

ОЛЕГ СЕРБІН,

заступник голови Ради молодих учених НБУВ,
старш. наук. співроб. НБУВ, канд. іст. наук

ЗГОРТАННЯ/РОЗГОРТАННЯ ІНФОРМАЦІЇ ЯК СПОСІБ УНІФІКАЦІЇ І ПРЕДСТАВЛЕННЯ ЗНАННЯ У ВЕБ-ОРІЄНТОВАНОМУ СЕРЕДОВИЩІ

У статті представлено аналітичний та методологічний огляд процесів згортання/розгортання. Проаналізовано аспекти динаміки розвитку форм організації інформації у веб-орієнтованому середовищі. Подано інформацію про використання принципів згортання/розгортання в межах веб-простору.

Ключові слова: уніфікація, згортання, розгортання, інформація, веб-середовище.

Розвиток знань завжди супроводжується розвитком способів передачі цих знань. Це виражається в еволюції форм закріплення, систематизації та представлення цих знань. Від перших способів закріплення інформації до сьогочасних змінилася не одна епоха розвитку людства. Звичайно піддалася кардинальним змінам форма, що символізує результат обробки інформаційних даних. Сьогодні ми можемо говорити про можливість роботи з інформаційними даними у веб-орієнтованому просторі. Інтернет-технології дають можливість працювати з електронними документами, специфічно обробляючи їх та формуючи з них нові інформаційні дані. Навіть у межах веб-простору форми організації інформації постійно змінюються, еволюціонуючи відповідно до загальної динаміки знань. Константним залишилося тільки розуміння значення способів уніфікації інформації в межах певної форми та роль змістовності інформаційних даних у проведенні вищезазначеного процесу. Адже зміст інформації не завжди залежить від складу інформаційного повідомлення. Саме змістовна складова відповідає за інформативність того чи іншого документа, враховуючи можливість аспектного передавання даних.

Цілком справедливо можна припустити, що представлення інформації у вигляді подання множини певних складових не обов'язково виражає або тільки зміст, або тільки склад. Найявність у структурі певного переліку демонструє у першу чергу своєрідну характеристику складу інформації. Що ж до змісту, то тут необхідно акцентувати увагу на тому, що структурно він може бути поданий у вигляді переліку. Мета подання інформації, цільова аудиторія, формує метаінформаційну скла-

дову такої інформації – «інформацію інформації». І ось тут дуже важливо бачити значення змісту інформаційного повідомлення, не акцентуючи першочергово увагу на складі останнього. У такому разі зміст характеризується не складом кінцевого вигляду інформаційного повідомлення, а зосереджений у метаінформаційній складовій загального його складу. Зрештою, зміст – це своєрідна потенція, радше ідея інформації, своєрідна ентелехія кінцевого результату. Усе питання в тому, як через склад показати зміст і змістовним поштовхом сформувати певний склад. Адже говорячи про одну й ту саму річ, ми можемо в різний спосіб передати зміст засобами різнокількісності складу повідомлення, про яке йдеться. Наприклад, термін «оптимізація» можна виразити визначенням даного терміна, а можна представити у вигляді індексу будь-якої зручної для нас інформаційно-пошукової мови. Таким чином, можна говорити про певну уніфікацію подання інформації. Тому тема дослідження процесів уніфікації задля представлення в найсучаснішому веб-орієнтованому просторі, є не просто на часі, а видається нам актуальною з огляду на вищесказане.

Об'єктом даної публікації є уніфікація інформації задля представлення у веб-орієнтованому просторі. Предметом публікації є згортання/розгортання інформації для представлення у веб-орієнтованому просторі. Метою публікації є вивчення операцій згортання/розгортання інформації для здійснення її уніфікації та представлення у веб-орієнтованому просторі.

