

УДК 004.891.3

к.т.н., доц. Муляр І.В. (ХмНУ)
д.т.н., доц. Шкуліпа П.А. (ОДАТРЯ)
Романовська І.А. (ХмНУ)
Ряба Л.О. (ВІКНУ)

РОЗРОБКА МЕТОДУ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ВИДАЧІ РЕЗУЛЬТАТУ ПОШУКУ В ІНФОРМАЦІЙНОМУ СЕРЕДОВИЩІ

У статті проведено аналіз принципів роботи пошукових систем, описані особливості роботи. Виділені основні методи роботи систем, та їх недоліки. Представлено класифікацію пошукових запитів, та проблеми при їх відповідності до шуканої інформації, описуються поняття релевантності та пертинентності, виданих документів, на запит користувача. В статті розглядається проблема задоволення потреб загально цільової аудиторії яка немає навичок у роботі з системами пошуку Збільшення пертинентності за умови модифікації алгоритму ранжування.

Ключові слова: пошукова система, запит, пертинентність, інформація, пошук.

Вступ. Чим далі людина крокує в майбутнє, тим більше поєднує своє співіснування з цифровими технологіями. Вже маже неможливо уявити існування людини без комп'ютерних технологій, а в свою чергу комп'ютери - без мережі. Глобальна мережа Інтернет та комп'ютерні технології представляють сьогодні один з найважливіших факторів науково-

технічного прогресу. Інформаційний простір та розвиток мережевих технологій активно розширюється з кожним днем. Тільки в індексованій частині нараховується понад 20 мільйонів документів і більше 200 мільйонів веб-сайтів. В даний час всі паперові публікації та документи оцифровуються та розміщуються в електронних сховищах різного призначення, активно використовуються електронні бібліотеки та архіви, пошукові машини та сервіси. В зв'язку з динамічним зростанням у об'ємі інформації, посилюється актуальність та значимість досліджень в області інформаційного пошуку, завдяки своєчасному отриманні потрібної інформації.

Організація швидкого і точного пошуку інформації в мережі Інтернет – одна з областей теорії інформаційного пошуку, яка активно розвивається. В якості інструмента інформаційного пошуку, використовуються пошукові системи, популярними з яких є: Google, Yandex, Rembler і тд.

В інформаційні пошукові системи пошук документів виконується по сформованому користувачем запиту. Запит користувача створюється з ключових слів який може вміщувати як і одне, так і декілька слів. Деякі пошукові системи пропонують розширений пошук, який включає декілька функцій, щоб полегшити пошук і уточнити ключові слова. Адже, різні пошукові системи мають свої переваги і недоліки.

Існує багато характеристик пошуку, але дві визнані основними - це повнота (recall) і точність (precision). Багато уваги в даний час відводиться такій смисловій характеристиці, як пертинентність. Ця характеристика інформаційно-пошукових систем означає відповідність отриманих в результаті пошуку документів інформаційним потребам користувача, а не формальному відповідності документа запиту.

Постановка задачі. Існує широкий спектр вирішень та перспективних напрямлень в області інформаційного пошуку. Починаючи від постановки глобальних розподілених інформаційних структур і пошукових систем, закінчуючи аналізом документів з допомогою семантики. Всі вони є важливими при вирішенні своїх задач. Але саме від методів аналізу залежить ефективність існуючих пошукових систем, так, як вони є основою в кожній пошуковій системі і в деякій мірі вони визначають можливість і обмеження цих систем.

На сьогоднішній день багато пошукових систем з організацією повнотекстового пошуку не враховують в достатній мірі семантику. Коли саме семантика обслуговує високу степінь пертинентності. Нажаль, не завжди користувач може чітко та викладно сформулювати набір ключових слів для пошуку інформації, щоб отримати бажаний результат.

Отже, нерідко можна зустріти користувача який слабо проінформований в темі яку він шукає, і починає вводити різні альтернативи, синонім та інші слова близькі до теми. Саме це ставить актуальність і значимість проблеми інформаційного пошуку. Проблема відображення інформаційних потреб користувачів які і підвищують пертинентність пошуку.

