

УДК 321.01:323

к.і.н., с.н.с. Любовець Г.В. (ВІКНУ)

ПОШУКОВІ СИСТЕМИ В МЕРЕЖАХ ІНТЕРНЕТУ

У статті розглядаються проблеми зростання масштабів і динаміки інформаційних потоків в сучасному інформпросторі. Різде збільшення обсягів інформаційних повідомлень та супутних чинників викликає потребу задоволення запитів користувачів через створення та розвиток адекватних пошукових систем.

У світі сегмент пошуковиків представлений досить широкою лінійкою пошукових систем як для пересічного користувача, так і для спеціалістів, які поглиблено вивчають різноманітні аспекти окремих наукових чи прикладних сфер за заданою тематикою.

Однак і такі пошукові системи не завжди відповідають необхідним критеріям. Тому виникає необхідність у створенні гібридизованих систем, де б відбувалось поєднання машинного та людського потенціалу в спеціальних регламентних алгоритмах.

Ключові слова: інформаційно-контентні потоки, пошукові системи, семантичні фільтри, моніторинг.

Поставлення завдання. Сучасне життя пересічного громадянина (не кажучи вже про спеціаліста) насичене різноманітною потоковою інформацією, яка надходить до нього із

різних джерел: у вертикально-горизонтальному рівні виробничих відносин від керівництва, колеґ, підлеглих, партнерів; в горизонтальному рівні – від засобів масової інформації, засобів масової комунікації, сусідів, друзів, знайомих тощо.

Однак не завжди ця інформація задовольняє потреби (як виробничі, так і повсякденні) індивіда. Для вирішення поставлених завдань сучасному фахівцю потрібно оперувати останніми даними за напрямом роботи, відслідковувати зміни в галузі (окремому секторі) і опиратися на новітні розробки. Тому виникає потреба оперативного та достовірного пошуку такої інформації. В сучасному інформаційному масиві (інформаційно-контентних потоках) відшукати необхідні повідомлення досить складно. В такому випадку на допомогу приходять пошукові системи загального та спеціального користування. В повсякденному житті потреби переважної більшості користувачів цілком задовольняють пошукові системи загального глобального користування. Спеціальні пошукові системи застосовуються у вузькоспецифічних сферах: академічні наукові розробки, дослідження сегментних сфер виробництва, креативні індустрії, конкурентна розвідка тощо.

Останнім часом досить стрімко розвиваються соціальні медіа та соціальні мережі. І нерідко саме там можна знайти необхідну інформацію або зробити відповідний запит (уточнення, обмін). Це спричинило розвиток моніторингових систем соціальних мереж.

Для повноцінного використання можливостей пошукових систем потрібно орієнтуватися в їхньому розмаїті, знати переваги і недоліки кожного із пошуковиків, грамотно застосовувати для пошуку необхідної інформації тощо.

Отже, на першому етапі підготовки до роботи з пошуковими системами необхідно знати основні пошуковики загального та спеціального користування, розрізняти їхні відмінності та напрям застосування, а також орієнтуватися в можливостях кожного з них.

Аналіз наукових досліджень і публікацій. Із зростанням масштабів інформаційних потоків виросла динаміка інформаційного пошуку. Надвеликі обсяги різноманітних і різномасштабних повідомлень, багаторазове і поліваріантне дублювання інформації, наявність спаму та шумової інформації - всі ці та багато інших чинників призводить до нагальної потреби мати адекватні пошукові системи, які би задовольняли запити користувачів. Варто зазначити, що на сьогодні існує достатня кількість різних пошуковиків, адаптованих під спеціальні потреби, які пропонують користувачам лідери у сфері інформаційного пошуку (Google, Yahoo, AltaVista, Bing, Яндекс, Rambler тощо). Для спеціалізованого пошуку, які мають на меті поглиблене вивчення окремих аспектів суспільних, економічних, природничих чи інших домінантів, пропонується низка пошуковиків заданої тематики. Дослідженням різноманітних аспектів розробки, функціонування та використання таких систем присвячені роботи І. Безсуднова, С. Брайчевського, А. Додонова, О. Дубинського, Д.Ланде, А. Снарського, В. Путятіна та інш.

Серед іншого, авторами досліджується проблематика своєчасного та ефективного отримання багатоаспектної та об'єктивної інформації з інформпростору для подальшого використання в управлінській діяльності, що, відповідно потребує сучасних високотехнологічних рішень.