Одразу слід зауважити, що у випадку, коли ми говоримо про представлення уніфікації складу, варто, на нашу думку, говорити про ущільнення або стиснення інформації. Означені операції зумовлюють, у свою чергу, спрощення знання, спрощення інформаційного повідомлення. З іншого боку, коли за мету стоїть уніфікація інформації, але не стільки за рахунок складу, а за рахунок представлення інформаційного повідомлення таким чином, щоб змістовна його складова в аспекті елементності зменшилась, а от в аспекті інформаційної повноти, себто змістовності, залишилася сталою, потрібно говорити саме про згортання/розгортання інформації. Спираючись на сьогоденні визначення цього явища, можна говорити, що це зміна фізичного обсягу документа в результаті його аналітико-синтетичної обробки, що супроводжується зменшенням [1, с. 211] або збільшенням його (документа) інформативності. Учений-інформацієзнавець Д. Й. Блюменау в одній зі своїх праць говорить про те, що «... це процедура мислення, яка спрямована на визначення рівня необхідної і достатньої інформативності повідомлення в залежності від комунікативного завдання...» [2, с. 180]. Ми можемо говорити про операцію згортання/розгортання інформації, першочерговим в якій є початковий процес мислення, тобто ми знову повертаємося до

зазначеної вище «ідеї інформації» або «інформації інформації». Цілком очевидним є належність мислинневої складової, що, у свою чергу, орієнтує нас на потребу глибинного вивчення питання змістовності такої операції, адже, на відміну від поданого вище визначення згортання/розгортання інформації, зрозуміло, що уніфікація змісту повідомлення не виражається уніфікацією тільки складу такого повідомлення. Тому потрібно говорити про згортання/розгортання інформації самої в собі, самої із себе і саму в себе. До того ж, розглядаючи уніфікацію змістовності, ми, таким чином, розглядаємо не просто технологічний чи суто інформативний процес, а науковий підхід представлення уніфікації змісту [3, с. 24], адже якість згортання в основному визначається тим, наскільки воно потенційно багате можливостями подальшого розгортання.

Розкриваючи суть процесу згортання/розгортання, варто вказати на помилкове бачення деяких дослідників змістовної складової означеного процесу [2, с. 180]. На нашу думку, тотожність згортання/розгортання визначенню «інформації про інформацію» не зовсім відповідає дійсності. І ось чому. Говорячи про «інформацію про інформацію», ми в першу чергу пояснюємо й окреслюємо вторинність результату. Тобто, подаючи, приміром, процес згортання, як оптимізацію інформації задля представлення її в оптимальному вигляді, ми не говоримо й не повідомляємо додаткових даних про означену інформацію. Ми просто подаємо цю інформацію структурно в іншому вигляді, оптимізуючи при цьому не змістовність, а зміст. Одразу необхідно зауважити про відмінність цих двох понять – «зміст» та «змістовність». Коли йдеться про оптимізацію змісту, то мається на увазі здійснення певних кроків щодо оптимального подання структури елементів, які складають дане конкретне інформативне повідомлення. Це здійснюється за рахунок кількісних характеристик. Коли ж йдеться про оптимізацію змістовності інформації, слід розуміти в першу чергу оптимізацію ідейних потенцій, що впливають на суть та здійснення процесу оптимізації вже змістовного компонента інформаційного повідомлення. Тобто, щоб оптимізувати за рахунок згортання, потрібно саму ідею згортання представити в оптимальному вигляді. І тому в даному випадку необхідно говорити не про «інформацію про інформацію», а про «інформацію інформації». Саме означення процесу згортання/розгортання через поняття «інформація інформації» розкриває суть цього процесу, вказуючи на його науковість.