Кожна пошукова система по-своєму виставляє ряд результатів, які б мали задовільнити потребу користувача. І в кожній системі ці ряди відрізняються, оскільки, різні алгоритми роботи, інші способи ранжування сторінок. Також немало важливим фактором являє собою авторитетність сторінки, яка кожної системи визначається по-своєму. Потрібно розробити і дослідити спосіб підвищення показників пертинентності інформаційного пошуку, заснованого на концепції інтерфейсної пошукової системи, що здійснює об'єднання і впорядковує відгуки на запити користувачів популярних пошукових систем мережі Інтернет. Для отримання поставленої мети потрібно модифікувати загальну формулу розрахунку релевантності пошуковими системами Інтернет за рахунок формалізації впливу власних і внутрішніх чинників ранжирування пошукових систем.

Виклад основного матеріалу досліджень. Актуальність проблеми пошуку і обробки текстової інформації сьогодні не викликає сумнівів. Щодня ми стикаємося з необхідністю її ефективного вирішення на роботі і вдома. Багато сфер нашої діяльності тісно пов'язані з електронними технологіями, які отримали за останні два десятиліття широке поширення в нашій країні. Так, для того щоб знайти необхідну книгу або ноти не обов'язково виходити з

дому, а для того, щоб дізнатися курс валют або акцій йти на біржу. Більшість інформації можна знайти в загальнодоступних мережах, однак більшість цих мереж містять інформацію невпорядковану, що призводить до необхідності її пошуку. Так, як більшість інформації представлено в тексті, найбільш часто встає перед користувачем завдання повнотекстового пошуку.

В даний час інформаційні ресурси тільки веб-простору складають понад двадцять мільярдів документів, до яких можливий вільний доступ будь-якого користувача. Природно, для того, щоб знайти необхідну інформацію цієї найбільшої розподіленої повнотекстової бази даних необхідно використовувати найпотужніші пошукові системи. Такі системи існують і конкурують один з одним. Можна виділити наступні основні методи пошуку інформації в Інтернеті, які в залежності від цілей і завдань шукають та використовуються окремо або в комбінації один з одним:

1. Безпосередній пошук з використанням гіпертекстових посилань

Оскільки, всі сайти в просторі WWW фактично виявляються пов'язаними між собою, пошук інформації може бути проведений шляхом послідовного перегляду пов'язаних сторінок за допомогою браузера. Хоча, цей повністю ручний метод пошуку виглядає повним анахронізмом у мережі, що містить більше 60 млн. вузлів.

2. Пошук з використанням пошукових машин

Як правило, застосування пошукових машин засноване на використанні ключових слів, які передаються пошукових серверів у якості аргументів пошуку. Якщо робити все правильно, то формування списку ключових слів вимагає попередньої роботи по складанню тезауруса.

3. Пошук із застосуванням спеціальних засобів

Цей повністю автоматизований метод може виявитися досить ефективним для проведення первинного пошуку. Одна з технологій цього методу заснована на застосуванні спеціалізованих програм - спайдерів, які в автоматичному режимі переглядають Web-сторінки, відшукуючи на них шукану інформацію. Фактично, це автоматизований варіант перегляду з допомогою гіпертекстових посилань (пошукові машини для побудови своїх індексних таблиць використовують схожі методи). Немає потреби говорити, що результати автоматичного пошуку обов'язково вимагають подальшої обробки. Застосування даного методу доцільно, якщо використання пошукових машин не може дати необхідних результатів (наприклад, в силу нестандартності запиту, який не може бути адекватно заданий існуючими засобами пошукових машин). У ряді випадків цей метод може бути дуже ефективний. Вибір між використанням спайдера або пошукових серверів, являє собою варіант класичного вибору між застосуванням універсальних, або спеціалізованих засобів.

4. Аналіз нових ресурсів

Пошук по новоствореним ресурсам може виявитися необхідним при проведенні повторних циклів пошуку, пошуку найбільш свіжої інформації або для аналізу тенденцій розвитку об'єкта дослідження в динаміці. Іншою можливою причиною може з'явитися те, що більшість пошукових машин оновлює свої індекси зі значною затримкою, викликаній гігантськими обсягами оброблюваних даних, і ця затримка зазвичай більша, ніж популярність теми, що цікавить. Це міркування може виявитися досить істотним при проведенні пошуку в вузькоспеціалізованій предметній області.

