

AZƏRBAYCAN RESPUBLİKASI
DAXİLİ İŞLƏR NAZİRLİYİ

POLİS AKADEMİYASI

"DİO-nun İNZİBATI FƏALİYYƏTİ " KAFEDRASI

"Müasir informasiya texnologiyaları" fənni üzrə

M Ü H A Z İ R Ə

(əyani təhsilalanlar üçün)

Mövzu № 5: "İnformasiya axtarışının müasir texnologiyaları".

Vaxt – 4 saat

Mühazirə - 2 saat

Seminar – 2 saat

AZƏRBAYCAN RESPUBLİKASI
DAXİLİ İŞLƏR NAZİRLİYİ

POLİS AKADEMİYASI

"DİO-nun İNZİBATI FƏALİYYƏTİ " KAFEDRASI

"Müasir informasiya texnologiyaları" fənni üzrə

M Ü H A Z İ R Ə

(əyani təhsilalanlar üçün)

Mövzu № 5: "İnformasiya axtarışının müasir texnologiyaları".

Vaxt – 4 saat

Mühazirə - 2 saat

Seminar – 2 saat

Tərtib etdi: "DİO-nun inzibati fəaliyyəti" kafedrası baş müəllimi,
polis polkovnik-leytenantı Heydər Heydərov

Mühazirə kafedranın iclasında müzakirə olunmuş və təsdiq edilmişdir.

Protokol № 09 " 30 " may 2019-cu il.

Dəyişiklik və əlavə edilib: " 20 " aprel 2020-ci il.

Bakı - 2020

Mövzu № 5: “İnformasiya axtarışının müasir texnologiyaları”.

PLAN:

1. İnformasiya axtarış texnologiyalarının əsasları.
2. İnternet – informasiya axtarış məkanı kimi.
3. İnternetdə informasiya axtarış sistemləri.

Ə D Ə B İ Y Y A T :

1. “İnformasiya, informasiyalaşdırma və informasiyanın mühafizəsi haqqında” Azərbaycan Respublikasının 3 aprel 1998-ci il tarixli Qanunu.
2. “Telekommunikasiya haqqında” Azərbaycan Respublikasının 14 iyun 2005-ci il tarixli Qanunu.
3. Balayev R.Ə., Əlizadə M.N., Musayev İ.K. İntellektual sistemlər və texnologiyalar. Bakı, “MSV NƏŞR” nəşriyyatı, 2016, Dərs vəsaiti, səh. 166-170.
4. İsrailova E.N., Həsənova N.Ə. “İnformasiya texnologiyaları və sistemləri”. Dərs vəsaiti. Bakı, Mütərcim nəşriyyatı, 2015. 312 s.
5. Əliquliyev R.M., Mahmudov R.Ş. İnternetin tənzimlənməsi problemləri: ekspress-informasiya / AMEA İnformasiya Texnologiyaları İnstitutu. – Bakı, İnformasiya Texnologiyaları nəşriyyatı, 2010. - 115 s.
6. Gündüz O.M., Əfəndiyev Q., Rüstəmov N. Kompüter: İnformasiya texnologiyalarının əsasları / - Bakı, Multimedia Mərkəzi, 2006. - 304 s.
7. Kərimov S.Q.İnformasiya sistemləri. Bakı, Elm nəşriyyatı, 2008 səh. 208-226
8. Kərimov S.Q., Həbibullayev S.B., İbrahimzadə T.İ. İnformatika. Dərslik. Bakı, 2011. Səh 409- 421.
9. Xəlilov M.S. İnformatika. Dərslik. Bakı, OYU Nəşriyyatı, 2009 Səh. 237-246
10. İnformasiya texnologiyaları KİV-in insanların ictimai-siyasi və mədəni həyata təsirini əhəmiyyətli şəkildə artırır // http://www.ict.az/az/index.php?option=com_content&task=view&id=2541&Itemid=122

GİRİŞ

Bəşəriyyət tarixində hər bir dövrün informasiya texnologiyaları və resursları mövcud olmuşdur. Zaman keçdikcə insanları əhatə edən informasiya ehtiyatlarının həcmi artdıqca, onların arasından lazım olan məlumatların axtarılıb tapılmasını çətin bir vəziyyətə saldı. İnformasiya axtarışı insanları ciddi düşündürən problemlərdən birinə çevrildi. Ona görə də informasiya axtarışı üçün daha təkmilləşdirilmiş və mükəmməl üsul və vasitələrin işlənilib hazırlanması istiqamətində daim ciddi işlər aparılırdı. Bu problemin aradan qaldırılması üçün informasiyaların müəyyən əlamətlər üzrə sistemləşdirilməsi, kartoteka sistemlərinin istifadə olunması geniş tətbiqini tapdı.

Lakin informasiya ehtiyatlarının həcmnin kəskin sürətlə artması, onların xarakterinin və formasının əhəmiyyətli dərəcədə dəyişməsi, informasiyanın saxlanması və təqdim olunması üçün yeni texnologiyaların meydana gəlməsi səbəbindən, mövcud axtarış üsul və vasitələri insanların informasiya tələbatlarını ödəməkdə aciz qalır.

Kompüter sistemlərinin meydana gəlməsi və daha sonralar İnternet qlobal şəbəkəsinin təşəkkül tapması informasiya massivlərinin həcmi maksimum dərəcədə artırdı. İnternet qlobal şəbəkəsi sistemində texnoloji yeniliklər öz tətbiqini tapdı. Qeyd olunan texnoloji yeniliklər arasında İnternet axtarış səhifələrini göstərmək mümkündür. Bu texnologiyaların müəyyən əlamətlər üzrə axtarışın tətbiqini həyata keçirməyə imkan verir.

İnternet şəbəkəsinin geniş yayılması, onun informasiya xidmətlərinin sürətli inkişafı virtual informasiya fəzasının, o cümlədən İnternet saytların və informasiya sistemlərinin, virtual informasiya saxlanclarının, virtual kitabxanaların və s. informasiya ehtiyatlarının yaranmasına gətirib çıxarmışdır.

Hələ ötən əsrin 80-ci illərində, internet şəbəkəsinin FTP, Gopher, E-mail, Telnet kimi xidmətlərinin bazasında kifayət qədər böyük həcmdə informasiya ehtiyatlarının yarandığı, toplandığı, emal olunduğu və ötürüldüyü vaxtlarda axtarış vasitələrinin işlənilib hazırlanmasına və tətbiqinə böyük tələbat yaranmışdı. İnternetdə ilk informasiya axtarışı vasitələrinin çox əvvəllər yaranmasına baxmayaraq, onların həqiqi mənada vacibliyi yalnız web-texnologiyanın, yeni dünya hörümçək torunun meydana gəlməsindən sonra daha ciddi hiss olunmağa başladı.

Hazırda İnternetdə, onun nəhəng web-fəzasında saytların sayı və informasiyanın həcmi elə həddə çatmışdır ki, keçilməz kiber cəngəllikləri xatırladan İnternetin nəhəng informasiya fəzasında lazım olan məlumatın tapılması ciddi problemə çevrilmişdir. İnternetdə sənəd axınlarının gücünün astronomik sürətlə artması və yuxarıda göstərilən digər problemlər intellektual informasiya-axtarış vasitələrinin işlənilib hazırlanmasını və tətbiqini tələb edir.

Əlbəttə, bu problemin həllində web-texnologiya bazasında qurulmuş informasiya-axtarış sistemləri xüsusi rol oynayır.

Daxili işlər orqanlarının informasiya resursları böyük həcmi verilənlər bazasından təşkil olunmuşdur. Bu məlumatlar içərisində statistik uçot sistemlərini, daktiloskopik sistemləri və digər faktoqrafik informasiya sistemlərini misal göstərmək olar. Qeyd olunan informasiya sistemləri sırasından tələb olunan məlumatları tapmaq üçün kompüter texnologiyalarında istifadə olunan ənənəvi metodlarla yanaşı, identifikasiya proqramlarından da geniş istifadə olunur.

Sual 1. İnformasiya axtarış texnologiyalarının əsasları.

Məlum olduğu kimi faktorqafik informasiya sistemləri əsasən strukturlaşdırılmış verilənlərin saxlanması və emalı üçün tətbiq edilir. Həmin verilənlər əksər halda informasiya obyektlərinin xassələrini təsvir edən ədədi, mətni və ya məntiqi qiymətlərdən ibarət olurlar. Praktikada bir çox hallarda isə informasiya strukturlaşdırılmış verilənlər massivləri şəklində deyil, təbii dildə ifadə olunmuş müxtəlif sənədlər şəklində (monoqrafiyalar, dərsliklər, məqalələr, dövri nəşrlər, normativ və hüquqi sənədlər, arxiv sənədləri və s.) ifadə olunur. Sənədli və ya tam-mətnli informasiya bazaları ilə işləmək üçün informasiya sistemlərinin digər sinfindən sənədli informasiya sistemlərindən istifadə olunur.

Sənədli və ya sənəd-yönlü informasiya sistemi sənədlərin saxlanması, axtarışı və emalını avtomatlaşdırmaq üçün nəzərdə tutulan və texniki, proqram, linqvistik və təşkilati vasitələrdən ibarət olan mürəkkəb kompleksdir. Sənədli informasiya sistemlərinin tipik nümayəndəsi informasiya-axtarış sistemləridir.

İnformasiya axtarış sistemi (İAS) təbii dildə sənədlərin toplanmasını, saxlanmasını və müxtəlif kriterilərə görə axtarışını təmin edir. Bu cür sistemlərdən həm müəssisə (korporasiya) səviyyəsində, həm də İnternet şəbəkəsində müxtəlif tip sənədlərin toplanması, sistemləşdirilməsi və axtarışı üçün istifadə edilir.

Sənədli informasiya sistemləri, o cümlədən informasiya-axtarış sistemləri tarixən faktoqrafik sistemlərə nəzərən daha erkən yaranmışlar. Bunun səbəbini dəqiq demək çətindir. Çox güman ki, burada strukturlaşdırılmış verilənlərin təsviri üçün modellərin nisbətən gec formalaşdırılması müəyyən rol oynamışdır. İnformasiya axtarış sistemi girişdə mətn formasında verilən və semantik(məna) strukturu daha mürəkkəb olan informasiya (sənədlər) ilə işləyir. Odur ki, İAS-ın verilənlər bazasına informasiya və ya sənədlər bazası demək daha düzgün olardı.

İstənilən İAS-ın əsas funksiyası sorğulara verilən cavablar əsasında istifadəçiləri informasiya ilə təmin etməkdən ibarətdir. Tələb olunan informasiyanın tapılıb istifadəçilərə çatdırılması sistemin yerinə yetirdiyi və informasiya axtarışı adlanan əsas əməliyyat vasitəsilə həyata keçirilir.

İnformasiya axtarışı istifadəçinin sorğusuna uyğun cavabları özündə əks etdirən sənədlərin və ya həmin sənədlərin göstəricilərinin tapılması prosedurasından ibarətdir. Faktoqrafik informasiya sistemlərində istifadəçinin sorğusuna cavab kimi konkret faktlar (verilənlər) təqdim edilir, sənədli sistemlərdə isə informasiya axtarışı nəticəsində istifadəçiyə onun sorğusuna uyğun sənədlər verilir. Bəzən istifadəçiyə sənədlərin özü yox, onların göstəriciləri (saxlandıqları yer və ya şəbəkə ünvanı (URL)) təqdim edilir.

Sistemdə informasiya axtarışı istifadəçinin informasiyaya olan tələblərini ödəmək üçün verdiyi sorğu əsasında aparılır. İnsanın praktik fəaliyyəti prosesində müəyyən informasiyaya olan ehtiyacı informasiya tələbatı adlanır. Alınan informasiyanın işlənməsi nəticəsində insanların informasiya tələbatı formaca və məzmunca daima dəyişir. Odur ki, informasiya tələbatını birmənalı ifadə və təsvir etmək mümkün deyil. Lakin informasiya tələbatı xüsusi hallar üçün və müəyyən vaxt anları ərzində ifadə oluna bilər. Müəyyən vaxt müddəti ərzində bu cür xüsusi hallar üçün təbii dildə ifadə olunmuş informasiya tələbatı istifadəçi tərəfindən sisteme verilən informasiya sorğusu vasitəsilə ödənilir. İnformasiya sorğusu düzgün tərtib edilməlidir, əks halda o istifadəçinin həqiqi informasiya tələbatını dəqiq ifadə etməyə bilər. İnformasiya axtarışı zamanı istifadəçinin həqiqi informasiya tələbatına deyilir, yalnız informasiya sorğusuna baxılır. Odur ki, sistemin göstəricilərinə təkcə informasiya tələbatına görə yox, həm də informasiya sorğusuna nəzərən baxılır.

Bu münasibətlərin ifadə edilməsi üçün informasiya-axtarış sistemlərinin nəzəriyyəsinə iki fundamental anlayış daxil edilmişdir: pertinentlik və relevantlıq. Pertinentlik dedikdə sənədin məzmununun istifadəçinin informasiya tələbatına uyğunluğu başa düşülür. İnformasiya tələbatını ödəyən sənədlərə pertinent sənədlər deyilir. Relevantlıq sənədin məzmununun informasiya sorğusuna uyğunluğunun göstəricisidir. Sorğuya cavab verən sənədlərə relevant sənədlər deyilir.

İnformasiya axtarışı prosesinin kompüter vasitəsilə avtomatlaşdırılması sənədlərin və informasiya sorğusunun əsas məzmununun formal təsvirini tələb edir. Sənədin məzmununun formal təsviri sənədin axtarış surəti (SAS), sorğunun məzmununun formal təsviri isə sorğu təlimatı (ST) və ya sorğunun axtarış surəti (SOAS) adlanır. Sənədlərin və sorğuların axtarış sürətlərinin tərtib edilməsi üçün xüsusi dillərdən istifadə edilir. Həmin dillərə informasiya-axtarış və ya sadəcə informasiya dilləri deyilir.

İnformasiya axtarışı prosesində sənədin axtarış sürətini sorğunun axtarış sürəti ilə müqayisə etməklə sənədin və sorğunun məzmunlarının uyğunluq dərəcəsi müəyyənləşdirilir. Bu cür müqayisə əsasında sənədin sorğuya relevant olub-olmaması təyin edilir.

Baxılan sənədin sorğuya cavab verib-verməməsi haqqında qərar SAS və SOAS arasında mənə oxşarlığının dərəcəsini təyin edən müəyyən qaydalar əsasında qəbul edilir. Bu cür qaydalar dəstinə mənə oxşarlığının kriterisi (MOK) deyilir. MOK aşkar və ya qeyri-aşkar verilə bilər.

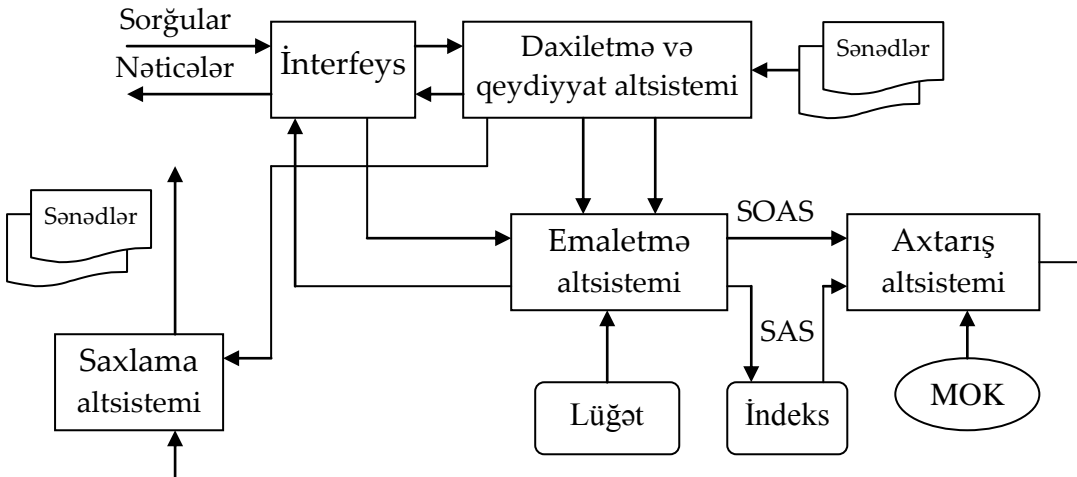
Əslində mənə oxşarlığı kriterisi relevantlıq anlayışına deyil, formal relevantlıq anlayışına, yəni SAS ilə SOAS arasında formal uyğunluğa əsaslanır. Sənədin məzmununun informasiya sorğusuna mənaca uyğunluğu, yəni faktiki relevantlıq yalnız insan tərəfindən sənədin və sorğunun araşdırılmasından sonra təyin edilə bilər.