У науковому згортанні/розгортанні можливо спостерігати риси експлікації. Коли ставиться за мету розкриття змісту інформації за рахунок варіацій з його компонентністю, але без зміни змістовності даної інформації – можна говорити про експлікацію як складову процесу наукового згортання/розгортання. Що ж до зменшення змістовності інформації, то, на нашу думку, такі процеси не допустимі в межах згортан-

ня/розгортання (і особливо розгортання), з огляду на втрату інформативності повідомлення. Це, у свою чергу, призведе не до оптимізації інформації, а до її хибної уніфікації, помилкової уніфікації інформативності. Тоді метаінформація не матиме місця, це значить, що втратить ся суть процесу згортання/розгортання, а отже, виникає великий ризик підміни понять сутнісного й кількісного. Така обставина призведе тільки до одного – інформаційного мовчання, або ж хибної інформаційної відповіді. І навіть таке поняття, що використовується певними дослідниками, як «ущільнення знань» не відповідатиме результату процесу згортання/розгортання. Адже наукове згортання/розгортання передбачає можливість зміни кількісних характеристик зі збереженням цілісності змістовного. Тобто абсолютно вірним науковим згортанням/розгортанням буде той процес, у результаті якого зміна структури не змінила змістовність.

Однак як відбувається процес експлікації в загальному згортанні/розгортанні, враховуючи те, що змістовна зміна спричиняє не «ущільнення знання», а втрату його змістовності? Саме в цьому і є особливість експлікаційності наукового згортання/розгортання. Розгортання відбувається таким чином, що кожен з елементів, відмежовуючись від загального, єдиного елемента, своєю опосередкованістю формує розуміння і про узагальнююче (пригадаємо вищевказану «інформацію інформації»). І так зберігається, а фактично й підсилюється цілісність картини про узагальнюючий елемент.

Кожен елемент, таким чином, є фактично опосередкованим інформаційним повідомленням про узагальнююче. З огляду на це, варто звернути увагу на роль опосередкованого елемента в семантичному розкритті змісту. Дійсно, семантична навантаженість зумовлює розуміння про зміст інформації, шляхом семантичного згортання/розгортання. У такому випадку можна говорити про подання інформації через оптимізацію представлення змісту. Розкриваючи значення певного терміна за рахунок введення більшої кількості змістовних компонентів, ми формуємо повідомлення значнішої інформативності й кількісно виражене декількома, а не одною семантичною одиницею. Тут варто звернути увагу на те, що в даному випадку, незважаючи й на розкриття сутнісного, потрібно говорити не про наукове, а про інформаційне розгортання. Адже інформація, що розкривається в результаті, є фактично вторинною і символізує «інформацію про інформацію». Отже, можна говорити про ще один вид згортання/розгортання – інформаційний. У результаті означеного процесу розкривається знову така не змістовність, а зміст, тому що змінюється в першу чергу кількісна характеристика, збільшується або зменшується кількість семем.

У процесі інформаційного згортання/розгортання можливе також представлення однієї і тієї ж інформації засобами іншої системи, відмінної від первинної. Якщо в аспекті семантичного чи лексичного згортання/розгортання здійснюється оптимізація тексту тими самими засобами, якими й подано первинну інформацію, то процес передавання інформації іншою системою потребує застосування вже засобів цієї нової системи. Враховуючи ж те, що будь-яка знакова система – це система сигналів або повідомлень, що інтерпретуються та трактуються однаково, стає зрозумілою вся гамма можливостей переходу й переведення інформації в аспекті згортання/розгортання. Усі системи знаків, на нашу думку, несуть у собі певне індексне вираження. Адже фактично кожен знак може бути представлено як індекс. Будь-яка нотація (нотна, літерна, цифрова) є певною індексною нотацією, що демонструє процес індексування в загальному згортанні/розгортанні. Тому передусім спробуємо поглянути на процес індексування не з позиції якоїсь вузькопрофільної потреби, а з позиції важливості систематизації інформації. І процес згортання/розгортання, таким чином, є відображенням індексування інформації. Адже аналізуючи певний інформаційний блок і ставлячи за мету оптимізувати його зміст або інформативну силу, ми таким чином використовуємо як інструмент операції з уніфікації складових даного інформаційного блоку. Фактично відбувається корегування інформації шляхом зміни первинного коду, який складається з певних одиниць (лексичних, семантичних, графічних). До того ж переведення одних інформаційних одиниць в інші задля оптимізації виражає не просто процеси згортання/розгортання, а показує уніфікацію цих одиниць у спосіб надання первинному елементу індексу у вигляді вторинного, отриманого таким чином, елемента. Розуміння процесу згортання/розгортання як індексування дає можливість говорити про значення його в систематизації інформації – зведення розрізнених знань про предмети (явища) об'єктивної дійсності в єдину наукову систему, встановлення їх єдності.

