

AXBOROT TEXNOLOGIYALARI SOHASIDA BILIMLARNI SHAKLLANTIRISH UCHUN INTERFAOL INTERAKTIV TIZIMNI ISHLAB CHIQISH USULLARI

Maxsudjon Baxramov

O'zbekiston davlat san'at va madaniyat institutining Farg'ona mintaqaviy filiali

ANNOTATSIYA

Hozirgi kunga kelib har bir sohada olib borilayotgan izlanishlar va tadqiqotlar ushbu sohaga tegishli ma'lumotlarning to'liqlik darajasi yuqori bo'lgan manbaini talab qilmoqda. Bunday manbaalar hajmi esa chegaralangan. Maqolada ushbu manbalar miqdorini barcha ko'rsatkichlar bo'yicha oshirish usullari yoritib berilgan.

***Kalit so'zlar:** axborot, intellektual izlanish, internet, tarmoq, ma'lumotlar ombori.*

ABSTRACT

In this day, research and studies in each field require a high level of completeness of information in this field. The size of such resources is limited. The article describes the ways to increase the amount of these resources on all indicators.

***Keywords:** information, intellectual research, internet, network, database.*

АННОТАЦИЯ

Сейчас исследования в каждой области требуют высокого уровня полноты информации в этой области. Размер таких ресурсов ограничен. В статье описаны способы увеличения количества этих ресурсов по всем показателям.

***Ключевые слова:** информация, интеллектуальное исследование, интернет, сеть, база данных.*

KIRISH

Mutaxassislikni chuqurlashtirish jarayoni va ma'lum faoliyat sohalarida yangi bilimlarning paydo bo'lish suratining tezlashishi va zarur bo'lgan ma'lumotlarni to'plashning tezkorligi va qulayligi ulardan foydalanishning samaradorligini baholashda eng muhim ko'rsatkichga aylanishini aniq tasavvurga keltiradi.

Hozirgi kunda har bir kasbda ishlash uchun zarur bo'lgan bilimlar miqdori yuqori darajada o'sib bormoqda, bilimlarga o'z vaqtida ma'lumotlar topish va yetkazish uchun esa qidiruv tizimidan tashqari to'g'ri axborotni topish uchun deyarli hech qanday vosita yo'q.

Axborotni izlash (angl. information retrieval) — axborotni qidirish jarayoni bo'lib, axborotga bo'lgan ehtiyojni qondirish uchun strukturalashmagan hujjatlardan zarur ma'lumotlarni izlash haqidagi fan sifatida rivojlanib kelmoqda.

MUHOKAMA VA NATIJALAR

Axborot izlash – bu ma'lum bir hujjatlar (matnli) to'plamidan oldindan belgilangan shartli mavzu (so'rov) yoki zarur (axborotga bo'lgan ehtiyojni qondirishga tegishli) ma'lumotlarni, faktlarni, xabarlarini izlash – aniqlash jarayoni. Qidiruv jarayoni ma'lumotlarni to'plash, ularga ishlov berish va taqdim etishga qaratilgan operatsiyalar ketma-ketligini o'z ichiga oladi.

Axborotni qidirish masalalari

• Axborotni qidirishning asosiy masalasi – foyalanuvchiga uning axborotga bo'lgan ehtiyojlarini qondirishga yordam berishdan iborat. Asosiy masalalar:

- Modellashtirish masalasi;
- Hujjatlarni klassifikatsiyalash;
- Hujjatlarni filtrlash;
- Hujjatlarni klasterizatsiyasi;
- Qidiruv tizimlari arxitekturasi va foydalanuvchi interfesini loyihalash;
- Axborotlarni ajratib olish, xususiyl holda hujjat annotatsiyasi va referatini tayyorlash;
- So'rov tillari va boshqalar.

Umumiy holda axborotni qidirish to'rtta bosqichdan tashkil topgan:

- axborotga bo'lgan ehtiyojni aniqlash va axborot so'rovini shakllantrish;
- mumkin bo'lgan axborotlar massivining egasini (manbasini) aniqlash;
- aniqlangan axborot massividan ma'lumotlarni ajratib olish;
- olingan axborot bilan tanishish va qidiruv natijasini baholash.