İnformasiya-axtarış sisteminin ümumi funksional strukturu

Tipik İAS əsasən beş altsistemdən ibarət olur:

- interfeys;
- daxiletmə və qeydiyyat altsistemi;
- emalətmə altsistemi;
- saxlama altsistemi;
- axtarış altsistemi.

İAS-ın ümumi funksional strukturu aşağıda göstərilmişdir



Şəkil. İAS-ın ümumi funksional strukturu

İstifadəçilər sistemlə interfeys vasitəsilə əlaqə saxlayırlar. Müasir İAS-da istifadəçilərin sistemlə rahat işləməsi üçün müxtəlif menyulardan, formalardan, düymələrdən və s. ibarət olan qrafik interfeysdən istifadə edilir.

Sistemin girişinə daxil olan sənədlər həm kağız formada, həm də müxtəlif formatlı elektron variantda ola bilər. Odur ki, daxilətmə və qeydiyyat altsistemi aşağıdakı əsas məsələləri həll etməlidir:

- kağız sənədlərin elektron surətlərinin hazırlanması. Bu, ya kağız sənədlərin klaviaturadan daxil edilməsi və ya onların skaner qurğusu vasitəsilə oxunub, xüsusi proqramın köməyi ilə tanınması ilə əldə edilir;
- elektron sənədlərinin çatdırılması kanallarına qoşulmanın təmin edilməsi(məsələn, internet şəbəkəsinə);
- elektron sənədlərinin qavranılması və lazım gəldikdə formatlarının çevrilməsi;
- elektron sənədlərinin qeydiyyatı. Bu, elektron sənədlərinə unikal identifikatorun mənsub edilməsi və adların sinxronlaşdırılması cədvəli vasitəsilə həyata keçirilir. Lazım gəldikdə köhnə adlar saxlanır.

Sistemə daxil olan sənədlərin hamısı heç bir dəyişiklik edilmədən sənədlər bazasında saxlanmaq üçün saxlama altsisteminə göndərilir. Sənədlər bazası adi halda HD disk yaddaşının kataloqlarında yerləşdirilmiş fayllar toplusundan ibarət ola bilər. Lakin sənədlər bazasının bu cür təşkili HD fəzasından səmərəli istifadəyə imkan vermədiyindən və faylların sayı çox olduqda sənədlərə müraciətə çox vaxt sərf olunduğundan, yaxşı hesab olunmur. Odur ki, sənədlərin saxlanması üçün informasiyanın sıxılmasından və sürətli axtarış vasitələrindən istifadə edilir. Bu halda saxlama altsistemində identifikator vasitəsi ilə sənədlərə müraciəti təmin edən standart və ya xüsusiləşdirilmiş arxivləşdirmə vasitələrindən, VBİS-dən və s. istifadə edilir. Sonra sənədlər emalətmə altsisteminə verilir. Bu altsistemin vəzifəsi hər bir sənəd və ya sorğu üçün axtarış sürətini tərtib etməkdən ibarətdir. Sənədin axtarış sürəti sənədin əsas mövzusunun ifadə edən əlamətlərdən (açar sözlər və ya deskriptorlardan) ibarət olur və indeksdə saxlanır.

Məntiqi baxımdan indeks sətirləri sənədlərə və sütunları əla-mətlərə uyğun təşkil edilmiş cədvəl kimi formalaşdırıla bilər. Belə cədvəlin xanalarında baxılan əlamətin baxılan sənəddə olub-olmamasından asılı olaraq 1 və ya 0 saxlanır. Aydın ki, bu cür cədvəl həddən artıq boşalmış vəziyyətdə olacaq. Sıfır qiymətlərinin hamısının saxlanması mənasızdır. Odur ki, praktikada cədvəlin sətirlərə və ya sütunlara görə bükülməsindən istifadə edilir. Sənədlərin bu cür təşkil formaları uyğun olaraq birbaşa (düz) və ya invers adlandırılır. Cədvəlin bükülməsi zamanı indeks mürəkkəbləşdiyindən, onun reallaşdırılması üçün VBİS vasitələrindən və ya xüsusi vasitələrdən istifadə oluna bilər.

Sistemin interfeysi vasitəsilə qəbul olunan informasiya sorğusu emalətmə altsistemi tərəfindən emal olunur və formalaşdırılmış SOAS axtarış altsisteminə ötürülür. Axtarış altsisteminin vəzifəsi qəbul edilmiş MOK-a əsasən indeksdə baxılan SOAS-a uyğun gələn sənədlərin axtarış sürətini tapmaqdan ibarətdir. Relevant sənədlərin identifikatorları axtarış altsistemindən saxlama altsisteminə və interfeysə göndərilir. İnterfeys istifadəçiyə onun sorğusuna cavab verən sənədlərin identifikatorları və göstəriciləri haqqında məlumat verir, saxlama altsistemi isə həmin identifikatorlara görə relevant sənədlərin özlərini istifadəçiyə çatdırır.

Sənədlərin məzmununun formal təsviri

Məlum olduğu kimi, təbii dil (TD) insanlar arasında informasiya mübadiləsi üçün universal işarə sistemidir. İnformasiya axtarış sistemində daxil edilən sənədlər təbii dildə ifadə olunduğundan, İAS-ın bütün fəaliyyət dövrü ərzində informasiyanın təsviri üçün TD-dən istifadə edilməsi daha əlverişli olardı. Əgər sorğu ilə sənəd arasındakı uyğunluğu insan təyin etsəydi, informasiyanın təsviri üçün TD-dən istifadə edilməsində heç bir problem olmazdı. Lakin müasir İAS-da bu əməliyyatı kompüter yerinə yetirdiyindən, informasiyanın əsas təsvir vasitəsi kimi TD-dən istifadə edilməsi ciddi problemlər yaradır. Bu problemlər kompüter texnologiyalarının tətbiqi ilə informasiya emalı baxımından təbii dilə xas olan aşağıdakı çatışmazlıqlardan irəli gəlir :

Mənanın ifadə vasitələrinin çoxluğu. Məlumatın mənasının ifadə olunmasının əsas vasitəsi təbii dilin leksikası hesab olunur. Lakin müxtəlif hallarda bu funksiyarı aşağıdakı elementlər də yerinə yetirə bilər:

- kontekst;
- sözlər arasındakı paradigmatik əlaqələr;
- sözlər arasındakı mətnə uyğun əlaqələr;
- məlumatın mətnində əvvəl işlədilmiş sözlərə, söz birləşmələrinə,
- cümlələrə və s. istinadlar.

Müxtəlif mənalılıq. Təbii dildə yazılmış məlumat semantika baxımından müxtəlif mənaya malik ola bilər. Müxtəlif mənalılıq təbii dildəki sözlərin sinonimiyası və çox mənalılığı səbəbindən yaranır.

Sinonimiya eyni anlayışı ifadə edən eyni və ya oxşar mənalı sözlərin varlığından yaranır. Həmin sözlər bir-birindən ya mənalıdakı, ya üslublarıdakı çalarlarla və ya hər iki amillə fərqlənir. Təbii dildə bəzi sözlər, həm də söz birləşmələri sinonimlər ola bilərlər. Sinonimiyaya aid misal: “kompüter”, “hesablama maşını”, “elektron hesablama maşını”.

Çoxmənalılıq təbii dildə ayrı-ayrı sözlərin birmənalı olmamasını göstərir. Çoxmənalılıq iki cür özünü göstərə bilər: polisemiya və omonimiya.

Polisemiya ümumi xassələrə və ya əlamətlərə malik olan müxtəlif əşyaların adlarının uyğunluğu deməkdir. Tipik ümumi xassələrə əşyaların oxşarlığı, bir-birinə çox yaxınlığı (fəza, vaxt və s. baxımından) və həmçinin eyni funksional vəzifələri aid edilə bilər. Məsələn, “avtomat” - avtomatik silah, “avtomat” -öz-özünə işləyən aparat (maşın).

Omonimiya öz aralarında ümumi xassələri olmayan müxtəlif əşyaların adlarının eyni olmasıdır. Məsələn. “qaz” - kimyada maddənin bir halı, qaz - quş növü, qaz - qazmaq felinin əmr forması.

Omonimiyaya omoqraf adlanan sözləri də daxil edirlər.

Omoqraf müxtəlif əşyaları ifadə edən, yazılışı eyni olan, lakin müxtəlif cür tələffüz olunan sözlərə deyilir. Məsələn, “alma” -meyvə növü, “alma” - almaq felinin inkar əmr formasıdır. İAS təbii dildə yazılı formada ifadə edilən məlumatla əməliyyat apardığından, dilin fonetikası əsas rol oynamır. Odur ki, omoqraflara da omonim sözlər kimi baxılır.

Ellipstiklik. Təbii dildə ifadə olunmuş bir çox mətnlərdə ellipslərə, yəni nəzərdə tutulan sözlərin buraxılmasına rast gəlinir. Ellipstiklik məlumatla bilavasitə işləyən insan üçün mənfi rol oynayır. Məlumatın kompüterdə emalı halında işə ellipstiklik daha xoşagəlməz sayılır.

İnformasiya-axtarış dilləri.

İnformasiyanın təsviri üçün təbii dildən istifadə edilə bilməməsi süni dillərin tətbiqinə ehtiyac yaratmışdır. Bu cür dillərə informasiya-axtarış və ya informasiya dilləri deyilir. İnformasiya-axtarış dili (İAD) – sistemə daxil edilən və axtarış üçün nəzərdə tutulan məlumatın əsas məzmununu təsvir etmək üçün xüsusişdirilmiş süni dildir. İAD təbii dil əsasında yaradılır, lakin ondan yığcamlığı, dəqiq qrammatik qaydaların mövcudluğu və semantik çox mənalığın olmaması ilə fərqlənir.

İnformasiya-axtarış dillərini iki tipə: təsnifat və deskriptor dillərinə ayırmaq olar. Bu tip dillər arasında prinsiplial fərq dilin cümlələrinin qurulma prosedurundadır.

Təsnifat dillərinin köməyi ilə məlumatlar siniflərə bölünür. Siniflər leksik vahidlərlə ifadə olunur. Bu tip dillərin leksik tərkibinə sadə anlayışları ifadə edən sözlərlə yanaşı, mürəkkəb anlayışları ifadə edən söz birləşmələri və ifadələr də əvvəlcədən daxil edilir. Bu cür İAD-da məlumatın məna məzmununu ifadə etmək üçün yalnız dilin leksik vahidlərindən, o cümlədən, mürəkkəb anlayışlardan istifadə edilir. Mürəkkəb sintaktik konstruksiyanın qurulması faktiki olaraq hazır leksik vahidlər dəstindən uyğun anlayışın (sözün, sözbirləşməsinin və ya ifadənin) seçilməsi ilə əldə edilir. Məsələn, təsnifat dili vasitəsilə polis sinfini belə təsvir etmək olar:

Polis. Daxili işlər orqanları.

Polis. Daxili işlər orqanları, İctimai asayiş.

Polis. Daxili işlər orqanları, Cinayətkarlıqla mübarizə.

Polis.

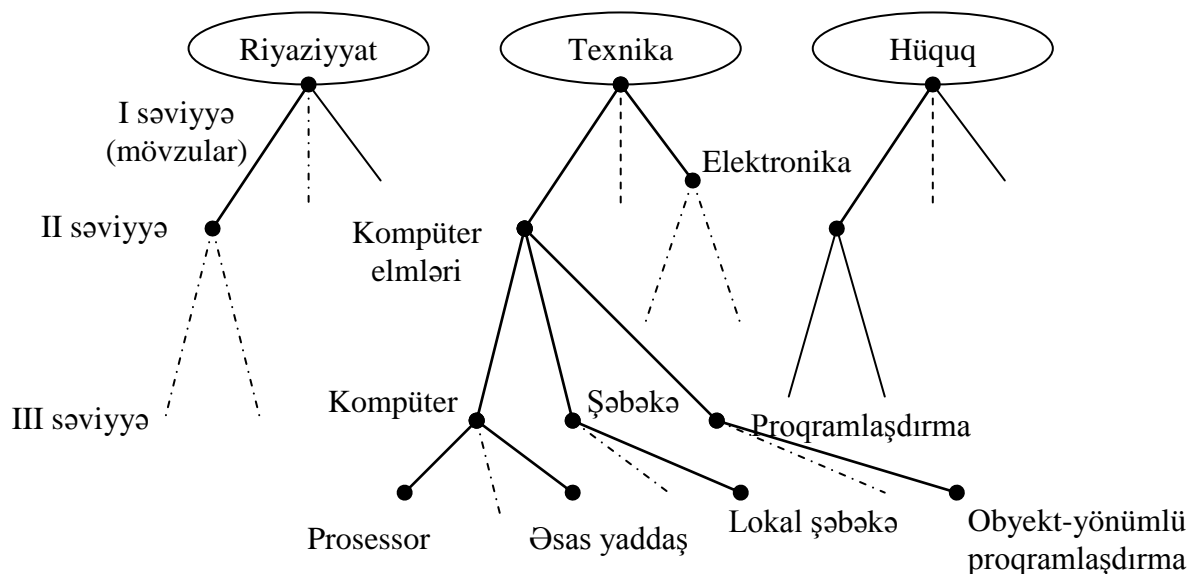
Polis. Daxili İşlər Nazirliyi.

Polis. Daxili İşlər Nazirliyi, Baş Cinayət Axtarış İdarəsi..

Polis. Daxili İşlər Nazirliyi, Baş Narkotiklərlə Mübarizə İdarəsi.

Təsnifat dilinin xüsusi halı leksik vahidləri tematik rubrikaların adlarından ibarət olan rubrikatordur.

Ümumiyyətlə, hər hansı mövzu sahəsinin rubrikatoru dedikdə ayrı-ayrı ağaclarından ibarət istiqamətlənmiş qraf başa düşülür. Ağacların yarpaqları rubrikatorlar-obyektlər adlandırılır və onlar baxılan mövzu sahəsinin konkret fraqmentləri haqqında bilikləri əks etdirirlər. Yarpaq olmayan bütün təpələr yarpaq təpələrin soy-növ tipli təsnifat ümumiləşdirilmələridirlər və yalnız informasiya axtarışı zamanı istifadə edilir. Misal üçün, elmi sahələrin texnikaya aid olan sinfini təsvir edən rubrikator-ağac şəklində göstərilmişdir.



Səkil 2. Rubrikatora aid misal

Adətən rubrikator ekspertlər qrupu tərəfindən mövzu sahəsi haqqında biliklər əsasında və istifadəçilərin informasiya tələbatını nəzərə almaqla hazırlanır.

Təsnifat dillərinin digər nümayəndəsi də biblioqrafik məlumatın təsviri üçün istifadə edilən “Universal onluq təsnifat” (UOT) dilidir. Burada rubrikalar sözlərlə deyil, tam onluq ədədlərlə təsvir edilir. Məsələn:

UOT 681.3(075)- informatika

UOT 681.3.06 - proqramlaşdırma.

Təsnifat dillərinin bir xüsusiyyətini də qeyd etmək lazımdır. Mürəkkəb anlayışlar əvvəlcədən, yəni məlumatın təsvirindən qabaq formalaşdırıldığından, onları təşkil edən sözlər də əvvəlcədən bir-biri ilə əlaqələndirilir. Odur ki, bu tip dillərə “əvvəlcədən koordinasiya edilən” dillər də deyilir.

İnformasiya-axtarış dillərinin digər tipi deskriptor dillərdir.