Не варто також забувати, що у вужчому розумінні індексування – це процес опису документів і запитів у термінах інформаційно-пошукової мови (ІПМ) [4]. Тому і процеси згортання/розгортання інформації засобами індексної нотації мають характер та риси індексування. Як відомо, під ІПМ ми розуміємо формалізовану штучну мову, призначену для індексування документів, інформаційних запитів і опису фактів з метою подальшого зберігання й пошуку [5]. Серед найвідоміших ІПМ є: класифікаційна, дескрипторна, фактографічна, документальна, ІПМ предметних рубрик, ІПМ ключових слів, вербальна ІПМ.

На наше глибоке переконання, вербальна ІПМ заслуговує уваги, враховуючи тематичну дефініцію даного дослідження, оскільки ця ІПМ ви-

користує для представлення своїх лексичних одиниць слова й вислови природної мови в їх орфографічній формі. Тобто, якщо штучні мови абсолютно закономірно використовуються в процесі оптимізації інформації задля її систематизації (саме для цього вони й створювались), то використання лексичних одиниць природної мови заслуговує окремої уваги. Саме вербальні ІПМ можуть бути тим інструментом, який дійсно наблизить реальну еволюцію наук до відображення її позицій у інформаційній систематизації знань. Адже знання динамічні, вони змінюються внаслідок цілої низки факторів, відтворюючи таким чином еволюцію знання людського. Тому вербальні ІПМ, серед ряду інших, покликані найкомпетентнішим чином забезпечити здійснення процесів згортання/розгортання інформації в індексну форму, для подальшої систематизації. У процесі аналітичного огляду вербальних ІПМ, на нашу думку, варто вказати на ту мову, засобами якої можливе здійснення індексування широкого кола людських ресурсів. Сьогодні таку можливість забезпечує веб-орієнтований простір. Саме в Інтернеті представлено найбільше інформації та її споживачів, тому тут можливе здійснення повномірної систематизації інформації засобами колективного розуму.

Доступ мільйонів користувачів до одного ресурсу зумовлює реалістичність можливостей щодо колективізації авторства інформаційного продукту [7, с. 104–110]. Коли йдеться про індексування, як результат згортання інформації, такий процес часто називають фолксономією, тобто народною класифікацією (індексуванням). «Фолксономія» – процес формування бази координатних одиниць (тегів) в інтерактивному просторі. Етимологічно термін «фолксономія» розкривається як поєднання термінів «таксономія»¹ (з грец. розміщення по порядку + закон) та «folk» (з англ. народ). Наявність складової, що виражає належність даного процесу як до продукту народно-колективного розуму, цілком виправдана. Адже дійсно, до цієї технології індексування інформації вдаються задля того, щоб залучити якомога більше людей різних професійних та освітніх рівнів. З іншого ж боку відомо, що «фолксономічне» формування індексційної бази не є за принципом формування ієрархічною системою. Так звана «фолксономія» – це суто вербальна інформаційно-пошукова мова, і в системі, яку вона утворює, зв'язки виражено суто лексично із суб'єктивним семантичним значенням для кожного користувача. Тому використання термінологічної частини «таксономія» у формування розглядуваного терміна є не зовсім виправданим, враховуючи неповномірну відповідність його до

¹ Таксономія (з гр. *taxis* розміщення по порядку + *nomos* закон) – теорія класифікації і систематизації складно створених областей дійсності, що мають зазвичай ієрархічну структуру [8, с. 484].