В основу роботи всіх пошукових систем покладені спеціальні алгоритми, які є модифікаціями основних підходів - моделей пошуку основу традиційних методів покладені три основні підходи, перший з яких базується на теорії множин (булева модель), другий - на векторній алгебрі (векторно-просторова модель), а третій - на теорії ймовірностей (імовірнісна модель). Ці підходи можуть примінятися на практиці і в канонічному вигляді, проте у них є спільний недолік, обумовлений припущенням, що зміст документа визначається безліччю слів і стійких словосполучень - термів, які входять в нього без урахування взаємозв'язків, більш того, вважаються незалежними. Звичайно ж, таке припущення веде до втрати змістовних відтінків, проте воно дозволяє реалізувати пошук і

групування документів формальними ознаками. Відомі такі основні недоліки традиційних моделей:

1. Булева модель - невисока ефективність пошуку, відсутність операторів, неможливість ранжування результатів пошуку.

1. Векторно-просторова модель пов'язана з розрахунком масивів високої розмірності і в канонічному вигляді малоприсаєдана для обробки великих масивів даних.

2. Імовірнісна модель характеризується низькою обчислювальною масштабованістю (т. е. різким зниженням ефективності при зростанні обсягів даних), необхідністю постійного навчання системи.

Крім представлених нижче, існують і інші моделі пошуку, наприклад, семантичні, в рамках яких робляться спроби організації смислового пошуку, за рахунок аналізу граматики тексту, використання баз знань, тезаурусів, онтологій, які реалізують семантичні зв'язку між окремими словами і їх групами [2].

Отже розглянемо самий алгоритм роботи системи. Користувач вибирає тип запиту («в мережі Інтернет», «в локальній БД», «в локальній мережі») і вводить пошуковий запит на природній мові.

1. Запит обробляється, з нього витягуються ключові слова. При цьому використовується розширення запиту за рахунок використання словника синонімів (Використовується словник синонімічних предикатів і словник синонімічних іменних груп), з запиту викидаються стоп-слова і т. д.

2. Перетворений таким чином запит відправляється, або на декілька пошукових машин Інтернет, або в локальну БД (пошук здійснюється серед поміщених в ній документів), або виконується в папках локальної мережі (в даний момент підтримуються такі формати, як HTML, TXT, MS Word, MS Excel і MS PowerPoint).

3. Знайдені документам обробляються і поміщаються повнотекстову базу даних системи.

4. Запит користувача і знайдені документи піддаються лінгвістичному аналізу, що включає морфологічний, синтаксичний і поверхневий семантичний аналіз, будуються семантичні образи запиту і документів, проводиться порівняння образів і обчислення семантичної релевантності знайдених документів запитом користувача

5. Знайдені документи сортуються по обчисленій релевантності. Низькорелевантні документи відкидаються.

Інформаційний пошук в системі виконується на основі введеного користувачем запиту на пошук необхідної інформації. Сама ця потреба в пошуку конкретної інформації носить звання інформаційної потреби. Як тільки користувач отримує якусь кількість інформації, то його інформаційна потреба постійно змінюється. В зв'язку з тим її неможливо однозначно виразити і описати. Але інформаційна потреба, в деякі проміжки часу може бути послідовна але фіксована. Отже інформаційна потреба в певні проміжки часу на природній мові являє собою інформаційний запит.

В свою чергу запити можуть відрізнятися за різними критеріями і нести за собою ті чи інші результати. Тож розглянемо класифікацію пошукових запитів [3].

Класифікація багатослівних пошукових запитів. Використовуючи пошук, користувачі можуть переслідувати різні цілі. Якщо вони хочуть знайти сайт організації, людини, фірми, послуги, то такий запит називається навігаційним. Навігаційні запити повинні приводити користувача до конкретної сторінки мережі Інтернет. Інформаційні запити дозволяють користувачам знайти конкретні відомості в мережі Інтернет. Вони служать для пошуку інформації, але зовсім не обов'язково призводять на конкретний сайт. А транзакційні запити використовують користувачі для здійснення певних дій через мережу Інтернет. Транзакційні запити демонструють намір користувача зробити конкретну дію - «купити», «продати», «зняти» і т. д.

За багатослівності або чисельним складом слів пошукового запиту прийнято виділяти однослівні і багатослівні (двох, трьох, чотирьох і т.д.) запити. Однослівні пошукові запити складаються з одного слова. Багатослівні пошукові запити складаються з декількох слів

Останнім часом спостерігається зростання числа багатослівних запитів. Це можна пов'язати зі збільшенням числа Інтернет сторінок і підвищенням навичок пошуку користувачів.