Deskriptor dillərində leksik vahidlər əvvəlcədən hər hansı tekstual əlaqələrlə bağlanmırlar. Mürəkkəb sintaksis konstruksiyalar-cümlələr və ya ifadələr-sənədlərin məzmunlarının təsviri zamanı leksik vahidləri birləşdirməklə (koordinatlaşdırmaqla) qurulur. Bu dillərdə hazır cümlələr və ifadələr yoxdur, odur ki, mürəkkəb anlayışların qurulmasında məhdudluqlar olmur. Faktiki olaraq çox da böyük olmayan sayda leksik vahidlərin köməyi ilə istənilən mənaya malik olan cümlələri formal təsvir etmək olur. Cümlənin sözləri arasında koordinasiya onun yazılışı zamanı aparıldığından, bu cür dillərə həmçinin « sonradan koordinasiya edilən dillər də deyilir.

Deskriptor dili mövzu sahəsini xarakterizə edən leksik vahidləri-sözləri və söz birləşmələrini özündə cəmləyən lüğətdən istifadə edilməsini nəzərdə tutur. Bir çox hallarda həmin lüğətdə toplanan leksik vahidlər arasında paradigmatik kontekstdən asılı olmayan məntiqi əlaqələr qeyd olunur. İnformasiya axtarışında bu əlaqələrdən ən çox istifadə olunanları ekvivalentlik və tabelilik əlaqələridir.

Ekvivalentlik əlaqəsi sözün müxtəlif formalarını (hallarını) və sinonimlərini ifadə edir. Həmin söz formalarını və sinonimlərin içərisindən baxılan anlayışı daha dəqiq və tam ifadə edən bir söz və ya söz birləşməsi seçilir və o, deskriptor (təsvir edici) kimi qəbul olunur.

Tabelilik əlaqəsi baxılan sözə görə hansı sözlərin daha ümumi və ya xüsusi olmasını ifadə edir və 2 növ əlaqəni özündə birləşdirir: “tam-hissə” (məsələn, “kompüter” – “prosessor”) və “soy-növ” (məsələn, “prosessor” - “Pentium-4”).

Başqa sözlə, tabelilik əlaqəsi anlayışların “Sinif-altsinif” münasibətlərini əks etdirir. İnformasiya axtarışında bu münasibətlərin nəzərə alınması mühüm əhəmiyyət kəsb edir və axtarış göstəricilərini xeyli yaxşılaşdırır.

Ekvivalentlik və tabelilik əlaqələrinin nəzərə alınması ilə leksik vahidləri (sözləri və söz birləşmələrini) özündə cəmləşdirən lüğətə deskriptorlar lüğəti deyilir. Deskriptorlar lüğəti bir növ tezaurus adlanan lüğətə oxşadığından, İAS nəzəriyyəsində ona informasiya axtarış tezaurusu və ya sadəcə olaraq tezaurus deyilir.

Tezaurus tematik əlamətlərə görə nizamlanmış sözlərdən və ifadələrdən ibarət olan lüğətə deyilir. Tezaurus özündə anlayışların adları lüğəti ilə təsnifat əlaqələri lüğətini birləşdirir. Beləliklə, İAS baxımından tezaurus baxılan mövzu sahəsini xarakterizə edən deskriptorları-sözləri və həmin deskriptorlarla paradigmatik əlaqələrlə bağlı olan sözləri özündə cəmləyir. Tezaurusda hər bir deskriptor və onunla bağlı olan sözlər «deskriptor məqaləsi adlanan vahid kimi

qəbul edilir. Deskriptor məqalələrinə aid misal göstərilmişdir. Burada deskriptorların altından xətt çəkilmişdir.

Qeyd edək ki, tezaurus bir qayda olaraq konkret problem (mövzu) sahəsi üçün yaradılır, çünki tezaurusun leksik tərkibi və leksik vahidlər arasındakı paradigmatik əlaqələr konkret problem sahəsinə aid olur.

Sənədin və sorğunun axtarış surətlərinin tərtib edilməsi zamanı onların əsas məzmununu xarakterizə edən sözlər (və ya söz birləşmələri) seçilir, sonra isə tezaurusdan istifadə etməklə həmin sözlər deskriptorlarla əvəz olunur. Beləliklə, deskriptor dilində təsvir edilən sənədin və sorğunun axtarış surətləri sənədin və sorğunun əsas məzmununu təyin edən deskriptorlardan ibarət olur. Bu cür deskriptor dilinə qrammatikasız dil deyilir. Qrammatikasız deskriptor dili sənədin məzmununu lazımi dəqiqliklə təsvir etməyə imkan vermədiyindən, informasiya axtarışının keyfiyyət göstəriciləri nisbətən aşağı olur. Odur ki, bir çox hallarda qrammatikalı deskriptor dilindən istifadə edilir.

Qrammatikalı deskriptor dilində sənədin və sorğunun axtarış surətinə deskriptorlarla yanaşı, qrammatik elementlər də daxil edilir. Əsasən iki cür qrammatik elementdən istifadə edilir: rol göstəricilərindən və əlaqə göstəricilərindən.

Rol göstəricisi axtarış surətində deskriptorun hansı qrammatik və ya semantik rola malik olduğunu təyin edir. Məsələn, rol göstəriciləri kimi “obyekt”, “subyekt”, “vaxt”, “hərəkət” (proses), “xassə”, “məkan” (yer) və s. istifadə edilə bilər.

Ümumi sözlər	Ekvivalent sözlər	Xüsusi sözlər (alt sinif)
Proqramlaşdırma dili	Prosedur dili İmperativ dil	Basic Fortran Pascal
Ümumi sözlər	Ekvivalent sözlər	Xüsusi sözlər (alt sinif)
Süni dil	Proqramlaşdırma dili Alqoritmik dil	Prosedur dili Funksional dil Məntiqi dil Obyekt yönlü dil
Ümumi sözlər	Ekvivalent sözlər	Xüsusi sözlər (alt sinif)
Dil	Süni dili Formal dil	Proqramlaşdırma dili İnformasiya-axtarış dili Beynəlxalq köməkçi dil

Şəkil 3. Deskriptor məqalələrinə aid misal

Rol göstəricilərinin təsvirini asanlaşdırmaq üçün axtarış surətinin mövqeli yazılış formasından istifadə etmək olar. Bu halda SAS cədvəl formasında təsvir edilir. Cədvəlin sütunları rollara uyğun götürülür, məsələn, 1-ci sütunda obyekt rolunda çıxış edən deskriptor, 2-ci sütunda xassə rolunda çıxış edən deskriptor, 3-cü sütunda proses rolunda çıxış edən deskriptor və s. yazılır.

Əlaqə göstəricisi SAS-i təşkil edən deskriptorlar arasında əlaqəni təyin edir, yeni baxılan deskriptorun hansı deskriptorla qarşılıqlı əlaqəli olmasını göstərir. Bu halda əlaqənin tipi göstərilir. Əlaqənin tipini baxılan deskriptorun rolu ilə təyin etmək olar. Bu halda qrammatik konstruksiya kimi rol və əlaqə göstəriciləri birlikdə tətbiq edilir. İndeksləşdirmə prosesinin mürəkkəbləşməsinə baxmayaraq bu cür qrammatik konstruksiya sənədlərin məzmununu daha dəqiq təsvir etməyə və axtarışın intellektual səviyyəsini artırmağa imkan verir.

Qeyd edək ki, qrammatikasız deskriptor dilindən istifadə edildikdə axtarış dəqiqliyinin aşağı olması nəticəvi informasiyanın izafiliyinə səbəb olur. Bu halda

izafi informasiyadan faydalı informasiyanın seçilməsinə müəyyən vaxt və əmək sərf olunur. Lakin bu əlavə səflər qrammatikalı deskriptor dilindən istifadə edildikdə indeksləşdirmənin mürəkkəbləşməsi ilə bağlı əlavə əmək və vaxt itkisi ilə müqayisədə çox olmur. Odur ki, informasiya-axtarış sistemlərinin əksəriyyətində qrammatikasız deskriptor dilindən istifadə olunur.

İndeksləşdirmə.

İAS-in girsinə daxil olan sənədlər təbii dildə ifadə edildiyindən, onların informasiya-axtarış dilinə çevrilməsi tələb olunur. Sistemdə tətbiq edilən informasiya-axtarış dilinin bu əməliyyatın yerinə yetirilməsinə böyük təsiri var. İAD kimi deskriptor dilindən istifadə edildikdə bu əməliyyata indeksləşdirmə, rubrikatordan istifadə edildikdə isə rubrikatlaşdırma deyilir. Qrammatikasız deskriptor dilindən və sərbəst, yəni leksik tərkibi məhdudlaşdırılmayan lüğətdən istifadə edildikdə, bu əməliyyata tammətnli indeksləşdirmə deyilir.

Sənədlərin təbii dildən informasiya dilinə çevrilməsi iki mərhələdə yerinə yetirilir:

1. Sənəddə bəhs edilən obyektlər, onların xassələri və onlar arasındakı əlaqələr haqqında sistemə məlum olan məlumatın seçilməsi məqsədilə sənədin məzmununun təhlili.

2. Ayrılan məlumatın informasiya dilində əks etdirilməsi, yəni həmin məlumatı İAD vahidləri ilə ifadə etməklə, sənədin axtarış surətinin tərtib edilməsi.

Mətnin məna baxımından məzmununun təhlili linqvistik və ekstralingvistik biliklərdən istifadə olunmasını tələb edir. Linqvistik biliklər müəyyən dil üçün ümumidir və bu günə qədər artıq kifayət səviyyədə formallaşdırılmışdır. Ekstralingvistik biliklər isə konkret mövzu sahəsindən asılıdır və onların formallaşdırılması mürəkkəb məsələlərdən biri sayılır. Odur ki, müasir İAS-da sənədin məna məzmununun təhlili əksər hallarda sözlərin və sözbirləşmələrinin *normallaşdırılması* məqsədilə aparılan linqvistik təhlillə reallaşdırılır. Sözlərin normallaşdırılması dedikdə onların kanonik formaya gətirilməsi (məsələn, isimlər üçün-adlıq və tək hala gətirilməsi və s.), sözbirləşmələrinin normallaşdırılması dedikdə isə onların tərkib hissələrinin normallaşdırılması və müəyyən ardıcılıqla yazılması (məsələn, əvvəlcə əsas söz, sonra isə asılı sözlər yazılır) başa düşülür. Normallaşdırılmış sözlərə və sözbirləşmələrinə çox vaxt *terminlər* deyilir.

Mətnin linqvistik təhlili

Mətnin linqvistik təhlili iki mərhələ: morfoloji və sintaktik təhlil ilə aparılır:

Morfoloji təhlildə məqsəd sözlərin əsaslarını hər bir söz forması üçün qrammatik kateqoriyaların (məsələn, nitq hissəsi, cins, hal və s.) qiymətləri ilə birlikdə təyin etməkdir. Sözü əsası dedikdə onun şəkilçisiz yazı forması başa düşülür.

Morfoloji təhlilin dəqiq və təxmini metodları mövcuddur. Dəqiq metodlar sözlərin və sözbirləşmələrinin əsasları lüğətdən istifadə edilməsini nəzərdə tuturlar. Təxmini metodlar isə söz formalarındakı hərflər birləşmələri ilə onların qrammatik qaydaları arasında eksperimental yolla qurulmuş əlaqələrə əsaslanır.

Dəqiq metodlarda söz formaları lüğətdən istifadə edilməsi morfoloji təhlilin bir sıra çətinliklərini (məsələn, saitlərlə samitlərin növbə ilə bir-birini əvəz etməsi) asanlıqla aradan qaldırmağa imkan verir. Bu yolla sözlərin əsaslarının və qrammatik qiymətlərinin təyin edilməsi lüğətdə axtarışa və uyğun informasiyanın seçilməsinə gətirilir. Morfoloji təhlil isə yalnız söz formasının lüğətdə tapılmadığı halda tələb

olunur. Lakin söz formaları lüğəti söz əsasları lüğətindən həcmcə 2-3 dəfə böyük olduğu üçün yaddaş və vaxt sərfi çoxalır.

Söz əsasları lüğətindən istifadə etməklə aparılan morfoloji təhlil sözün əsasını düzgün ayırmaq məqsədilə yerinə yetirilən flektiv təhlilə əsaslanır. Bu yolla aparılan morfoloji təhlildə əsas çətinlik söz əsaslarının omonimiyası halı ilə bağlıdır. Onun qarşısını almaq üçün ayrılan əsasla şəkilçi arasında uyğunluğun olub-olmaması yoxlanılır.

Morfoloji təhlilin təxmini metodlarının əsasında aşağıdakı hipotez durur: son hərflərinə və hərf birləşmələrinə görə sözün qrammatik sinfini praktik olaraq birmənalı təyin etmək mümkündür. Sözün əsası belə təyin edilir: sözün sonundan hərflər ardıcıl olaraq bir-bir ayrılır və alınmış hərf birləşmələri baxılan qrammatik sinfə uyğun şəkilçilər siyahısında müqayisə edilir. Uyğunluq alınan kimi, sözün qalan hissəsinin əsas olduğu haqda nəticə çıxarılır. Azərbaycan dilində mətnlərin morfoloji təhlilini aparmaq üçün rus dilinə nisbətən daha çox hərfin (orta hesabla 3-4 hərfin) ayrılmasına ehtiyac olur.

Morfoloji təhlilin yerinə yetirilməsi nəticəsində qrammatik informasiya birmənalı təyin olunmaya bilər. Bu halda birmənalılığın əldə edilməsi üçün sintaktik təhlil tətbiq edilir.

Sintaktik təhlilin vəzifəsi lüğətdə əks olunan informasiya əsasında cümlənin qrammatik araşdırılmasını yerinə yetirməkdir. Nəticədə cümlənin mübtədasi, xəbəri, tamamlığı və s. ayrılır və onlar arasında asılılıq ağacı şəklində idarəetmə əlaqələri göstərilir.

İstənilən sintaktik təhlil vasitəsi iki hissədən ibarət olur: konkret dil haqqında biliklər bazasından və sintaktik təhlil alqoritmindən, yəni həmin biliklər əsasında mətni emal edən standart operatorlar dəstindən.

Qrammatik biliklərin mənbəyini morfoloji təhlil nəticəsində alınmış verilənlər və müxtəlif cədvəllər təşkil edirlər. Sintaktik təhlilin aparılması üçün müəyyən qanunauyğunluqların əldə edilməsi məqsədilə mətnlərin insan tərəfindən empirik emalının nəticələrindən ibarət olan həmin cədvəllər standart şəkildə doldurulur. Bu cədvəllərin əsasını hər bir leksik vahid üçün digəri ilə mümkün əlaqə variantlarını göstərməklə leksik vahidlər siyahılarından ibarət olan konfigurasiyaların toplusu və ya sintaktik və semantik-sintaktik valentliyin yığımları təşkil edirlər. Sintaktik təhlilin praktik reallaşdırılması zamanı həmin cədvəllərin emalı qaydalarının onların məzmunundan asılı olmamasına nail olmaq lazımdır. Bu halda cədvəllərin məzmununun dəyişdirilməsi tələb olunduqda təhlil alqoritmi yenidən qurmağa ehtiyac olmur.

İndeksləşdirmənin ümumi qaydaları

Sənədlərin və sorğuların insan tərəfindən əllə indeksləşdirilməsi birmənalı nəticəyə gətirib çıxarmır: müxtəlif icraçılar eyni sənədi lüğətin köməyi ilə müxtəlif cür indeksləşdirə bilərlər. Odur ki, axtarışı yaxşılaşdırmaq üçün sənədlərin və sorğuların məzmununun təbii dildən informasiya dilinə çevrilməsi əməliyyatının avtomatlaşdırılması problemi ilk İAS-ın yarandığı andan qarşıya qoyulmuş və bu problem artıq öz bəllini tapmışdır.

Hər bir təbii dil özünə məxsus qrammatikaya malik olduğundan, bütün dillər üçün yararlı olan ümumi indeksləşdirmə alqoritmini yaratmaq mümkün deyil. Bundan əlavə, hər bir təbii dil üçün mövzu sahələrinin dil xüsusiyyətləri fərqləndiyindən, bütün sahələr üçün tətbiq edilə bilən ümumi indeksləşdirmə alqoritmini qurmaq da özünü

doğrultmur. Hər bir mövzu sahəsinin dil xüsusiyyətlərinə və ona uyğun hazırlanmış tezaurusa görə indeksləşdirmə alqoritmi qurulur. İndeksləşdirmə alqoritminin qurulmasında indeksləşdirmənin ümumi qaydalarını da nəzərə almaq lazımdır. Qısa şəkildə həmin qaydalara baxaq.