Вступ. На першому етапі функціонування інформаційно-пошукових систем у світі (приблизно середина минулого століття) перед ними ставились досить прості задачі:

- знайти необхідну інформацію,
- зберегти її в початковому вигляді (оригінал),
- забезпечити можливість додавання нових матеріалів (видалення непотрібних),
- мати функцію витягу потрібної інформації за потребою,
- систематизації у відповідних блоках (папках).

З часом вимоги до рівня пошукових систем значно зросли. До характеристик додалися такі вимоги як швидкість пошуку інформації, повнота, релевантність, ранжування результатів, агрегація, оперативність оновлення баз даних тощо.

Такі системи були запропоновані провідними світовими ІТ-компаніями та корпораціями, зокрема Google, Yahoo, AltaVista, Bing, Яндекс, Rambler тощо.

Основна частина. Коротко розглянемо деякі з відомих пошуковиків загального користування.

Лідером серед пошукових систем є Google. Його "працездатність" перевищує 40 мільярдів запитів в місяць. До того ж, пошук може вестись відповідними підсистемами, кількість яких тільки зростатиме з часом.

Найзручнішими для проведення традиційного моніторингу є, на наш погляд, наступні:

Google Alerts, розширений пошук Google, Google Trends, Google Correlate, Google Fusion Tables, Google Tables, Google Public Data Explorer, Google Maps – Streetview, Google Images.

Ще одна перевага - для пошуку може застосовуватись десять мов. Окрім цього, система перекладу дозволяє швидко перекласти необхідний матеріал, що дає змогу зекономити час при традиційному моніторингу інформаційного простору у визначеному сегменті та за вказаною тематикою. Це, своєю чергою, дає змогу пришвидшити аналітичну розробку заданої теми.

З цих десяти пошукових підсистем 6 доступні українською мовою. З них найбільший інтерес представляють такі.

Google Alerts надає можливість постійного оновлення (за визначеною замовником частотою - щодня, щотижня або щомісяця) тематичної інформації за вказаними напрямками (наприклад, новини) та доступною мовою з переліку можливих. За необхідності, такий пошук здійснюється в автоматичному режимі і потрібна інформація надходить на заданий аккаунт Gmail.

Розширений пошук Google досить зручний для моніторингу, наприклад, при підготовці журналістських матеріалів. Тут застосовуються такі критерії як "0", "+" і "-", є можливість використовувати для пошуку доменні імена (gov, ua, ru, gt тощо). Підсистема дозволяє також використовувати пошук за типом файлу (doc, rtf, pdf).

За допомогою **Google Trends** можна відслідковувати теми, в яких зацікавлений замовник та одночасно перевіряти ефективність власних публікацій. Таким чином, здійснюється так званий "власний моніторинг".

Якщо ви працюєте з базами даних (що є зараз однією з умов успішного виконання аналітичної роботи), тоді краще всього працювати з **Google Public Data Explorer**. Ця підсистема Google дозволяє отримувати дані з бази Організації об'єднаних націй та Світового банку і вибудовувати відповідні графічні конструкції. Однак ефективність при цьому можлива при користуванні англійською.

З квітня 2015 року Google вивів свої пошукові алгоритми на новий рівень – наступною новинкою стає ранжування сайтів, які добре зарекомендували себе на мобільних пристроях. Але тут існують певні перестороги до подачі матеріалів на таких сайтах. До них відносяться, передусім, розмір тексту, простір між посиланнями та величина матеріалу, тобто чи поміщається він на екрані пристрою. Власники сайтів мають можливість перевірити комутацію матеріалів сайту з мобільними пристроями на сторінці Mobile Friendly Test.

Такий підхід заставляє низку сайтів переналаштовувати внутрішні композиції матеріалів з урахуванням нововведення Google.

Ми детально зупинились на розгляді можливостей пошукової системи Google як основного інструментарію в роботі з моніторингом світового інформаційного простору **первинного рівня**.