семантичного навантаження терміна, що покликаний виражати процес індексування інформації колективним розумом. Безперечно, певний порядок у процесі формування «фолксономічної» системи є, але він зовсім не характеризується ієрархічними принципами створення. Тому тут ідеться про процес індексування інформації в чистому вигляді, тобто присвоєння індексів, формування певного списку. Що ж до іншої складової терміна вираження індексування колективним розумом, то тут, на нашу думку, необхідно говорити про використання терміна *populus* – (з лат.) народ. Таким чином, термін, що виражав би процес індексування інтерактивних документів за допомогою колективного розуму, утворюється поєднанням семантичних складових двох термінів – *populus* – (з лат.) народ та *index* – (з лат.) список, реєстр, покажчик. У результаті матимемо – *populindex*, тобто процес популіндексації – надання індексаційних координат інформаційним одиницям засобами колективного розуму. Тут слід зазначити багатогамну одноаспектність даного терміна щодо тлумачення його змісту як способу систематизації об'єктів, на основі зв'язків останніх з одним або декількома ключовими словами – тегами. Відповідно до суб'єктивного бачення функцій, значення, можливостей тощо властивих рис об'єкта, користувач присвоює певне ключове слово, за яким характеризує, і як наслідок здійснює процес пошуку даного об'єкта. Таким чином, ми спостерігаємо здійснення присвоєння певного індексу, у вигляді конкретного тегу, тому чи іншому поняттю. Очевидно, що ці індекси-теги абсолютно різні з лексичної точки зору, а із семантичної – однакові. Можна зробити висновок, що засобами популіндексаційної ІПМ можна здійснювати процеси згортання/розгортання, унаслідок чого семантична складова може бути збережена, а лексична, скоріше за все, може бути змінена. Адже внаслідок певної свободи в підходах індексування з популіндексаційною ІПМ, користувач не обмежений певними рамками, і тому індекс, що утворюється, може бути логічно зрозумілим тільки цьому користувачеві.

Процеси згортання інформаційного блоку в індекс-тег спрямовані в першу чергу на уніфікацію інформації для оптимізації систематизації. Оскільки систематизацією є присвоєння даним або документам індексів відповідно до правил вибраної інформаційно-пошукової мови, то й пошук цих даних відбуватиметься відповідно до структури засистематизованої інформації, а значить, і процеси розгортання інформації забезпечуватимуться представленням індексів-тегів у доступній формі для подальшої роботи з ними. Такою формою представлення є сформована відповідно до кількісної характеристики звернень за одним і тим самим тегом «хмара тегів». «Хмара тегів» – це візуальне вираження множинності звернень за допомогою різних ключових слів до об'єктів. Зо-

крема, в аспекті створення CMS² використовують систему анотування і пошуку на базі тегів («хмара тегів»). «Хмара тегів» з підтримкою морфології або морфологічного пошуку – це одне з останніх технологічних нововведень. У CMS включені механізми морфологічного розбору й повнотекстового індексування інформаційно значущого наповнення інформаційного ресурсу, на базі якого проводяться повнотекстовий пошук і подальша побудова «хмари тегів». Відповідно, у процесі формування «хмари тегів» здійснюється оптимізація конкретного сайту або порталу, на якому й представлено таке групування тегів-індексів. Таким чином, ми знову доводимо те, що уніфікація інформації, шляхом її правильної систематизації, покликана забезпечити оптимальність пошуку заіндексованої інформації. І далі вже йдеться про пошукову оптимізацію (англ. search engine optimization, SEO) – комплекс заходів, що спрямовані на підвищення позицій сайту в результатах видачі пошукових систем за певними запитами користувачів (зазвичай, чим вище позиція сайту в результатах пошуку, тим більше зацікавлених відвідувачів переходить на нього з пошукових систем) [9]. Пошукова система враховує такі параметри сайту при обчисленні релевантності сайту: частота ключових слів (складні алгоритми сучасних пошукових систем дають можливість проводити семантичний аналіз тексту, щоб зменшити пошуковий шум, коли ключове слово зустрічається дуже часто (більше 7–8 % від усього вмісту) або дуже рідко (1–3 %) на сторінці. Оптимальна щільність ключових слів: 5–7 % від текстового вмісту веб-сторінки, при цьому ключові слова (фрази) мають зустрічатися в тексті як мінімум 3–4 рази); індекс цитування сайту або кількість веб-ресурсів, що посилаються на даний сайт (багатьма програмно-апаратними комплексами з веб-інтерфейсом, що надають можливість пошуку інформації в Інтернеті, не враховуються координатні посилання. Важливо також, щоб посилання із сайтів були схожої тематики з тематикою сайту, що оптимізується. З вищенаведеного зрозуміло, як оптимізація інформації, шляхом згортання/розгортання, у вигляді індексів-тегів, приводить до оптимізації пошуку інформації, а значить, і до оптимізації і підвищення рівня веб-ресурсів та їх релевантності.