İndeksləşdirmə prosesi aşağıdakı ardıcılıqlarla yerinə yetirilən mərhələlərdən ibarətdir:

- indeksləşdirmə obyektini kimi sənədin təhlili və məzmununun təyin edilməsi;
- sənədin məzmununu xarakterizə edən anlayışların seçilməsi;
- anlayışların ifadə edilməsi üçün indeksləşdirmə terminlərinin seçilməsi;
- indeksləşdirmə terminlərindən ibarət olan SAS-in tərtib edilməsi.

Göstərilən mərhələlər texnoloji proseduraların tərkibində birləşdirilə bilər.

Sənədin təhlili zamanı indeksatora (indeksləşdirməni aparən şəxsə) sənədlə tam həcmdə tanış olmağa şərait yaradılmalıdır. Bu mümkün olmadıqda indeksator indeksləşdirmənin aşağıdakı əsas mənbələrini öyrənməlidir:

- sənədin arayış komponentləri: başlığı (adı), annotasiyası, referatı, mündəricatı, ön söz, nəticə hissələri və s.;
- giriş;
- hissələrin və bölmələrin adları;
- bölmələrin və paraqrafların ilkin cümlələri;
- illüstrasiyalar, sxemlər, cədvəllər və onların adları;
- mətnə poliqrafik vasitələrə ayrılan sözlər və ifadələr.

Sənədin yalnız başlığına (adına) görə indeksləşdirmə lazımi nəticə vermir. Referata və annotasiyaya görə indeksləşdirmə zamanı onların sənədin məzmununa adekvat olub-olmamasına nəzər yetirmək lazımdır.

Qeyri-mətni (audiovizual, qrafik və s.) sənədlərin təhlili zamanı onları xarakterizə edən mətni komponentlərə (adlarına, qısa təsvirlərinə və s.) görə indeksləşdirmə aparıla bilər. Əgər mətni material kifayət etməzsə, sənədin özü ilə tam tanış olmaq gərəkdir.

SAS-da əks olunan xarakteristikaların və anlayışların sayı onun tamlığını əks etdirir və indeksləşdirmənin vacib göstəricisi sayılır. Sistemin istifadəçiləri üçün qiymətli ola bilən bütün anlayışlar SAS-da Öz əksini tapmalıdır. SAS-da istifadəçilərin maraqlı dairəsi çərçivəsində bir neçə mövzu ola bilər.

Anlayışların seçilməsində əsas kriteri kimi sənədin məzmununun ifadə olunması və ya onun axtarışı üçün anlayışın potensial əhəmiyyəti götürülür. Bu zaman diqqəti tipik sorğulara yönəltmək lazımdır. Bunun üçün:

- a) İAS-ın istifadəçiləri tərəfindən ən çox istifadə edilən anlayışlar seçilməlidir;
- b) istifadəçilərlə əks əlaqə əsasında tezaurusun leksik tərkibini və İAD-ın qrammatik qaydalarını dəqiqləşdirmək lazımdır.

Hər bir sənəd üçün müəyyənləşdirilən indeksləşdirmə terminlərinin sayı sənədin əhatə etdiyi əsas məlumatların sayı ilə təyin edilir. Hər bir İAS üçün qəbul edilmiş indeksləşdirmə tamlığı onun funksional təyinatı ilə müəyyənləşdirilir. Sənədin həcmi də indeksləşdirmə tamlığına ciddi təsir edir. Bu amilləri nəzərə almaqla ekspert yolu ilə sənəddən anlayışları seçmək lazımdır.

SAS-ın spesifikliyi (dəqiqliyi) sənədin anlayışlarının indeksləşdirmə terminləri ilə nə dərəcədə düzgün əks olunması ilə təyin edilir və indeksləşdirmə keyfiyyətinin parametrlərindən biri hesab olunur. Anlayışın daha geniş anlayışı ifadə edən terminlə əvəz edilməsi spesifikliyi zəiflədir. Daha geniş anlayışları ifadə edən terminlərdən xüsusi hallarda istifadə etmək olar:

a) əgər həddən artıq spesifik terminin istifadəçilərə aydın deyilsə;

b) əgər anlayış sənəddə tam açılmayıbsa və ya sənədin məzmununu ifadə etmək üçün köməkçi rol oynayırsa.

İndeksləşdirmə terminlərinin seçilməsi prosesində sənədin məzmununu xarakterizə edən anlayışlar konkret İAD-ın qaydaları ilə müəyyənləşdirilmiş leksik vahidlərlə (deskriptorlarla və ya açarlı sözlərlə) təsvir olunurlar. Yeni anlayışlara ekvivalent terminlər olmadıqda onları təsvir etmək üçün baxılan sahə üzrə informasiya-arayış mənbələri (lüğətlər, ensiklopediyalar, arayış kitabçaları, təsnifat cədvəlləri, informasiya-axtarış tezaurusları, terminoloji standartlar və s.) əsasında yeni terminlər yaradılır.

Baxılan problem sahəsi üçün informasiya-axtarış tezaurusu mövcud olduqda indeksləşdirmə terminləri həmin tezaurus əsasında seçilir. Tezaurusdan istifadə edildikdə sənədin axtarış surətinə ümumi terminləri daxil etməyə ehtiyac olmur. Bu halda həmin terminlərin sorğunun axtarış surətinə daxil edilməsi və axtarış prosesində nəzərə alınması daha yaxşı nəticə verir.

Tezaurusda olmayan, lakin SAS-ın tərkibində olması tələb edilən anlayışları iki üsulla ifadə etmək olar:

- tezaurusa və SAS-a daxil edilən yeni spesifik terminlə;
- tezaurusda olan daha ümumi terminlə, Bu halda spesifik termin gələcəkdə tezaurusa daxil edilməsi üçün xüsusi siyahıya salınır.

Yeni anlayışlar tezaurusda onlara daha yaxın olan leksik vahidlərlə ifadə olunur və ya onların tezaurusda yeni leksik vahidlər kimi ifadə olunmasının gərəклиyi araşdırılır.

İndeksləşdirməni sənədin mətnindən sərbəst götürülən açar sözlərlə apardıqda həmin sözlər kanonik formaya gətirilməlidir.

Sənədin axtarış surəti tezaurus əsasında və ya sərbəst seçilən və İAD-ın qrammatik vasitələri ilə təşkil olunmuş indeksləşdirmə terminləri dəstindən ibarət olur.

Normallaşdırma dərəcəsinə görə SAS-a iki tip termin daxil edilə bilər: deskriptorlar və açar sözlər. Terminin deskriptor və ya açar söz olması tezaurus vasitəsilə müəyyənləşdirilir. İndeksləşdirmə terminləri SAS-da təbii dilin orfoqrafik qaydaları ilə yazılmalıdır.

Əgər tezaurusda deskriptorlara şərti kodlar verilibsə, SAS-da onların həmin kodlarla yazılışı mümkündür. Bu halda İAS deskriptorların kodlarına görə onların orfoqrafik formalarının axtarışını avtomatik təşkil etməlidir.

Əgər indeksləşdirmə qrammatikalı deskriptor dilində aparılırsa, əlavə olaraq rol və əlaqə göstəriciləri vasitəsilə indeksləşdirmə terminlərinin rolu və onlar arasında əlaqələr göstərilə bilər. Bəzi hallarda isə lazım gələndə terminlərin semantik və morfoloji xarakteristikaları və həcmnin informasiya çəkilişi də qeyd oluna bilər.

Terminin semantik xarakteristikası onun aşağıdakı leksikoqrafik kateqoriyalardan birinə aid edilməsi ilə təyin olunur:

Elmi-texniki anlayış: xüsusi ad, indentifikator; parametrin adı; parametrin qiyməti; ölçü vahidi.

Terminin morfoloji xarakteristikası onun aşağıdakı leksikoqrafik kateqoriyalardan birinə aid edilməsi ilə təyin olunur:

Törəmə söz; mürəkkəb söz; sözbirləşməsi; abreviatura; sözün fraqmenti.

Morfoloji xarakteristikalarından İAS-da leksik vahidlərin formal əlamətlərə görə mənə təhlilini reallaşdırmaq üçün istifadə edilir.

Terminin informasiya çəkisi baxılan anlayışın baxılan sənəd üçün nə dərəcədə əhəmiyyətli olmasını göstərir, məsələn: sənədin əsas mövzusunı ifadə edən anlayış; sənədin əlavə mövzusunı ifadə edən anlayış; sənədin məzmununu təsvir etmək üçün istifadə edilən anlayış.

Bəzən indeksləşdirmə prosesinin təşkilində istifadə edilməsi üçün terminin alınma üsulunu da göstərirlər: Avtomatik indeksləşdirmə nəticəsində alınan termin; indeksator tərəfindən daxil edilən, lakin sənəddə olmayan termin; sənəddə olmayan, lakin tezaurus əsasında daxil edilən termin.

Tezaurus əsasında indeksləşdirmənin avtomatlaşdırılması

İndeksləşdirmənin avtomatlaşdırılmasında məqsəd indeksləşdirmə əməliyyatına sərf olunan material və insan resurslarını minimuma endirmək və həmçinin nəticələrin stabilliyini və bir-mənalılığını təmin etməkdən ibarətdir.

Avtomatlaşdırılmış indeksləşdirmə (Aİ) iki informasiya mənbəyinə görə: ilkin sənədin mətninə və sənədin başlığına, annotasiyasına, referatına görə aparıla bilər.

Avtomatlaşdırılmış indeksləşdirmə aşağıdakı mərhələlər ardıcılığı ilə yerinə yetirilir:

- a) sənədin informativ hissələrinin ayrılması;
- b) mətnin sözlərinin identifikasiyası və onların normallaşdırılmış şəkllə salınması (morfoloji təhlil və sintez);
- c) mətnin açar sözlərinin siyahısının tərtib edilməsi;
- d) tezaurusdan deskriptorların seçilməsi;
- e) sənədin axtarış sürətinin tərtib edilməsi.

Aİ texnologiyası indeksləşdirilən sənədin daha informativ fraqmentlərinin təyin edilməsini və indeksatora və ya proqrama təqdim edilməsini nəzərə almalıdır.

Mətnin sözlərinin indektifikasiyası dəqiq və ya təxmini morfoloji təhlil metodları ilə qurulan alqoritm və proqramlarla yerinə yetirilir. Sözlərin normallaşdırılması isə adətən söz-formaları lüğətindən istifadə edilməklə reallaşdırılır.

Açar sözlərin siyahısının tərtib edilməsi baxılan təbii dilin qrammatik kateqoriyalarının ahəngliyi qaydalarını nəzərə almaqla aparılan sintaksis təhlil nəticəsində yerinə yetirilir. Mətnin sintaksis təhlili aşağıdakı məsələləri həll edir:

Verilmiş kriterilərə görə mətnin fraqmentlərə ayrılması; Mətnin söz-formaları arasında sintaksis asılılığın təyin edilməsi; Sözbirləşmələrinin müəyyənləşdirilməsi; Ayrılan açar sözlərin normallaşdırılması.

Aİ sənədin axtarış sürətinin sərbəst açar sözlərdən və ya baxılan problem sahəsi üçün qurulmuş tezaurusun deskriptorlarından təşkil olunmasını təmin etməlidir.

İndeksləşdirmə tezaurusun deskriptorları ilə aparıldıqda SAS-ın tərtibi mərhələsində açar sözlər tezaurusun uyğun deskriptorları ilə əvəz edilir. Bu zaman SAS seçilən deskriptorlara görə yuxarı səviyyədə duran, yəni daha ümumi anlayışları ifadə edən deskriptorlarla zənginləşdirilə bilər. Avtomatlaşdırılmış indeksləşdirmə proqramı SAS-a tipik qrammatik vasitələri daxil etməyi nəzərə almalıdır.

Aİ sistemi modul strukturuna malik olmalı və tipik proqram və texniki vasitələrin tətbiqi ilə reallaşdırılmalıdır.

Sual 2. İnternet – informasiya axtarış məkanı kimi.

İnternetin tarixi. İnternet təxminən 50 illik keçmişə sahibdir. Bu gün rabitədə vazkeçilməz kimi görünən İnternet 50 ilə qədər öncə xəyal belə etmirdilər. 1960-cı illərdə bilgisayar mütəxəssislər üçün haqq-hesab yerinə yetirən hesab maşınından başqa bir şey deyildi. Ancaq onun əlaqə, rabitə vasitəsi kimi perspektivini görenlər də az deyildi. İnternetin köklərini Amerikanın ən böyük ali təhsil müəssisələrindən olan Massaçüets Texnoloji institutunda cənab J. Likliderin 1962-ci ildə təşkil etdiyi müzakirələrdə tapmaq olar. Liklider iddia edirdi ki, bütün dünyanı əhatə edən bir sistem qurmaq olar, həmin sistemə hər kəs qoşula, istədiyi bilgini ala, yaya bilər. Düşündürücü ideyalar sahibi çox keçmədi ki, Amerikanın Hərbi Araşdırma Layihəsinin (ARPA-Defense Advanced Research Project Agency) bilgisayar bölməsinin rəhbəri oldu.

Bilgisayarların ilk dəfə bir-biri ilə "danışmasını" isə Federal Təhlükəsizlik Bürosunun araşdırmaçıları Lavrens Roberts və Tomas Merrill 1965-ci ildə gerçəkləşdirib. 1966-cı ilin sonundan L.Roberts Amerika Hərbi Araşdırma Layihəsində "Arpanet" adlı layihə üzərində çalışmaya başlayıb. Layihə çərçivəsində ilk dördtərəfli bağlantı 1969-cu ildə mümkün olub, bəldliklə İnternetin ilk şəklə ortaya çıxıb. "Arpanet" layihəsini dörd mərkəz - Los Anceles (UCLA), Santa Barbara (UCSB) Universitetləri, Stanford Araşdırma İnstitutu və Yuta Universiteti gerçəkləşdirib. Qısa zamanda digər mərkəzlər də Arpanet şəbəkəsinə bəllanıblar.

1972-ci ildə isə Arpanet şəbəkəsi daxilində ilk dəfə olaraq elektron poçtdan (e-mail) istifadə olunub.

İnternet haqqında 2 geniş yayılmış fikir var:

1. İnternetdə, demək olar ki, hər şey var.
2. Ancaq bu hər şeyin içindən konkret nə isə tapmaq, demək olar ki, mümkün deyil.

Bu sözlər Konfutsinin fikri ilə üst-üstə düşür: "Qaranlıq otaqda qara pişiyi tapmaq çətindir. Xüsusilə əgər pişik otaqda yoxdursa".

Həqiqətən də İnternetdə çox böyük həcmdə informasiya toplanıb. Bu qədər informasiyanın içindən axtardığınız konkret məlumatı tapmaq üçün nə axtardığınızı dəqiq bilməklə yanaşı, bunu axtarış sistemlərinə də verməlisən, İnternet nəhəng kitabxanadır. Bütün kitabxanalarda olduğu kimi, burada da axtarış aparatlarından istifadəni bacarmaq lazımdır. Www-nun köməyi ilə İnternetdəki informasiya və xidmətlər qovluğu artıq onlarla cild kitab həcmi keçmişdir.

Bəlkə də İnternetin ən lazımlı cəhəti onda axtarış serverlərinin (Search engines) olmasıdır. Bunlar bütün şəbəkədəki resursları tapıb onları indeksləşməyi bacaran ayrılmış kompüterlərdir, Bundan sonra sizi maraqlandıran mövzuya dair açar sözü, ifadəni yazıb serverə göndərirsiniz. O isə, öz növbəsində, bu sözlərə uyğun gələn saytların siyahısını sizə təqdim edir. İnternetdə nə isə axtarıb tapmağın ən asan yolu budur. Buna nümunə olaraq İnternetin ən məşhur axtarış mexanizmlərindən biri olan "AltaVista"-nın saytını misal göstərmək olar.