Axborotni ajratib olish. Axborotni ajratib olish (angl. information extraction) — bu komp'yuterda tayyorlangan aniq strukturaga ega bo'lmagan yoki kuchsiz strukturalashgan hujjatlardan aniq strukturaga ega bo'lgan ma'lumotlarni avtomatik ajratib olish yoki qurish. Axborotni ajratib olish tabiiy tildagi matnlarni qayta ishlab bilan bog'liq bo'lib, axborotlarni qidirishning bir ko'rinishi hisoblanadi.

Axborotlarni ajratib olishning maxsus masalalari:

- Nomiqlangan (ya'ni, muhim, tayanch) elementlarni tanish, masalan: odamlar ismi, tashkilot nomlari, geografik nomlar, vaqt va pul birliklari va h.k.

- Anaformalar va o'zaklarning aniqlanishi: bitta va aynan shu ob'etga tegishli (daxldor) bog'liqliklarni aniqlash. Bunday bog'lanishlarning xos holatlari – olmoshli anaforalar.

- Atamalarni (terminlarni) ajaratish: berilgan matn uchun kalit so'zlari va iboralarni topish.

- Avto (referat tuzish): matndan semantik, hissiy, baholovchi ma'lumotlarni ajratib olish. Bu generativ va deklarativ bo'ladi.

Samaradorlikni baholash. AQT tomonidan topilgan hujjatlar so'rovga qanchalik mos kelishini baholashning ko'plab usullari mavjud. Afsuski, so'rovlarning muvofiqlik darajasi, boshqacha aytganda relevantnosti, sub'ektiv tushuncha hisoblanadi, muvofiqlik darajasi esa so'rov bajarilishi natijalarini baholovchi aniq insonga bog'liq.

Umumjahon internetni universal axborot manbaiga aylantirish muqarrar ravishda kontent bilan ishlash uchun yangi texnologiyalarning paydo bo'lishiga olib keladi.

So'nggi yillarda axborot izlash va kontentni boshqarish texnologiyalari jadal sur'atlar bilan rivojlanib, intellektual xususiyatga ega bo'ldi. Masalan, foydalanuvchi didiga mos keladigan Google (www.google.com) istalgan turdagi ma'lumotlarni qidirishga mo'ljallangan maxsus tizim. Dastlab qidirilgan so'rovlarga asosan, murojaatlarda imtiyozlarni belgilaydi va eng to'g'ri takliflarni taqdim etadi. Qidiruv samaradorligini oshirishning yana bir yondashuvi maxsus veb-saytlarda hujjatlarni izlashga mo'ljallangan maxsus qidiruv tizimlarida amalga oshiriladi. Misol tariqasida, siz ilmiy ma'lumotlar (www.ziyonet.uz) bo'yicha ishlaydigan Ziyonet tizimini tanlashingiz mumkin. Murakkab intellektual izlanishlar ko'plab muammolar uchun samarali yechimga aylanishi mumkin. Internetning yangi avlodini yaratish g'oyalari 1999 yilda Tim Berns-Li tomonidan ilgari surildi. Ulardan biri - "ontologiyalar" deb ataladigan hujjatlarni taqdim qilish uchun XML-ga asoslangan formatlash vositalaridan foydalanish. Ular "aqlli agentlar" deb ataladigan ilovalar tomonidan o'qishga odatlangan kontentdan tarkib topgan bo'lishi kerak. Ko'plab sanoat standartlari va til vositalari yillar mobaynida ishlab chiqilgan bo'lsa-da, semantik veb-dasturni amalga oshirish hali juda uzoq vaqt talab qiladi.

Axborot-qidiruv tizimining rivojlanish istiqboliga bag'ishlangan barcha ishlarga qaratilgan umumiy xulosa shundaki, ushbu yo'nalishning ochiq-oshkora axborot uchun veb-makon va nazariy tadqiqotlardagi ayrim yutuqlari uchun ustuvor ahamiyatga ega bo'lishiga qaramasdan, biz radikallikni keltiradigan amaliy yechimlar paydo bo'lishini kutishimiz kerak, tezkor va samarali intellektual ma'lumot olish

sohasida uzilishlar bunga arziydi. Shu bilan birga, axborot to'plash mexanizmlari to'plangan ma'lumotlarning to'planishiga asosan, axborot almashish muammolari, axborotni izlab topish va aksincha, jamiyat faoliyatining muayyan sohalarida optimallashtirish qarorlarni qabul qilish muammosi tobora keskinlashib bormoqda. AQShning "Axborotni aniqlab bo'lmasligi uchun to'lov" hisobotida quyidagi ma'lumotlar mavjud.