За класифікацією чіткості формулювання

За чіткістю формулювання всі пошукові запити можна умовно розділити на чіткі і не чіткі запити.

Чіткими прийнято називати запити, які містять назви або точні описи шуканих об'єктів. На противагу чітким запитам, пошукові запити; з яких явно не зрозуміло, що саме шукає користувач, прийнято називати нечіткими. Часто користувачі не можуть чітко сформулювати свій запит через невміння роботи з пошуковою системою або недостатнього розуміння власної проблеми.

3) Класифікація по конкурентності запиту

Висококонкурентні запити - це ті, для потрапляння на першу сторінку яких, буде потрібно витратити чимало сил, грошей і часу. При цьому, велика кількість сайтів, під впливом оптимізаторів, борються за право бути в лідерах в цій пошуковій видачі. Низькоконкурентні запити є легкою здобиччю для оптимізаторів, тому що по них немає високої конкуренції, а значить потрапити в топ по видачі цим запитам порівняно легко і не складно.

4) Класифікація на основі частоти запиту

Ще один вид умовної класифікації - класифікація на основі частоти запиту, яка враховує, як часто користувачі використовують даний пошуковий запит. Низькочастотні запити характерні порівняно низьким числом використання даного запиту. Зазвичай, вимірюються в числі показів результатів пошукових запитів за місяць. Дана класифікація умовна, так, як для одної людини, вважається малим числом, наприклад, один мільйон запитів на місяць, інший може назвати це число дуже великим. Інший тип запитів - високочастотні, здатні приносити великий трафік на сайт, але конкуренція за ним, як правило, значно вища. Однак, великий попит високочастотним запитам у користувачів зовсім не гарантує сайту благополучне існування.

5) Класифікація комерційної привабливості запиту

Запити можуть мати високу і низьку комерційну привабливість. Залежно від цього часто визначається вартість просування по запиту. Оскільки, запити з високою комерційною привабливістю, як правило, висококонкурентні.

Як вказується вище запит може бути неправильно сформульований споживачем і відводити його від істинної інформаційної потреби в момент звернення до системі. Таким чином, при проведенні інформаційного пошуку в системі фактично розглядається не інформаційна потреба користувача, а тільки інформаційний запит, у відповідь на який і видаються ті чи інші документи. Отже, реакцію системи необхідно розглядати не тільки по відношенню до інформаційної потреби, але і по відношенню до інформаційного запиту. Для вираження цих відносин в теорії Діпсі введені два фундаментальних поняття: пертинентність і релевантність. Під пертинентністю розуміється відповідність змісту документа інформаційної потреби споживача. [1] Документи, зміст яких відповідає інформаційної потреби, називають пертинентними. Досягнення високої пертинентності – основне поле конкурентної боротьби сучасних пошукових систем. Релевантність, представляє собою, відповідність змісту документа інформаційного запиту в тому вигляді, в якому він сформульований, а документи, зміст яких відповідає запиту споживача, носять назву релевантних.

Умовну формулу для обчислення показника релевантності можна уявити в такий спосіб [6]

$$Ra(x) = m - Ta(x) + p - La(x) \setminus F(PRa)$$

де $Ra(x)$ - підсумкове значення релевантності документа а запитом x ;

$Ta(x)$ - значення релевантності коду документа a запитом x ;
 $La(x)$ - релевантність тексту посилань з інших документів на документ a запитом x - цей показник для стислості, можна назвати релевантність посилань,
 PRa - показник авторитетності сторінки a щодо запиту x ;
 $F(PRa)$ - деяка монотонна нефіксована функція, причому $F(0)=1$; допускають спрощення, що $F(PRa)=(L+q-PRa)$;
 m, p - деякі коефіцієнти, що дозволяють апроксимувати формулу релевантності під кожну ПС.