İnternetdə axtarış aparmağın digər üsulu da axtarış qovluqlarından (İnternet directories) istifadədir. Qovluqlarda mövzulara görə ayrılmış şəbəkə resurslarına əsasən də www sənədlərə istinadlar yığılır. Bu istinadlar avtomatik olaraq yığılır, hər qovluğun inzibatçısı tərəfindən həyata keçirilir. Bu da istifadəçilər üçün rahatlıq yaradır. Belə ki, artıq bütün İnternetdə deyil, konkret mövzuya aid qovluqda aparılan axtarış daha səmərəli olur. İnternet qovluqlarla işləyən axtarış saytları əlavə xidmətlər də göstərir: açar sözə görə axtarış, son daxil olunmuş resursların siyahısı,

təsadüfi istinadlar siyahısı, yeni daxil olan resurslar haqqında elektron poçta xəbər göndərmək və s. Bütün bunlar İnternet qovluqları ilə işləməyi daha da sadələşdirir. Bu üsulla işləyən axtarış saytlarından ən məşhuru Yahoo.com saytının əsas səhifəsi sonrakı şəkildə göstərilib.

Axtarışın xüsusi vasitələrlə aparılması zamanı xüsusi proqramlardan istifadə edilir. Belə proqramlardan "Copernic 2000 Pro" sistemini göstərmək olar. Bu proqramın əsas imkanları:

- 250-dən çox axtarış serverinə müraciət;
- Məntiqi operatorların tam lüğətindən istifadə edilir;
- Axtarış nəticələrinin yadda saxlanması;
- Nəticənin relevantlıq səviyyəsi göstərilir;
- Təkrar olunan və "ölü" linklər silinir;
- Web, News və Email ünvanlarına görə axtarış imkanı;
- 32 meyara görə mövzu yönümlü axtarış imkanı;
- Açar sorğunun orfoqrafik yoxlanması.

Www-da axtarışın əsasları

İnternetdə tez-tez axtarış aparılmaq üçün "Explorer 6"da ünvan sətirində URL əvəzinə axtarış meyarını yazaraq axtarış aparmaq imkanı var. Təəssüf ki, susma halına görə axtarış MSN-də həyata keçirilir, bu zaman çoxu "Google"a üstünlük verir. Bəs axtarış mexanizmini necə dəyişmək olar? Hər hansı bir sayta daxil olmadan digər axtarış mexanizmindən istifadə etmək imkanı varmı?

Əlbəttə, belə bir imkan var. İnternet Explorer proqramını açın və View→Explorer Bar→Search menyularını seçin. Nəticədə sol tərəfdə bir panel açılacaq. Əgər onda axtarışın animasiyalı piktoqramı varsa, onda Change Preferences?, sonra da Change İnternet Search Behaviouru daxil edin. Ekranda məşhur axtarış saytlarının adları əks olunacaqdır. Siz bunlardan istədiyinizi seçə bilərsiniz.

Susma halına görə axtarış aparmaq üçün yalnız sol tərəfdə çıxan paneldən istifadə edə bilərsiniz. Ünvan sətirinə açar sözlərin daxil edilməsi zamanı axtarış əvvəlki kimi MSN-də aparılacaq. Yox, əgər panel açıq qalarsa, onda istənilən axtarış mexanizmini sayta girmədən istifadə edə bilərsiniz. Yandakı paneldə axtarış meyarlarının konkretləşdirilməsi üçün təkliflər olur.

Əgər siz ünvan sətirindən axtarışı həyata keçirtmək istəyirsinizsə, onda Change İnternet Search Behaviour funksiyasından sonra With Classic İnternet Search rejimini seçin. Sonra İnternet Explorer pəncərəsini bağlayın və yenidən açın. Yan panelin xətkəşində Customize bəndi əmələ gələcək. Customize Search Settings adlı dialoq pəncərəsini açmaq üçün bu düyməni sıxın. Açılmış pəncərədə Autosearch Settings düyməsini sıxmaqla ikinci dialoq pəncərəsini açırırsınız. Bu pəncərənin köməyi ilə ünvan sətiri vasitəsilə aparılan axtarış mexanizmini susma halına görə müəyyənləşdirmək mümkündür, Əgər İnternet Explorer-in siyahısında lazımi axtarış mexanizmləri yoxdursa, onda onları əlavə etmək olar.

Axtarış səhifələrinin hər birində ətraflı axtarış linki mövcuddur. Bu istinaddan adətən o zaman istifadə edilir ki, axtarılan mövzu haqqında informasiya həddindən çox olur. Bu zaman ətraflı axtarışın köməyi ilə verilən şərtləri daha da konkretləşdirərək çoxsaylı resursların içindən verilən şərti ödəyənləri tapmaq olur. Ətraflı axtarış zamanı əsasən resursların hansı dildə olması, açar sözlərin səhifə

başlığında, alt başlıqda, yoxsa bütün səhifədə axtarılması, hansı domen genişlənməli saytlarda axtarmalı (.com, .org, .az), əgər şəkil axtarılsa, onun tipi (jpeg, bmp, gif), ölçüləri (lap böyük, orta, kiçik), rəngli olub-olmadığı göstərilir. Bu göstərilənlərdən istədiyiniz üzrə axtarışı davam etdirə bilərsiniz. Bunun üçün onları qeyd və ya seçim edib "Axtar" (Search) düyməsini vurmaq lazımdır.

Müxtəlif axtarış sistemlərində axtarış nəticələrinin müxtəlifliyinin səbəbi həm onların axtarış alqoritmlərinin müxtəlifliyi, həm də axtarışı həyata keçirmək üsullarının fərqliliyidir. Axtarış alqoritmləri haqqında söhbət açmaq bu kitabın mövzusunda kənar olduğu üçün ikinci fərq olan axtarış tiplərinin fərqliliyi ilə tanış olaq. Yuxarıda qeyd olunmuşdur ki, İnternetdə axtarış əməliyyatları axtarış mexanizmləri və axtarış qovluqları vasitəsilə aparılır. Məsələ ondadır ki, axtarış mexanizmlərində resursları serverin verilənlər bazasına robotlar yığır, axtarış qovluqları üsulu ilə işləyən sistemlərdə isə resursları serverin verilənlər bazasına sistemin inzibatçıları yığır. Bu səbəbdəndir ki, axtarış mexanizmlərində resurs daha çox olur, ancaq onların arasında "lazımsızların" sayı da həddən artıq ox olduğu zaman axtarış qovluqları ilə işləyən Yahoo.com kimi axtarış sistemlərində resurslar sayca az olsa da, lazım olan resurs sayı da çox olur.

Axtarış sistemləri linklərin səmərəliliyinə fikir vermədən onları indeksləyir. Buna görə də bu sistemlər İnternetin "künc-bucağ"ındakı bütün resursları tapır. Digər tərəfdən, əgər sorğunu düzgün formalaşdırmamısınızsa, onda axtarış uğurlu nəticə verməyəcək. Axtardığınız resurs haqqında bildiyiniz bilgilərə əsasən İnternet-qovluqlardan istifadə etsəniz, daha səmərəli nəticə əldə edə bilərsiniz. Məsələn, qəzetdə XYZ ifadəsinə rast gəlinirsiniz. Bu haqda informasiyanı axtarış serverlərindən rahat tapa bilərsiniz. Müxtəlif axtarış mexanizmlərindən istifadə ilə daha dəqiq resurs əldə edə bilərsiniz. Müxtəlif axtarış serverlərinin açar sözlərini qiymətləndirməsi, relevantlıq əmsalı müxtəlif olur. Bu iki tip axtarış formasından vaxtında istifadə etmək bacarığınız nə qədər çox olarsa, İnternetdə axtarırlarınızın səmərəsi bir o qədər artmış olar.

İnternetdə informasiya axtarışının təşkili

İnformasiyanın axtarışı probleminə İnternetdə tez-tez rast gəlinir. "World Wide Web"-in əsasını təşkil edən prinsiplər sayəsində onun fantastik sürətlə inkişaf etməsinə baxmayaraq, lazımi informasiyanın axtarışının çətinləşməsi bu inkişafın yaratdığı problemlərdəndir. Çox vaxt nizamlanmamış halda nəticə kimi tapılan səhifələrin içindən lazım olanını tapmaq günü-gündən çətinləşir. Hal-hazırda İnternetdəki informasiyanın biliklərə çevrilməsi istiqamətində bir neçə yanaşma mövcuddur. Bunlardan biri açar sözlərdən istifadə etməkdir. "Yahoo", "Lycos", "Infoseek" kimi axtarış maşınları bu üsuldən istifadə edir. Bütün axtarış sorğuları axtarış sisteminə məlum olan sənədlərdəki sözlərin (term) müəyyən mənasını daşıyan indeks əsasında cavablanır. İndekslərin yaradılmasında istifadə üçün sənədlər haqqında informasiyanı yığmaq məqsədilə **crawlers** adlı robotlardan istifadə edilir. Bu robotlar İnternet resurslarını gəzərək öz verilənlər bazasına yeni resurslar haqqında məlumatlar daxil edir. Bu zaman müxtəlif problemlər - indeksləmə obyektinin seçilməsi, informasiya fəzasında bərabər hüquqlu indeksləmə təminatı, eləcə də hansı sorğular zamanı hansı informasiyanın çatdırılması problemləri yaranır.

Bəzi müasir axtarış sistemləri də var ki, onlarda ("Metacrawler", "WebSeek" və s.) axtarış paralel olaraq bir neçə axtarış maşınından istifadə edərək indekslərdən

ibarət metaqovluqlar və verilənlər bazaları yaradırlar. Göründüyü kimi, bu da açar sözlərdən istifadəyə əsaslanır.

Digər tərəfdən, web-fəzanın ölçülərinin çox nəhəng olması indeksləmənin tam həyata keçirilməsinə mane olur. Bunun nəticəsində axtarışlar vaxtı istifadəçi üçün maraqsız informasiyalar yaranır.

İnternetin intellektuallaşması istiqamətində başqa bir yanaşma HTML standartına yeni teqlər əlavə etməkdir. Bu teqlər (<META>) HTML-in 2.0 versiyasında daxil edilib. Əgər işlənən sənəddə <META> və </META> teqləri vasitəsilə sənədin tərkibi haqqında açar sözlər varsa, bu sənədin lazım olan vaxtda tapılma şansı çoxalır. Bu teqlərdən sənədin tərkibi yox, diqqəti cəlb etmək məqsədilə istifadə geniş tətbiq edilmədiyindən burada da problemlər yaranır. Digər əsas cəhət isə HTML sənədlərin standart "axtarış" xassəsinin olmamasıdır.

İnternetdə informasiya axtarışının resursları

İnternetdə olan resursların əksər hissəsi HTML (Hyper Text Markup Language) əsasında yaradılmış WWW-sənədlər şəklindədir. Ümumiyyətlə, İnternetdəki resursları aşağıdakı kimi bölmək olar:

Axtarış üsuluna görə

- Web-səhifələr;
- Verilənlər bazası;
- Fayl serverləri;
- Xəbər qrupları.

Dil meyarına görə

Coğrafi yerinə görə.

Resursların xarakteri **və** növünə görə bölünməsi daha əhəmiyyətlidir:

- Mövzu yönümlü informasiya;
- Elmi məqalə;
- Kommersiya tipli informasiya;
- Lüğəti informasiya;
- Xəbərlər;
- İkinci dərəcəli informasiya.

İnformasiya axtarış sistemlərinin xüsusiyyətləri:

İnternetdə axtarış aparmağın üsulu var:

1. Hipermətn linklər vasitəsilə;
2. Axtarış sistemlərindən istifadə etməklə;
3. Xüsusi vasitələrlə axtarış aparma.

İnformasiya axtarış sistemləri iki tipdə olur:

- Axtarış serverləri (robotları);
- Axtarış qovluqları.

Axtarış serverləri (robotları) iki cür olur:

• **Axtarış sistemləri** - İnternetdəki dəyişiklikləri güdən və yeni resursları bazasına əlavə edən sistemlər;

• **Meta** - axtarış sistemləri - bir neçə axtarış sisteminin bazasında eyni zamanda axtarış aparan sistemlər.

Xarakterik cəhətləri:

- axtarışın sadəliyi;
- yüksək sürət.

Axtarış qovluqları üç yerə bölünür:

- **Ümumi qovluqlar** - müxtəlif mövzulara dair materiallar;
- **Mövzu yönümlü qovluqlar** - konkret mövzuya dair materiallar;
- **Resurslar siyahısı** - axtarış aparmaq üçün qovluq. siyahıları.

Xarakterik cəhətləri: əllə yığılır, iyerarxik formada düzülür, az "zibil" olur, meyarlara görə axtarışın sadəliyi.

Müasir dövrdə hər hansı mütəxəssisin, rəhbərin, iş adamının və ya təşkilatın fəaliyyəti lazımi informasiyanın əldə etməsi və ondan səmərəli istifadədən asılıdır. İnfomasiyaya tələbatı ödəmək üçün kompüter texnologiyalarında infomasiya axtarış sistemi istifadə olunur. İnfomasiya axtarış sistemi mürəkkəb insan-maşın sistemi olub infomasiyanın saxlanması və axtarılmasını həyata keçirir. Bir çox halda istifadəçi bütün infomasiya "dənizi"ndən yalnız ona lazım olan infomasiyanı əldə etmək istəyir. Çox halda sistem bunun öhdəsindən gələ bilmir. Burada problem istifadəçilərin müxtəlif sahələrə mənsub olması ilə daha da qəlizləşir. İstifadə olunduğu sahələrdən asılı olaraq sözlər müxtəlif mənalar daşıyır.

Elektron sənəd texnologiyasının inkişafı emal tələb edən sənədlərin sayını artırır. Bu da infomasiyaya müraciətin təşkil olunma vasitələrini tələb edir. Hal-hazırda insanın fəaliyyət sahəsində çox böyük həcmdə infomasiya yığılıb. İnfomasiyanın belə artması avtomatlaşdırılmış üsulların, verilənləri emal edən və saxlayan sistemlərin tətbiq olunmasını tələb edir. Bununla əlaqədar infomasiya materialları elektron formata keçirilir. Aydın ki, infomasiyanın öyrənilməsi və analizi üçün lokal və ya uzaq məsafədə olan infomasiya mənbələrinə çıxış olmalıdır. Bu problem İnternet/İntranet texnologiyasının köməyi ilə həll olunur. Digər tərəfdən verilənlərin inteqrasiyası problemi yaranır. Hətta bir serverdə yerləşmiş müxtəlif kolleksiyalar, bir mövzuya aid nəşrlər belə müxtəlif məntiqi girişli olur və müxtəlif mənbələrdəki verilənlərlə əlaqə yarada bilmir.

Verilənlərin inteqrasiyası məsələsinin bir həlli də onları yenidən strukturlaşdırmaqdır. Bu üsul vasitəsilə müxtəlif mənbələrdən bütün verilənlərin yenidən lazımi qaydada strukturunu yaratmaq, unikal identifikatorlar təşkil etmək olar. Ancaq bu, çox çətin, uzun müddət tələb edir və səmərəli deyil.

Sənədlərin təsnifatı da çıxış yollarından biridir. Sənədlərin təsnifatının təşkil olunması axtarışa çox kömək edir, xüsusilə əgər istifadəçi axtardığı infomasiya üçün düzgün açar söz seçə bilmirsə, mövzunu dəqiqləşdirərək tapmaq imkanı yaranır. Bu cür axtarış daha dəqiq olur. Ancaq sənədlərin məzmunlu axtarışı zamanı həmin sözlərin özləri olmasa da, onlarla əlaqəsi olan sənədlərin siyahısını tapır. Məsələn: "kiçik zərrəciklər" ifadəsini yazdıqda, bizə "mikroblar" sözü daxil olan sənədlər də siyahıda göstərilir.

Başqa bir yol hazırda mövcud olan verilənləri saxlama sistemlərindən (data warehouse) istifadə etməkdir. Belə sistemlər, sorğuları müxtəlif mənbələrə göndərməklə və sonra nəticələri yığıb süzgecdən keçirməklə özlərini metaaxtarış sistemləri kimi apararaq infomasiya mənbələrinin inteqrasiyasını təmin edir. Bu halda verilənlər arasındakı semantik əlaqələr nəzərə alınmır. Bu da, öz növbəsində, əlaqəli infomasiyaların tapılması zamanı çətinliklər yaradır və sistemin səmərəliliyini azaldır.