Звичайно, такі пошукові системи як **Yahoo!** (американська компанія, друга в рейтингу серед пошукових систем світу), **Yandex** (найбільша на сьогодні російська ІТ-компанія, яка задовольняє потреби російськомовного інтернету, є четвертою в рейтингу пошукових систем світу за кількістю оброблених запитів), **Rambler.Ru** (заснована як перша пошукова система російськомовного інтернету, тривалий час очолювала російський ринок пошуковиків) та інші теж цілком можуть задовольнити потреби моніторингу інформаційного простору. Використання тієї чи іншої пошукової системи залежить від поставленого завдання, визначеної тематики та, зрештою, особистих симпатій користувача.

Світовими лідерами пошуку однозначно залишаються Google, Yahoo, Baidu, AltaVista, Bing, Яндекс, Rambler та ще деякі. Однак пошукових систем достатньо у кожній країні. Не маючи змоги чи не ставлячи перед собою мету вийти на міжнародний рівень такі пошуковики намагаються укріпити свої позиції в сегментному полі. Лідером тут є США, Німеччина, Канада, Австралія, Франція, Великобританія, Іспанія тощо. Досить бурхливо розвивається цей напрям останнім часом у Китаї. Достатньо сильні позиції займає Росія, що є важливим для всього пострадянського простору.

Україна тривалий час за даним напрямом знаходилась в полі тяжіння Росії. Ситуація загалом не сильно покращилась за останні роки. Але поступово ця ніша стає все більше українською. Починаючи, власне, від розвитку україномовного сегменту світових лідерів пошукових систем і закінчуючи становленням вітчизняного розробника пошуковиків.

З українських пошукових систем для проведення моніторингу найдоцільніше використовувати системи META.UA, UAport, Bigmir)net. Кожна з них має свої переваги. Так, META.UA проводить повнотекстовий пошук, в який заведено особливості української та російської морфології. Аналізується весь сегмент українського інтернету, за потреби задається тематична рубрика тощо. UAport може вести пошук кількома мовами, опирається на власні архіви української преси. Bigmir)net виконує пошук через Google або Яндекс.

Однак і такі визнані "метри" інформаційного пошуку як Google, Yahoo, AltaVista, Bing, Яндекс та інші на даному етапі вже не в змозі справитись з динамічним зростанням інформаційних потоків і, відповідно, задовольнити потреби замовників.

Серед проблем, які постають останнім часом перед інформаційно-пошуковими системами, наприклад, неможливість забезпечення індексації мережевого простору та "бачення" всіх веб-сторінок.

Відтак ці та інші проблеми на даному етапі вирішують метапошукові системи. Вони ведуть пошук одночасно в декількох пошукових системах, але з врахуванням особливостей кожної з них. При цьому власних баз даних такі пошуковики не мають. Це вигідно з точки зору економії часу, оскільки не потрібно робити запит по кожній пошуковій системі окремо. З технологічного боку перевагою таких пошуковиків є "можливість підтримки індексних даних в актуальному стані, оскільки кожна локальна пошукова система охоплює власний фрагмент веб-простору" [1].

Поглиблений та спеціалізований моніторинг замовленої тематики в інформаційному просторі. Наведені вище приклади роботи інформаційно-пошукових систем, за допомогою яких здійснюється традиційні моніторинги заданої тематики в інформаційному просторі, задовольняють потреби більшості замовників. Однак низка організацій, як державного сектору, так і бізнесу, потребує більш глибокого, сегментного моніторингу для проведення всебічного аналізу стану справ у визначеній сфері та перспектив її розвитку. Відомі пошуковики з такими задачами справитися вже не можуть. Тим більше, що кількість сайтів і персональних блогів в мережі, за свідченням англійської компанії Netcraft, у вересні 2014 року перетнула рубіж у мільярд [2]. Коли Netcraft починала вести аналітику розвитку світової мережі у 1995 році, на той час ця цифра складала лише 18 957 сайтів. Після 2000 року, щоб забезпечити підрахунок змін, що відбуваються у цьому сегменті, компанія змушена була запустити метричну систему, яка працює щоденно протягом року.

Відтак на зміну простим пошуковикам приходять спеціалізовані системи і сервіси, для яких розробляється відповідне програмне забезпечення. До таких професійних пошуковиків, які проводять моніторинг інформаційного простору відносяться, серед інших, такі [3].