В даний час, ще не розроблено алгоритмів, які дозволяють нам адекватно оцінити пертинентність результатів пошуку. Історично алгоритми пошукових систем були націлені на збільшення релевантності пошуку, від якої в деякій мірі залежить і пертинентність. Але ця залежність не має виражений характер. Тому для оцінки пертинентності не обійтися без залучення експертів - фахівців у вузьких предметних областях, думки яких дозволяють формувати пертинентні набори документів з відгуків різних пошукових систем. Кількісний аспект поняття пертинентності результатів пошуку традиційно визначається, як відношення кількості інформації корисної для користувача до загальної кількості інформації, отриманої користувачем. Оцінка саме кількості, а не обсягу інформації, що міститься у видачі пошукової системи, є не тривіальним завданням. Тому без істотного збитку для цілей дослідження показник пертинентності $\sigma(\omega)$ результатів пошуку, представлених конкретною пошуковою системою, можна визначати як відношення числа N_{pert} пертинентних (корисних для користувача) документів у видачі до загальної кількості N документів цієї видачі [4]:

$$\sigma(\omega) = N_{pert} / N$$

де N - деяке фіксоване кількість документів, розміщених на видачу першими. Зазвичай $N=100$

Оскільки, факт корисності документа користувачу може бути встановлено тільки самим користувачем, то визначення значення N_{pert} встановлюється експертним шляхом в процесі тестування і налаштування пошукової системи, і користувачами в процесі експлуатації цієї системи.

Сформулюємо вимоги, яким повинен задовольняти показник пертинентності видачі для того, щоб його зручно було використовувати для порівняння декількох видач. Показник повинен відображати наступне:

- 1) фінальний результат перегляду видачі, а саме - ступінь задоволення інформаційної потреби користувача-, досягнуту в процесі і ознайомлення з представленими документами;
- 2) тривалість роботи користувача з виданням до прийняття рішення по завершенні процесу її вивчення.

Ранжування застосовується у випадках, коли неможлива або недоцільна безпосередня оцінка. При цьому ранжирування об'єктів містить лише інформацію про те, який з них кращий, і не містить інформацію про те, наскільки або в скільки разів один об'єкт переважає іншого.

Метод простого ранжування полягає в тому, що система має об'єкти ранжирування (наприклад, критерії) в порядку спадання їх значущості (скажімо, для альтернатив це спадання переваги). Ранги позначаються цифрами від 1 до n , де n - кількість рангів. Сума рангів при цьому буде дорівнює сумі чисел натурального ряду [5]:

$$S = n * (n + 1) / 2.$$

Фактори, що впливають на ранжування [4], діляться на статичні, динамічні і власні: Статичні фактори не залежить від запиту до ПС, наприклад авторитетність сторінки. Динамічні фактори залежать від запиту і. поділяються на внутрішні (Організація документа) і зовнішні (ранжирування посилань). На значення впливає розташування шуканого тексту запиту а на сторінці і частота народження слів запиту в даному документі, тобто частка слів з тексту запиту в загальній кількості слів документа. Більший вагу мають слова, розташовані

ближче до початку документа. особливо цінується точне входження тексту запиту в текст документа. Слід зазначити, що пошук по стоп-словами (сполучники, вигуки, прийменники; частки) не проводиться, але вони враховуються при ранжируванні документів, що може серйозно впливати на результат видачі.

Отже можемо зазначити що системи пошуку видають результати на запит в деякій послідовності вказуючи спираючись на авторитетність та деякі інші фактори і видати релевантні результати і в якійсь мірі вони є пертинентними. Показник авторитетності сторінки відноситься до статичних факторів ранжування, що не залежать від тексту запиту. Він відображає вплив авторитетності сторінок із зовнішніми посиланнями на документ.

В результаті нашою метою роботи є дослідження способу підвищення показників пертинентності інформаційного пошуку, наприклад, інтерфейсу інформаційно пошукових систем. Потрібно об'єднати результати різних популярних пошукових систем і об'єднати впорядкувати результати на запит користувача. Для досягнення мети потрібно описати основний зміст факторів ранжування пошукових систем на формальній мові. Також модифікувати загальну формулу розрахунку релевантності пошуковими системами Інтернет за рахунок формалізації впливу власних і внутрішніх чинників ранжирування пошукових систем. Оскільки, пошукові системи є конкурентами, а їхні результати на запити є різним і кожен в свою чергу пертинентними, можливо створення, за допомогою якого, потрібні результати будуть проходити обробку модифікованого алгоритму ранжування.

Висновки. На основі методів та алгоритмів пошукових систем розглянуто принцип роботи користувача з пошуковою системою. Наведено ряд класифікацій запитів, які в свою чергу поділяються за цілями, частотою, чіткістю і т.д. Перераховані недоліки моделей пошукових систем. Представленні поняття релевантності та пертинентності. Представлений метод ранжирування документів та його фактори.

Виявлений ряд проблем, породжених вимогами до пошуку даних та стрімким збільшенням об'ємів самих даних, в якому користувач на даний момент найбільш потребує удосконалення систем пошуку надвеликих інформаційній павутинні.