Yuxarıda qeyd olunanları nəzərə alaraq müəyyən tətbiq sahəsi çərçivəsində məzmunlu axtarışı həyata keçirə bilən intellektual axtarış sistemi modeli yaratmaq təklif olunur. Sistem müxtəlif növlü infomasiya mənbələri ilə əlaqədə olaraq onların inteqrasiyasına şərait yaradır. Məzmunlu axtarışı həyata keçirmək üçün sistemə tətbiq sahəsini təsvir edən biliklər bazası əlavə edilir.

Biliklər bazası məzmunlu axtarış üçün sistemdəki informasiyanın konseptual səviyyəsini özündə saxlayır. Tətbiq sahəsinin müxtəlif anlayışlarının izahını və onlar arasında əlaqəni göstərmək üçün biliklər bazasında ontologiyadan istifadə edilir. Ontologiya məfhumuna bir çox anlayışlar daxildir: obyekt yönümlü proqramlaşdırmada siniflərin iyerarxiyası, konseptual xəritələr (concept maps), semantik şəbəkələr və s. Ontologiya sadə halda tətbiq sahəsinin anlayışlarını ifadə edən terminləri və onlar arasındakı semantik əlaqələri əks etdirən lüğət kimi formalaşdırılır. Ontologiya müxtəlif anlayışların semantik şəbəkəsi kimi izah oluna bilər. Bu da müxtəlif anlayışlar arasında əlaqələr yaratmağa imkan verir.

Ontologiyanın bir neçə tipi mövcuddur:

1. Mövzu yönümlü (Domain-oriented)
2. Tapşırıq yönümlü (Task-oriented)
3. Əsas texniki ontologiya (Basic technical ontology)
4. Ümumi ontologiya (Generic ontologies)

Bunlardan sənayedə ən çox istifadə olunanı mövzu yönümlü və əsas texniki ontologiyadır. Tapşırıq yönümlü ontologiya isə tətbiqi proqramların istifadə etdikləri terminləri saxlayır.

Ontologiyanın qurulması ardıcılığı sadə halda aşağıdakı kimidir:

1. Konseptlərin seçilməsi - tətbiq sahəsinin əsas anlayışları;
2. Konseptlər arasındakı əlaqənin yaradılması - əsas anlayışlar arasındakı əlaqələrin, münasibətlərin müəyyənləşməsi;
3. Qurulmuş ontologiya ilə mövcud olanlarını müqayisə etmək digər bilik sahələri ilə paralellər aparmaq.

Konseptlərin mövzulara görə təsnifatlanmış semantik lüğətini yaradarkən onların hər birinin sinonimlərini, daha ümumi/xüsusi mənalarını, eləcə də mətnlərdə semantik əlaqəsi olan və onların aid olduqları yuxarı səviyyədəki terminləri nəzərə almaq lazımdır.

Biliklərin idarə edilməsi sistemlərində ontologiyanın tətbiq edil-məsinin üstün cəhətlərindən biri də tətbiq sahəsinin öyrənilməsinə sistemli yanaşmadır. Tətbiq sahəsindən, proqramdan asılı olaraq informasiya resursları bir-birindən fərqlənir, ancaq bir qayda olaraq sifarişçi haqqında informasiya, xəbərlər, məktublar, təlimat və s. özündə birləşdirir. Bütün bu resurslar müxtəlif yerlərdə - verilənlər bazasında, biliklər bazasında, kartoteka bloklarında və ya digər yerdə yerləşə bilər. Ümumilikdə, ontologiyada nəzərdə tutulan informasiya resurslarının növləri aşağıda göstərilib;

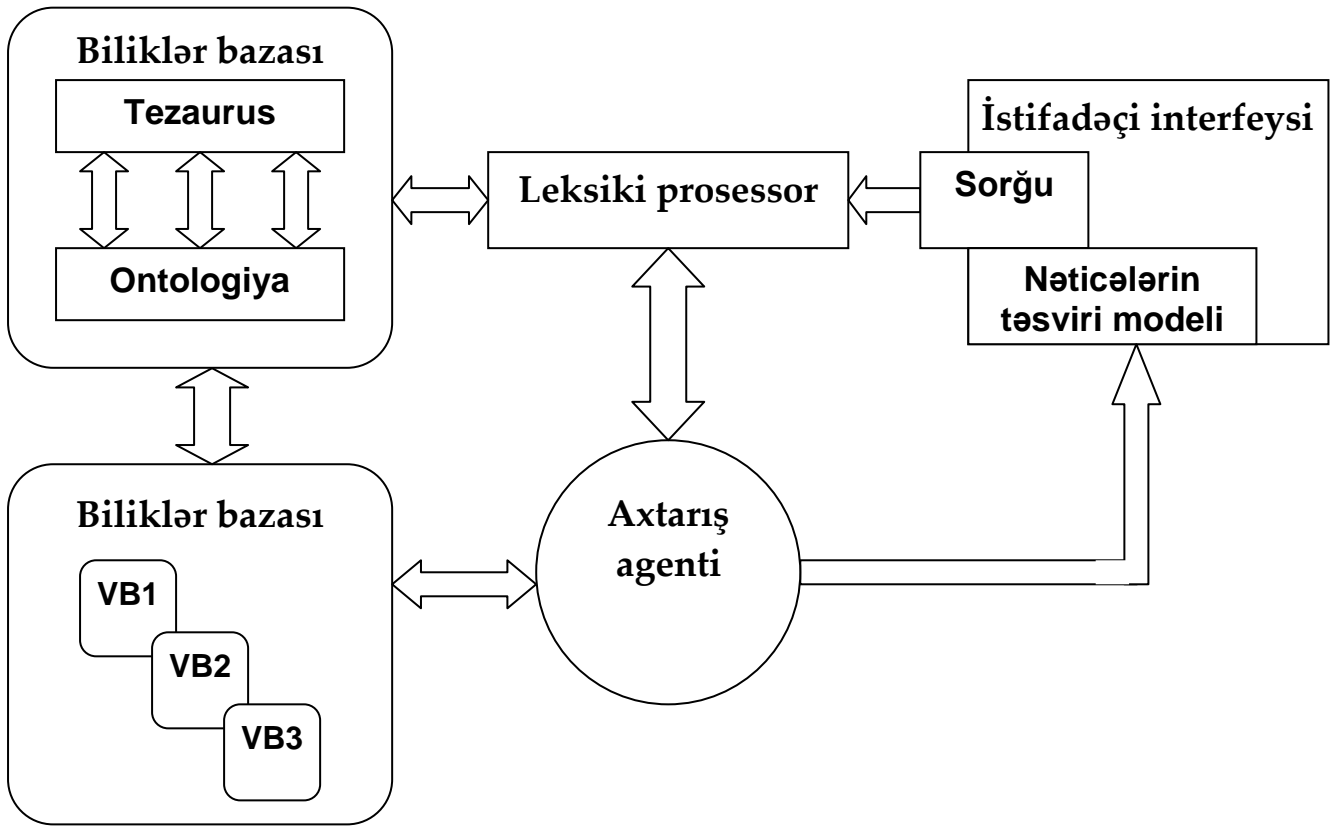
- Cizgilər, xəritələr, eskizlər;
- Normativ-hüquqi informasiya;
- Mövcud elektron arxivlər, verilənlər bazası və əvvəlki proqram təminatının qismən avtomatlaşdırılması;
- Təşkilati struktur və vəzifə bölgüsü;
- Mütəxəssislərin istifadə etdiyi sözlər;
- Texnoloji sənədlər;
- İstehsal texnologiyası və biznes proseslər.

Ontologiyadan istifadə edərkən aşağıdakılar əldə edilir:

1. Sistemlilik - ontologiya tətbiq sahəsini tam olaraq öyrənməyə kömək edir;
2. Yeknəsəqlik – eyni bir formada təqdim olunan material daha yaxşı qəbul edilir;
3. Elmilik - ontologiyanın yaradılması bütün məntiqi əlaqələri tam

yaratmağa imkan verir.

Bir də qeyd etmək lazımdır ki, ontologiya bilikləri idarəetmə sistemlərinin yaradılması zamanı tək cə məqsəd yox, həm də vasitə olmalıdır. Bütövlükdə sorğunun emalı zamanı axtarış sistemi belə işləyir:



Sorğunun emalı zamanı axtarış sistemi

İstifadəçinin sorğusu leksiki prosessorda tətbiq sahəsinin terminlərinə çevrilir. Bundan sonra sorğunun mənasını dəqiqləşdirmək üçün ontologiya əsasında anlayışlar zənciri və semantik əlaqələr şəbəkəsi yaradılır. Belə nəticə alınır ki, sorğu axtarılan anlayışların qiymətləndirilmiş və məhdud fragmentidir. Konkret verilənlər bazasındakı informasiyaya müraciət etmək üçün xüsusi axtarış agenti yaradılır. Bu üsulla formalaşdırılmış sorğu əsasında axtarış aparılır. Xüsusi interfeysdən istifadə edərək axtarışın nəticələrinə baxmaq olar.

Sorğuların təşkili zamanı yaranan problemlər

İnternetdə hər hansı bir informasiyanı axtararkən müxtəlif üsullardan istifadə edirik. Üsulundan asılı olmayaraq biz axtarışı sorğularla həyata keçiririk. Sorğuların düzgün təşkili axtarışın uğurlu nəticələnməsinə təsir edən əsas amildir. Lakin sorğuların təşkili zamanı müxtəlif problemlər çıxır. Bu problemlərin bir hissəsi əksər axtarış serverlərində olan ətraflı axtarış imkanı ilə istifadəçilərlə dialoq rejimində həll olunur. Ancaq tam olaraq deyil. Digər tərəfdən də, müxtəlif axtarış mexanizmlərinin sorğuları qavraması səviyyəsinin müxtəlif olması daha da problemi böyüdür. Hər bir axtarış sisteminin özünün sorğuları emal etmə leksiki xüsusiyyəti var.

Sorğu dillərinin strukturu

Sorğu - kompüterə istədiyimiz tapşırığı vermək üçün xüsusi qaydalar çərçivəsində olan söz və ya söz birləşməsindən ibarət əmrdir. Adətən kobud axtarış

aparmaq üçün açar sözləri aralıqla, vergülsüz yazmaq kifayət edir. Bu zaman axtarış hər sözə görə aparılır. "+" işarəsi sənəddə olması vacib olan sözlərin qarşısında qoyulur. Məsələn: "+azerbaijan+footbail" yazdıqda yalnız "azerbaijan football" kəlmələri olan sənədləri tapacaq. "-" işarəsindən isə əksinə, sənəddə olması vacib olmayan sözlərin qarşısında istifadə edilir. "|" işarəsi sinonim sözləri tapmaq üçündür. Məsələn, "fotoqrafiya" sözü əvəzində "fotoqrafiya|foto|foto şəkil" yazmaq olar. Sorğu dillərinin ümumi strukturu aşağıdakı kimidir:

Məntiqi operatorlar; Qruplaşdırma operatorları; Sözüün hissəsinin əvəzi operatorları; Morfologiya.

Sorğuların yaradılması üsulları

Yuxarıda qeyd olunduğu kimi, müəyyən tətbiq sahəsi çərçivəsində məzmunlu axtarışı həyata keçirə bilən intellektual axtarış sistemi modeli yaratmaq təklif olunur. Sistem müxtəlif növlü informasiya mənbələri ilə əlaqədə olaraq onların integrasiyasına şərait yaradır. Məzmunlu axtarışı həyata keçirmək üçün sistemə tətbiq sahəsini təsvir edən biliklər bazası əlavə edilir.

Biliklər bazası məzmunlu axtarış üçün sistemdəki informasiyanın konseptual səviyyəsini özündə saxlayır. Tətbiq sahəsinin müxtəlif anlayışlarının izahını və onlar arasında əlaqəni göstərmək üçün biliklər bazasında ontologiyadan istifadə edilir.

İnternetdə informasiya axtarışının xüsusiyyətləri.

İstənilən tip informasiyanın toplanması, saxlanması və ötürülməsi üçün İnternet şəbəkəsinin informasiya xidmətləri çox cəlbedici və əlverişlidir. Lakin informasiya resurslarının sayının ildən-ilə çoxalması, informasiya massivlərinin həcmnin və verilənlər bazalarının ölçülərinin böyüməsi nəticəsində tələb olunan informasiyanın axtarılıb tapılması get-gedə mürəkkəbləşir və problem xarakteri alır.

Lazımi informasiyanın əldə edilməsi üçün istifadəçilər hər dəfə çoxlu vaxt, şəbəkə və material resursları sərf etməklə müxtəlif informasiya serverləri ilə əlaqə yaratmalı, bir informasiya mənbəyindən digərinə keçməli və bəzi halda bütün şəbəkəni gözdən keçirməlidirlər. Bu məqsədlə istifadə edilən instrumental vasitələr lazımi informasiyanın axtarışına və seçilməsinə xeyli kömək edə bilərlər.

İnternet-in informasiya fəzası öz tətbiqinə görə paylanmış informasiya sistemidir. Lakin bütün resurslar avtonom VB-də, lokal və ya korporativ VB-də yerləşən və VB-nin və ya şəbəkənin administratorunun nəzarəti altında olan ənənəvi İS-dən fərqli olaraq, İnternet-in informasiya fəzası şəbəkənin üstünlükləri, çatışmazlıqlar və xidmətləri ilə bağlı olan aşağıdakı xüsusiyyətlərə malikdir:

1. ***İnformasiya fəzasının ölçüsü.*** İnternet-in minlərlə qovşağında yerləşdirilmiş informasiyanın həcmi çox böyükdür. Odur ki, axtarış serverləri informasiya fəzasının bütün resurslarını əhatə edə bilməzlər. İnformasiya resursları çox vaxt paylanmış şəkildə yerləşdirilir, onların bir hissəsi bir serverdə, digər hissələri isə başqa serverlərdə saxlanılır.

2. ***Nizamsızlıq və systemsizlik.*** İnternet-da informasiya resurslar nizamsız yerləşdirilir. Onların yaradılmasında, toplanmasında və saxlanılmasında qayda-qanun yoxdur. Odur ki, informasiya əsasən parçalanmış formada olur və bütün dünya üzrə səpələnmiş müxtəlif şəbəkə qovşaqlarında yerləşir. Yeni informasiya resurslarının yaradılması və mövcud resursların dəyişdirilməsi asan olduğundan, istənilən istifadəçi öz saytını və ya səhifəsini yarada və orada istənilən informasiyanı yerləşdirə bilər. Bu baxımdan İnternet-də informasiyanın yerləşdirilməsi və

paylanması təsadüfə xarakter daşıyır. Ümumiyyətlə, şəbəkə qovşağının informasiya təminatı informasiya resurslarının, səhifələrin və saytların sahibləri tərəfindən təşkil edilir, odur ki, onları sistemləşdirərək mümkün olmur.

3. Natamamlıq, izafilik və ziddiyyətlilik. İnternet-də informasiya resursları avtonom, bir-birindən asılı olmadan, müxtəlif vaxtlarda və yerlərdə yaradıldığından, informasiyanın natamamlığı, bəzi hallarda isə əksinə, izafiliyi və ziddiyyətliliyi özünü göstərir. Bütün bunları nəzərə almaq, aradan qaldırmaq mümkün deyil. Bu isə bəzən qeyri-peşəkar istifadəçilərə İnternet-ə inamsızlıq yaranmasına səbəb olur.

4. Dillərin və kodlaşdırmanın müxtəlifliyi. İnternet-in informasiya resursları müxtəlif adamlar, müxtəlif sahələrdə və ölkələrdə, müxtəlif kompüterlərdə və sistemlərdə və müxtəlif məqsədlər üçün yaradıldığından, dil və kodlaşdırmanın müxtəlifliyi (ələxüsus milli resursların) əlavə problemlər yaradılır.