В ніші програмних продуктів:

Sopernic Agent (використовує до 1000 пошукових систем, об'єднує результати, показує найрелевантніші результати); Avalanche (веб-моніторинг базується на технології концепції "розумних" папок, автономному інтелектуальному пошуковому роботі та вбудованій базі даних, яка носить назву "персональної енциклопедії"), WebSite-Watcher (проводить моніторинг веб-сайтів, форумів, локальних сайтів, забезпечує фільтрацію інформації та зручну візуалізацію результатів).

В якості **сервісних рішень** наводяться приклади WatchThatPage (здійснює в автоматичному режимі збір нової інформації з ресурсів, що моніторяться), Newspaper Map (об'єднує геолокацію і інформаційно-пошукову систему по медіаресурсах, можна вибирати регіон, список он-лайн версій ЗМІ, мову, підтримує російську), WebGround (агрегує російськомовні новини).

Досить цікавими є пошуковики для моніторингу інформаційного простору російського походження.

RCO Zoom (http://www.rco.ru/product.asp?ob_no=7348) - російська пошуково-аналітична система з функціями контент-аналізу в реальному часі і транзакційного сховища документів.

За свідченням розробників, система дозволяє працювати з величезними масивами текстової інформації в реальному часі (обсяг бази - до сотень гігабайт, час пошуку та обробки - секунди).

Засіб відображення - інформаційний портрет - дає можливість отримати ключові слова, формулювати і перевіряти гіпотези, розділяти об'єкти, виділяти статистичні інваріанти в першому наближенні.

При цьому ситуація з інформаційної надмірністю повністю змінюється: чим більше інформації, тим краще.

Вбудовані механізми контент-аналізу забезпечують проведення контент-аналізу в реальному масштабі часу, автоматичне виділення головних тем - тегів документів, визначення тематично і текстуально схожих документів, автоматичну класифікацію та кластеризацію документів.

Морфологічний модуль системи підтримує російську і англійську мови, легко може бути підключена підтримка інших мов, інтегрований з таблицями синонімів, відповідними предметній галузі.

Ще один російський ресурс - "**Медialogия**" (<http://www.mlg.ru>) - розробник автоматичної системи моніторингу та аналізу ЗМІ в режимі реального часу. Вона оперує базою джерел інформації, що дозволяє здійснювати моніторинг більше 24000 ЗМІ та понад 92 млн. соціальних медіа-ресурсів.

Система складається з бази даних ЗМІ та автоматизованого аналітичного модуля, який дозволяє проводити самостійний пошук та аналіз за кількісними та якісними характеристиками за будь-який заданий період. Результати доступні миттєво.

Оперативне надходження контенту 24 години на добу, в тому числі транскрипти федеральних телеканалів доступні через 1,5 години після виходу сюжету в ефір.

Система **PolyAnalyst** (<http://megaputer.ru/polyanalyst.php>, російськомовний продукт американської компанії Megaputer Intelligence, працює в 30 країнах) дозволяє отримати необхідні знання з великої кількості текстових та структурованих даних, передає ці знання в доступній розумінню формі та у вигляді оперативно застосовуваних моделей для прийняття відповідальних рішень, переважно в бізнесовій сфері.

PolyAnalyst дозволяє створювати багатоступінчасті сценарії аналізу даних і форми звітів за допомогою простого drag-and-drop інтерфейсу.

PolyAnalyst надає користувачам результати, отримані з прогнозів, в доступних для розуміння формах, які полегшують прийняття зважених рішень.

Інструментарій PolyAnalyst пропонує клієнтам наступний функціонал: категоризацію, кластеризацію, прогнозування, аналіз зв'язків, знаходження ключових слів і пошук сенсу, виявлення закономірностей, знаходження аномалій.

Загалом в Росії є достатньо інформаційно-пошукових систем, які не лише забезпечують потреби моніторингу інформаційного простору, а й є інструментами конкурентної розвідки в мережі інтернет.

Серед вітчизняних повнофункціональних систем варто відзначити **X-SCIF** [4]. Вона, серед іншого, забезпечує "моніторинг інформації на заданих користувачем веб-сайтах (веб-сторінках) в мережі інтернет (інтранет) за визначеними темами; пошук нових джерел

інформації в мережі інтернет за визначеними користувачем тематиками з наступним їх моніторингом; створення і збереження складних запитів за визначеними темами у вигляді списку-каталогу або рубрики для наступного проведення автоматичного моніторингу, пошуку або контент-аналізу; форматування відібраної інформації за єдиним взарцем та завантаження її до сховища..."