² Система керування вмістом (СКВ; англ. Content Management System) – комп'ютерна програма або система, що використовується для забезпечення й організації сумісного процесу створення, редагування й керування текстовими та мультимедійними документами (вмістом або контентом). Цей вміст розглядається як неструктуровані дані наочного завдання в протилежність структурованим даним, що зазвичай перебувають під управлінням сукупності програмних і лінгвістичних засобів загального або спеціального призначення, які й забезпечують керування створенням і використанням баз даних [6, с. 681–708].

Таким чином, підсумовуючи, можна констатувати, що процес здійснення згортання/розгортання не просто складний, а фактично складно-складений. Адже розуміючи значення самої ідеї даного процесу, можна говорити про наявність у першу чергу метаінформаційної складової як відправної точки. Ця метаінформаційна складова є по суті «інформацією інформації», що характеризує глибинність змісту, себто змістовності. Уніфікація ж змістовності можлива тільки в плані вдосконалення, шляхом розкриття усіх її граней та дотичних. Що ж до змісту, то тут оптимізація можлива як у семантичному розумінні, так і в лексичному. Нагадаємо, що оптимізацію треба розглядати не тільки як певний процес, а й як об'єкт процесу, з огляду на те, що саме в окресленості форми ми можемо читати змістовне навантаження процесу в цілому. І тому тут вже йдеться про «інформацію про інформацію». Отже, уніфікація інформації засобами згортання/розгортання є фактично передумовою і професійним поясненням значення та потреби систематизації інформаційних даних для адекватного представлення знання у веб-орієнтованому просторі відповідно до плинності та динамічності цього наукового знання, як рушійної сили еволюційності розвитку класифікації й організації інформації.

Список використаних джерел

1. *Блюменау Д. И.* Информационный анализ/синтез для формирования вторичного потока документов / Д. И. Блюменау. – СПб. : Профессия, 2002. – 240 с.
2. *Блюменау Д. И.* Информация и информационный сервис / Д. И. Блюменау. – Л. : Наука, 1989. – 192 с.
3. *Блюменау Д. И.* Проблемы свёртывания научной информации / Д. И. Блюменау. – Л. : Наука, 1982. – 166 с.
4. ГОСТ 7.74-96. СИБИД. Информационно-поисковые языки. Термины и определения. – Изд. офиц. – М., 1997. – 35 с.
5. ДСТУ 2398-94. Інформація та документація. Інформаційні мови. Терміни та визначення. – Видан. офиц. – К., 1994. – 73 с.
6. *К. Дж. Дейт.* Введение в системы баз данных = Introduction to Database Systems / Дейт К. Дж. – 8-е изд. – М. : Вильямс, 2006. – 1328 с.
7. *Сербін О. О.* Бібліотечно-бібліографічні класифікації: історична еволюція та сучасні тенденції розвитку / О. О. Сербін ; наук. ред. О. С. Онищенко ; НАН України, Національна бібліотека України ім. В. І. Вернадського. – К. : НБУВ, 2009. – 139 с.
8. Словарь иностранных слов. – 14-е изд., испр. – М., 1987. – 608 с.
9. High accessibility is effective search engine optimization [Electronic resource]. – Mode of access: <http://www.alistapart.com/articles/accessibilityseo>. – Title from the screen.