5. Terminologiyanın müxtəlifliyi. Müxtəlif sənədlərdə müxtəlif müəlliflər tərəfindən istifadə edilən terminologiya sənədin növündən və xarakterindən, təyinatından, yaranma mənbəyindən, tətbiq edilən elmi-metodik yanaşmadan, müəllifin fərdi bacarığından, biliyindən və təcrübəsindən asılı olaraq dəyişir və müxtəlif olur. Bütün bunlar İnternet-də informasiya axtarışını və mövzu sahəsinə görə avtomatik təsnifat aparılmasını çətinləşdirir və əlavə vasitələrin (çoxdilli arayış lüğətlərinin, tezaurusların, assosiativ sözlər lüğətlərinin və s.) tətbiqini tələb edir.

6. İnformasiyanın əhəmiyyətliyi və həyat dövrü. Bir çox hallarda İnternet-də müəllifin özündən başqa digər istifadəçilər üçün əhəmiyyətli olmayan sənədlər nəşr edilir. Bəzən isə yeni nəşr olunan informasiya istifadəçilər üçün müəyyən əhəmiyyət kəsb etsə də, vaxt keçdikcə o köhnəlir və əhəmiyyətini itirir (baxmayaraq ki, o uzun müddət İnternet-də saxlanır). Elə hallar da olur ki, informasiya İnternet-ə çıxarılan andan maraqsız və köhnəlmiş olur. Bəzi hallarda web-səhifələr, saytlar və serverlər ayrı-ayrı adamlar və təşkilatlar tərəfindən yaradılır, lakin sonradan onlara xidmət edilmir, yəni informasiya yeniləşdirilmir və modifikasiya olunmur. Yeni onlar bir növ yiyəsiz «informasiya zibilinə» çevrilir. Bu isə informasiya-axtarış vaxtına və keyfiyyətinə mənfi təsir edir.

7. Hiperistinadlardan istifadə edilməsinin xüsusiyyətləri. Web və digər texnologiyalar sayəsində İnternet-də yerləşdirilən informasiya resurslarının həcmnin sürətlə artması İnternet-i nəhəng informasiya anbarına çevirmişdir. Bu anbarda lazımi istiqamətdə hərəkət etmək və tələb olunan sənədləri axtarıb tapmaq ciddi problemlər yaradır.

Hipermətnlərə və hiperistinadlara əsaslanan texnologiya həmin informasiya anbarında lazımi informasiyanın tapılmasında istifadəçiyə xeyli kömək edir. Lakin hiperistinadlar bir çox hallarda istifadəçini lazımi yerə istiqamətləndirmirlər. İstifadəçi bəzən onun üçün maraqsız səhifələrə gəlib çıxır və ya gərəksiz sənədlər arasında hərəkət etməklə vaxt itirir.

Bunun əsas səbəbi ondan ibarətdir ki, hiperistinadlar sənədləri hazırlayan müəlliflərin və ya operatorların subyektiv fikirlərinə görə yaradılır. Həmin istinadlar bəzən heç kimə lazım olmayan, qiymətsiz, köhnəlmiş, bəzən isə silinmiş və ya dəyişdirilmiş sənədlərə istiqamətlənilirlər.

Bununla yanaşı, digər sahələrdən və sistemlərdən fərqli olaraq, web xidməti düzgün olmayan istiqamətlərdə keçiddən addım-addım geri qayıtmaq imkanını nəzərə alır, yəni lazımi yerə qədər geri qayıdıb, istiqaməti dəyişmək olar.

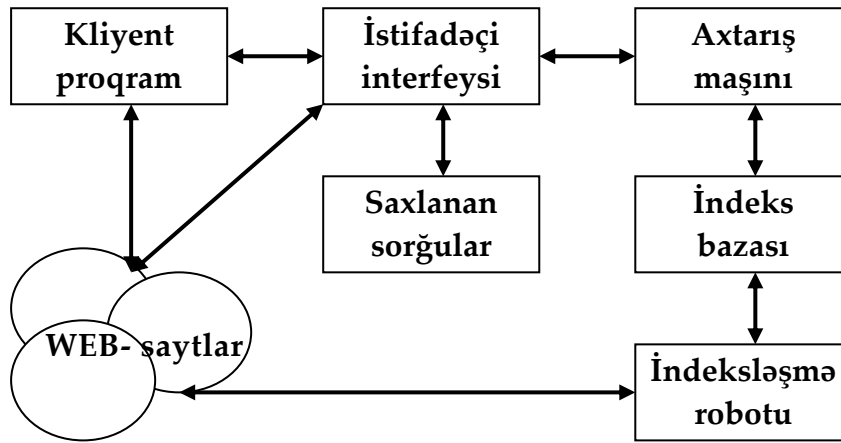
İnternet-də informasiya-axtarış sisteminin ümumi funksional strukturu

Yuxarıda qeyd etdiyimiz kimi, İnternet-də informasiya müxtəlif qovşaqlarda yerləşən saytlarda və səhifələrdə saxlanır və əksər hallarda istifadəçi ona lazım olan informasiyanın saxlandığı serverin və saytın ünvanını bilmir. İstifadəçilərin İnternet-də lazımı informasiyanın asan və tez axtarılıb tapmaları üçün hazırda çoxlu sayda axtarış sistemləri mövcuddur.

Mövcud axtarış sistemlərinin hamısında sənədlərin axtarışı onların məzmununu səthi xarakterizə edən açar sözlərlə aparılır. Burada xüsusi lüğətlərdən, tezauruslardan demək olar ki, istifadə edilmir. Sənədlərin məzmununu təyin edən açar sözlər xüsusi proqram (robot) vasitəsilə seçilib, sənədlərin axtarış sürətləri tərtib edilir. Sorğuda açar sözlərlə ifadə edilir.

İnternet-də informasiya axtarışı üçün mövcud olan sistemlərin müxtəlif komponentlərinin təşkili prinsipləri eynidir, onlar bir-birindən yalnız tətbiq olunan proqram həlləri ilə fərqlənirlər.

İnternet-in informasiya-axtarış sisteminin tipik funksional strukturu aşağıdakı şəkildə verilmişdir.



İNTERNET-də informasiya-axtarış sisteminin tipik funksional strukturu

Kliyənt proqramı - interfeysin pəncərəsini ekrana çıxarmaq, lazımı İAS-ı seçib, ona bağlanmaq və konkret informasiya resursuna baxmaq üçün istifadə olunur. Kliyənt proqram rolunu web-brauzerlər (İnternet Explorer, Netscape Communicator və s.) yerinə yetirirlər. Kliyənt proqramı web, Gopher sənədlərinə, FTP - fayllarına, elektron poçt göndərişlərinin siyahılarına və Usenet xəbərlər qruplarına baxışı təmin edir. Bütün bu informasiya resursları İAS-ın axtarış obyektləri ola bilərlər.

İstifadəçi interfeysi - istifadəçinin İAS-la ünsiyyətini-sorğunun tərtib edilməsini və sorğu nəticələrinə baxışı təmin edir. Sorğunun nəticələri adətən sorğuya cavab verən sənədlərin müəyyən fraqmentləri ilə birlikdə URL ünvanları (linklər) siyahısından ibarət olur.

Axtarış maşını - istifadəçinin sorğusunu formal şəkllə çevirən, həmin sorğuya görə axtarış aparın və alınmış nəticələri interfeysə ötürən proqramdır. Açır sözlərə görə axtarış «İndeksləmə robotu» proqramı tərəfindən yaradılmış indeks bazasında aparılır.

İndeks bazası - informasiya resurslarının axtarış sürətlərini (indekslərini) saxlayır. Burada hər bir sənədin axtarış sürəti (indeksi) ilə yanaşı onun URL ünvanı da saxlanır.

İndeksləmə robotu (spider) - İnternet-in informasiya resurslarına ardıcıl baxmaqla yeni yaranan sənədləri təyin edən, onları indeksləyib axtarış sürətlərini hazırlayan proqramdır. Bu proqram İnternet-də informasiya resurslarının vəziyyətini əks etdirən informasiyanın əsas mənbəyi olaraq indeks bazasının aktual vəziyyətini təmin edir. Bu proqrama başqa sözlə, web şəbəkəsinin skaneri və ya proqram «hörümçəyi» (crawler) deyilir. Hər bir İAS yalnız ona məxsus üsullarla informasiyanı təhlil edən xüsusi indeksləmə robotuna malik ola bilər. Web-sayt indeksləmə robotu tərəfindən indeksləşdirildəndən sonra, o, proqram «hörümçəyinin» ona «baş çəkməsi» haqda xüsusi əlamətlə qeyd olunur. Proqram «hörümçəyinin» şəbəkəni sonradan «gəzməsi» zamanı həmin sayta baxılır.

Saxlanan sorğular. Sorğular istifadəçinin şəxsi verilənlər bazasında saxlanıla bilər. Hər bir sorğunun saxlanması üçün müəyyən vaxt tələb olunur. Odur ki, yaxşı nəticəsi olan sorğuların həmin bazada saxlanması və sonradan onların hazır şəkildə götürülüb emal edilməsi vacib əhəmiyyət kəsb edir.

Web-saytlar bütün İnternet-i əhatə edir, daha doğrusu, onlar «İndeksləmə robotu» tərəfindən baxılması təmin olunan bütün informasiya resurslarıdır.

İnternet-in informasiya fəzasında axtarışın təşkili

Web-səhifələrinin sayı durmadan artır və İnternet-də yerləşdirilən informasiyanın həcmi hər yarım ildə təxminən iki dəfə çoxalır. Lakin informasiya-axtarış sistemlərinin imkanları tələb olunan səviyyədə geri qalır və lazımi tədbirlər görülmədən bu cür nəhəng informasiya fəzasında axtarışı lazımi sürətlə və keyfiyyətlə aparmaq olmaz. Bunun əsas səbəblərindən biri İnternet-də informasiya resurslarının nizamsız, sistemləşdirilməmiş şəkildə saxlanmasıdır, ona görə ki:

- informasiya resursları ixtiyari formada və bir-birindən asılı olmadan yaradılır;

- hər bir informasiya resursu ən yaxşı halda onu xarakterizə edən açar sözləri (hitlər) və tezlik göstəriciləri ilə təsvir edilir;

- informasiya resursları arasında istinadlar (linklər) intuitiv təşkil edilir, odur ki, yaxşı halda yalnız bir-birinə yaxın qonşu olan (2-3 keçidə qədər) resurslar arasında məntiqi əlaqə olur.

İnformasiya resurslarının semantik bağlılığı və onların mövzuya görə yaxınlığı praktik olaraq nəzərə alınmır. Odur ki, informasiya axtarışını istiqamətləndirmək mümkün olmur. Bu səbəbdən də İnternet-də informasiya-axtarış sistemlərinin səmərəliliyi 30%-dən çox olmur.

Böyük sistemlərin, o cümlədən, İnternet-in informasiya fəzasının çox qarışıq olmasına baxmayaraq, axtarışın səmərəli təşkil edilməsi məqsədilə onu sistemləşdirmək olar. Bunu informasiya fəzasını semantik baxımdan zəif əlaqəli axtarış zonalarına bölməklə əldə etmək olar. Həmin zonalar ayrı-ayrı axtarış sistemlərinin əhatə dairələrini təşkil edirlər. Bir zonaya daxil olan informasiya resursları məzmunlarına görə bir-birinə yaxın olur. İnformasiya resurslarının zonalara bölünməsi informasiya axtarışını qismən də olsa istiqamətli aparmağa və istinadların (linklərin) məzmunlu olmasına imkan yaradır.

İnternet-in informasiya fəzasının çoxsəviyyəli iyerarxik strukturla zonalara bölünməsi praktik baxımdan daha səmərəli sayılır. Hazırda üçsəviyyəli iyerarxik strukturdan daha çox istifadə edilir. İyerarxiyanın 1-ci (aşağı) səviyyəsində məzmununa və ya coğrafi əlamətlərinə görə müəyyənləşdirilmiş ayrı-ayrı axtarış serverlərinə və ya server qruplarına uyğun gələn *lokal Web-sahələr* təyin edilir.

İyerarxiyanın 2-ci səviyyəsində çoxlu qrup və ya birləşmiş Web-saytlardan ibarət olan ərazi informasiya resursları təyin edilir. Nəhayət 3-cü səviyyədə həm mövzuya, həm də əraziyə görə ayrılmış informasiya resurslarını özündə birləşdirən global informasiya resursları təyin edilir.

İnformasiya fəzasının üçsəviyyəli iyerarxik strukturla əks etdirilməsi, yeni informasiya resurslarının axtarış zonalarına ayrılması üçün formal modellər təklif olunmuşdur. Lakin istənilən halda informasiya fəzasının mövzuya görə zonalara ayrılması üçün bir-biri ilə bu və ya digər dərəcədə məna və ya məntiqi əlaqələrlə yaxınlıqları olan informasiya resurslarının müəyyənləşdirilməsi tələb olunur. Bu isə asan məsələ deyil. Bunun mümkün yollarından biri bütün informasiya fəzasında ümumiləşdirilmiş mövzulara görə ilkin axtarış aparmaq və alınmış nəticələrə görə axtarış zonalarını formalaşdırmaqdır. Ərazi mənsubiyyətlərinə görə informasiya resurslarını zonalara ayırmaq elə də çətin məsələ deyil. Lakin bu halda da informasiya resurslarının ilkin emal olunması və ya onlarda ərazi mənsubiyyətini göstərən qeydlərin aparılması lazım gəlir

Müasir axtarış sistemlərinin indeksləyici robotları (hörümçəkləri) İnternet fəzasını bütövlükdə «gəzirlər», yeni həm ərazi, həm də mövzuya görə paylanmış axtarış zonalarını (onların Web-serverlərini) əhatə edirlər.

Sual 3. İnternetdə informasiya axtarış sistemləri.

Yuxanda qeyd olunduğu kimi, nə «Gopher» iyerarxik modeli, nə Web hipermetn modeli ümumi həcmi terabaytlarla ölçülən milyonlarla müxtəlif tipli sənədlərdən ibarət olan İnternet-in nəhəng informasiya anbarında informasiya-axtarış problemini həll etmirlər. Hazırda bu problemin həlli üçün yeganə yol açar sözlərlə sürətli informasiya axtarışını yerinə yetirən informasiya-axtarış sistemlərindən ibarətdir.

«Gopher» informasiya-axtarış sistemindən istifadə edildikdə, tələb olunan informasiyanı tapana qədər kataloqlar ağacında uzun-uzadı axtarış aparmaq lazım gəlir. Həmin kataloqlar xüsusi xidmət tərəfindən daima yeniləşdirilməli və dəstəklənməlidirlər. Onların mövzuya görə bölüşdürülməsi istifadəçilərin informasiya tələblərinə uyğun olmalıdır.

İnternet-in pərakəndəliyi, istifadəçilərin maraqlarının müxtəlifliyi və saytın çoxluğu səbəbindən bəzən istifadəçini maraqlandıran mövzuya aid kataloq olmur və istifadəçi lazımı sənədlər tapa bilmir. Odur ki, Gopher-serverlər fəzasında (GopherSpace) axtarış aparmaq üçün «Veronica» adlı informasiya-axtarış proqramı yaradılmışdır.

Hadisələrin analogi inkişafı www-də baş vermişdir. Hələ 1988-ci ildə Frank Halaz böyük hipemətn şəbəkələrində informasiya axtarışının təşkilini bu tip sistemlərin gələcək nəsli üçün başlıca məsələ adlandırmışdır. Bu sahədə real nəticələr 1992-ci ildən başlayaraq özünü göstərdi. 1994-cü ildə www-yə həsr olunmuş II konfransda Kolorado universitetində Web üçün yaradılmış «www Worm» adlı informasiya-axtarış sistemi haqqında məruzə edildi və həmin sistem ən yaxşı naviqasiya vasitəsi kimi mükafat aldı. Həmin konfransda daha iki sistem nümayiş etdirildi: Microsoft kompaniyasının yaratdığı «Lycos» və Amerika On-Line kompaniyasının məhsulu olan «WebCrawler».