Також система може забезпечувати екстрагування сутностей з відібраної інформації, виявляти приховані зв'язки між об'єктами, аналізувати емоційне забарвлення документів тощо.

Безпосередньо для моніторингу застосовується підсистема X-Stream, яка відпрацьована на базі технології InfoStream компанії ELVisti. X-Stream моніторить в автоматичному режимі понад 7000 джерел. База архіву (ведеться з 1996 року) становить більше 85 млн документів. За територією ця підсистема охоплює весь сегмент україно-, російсько- і англійських видань України, Росії, а також певний перелік ЗМІ зарубіжжя.

InfoStream (<http://infostream.ua>) є одним з лідерів серед контент-моніторингу інтернет-ЗМІ (комплекс контент-моніторингу базується на технології Text Mining). Як повідомляють розробники, система InfoStream забезпечує доступ до оперативної інформації з єдиного інтерфейсу в пошуковому режимі з урахуванням можливого дублювання та семантичної близькості документів, мовних версій, розмірів документів їх цифрової насиченості тощо; доступ до унікального ретроспективного фонду, що перевищує 80 млн. записів; підтримку аналітичної роботи в режимі реального часу: побудова сюжетних ланцюжків, дайджестів, діаграм зустрічальності і таблиць взаємозв'язків понять, медіа-рейтингів.

Infostream надає доступ до понад 6 тисяч джерел з України, країн СНД тощо (з них: інформаційні агентства – 324, теле- та радіо канали – 239, газети – 334, журнали та щотижневики – 368, офіційні джерела – 375, сайти асоціацій, компаній – 714, Інтернет-видання – 2665).

Корпорація **"Медіа-простір"** (<http://media-prostir.kiev.ua>) забезпечує аналіз інформаційного простору України (більше 550 ЗМІ) за допомогою 25 регіональних (обласних) офісів. Вони не лише відслідковують потоки місцевого рівня, а й створюють власноруч електронні варіанти окремих матеріалів тих друкованих видань, які не мають власних інтернет-версій.

В основі моніторингу компанії – відстеження, систематизація, накопичення та обробка матеріалів центральних і регіональних ЗМІ.

За інформацією розробників, систематизація інформаційних повідомлень здійснюється за чотирма прив'язками:

предметні сфери (економіка, політика, соціальна сфера, гуманітарна сфера);
політичні суб'єкти (партії, виборчі блоки і об'єднання, громадські організації);
особистості (посадові особи, народні депутати, лідери політичних партій та громадських організацій);

території (великі міста, області, територіально-регіональні одиниці).

Інформаційні технології корпорації "Медіа-простір" надають можливість в процесі моніторингового аналізу "занурюватись" до третього рівня (наприклад, економіка - Міністерство транспорту - українські залізниці) та здійснювати перехресну прив'язку (наприклад, партія - економіка чи посадова особа - територія - соціальна сфера).

ІА «Контекст Медіа» - компанія, що займається збором, систематизацією й аналізом інформації. В щоденному режимі моніториться майже 14 000 джерел.

Застосовуються методики збору та обробки інформації, а також **частково** технології контекстного аналізу власних розробок.

Одна з послуг - моніторинг соціальних медіа кількістю понад 1100.

В інформаційному агентстві УНІАН для моніторингу (УНІАН-монітор) застосовується система **Web-Observer**, створена спеціалістами компанії "Finport Technologies Inc." і "АналитикСофтЛаб".

«Web-Observer» проводить сканування інформації з визначених джерел, аналізує

оновлення та обробляє їх, виділяючи лише потрібну інформацію. Згодом ця інформація структурується, індексується і зберігається у власній базі даних. Користувач системи може самостійно проводити пошук необхідних документів в накопиченій базі даних, отримувати потрібні документи, власноруч створювати звіти (статичні) та інформаційні дайджести, налаштовувати персональний новинний потік і отримувати повідомлення про знаходження нової інформації.

Основні функції рішення «WebObserver»: оперативна виїмка, розбір, індексація і збереження нових публікацій з визначених джерел; смислове сортування накопичених даних за рубриками; автоматизоване створення звітів і дайджестів з потрібних тем, експорт обраних публікацій, оцінка інформаційної активності компаній або персон на основі статистичних звітів; ведення інформаційного архіву за весь період роботи системи, формування різних досьє; ефективний пошук та обробка інформації в накопиченому архіві.