O'rta biznes korxonalarida axborot bilan ishlashning noqulayligi tufayli vaqt yo'qotilishi oqibatida yetkazilgan bevosita zararlari ma'lumotlar bilan ishlayotganlar uchun 2,5-3,5 ming dollarni tashkil qiladi va bu yo'qotishning yagona manbai emas. Qidiruv mexanizmlari bilan ishlagan har bir kishi uchun ma'lum bir limit mavjud bo'lib, u odamlarga axborot izlashni to'xtatadi va ular yetarlicha ma'lumotga ega bo'ldim degan xulosaga keladi. Xabardorlik darajasining kamligi noto'g'ri qaror qabul qilishga, noto'g'ri baho berishga yoki xatarlarni noto'g'ri baholashga olib keladi. Noto'g'ri asosda qabul qilingan qarorlar qabul qilinishidan kelib chiqadigan bevosita yo'qotishlar miqdori aniqlanmaydi, ammo ular sezilarli darajada buyukroq buyruqlar va qarorlar chiqarishdan o'zlari ongli ravishda bexabar xolda o'zlarini to'xtatadilar.

Bunga shuni qo'shib qo'yish kerakki, qidiruv natijalari tez-tez zerikarli 404 "Eslatib bo'lmaydigan" imzo qo'yilgan xabarlar bilan "Sahifa ko'rsatilmaydi" yoki "Siz izlayotgan sahifa hozircha mavjud emas." kabi qidiruv natijalarini taqdim etadi. Bu esa izlanuvchini ma'lum bir miqdorda o'z maqsadidan chekinishiga olib keladi. Xozirgi davr mobaynida internet tarmog'ida foydalanilmayotgan ma'lumotlar xajmi o'sdi (mos ravishda 3,8%, 10% va 15%). Shuningdek, "axborotning internetda yo'qolib qolishi" kabi xolatlar xam uchrab turadi.

Ilmiy jamoatchilikning aksariyat a'zolari axborot texnologiyalari vositalariga tobora ortib borayotgan qaramlik haqida juda keng o'ylashmoqda va bu ularni tashvishga solmoqda. Internet sahifasining o'rtacha umr davomiyligi bugungi kunda 100 kundan oshmaydi - bunday manbalarda ilm-fan va madaniyatga tayanib bo'lmasligi aniq.

XULOSA

Olib borilgan izlanishlardan so'ng, qaror qabul qilish uchun zarur bo'lgan axborotni qidirishni amalga oshiradigan loyihalar va tadqiqotlarning eng yuqori sur'atini hisobga olib, quyidagi qiziqarli fikrga kelindi: "Axborot bazasidan axborot olishning tezligi ularning samaradorligini baholashda eng muhim ko'rsatkich bo'ladi. Prinsipial jihatdan ma'lumotlarning to'liqligi yoki olingan ma'lumotlarning 100%

ishonchliligidan ham muhimroq". Shu sababli, tarmoqning katta xarajatlariga ko'ra, o'z-o'zidan shakllanadigan axborot makonida, portal texnologiyalari, maxsus bilimlar omborlari va korporativ ma'lumot portallari (korporativ ma'lumot portallari) ga asoslangan korporativ bilimlar va axborot xizmatlarining tashkil etilgan tarmoqni Bill Enmon, "axborot fabrikalari" deb ataydi. Biz esa ushbu "axborot fabrikalari" ning ma'lumotlar aniqliliga va kerakligi jihatdan samaradorligini oshirish zarur deb xisoblaymiz. Buning uchun quyidagi asosiy yo'nalishlarni rivojlantirish zarur:

- axborot izlash usullarini takomillashtirish;
- tizimlashtirish va yanada aniq ifodalashda korporativ axborotni kataloglash;
- korporativ ma'lumotlar omborlarini yaratish (ma'lumotlar ombori);
- bilim almashish mexanizmlarini ishlab chiqish va kelishilgan qarorlar qabul qilish.

REFERENCES

1. Арипов М.М. Метод эталонных уравнений для решений нелинейных краевых задач. – Ташкент: ФАН, 1988. 137б.
2. Дьяконов В.П., Абраменкова И.В Mathcad 7.0 в математике, физике и в Internet . Нолидж 1998 г.