ЛІТЕРАТУРА:

1. Струнгар А. Пертинентність і релевантність інформаційних ресурсів при пошуку інформації в електронних бібліотеках. / А. Струнгар. Наукові праці Національної бібліотеки України ім. В. І. Вернадського. – 2014.

2. Ланде Д.В. Інтернетка : Навігація в важких мережах: моделі та алгоритми / Д.В. Ланде, А.А. Снарський, І.В. Безсудов. – М: Книжний дом «ЛИБРОКОМ», 209. – 264с.

Терехов. А.А. Разработка методов и инструментальных средств повышения пертинентности поиска современных информационных средах: дис. кандидата техн. наук: 05.13.17 / Терехов Алексей Андреевич. – Р., 2010. – 165 с.

3. Методи оцінки і вибору альтернатив [Електронний ресурс] / Навчальні матеріали для студентів. - 2013. Дата оновлення 01.03.2014.: http://studme.com.ua/1274071412698/menedzhment/metody_otsenki_vybora_alternativ.htm(дата звернення: 28.04.2016)

4. Сегалович И.В. Как работают поисковые системы / И.В. Сегалович. – Мир Internet, 2002. – № 10.

REFERENCES:

1. Strungaru A. Pertinentnist i relevantnist informatsiynih resursiv at Poshuk Informácie in Electonics bibliotekah. / A. Strungaru. Naukovi pratsi Natsionalnoi Book Cabinets Ukraine IM. V. I. Vernadskogo. - 2014.

2. Lande DV Internetka: Navigatsiya in vazhkih trammel: modeli that algorithm. / DV Lande, AA Snarsky, I. V. Bezsudov. - M: bookshop "LIBROKOM" 209 - 264s.

3. Terekhov. A. Development of methods and tools to enhance the pertinence sought for modern IT environments: Dis. the candidate tehn. Sciences: 05.13. 17 / Alexey Terekhov. – Р., 2010. – 165.

4. i Metodi otsinki Vibor alternatives [Electron resource] / Navchalni materiali for studentiv. - 2013 date 01.03.2014. :

http://studme.com.ua/1274071412698/menedzhment/metody_otsenki_vybora_alternativ.htm (дата до updates zvernennya: 04/28/2016)

5. Segalovich I.V. How Search Engines Work / I.V. Segalovich // World Internet.2002. № 10.

Рецензент: д.т.н., проф. **Ленков С.В.**, начальник науково-дослідного центру Військового інституту Київського національного університету імені Тараса Шевченка

к.т.н. Муляр И.В., д.т.н., доц. Шкулипа П.А., Романовская И.А., Ряба Л.А.
РАЗРАБОТКА МЕТОДА ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ВЫДАЧИ
РЕЗУЛЬТАТОВ ПОИСКА В ИНФОРМАЦИОННОЙ СРЕДЕ

В статье проведены анализ принципов работы поисковых систем, описаны особенности работы. Выделены основные методы работы систем, и их недостатки. Представлена классификация поисковых запросов, и проблемы при их соответствии искомой информации, описываются понятие релевантности и pertinентности, выданных документов по запросу пользователя. В статье рассматривается проблема удовлетворения потребностей общеобразовательных целевой аудитории которая не имеет навыков в работе с системами поиска. Увеличение pertinентности при модификации алгоритма ранжирования.

Ключевые слова: поисковая система, запрос, pertinентность, информация, поиск

Ph.D. Mulyar I.V., Ph.D. Shkulipa P.A., Romanovska I.A., Ryaba L.A.
DEVELOPMENT METHODS TO INCREASE THE EFFECTIVENESS OF EXTRADITION
SEARCH RESULTS IN THE INFORMATION ENVIRONMENT

The article analyzes the principles of search engine described features work. Dedicated basic methods of operation systems and their shortcomings. The classification of search queries and problems of their compliance with the required information describes the concept of relevance and pertinence of documents issued to the user's request. The article considers the problem of satisfying the needs of a comprehensive target audience that does not have the skills to work with the search engines. Increasing modification conditions with pertinence ranking algorithm.

Keywords: search engine query pertinence, information, search

НАЦІОНАЛЬНИЙ
УНІВЕРСИТЕТ
ІМЕНІ
ТАРАСА ШЕВЧЕНКА
КИЇВ