Моніторинг соціальних медіа. Світовий інформаційний простір не так давно поповнився ще одним чинником - соціальними медіа. З них найбільший інтерес для нас представляють соціальні мережі, зокрема, їх аналіз. Як вважає Д. Ланде, "як і вся інформаційна мережа WWW, так і її окремі фрагменти і навіть сайти несуть значне соціальне навантаження, яке дозволяє порівнювати їх на змістовному рівні з соціальними мережами, утвореними відносинами людей або цитуванням в науці.

...Багато мережевих служб, які дозволяють людям встановлювати зв'язки в Мережі, автоматично формують соціальні мережі. Крім того, сьогодні бурхливо розвинувся спеціальний сервіс з ціленаправленої побудови соціальних мереж у веб-просторі" [5].

Слід зазначити, що соціальні мережі відносяться до так званих "складних мереж". В Україні цей сегмент розвивається останнім часом з надзвичайною швидкістю, формуючи нове комунікаційне середовище рівноправних громадян країни.

Соціальні медіа не мають чіткого визначення. Їх ділять (класифікують) по різному. Наприклад, в одному випадку (Andreas Kaplan і Michael Haenlein) до них відносять спільні проекти, блоги і мікроблоги, контент-спільноти, соціальні мережі, віртуальні ігрові світи, віртуальні соціальні світи. В іншому (О. Додонов, Д. Ланде, В. Путятін) – соціальні мережі, блоги, форуми, сайти відгуків, сервери фото- і відеохостингу, віртуальні служби знайомств і геосоціальні мережі.

Втім, у будь-якому випадку варто визнати, що поява соціальних медіа зруйнувала усталені погляди на комунікаційні процеси у суспільстві, які формувались на протязі попереднього часу, зокрема у 20 столітті.

Точка географічної прив'язки більше не відіграє визначальної ролі в позиціонуванні індивіда або корпорації. Тому комунікація, яка відбувається в мережі, за своїми масштабами набуває планетарного значення.

Останнім часом саме користувачі соціальних медіа стають ньюсмейкерами не тільки окреслених інформаційних полів, але й міждержавних інформаційних просторів. Публікації в соціальних мережах мають значну перевагу перед іншими першоджерелами, перш за все, в оперативності донесення інформації.

Публікації авторитетних авторів отримують резонанс у суспільстві та увагу перших осіб політикуму і державного менеджменту. Інформації, які ніким не цензурюються, але найбільше характеризуються правдивістю персоналізованого викладу та мають в таких випадках можливість показувати найнеприглядніші аспекти подій, фактів, діяльності осіб, призводять до скандалів високого рівня, які можуть мати непередбачувані наслідки.

Відтак моніторинг соціальних медіа стає ще однією необхідністю розвитку цього напрямку.

А. Додонов, характеризуючи архітектуру організації систем моніторингу соціальних медіа, говорить про його три рівні: перманентний, періодичний та стратегічний. Інтерес для нас представляє третій рівень моніторингу, на якому аналізується весь інформаційний простір соціальних медіа в постійному режимі [6].

Серед існуючих систем моніторингу соціальних медіа вчений виділяє канадський

соціальний аналітичний сервіс Radian6, який надає ці послуги в режимі real time, та динамічний польський стартап Brand24.

На пострадянському просторі в сегменті російськомовних соціальних медіа послуги щодо моніторингу пропонують достатньо компаній. Наприклад, YouScan (<https://youscan.ru>), Buzzlook (<http://buzzlook.ru>), IQBuzz (<http://iqbuzz.ru>), SemanticForce (<http://www.semanticforce.net>), BrandSpotter (<http://brandspotter.ru>), Poppler (<http://poppler.ru>), Babkee (<http://www.babkee.ru>) тощо.

В Україні сегмент моніторингу соціальних медіа розвивається значно повільніше. Лідерами українського ринку залишаються російські YouScan, SemanticForce.

Як вже зазначалось, такі послуги надає також ІА «Контекст Медіа».

Висновок. Наведені вище приклади свідчать про активний розвиток пошукових систем загального та спеціального користування, а також про зростання ролі моніторингу соціальних мереж.

