціонального реферування є те, що він може використовуватися лише по відношенню до текстів певної, досить вузької предметної сфери, а перевагою – його застосування не лише для реферування наукових, але й газетних текстів різного обсягу.

Загальними недоліками обох підходів до автоматичного реферування є складність складання словників та дискретність отримуваних текстів рефератів.

Отже, референт-професіонал у своїй практичній діяльності може використовувати цілу низку методів згортання інформації, але слід розуміти, що деякі з них можуть додавати до тексту вторинного документа велику частку суб'єктивізму. Наприклад, методи перефразування та інтерпретації. Стосовно методів автоматичного реферування – їх лише починають опановувати українські фахівці.

Темою подальшого дослідження може бути джерелознавчий пошук та ретельніше вивчення практичного досвіду зарубіжних колег приводу автоматичного реферування.

Використана література

1. Бірбраер В. До концепції вітчизняної системи реферування та видання реферативної інформації / В. Бірбраер, Л. Костенко // Вісн. Кн. палати. – 1998. – №3. – С. 25-26.
2. Вівчарик М. Український реферативний журнал «Політика. Політичні науки» в системі інформаційного забезпечення державотворчого поступу України / М. Вівчарик, Л. Кочубей // Вісн. Кн. палати. – 1998. – 34. – С. 20-21.

3. Кочубей Л. Відділ реферування інформації Книжкової Палати України: сучасний стан та перспективи /Л. Кочубей // Вісн. Кн. палати. – 1998. – №10. – С. 6-7.
4. Подпольный Д. Ф. Об определении значимости слова / Д.Ф. Подпольный // Науч. и техн. информация. Сер. 2. – 1989. – №2. – С. 9-15.
5. Сорока М, Система реферування українських наукових видань (основні концептуальні положення) / М. Сорока // Бібл. вісн. – 1999. – №3. – С. 5-7.
6. Скороходько Э.Ф. Информационная ценность термина в научном тексте (индикативные и информативные маркеры) / Э.Ф. Скороходько // Науч. и техн. информация. Сер. 2. – 1998. – №9. – С. 13-22.
7. Скороходько Э. Ф. Семантическая продуктивность и семантическая емкость слова в общеупотребительной и терминологической лексике: некоторые количественные характеристики / Э. Ф. Скороходько // Науч. и техн. информация. Сер. 2. – 1997. – №2. – С. 1-13.
8. Скороходько Э. Ф. Роль системо- и текстообусловленных характеристик термина в частотном индексировании научных текстов / Э. Ф. Скороходько // Науч. и техн. информация. Сер. 2. – 2002. – №8. – С. 1-6.
9. Скороходько Э. Ф. Термины, выражающие новые знания / Э. Ф. Скороходько // Науч. и техн. информация. Сер. 2. – 1999. – №4. – С. 13-22.