Вимоги, які висувались до можливостей пошуковиків в середині минулого століття, стали історією. Сучасні пошукові системи – це мільйони запитів щодня і мільярди щомісяця. Тому прості (загального користування) пошуковики можуть задовольнити лише невибагливі потреби пересічних користувачів.

Але навіть прості користувачі дедалі все більше висувують вимог до рівня параметрів пошукових систем, оскільки вони, наприклад, бажають «бачити» все поле комунікаційно-контентних процесів за відповідним запитом.

Тут на допомогу приходять метапошукові системи. Вони здійснюють пошукові дії одночасно в декількох пошукових системах, враховуючи особливості кожної з них.

Однак для поглибленого вивчення проблематики визначеного характеру використовуються спеціалізовані (сегментні) пошукові системи. Це вже моніторингові сервіси із спеціальним програмним забезпеченням. Вони відносяться до професійних пошуковиків, що здійснюють не просто моніторинг інформаційного простору за визначеними семантичними фільтрами, а вже діють як гібридизовані системи поєднання машинного та людського інтелектуально-прогнозного потенціалу в спеціальних регламентних алгоритмах.

ЛІТЕРАТУРА:

1. Додонов А.Г., Компьютерные сети и аналитические исследования /А.Г.Додонов, Д.В. Ландэ, В.Г. Путьтин. - К.: ИПРИ НАН Украины, 2014., с.162.
2. Количество сайтов в интернете перевалило за миллиард. Режим доступа: <http://techno.bigmir.net/technology/1577735-Kolichestvo-sajtov-v-internete-perevalilo-za-milliard>
3. Додонов А.Г., Компьютерные сети и аналитические исследования /А.Г.Додонов, Д.В. Ландэ, В.Г. Путьтин. - К.: ИПРИ НАН Украины, 2014., с.46-47.
4. А.Г.Додонов, Д.В. Ландэ, В.В. Прищепа, В.Г. Путьтин. Конкурентная разведка в компьютерных сетях. – К.: ИПРИ НАН Украины, 2013. с.54.
5. Ландэ Д.В., Снарский А.А., Безсуднов И.В. Интернетика: Навигация в сложных сетях: модели и алгоритмы. М.: Книжный дом "ЛИБРОКОМ", 2009. - с.166.
6. Додонов А.Г., Компьютерные сети и аналитические исследования /А.Г.Додонов, Д.В. Ландэ, В.Г. Путьтин. - К.: ИПРИ НАН Украины, 2014., с.81.

Рецензент: д.філол.н., проф. Горевалов С.І. завідувач кафедри кіно-, телемистецтва Інституту журналістики КНУ імені Тараса Шевченка

ПОИСКОВЫЕ СИСТЕМЫ В СЕТЯХ ИНТЕРНЕТА

В статье рассматриваются проблемы роста масштабов и динамики информационных потоков в современном информпространстве. Резкое увеличение объемов информационных сообщений и сопутствующих факторов вызывает потребность удовлетворения запросов пользователей через создание и развитие адекватных поисковых систем.

В мире сегмент поисковиков представлен достаточно широкой линейкой поисковых систем как для рядового пользователя, так и для специалистов, углубленно изучают различные аспекты отдельных научных или прикладных сфер по заданной тематике.

Однако и такие поисковые системы не всегда соответствуют необходимым критериям. Поэтому возникает необходимость в создании гибридных систем, где бы происходило сочетание машинного и человеческого потенциала в специальных регламентных алгоритмах.

Ключевые слова: информационно-контентные потоки, поисковые системы, семантические фильтры, мониторинг.

Ph.D. Lubovets G.V.

INFORMATION RETRIEVAL SYSTEMS IN THE WEB OF THE INTERNET

The article scrutinizes the problems of information streams volume and dynamics increase in the modern information environment. A sharp increase in the bulk of informational messages and concomitant factors brings the need to satisfy the users demands by mean of appropriate retrieval systems development.

In the world the retrieval segment is represented by a wide series of retrieval systems which are assigned both for a custom user and for experts who study different aspects of particular scientific or applied areas deeply.

However these retrieval systems don't meet the proper criteria. Therefore this brings the rise of necessity for development of hybrid systems with combined machine and human potential in special algorithm routine.

Keywords: information and content streams, information retrieval systems, semantic filters, monitoring.