Sonrakı illərdə web üçün çoxlu sayda informasiya-axtarış sistemləri yaradılmışdır. Həmin sistemlər üzrə İnternet-də xüsusi kataloqlar təşkil edilmişdir. Məsələn: Search Kit (www.alf.ru/scarch), BUKI (www.rinet.ru/buki) və s. Həmin kataloqlarda axtarış sistemlərinin adları, URL ünvanları və onların müqayisəli xarakteristikaları ətraflı əks olunur. Web üçün axtarış sistemləri arasında daha yaxşı axtarış imkanlarına malik olan və ona görə də daha geniş tətbiq edilənləri aşağıdakılardır:

- xarici axtarış sistemləri; Altavista, Deja, Fast, Directhit, Snap, Google, Northernlight, Oingo, Opentext, infoseek, Wais, Yahoo;
- rus axtarış sistemləri: Aport, Rambler, Yandex.

Bu sistemlərdən bəzilərini qısaca nəzərdən keçirək.

Altavista (www.altavista.com) - ən böyük axtarış portallarından biri olub, təqdim etdiyi servislərin sayına görə axtarış sistemləri arasında liderlik edir, 30-a qədər dildə (rus və türk dilləri də daxil olmaqla) informasiya axtarışı apara və tapılan sənədləri lazımı dilə çevirə bilir. Bu sistem yazıldığı dildən asılı olmayaraq bütün WEB **səhifələrini** indeksləyir. Onun indeks bazasında 500 000 000-dən çox səhifənin indeksi toplanıb (2005-ci ilə qədər).

Altavista-nın sorğu dili ən güclü dillərdən biri hesab olunur. Bu-rada sorğunun genişləndirilməsi, yəni mürəkkəb sorğu formalaşdırmaq imkanı var. Açar sözləri «and», «or», «not» operatorları ilə əlaqələndirmək, ifadələrə görə axtarış aparmaq, beşə qədər istənilən hərfi əvəz edən «*» metasimvolundan istifadə etmək mümkündür. Bundan əlavə, məzmunlu axtarışı reallaşdırmaq üçün terminlərin bir-biri ilə yanaşı yerləşməsinə tələb edən «NEAR» (yanaşı) operatoru da mövcuddur.

İfadələrə görə axtarış aparmaq üçün həmin ifadələri özündə cəmləyən kifayət qədər böyük lüğətə malikdir.

Bütün bunlarla yanaşı, axtarış apararkən sorğuda açar sözün sənəddə rast gəldiyi sahənin adını - hiperistinad (link), applet, hostların adları, şəkillərin adları, mətn, başlıq, URL - də vermək olar. Xəbərlər qruplarında axtarış aparmaq mümkündür. Verilən dildə sənədlərin axtarışını aparmaq olar. Lakin bu halda digər dillərdəki səhifələrə baxmaq mümkün olmur. Tapılan sənədlərin ingilis dilindən fransız, alman, italyan, ispan və portuqal dillərinə və əksinə, həmin dillərdən ingilis dilinə avtomatik tərcüməsi də nəzərə alınır.

Google (www.google.com) - digər sistemlərdən fərqli axtarış alqoritmindən istifadə edir, çox sadə interfeysə və yüksək relevantlıq dərəcəsi ilə ölçülən yaxşı axtarış nəticələrinə malikdir. Axtarış zamanı sorğunun axtarış sürətini sənədin axtarış sürətinə (İndeksinə) tam daxil olması ilə yanaşı sənədə digər serverlərdən edilən istinadların sayı da nəzərə alınır. İstinadların sayı çox olan sənədlərə üstünlük verilir və onlar axtarış nəticələrinin siyahısının lap əvvəlində təqdim edilir.

Google sisteminin maraqlı xüsusiyyətlərindən biri də onun interfeysində ənənəvi axtarış mexanizmini işə salan «Google Search» düyməsi ilə yanaşı, sorğuya maksimal cavab verən sayta müraciət etmək üçün «I'm Feeling Lucky» düyməsinin də nəzərə alınmasıdır. Google-də müxtəlif dillərdə, o cümlədən, rus, azərbaycan dillərində axtarış aparmaq imkanı var. Xəbərlər qruplarında da axtarış aparmaq mümkündür.

Bütün bu deyilənlər Google sistemini məşhurlaşdırmış və son 3 ildə axtarış sistemləri arasında ən məşhur etmişdir. Hazırda Google sistemində 3 milyarda qədər indekslənməmiş WEB səhifə əhatə edilmişdir.

Directhit (www.directhit.com) -- axtarış sistemləri ailəsində həm sadə, həm də güclü sistem hesab olunur. Onun sadəliyi ənənəvi axtarış sistemlərində olduğu kimi, açar sözlərlə axtarışın aparılması, sadə və aydın interfeysə malik olması ilə təyin olunur. Sorğudakı açar sözlərə uyğun gələn sənədlərin içərisində daha çox istinad edilən və daha çox baxılan (yəni baxılma müddəti daha çox olan) sənədlərə üstünlük verilir və onlar çıxış siyahısının əvvəlində yerləşdirilir.

Sorğudakı sözlərə, istinadların sayına və baxılmaların çoxluğuna görə seçilmiş sənədlərin siyahısı ilə yanaşı, sorğuya yaxın mövzular (sözlər) də ekrana çıxarılır. Həmin sözlər (Related Searches) sorğudakı sözlərə «sinonimlik», «assosiativlik» və «sınıf-altsinif» («soy-növ» və «tam-hissə») əlaqələrinə görə müəyyənləşdirilir. Bütün bunlarla bərabər, çıxışda reytingi yüksək olan sənədlərin məşhurluğu haqqında əyani formada məlumat verilir.

Snap (www.snap.com) ilkin axtarış üçün nəzərdə tutulub və bir sıra cəhətlərinə görə Directhit sisteminə oxşayır. Burada da saytların populyarlığı və oxşar mövzular istifadəçilərin rəyləri ilə (istinadların sayı və baxılma müddəti) müəyyənləşdirilir. Bəzi xüsusiyyətlərinə görə Snap sistemi Directhit sistemindən müsbət mənada fərqlənir. Məsələn, sorğuya cavab kimi verilən saytların və oxşar mövzuların (Related Searches) siyahıları ilə yanaşı, oxşar kateqoriyaların (Related Categories) siyahısı da ekrana çıxarılır. Həmin siyahıda SNAP-ın tematik kataloqunda tapılan saytların rast gəldiyi bölmələr göstərilir. SNAP-ın kataloqu xüsusi redaktorlar (insanlar) tərəfindən hazırlanır. Onlar ən məşhur saytları seçib, onları təsvir edir və kataloqun uyğun bölmələrinə yerləşdirirlər. Kataloqun bölmələri həmçinin tapılan sayt haqqında qısa informasiyada da

qeyd olunur (əgər sayt tematik kataloqa daxil edilibsə). Bu cür saytlar «TopWeb Sites» (ən yaxşı Web saytlar) kateqoriyasında təsvir edirlər.

Beləliklə Snap ilkin axtarış üçün kifayət qədər universal sistemdir. İlkin axtarış zamanı sorğuda bir söz göstərməklə, onunla bağlı olan məşhur saytları, tematik kataloqun bölmələrini, mövzuya yaxın (Related Searches) sözləri və ifadələri əldə etmək olar. Öz kataloqunda bir şey tapmadıqda, o, güclü «İnktomi» indeksinə müraciət edir.

Yahoo (www.yahoo.com) İnternet-də istifadə edilən ilk axtarış sistemlərindən biridir. Hazırda Yahoo bir sıra informasiya-axtarış vasitələri istehsalçıları ilə əməkdaşlıq edir və onun müxtəlif serverlərində müxtəlif proqram təminatından istifadə edir. Yahoo-ya tematik kataloq kimi də baxmaq olar, ona görə ki, onun tematik kataloqu ən böyük həcmə malikdir və hazırda kataloqda milyondan çox səhifənin və saytın ünvanları toplanıb.

Bütün tematik kataloqlar kimi, Yahoo da ağacvarı strukturla təşkil edilmişdir. Ən yuxarı səviyyədə əsas rubrikalar (məsələn, «Mədəniyyət və incəsənət», «Xəbərlər və kütləvi informasiya vasitələri», «Biznes və iqtisadiyyat», «Kompüter və İnternet» və s.) təsvir edilir. Rubrikalar bölmələrə, bölmələr altbölmələrə və s. ayrılır və ən aşağı səviyyədə saytların təsviri və onlara istinadlar (linklər) saxlanır. Hər bir saytın təsviri xüsusi redaktor (insan) tərəfindən aparılır. Onların sayı 50-dən artıqdır. Yahoo-nun kataloqu əl üsulu ilə tərtib edilir, odur ki, onun keyfiyyəti yüksəkdir. Hazırda yahoo ən böyük kataloq olmaqla yanaşı ən çox müraciət olunan sistemdir. Hər ay Yahoo-ya 40 milyondan çox insan müraciət edir.

Yahoo sisteminin informasiya-axtarış dili kifayət qədər sadədir. Daxil edilən sözlər boşluqla ayrılır. Onlar arasında AND və OR operatorları yazmaq olar. Çıxışda sənədlərin sorğuya uyğunluq dərəcəsi göstərilir, lakin tapılan sənədlərdə sorğuya uyğun sözlərin altından xətt çəkilir. Bu zaman leksikanın normallaşdırılması və ümumi sözlərin təhlili aparılmır. Çıxış siyahısında sənədlərin relevantlığa görə nizamlanması tapılan sənədlərdə sorğunun sözlərinin sayına görə aparılır.

Oingo (www.oingo.com) sistemi, sözü adi simvollar ardıcılığı kimi qəbul edən bir çox axtarış sistemlərindən fərqli olaraq, sözlərin mənalarını «başa düşür». İlkin axtarışda sorğunun sözlərinə digər sistemlərdə olduğu kimi, simvollar ardıcılığı kimi baxılır. Axtarış nəticələri iki siyahı ilə ekrana çıxarılır: tapılan WEB-saytların siyahısı və kataloqun bölmələrinin siyahısı. Uyğun düyməni basmaqla açılan üçüncü siyahıda isə sorğudakı sözün bütün mənaları əks etdirilir. Məsələn, sorğuda verilmiş «axtarış» sözünün müxtəlif mənalarının siyahısında aşağıdakılar göstərilir: 1) «informasiya axtarışı», 2) «insan axtarışı», 3) «mal axtarışı», 4) «cinayətkarın axtarışı», 5) «ev axtarışı» və s. Bu siyahıdan lazımı variantı seçdikdən sonra «Search Again» düyməsini basmaqla axtarış yenidən təkrarlanır və «axtarış» sözünün seçilən mənasına görə yeni nəticələr alınır.

Əgər seçilmiş mənaya görə axtarış nəticələri istifadəçini qane etmirsə, o, sorğunu dəqiqləşdirən bir neçə sözdən istifadə etməklə, yenidən axtarış apara bilər. Məsələn, əgər istifadəçi «mal axtarışı» variantını seçibsə və o konkret olaraq müəyyən modelə aid avtomobil axtarırsa, sorğunu belə ifadə edə bilər: avtomobil, mercedes, C-model, 2005.

Bu halda da sistem həmin sözlərə görə axtarış nəticələrini iki siyahıda verir və 3-cü siyahıda hər bir sözün mənə variantlarını əks etdirir. Əgər yeni nəticələr

istifadəçini qane etmirsə, o ayrı-ayrı sözlərin məna variantlarını seçib, axtarışı yenidən təkrarlaya bilər.

Oingo-da sorğu dili demək olar ki, yoxdur. Təkcə «+» işarəsindən istifadə etmək olar. Bu halda həmin sözün sənəddə mütləq olması tələb edilir.

Əgər Oingo verilmiş sorğuya görə öz kataloqunda heç bir şey tapa bilmirsə, o, Altavista sisteminin indeksinə müraciət edir.

Aport (www.aport.ru) Rusiyanın «Runet» adlanan 3 axtarış serverindən biridir. Onun indeks bazası o qədər də böyük deyil və operativliyi də yüksək deyil. Lakin APORT sistemi bəzən digər sistemlər tərəfindən tapılmayan sənədləri tapa bilir. *Bu* sistemin digər üstün cəhəti ondan ibarətdir ki, o, sənədin indeksinə görə onun ilkin mətnini bərpa edə bilər (hətta sənəd internet-dən kənarlaşdırılsa da).

Hazırda Aport-un indeks bazasında 2 milyondan artıq sənədin indeksi toplanmışdır. Sistem axtarış üçün daxil edilən sözdəki səhvləri düzəldir və müxtəlif söz formalarına görə axtarış apara bilər. and (&), or (|), not məntiqi operatorları, mötərizəyə alınmış məntiqi qrupları, ifadələrə görə axtarışı dəstəkləyir, sözlər arasındakı məsafənin məhdudluğunu, sözlərin və ifadələrin sayını nəzərə ala bilər. URL-ə görə axtarış apara bilər və sənədin yaranma tarixinin qəbul edilən qiymətini nəzərə ala bilər.

Bütün bunlarla yanaşı, APORT həm sorğunu, həm də sorğuya görə alınan nəticələri ingilis dilindən rus dilinə və əksinə çevirə bilər. Bu sistemdə də sözün sənəddə yerləşdiyi sahənin adına görə axtarış aparmaq imkanı var.

Rambler (www.rambler.ru) - çoxlu sayda axtarış servislərinə, o cümlədən: «Rambler Top 100» reytingi, müxtəlif mövzulara (məsələn, şəbəkə mağazaları, hədiyyələr, iş, hüquq, kompüter və s.) aid kataloqlar, faylların axtarışı, müxtəlif suallar üzrə arayış sistemi və s. malik olan çoxfunksiyalı sistemdir. Rambler-in bir çox serverləri ayrıca axtarış serverləri kimi də istifadə edilir.

Axtarış keyfiyyətinə görə Rambler digər tanınmış sistemlərdən geri qalmır. Burada da and, or, not məntiqi operatorlar, məntiqi qruplar, bir simvolu əvəz edən «?» və bir neçə simvolu əvəz edən «*» metasimvollar dəstəklənir. Axtarış üçün sorğu formasında axtarışın harada (sənəddə, sənədin əvvəlində, başlığında, adında, URL-də) aparılmasını, axtarılan sənədin dilini (rus, ingilis və digər), sözlər arasındakı məsafəni, sənədlərin yaranma tarixlərinin intervalını (*nə* vaxtdan *nə* vaxtadək) göstərmək olar. Sənədləri həm onların relevantlıq dərəcəsinə görə, həm də yaranma tarixlərinə görə nizamlamaq mümkündür.

Yandex (www.yandex.ru) - rus axtarış sistemləri arasında ən məşhur, indeks bazası ən böyük və axtarış imkanları ən yaxşı olan sistem hesab olunur. Onun indeks bazasında təkcə Rusiya saytları deyil, MDB ölkələrinin və digər ölkələrin də saytlarını əhatə edir.

Yandex-in əsas üstünlüyü sorğudakı sözlərin bütün formalarına görə axtarış apara bilməsidir. Hətta lüğətdə olmayan sözlər üçün də onların söz formalarını tərtib edə bilər. Sistem and, or, not məntiqi operatorları, məntiqi qrupları, ifadələrə görə axtarışı dəstəkləyir. Axtarışı açar sözlərin bütün formalarına görə və ya konkret verilmiş formaya görə aparmaq olar. Sözlərin arasındakı məsafəni onların ardıcılığını nəzərə almaqla təyin etmək mümkündür. Sənədlərin başlıqlarına və onlardakı istinadlara (linklərə) görə xüsusi axtarış da aparmaq imkanı var. Bundan əlavə, əvvəlki sorğuda tapılmış sənədlərə oxşar sənədlərin axtarışını da aparmaq olar. Axtarış üçün kataloqun konkret bölməsini seçməklə, axtarış fəzasını daraltmaq mümkündür. Axtarışın nəticələri istifadəçini qane etmirsə, axtarışı Altavista sisteminə davam etdirmək olar. Bu halda Yandex-də emal olunmuş sorğu hazır şəkildə Altavista-ya ötürülür.

Yandex ailəsinə axtarış sistemindən əlavə istifadəçilərə öz web-saytlarını yerləşdirmək üçün pulsuz disk yaddaşı və əlavə elektron poçt qutusu təqdim edən «Narod.Ru» adlı «virtual şəhər» də daxildir.

Metaaxtarış sistemləri. Göründüyü kimi, hər bir axtarış mexanizmi özünə məxsus müsbət və mənfi cəhətlərə malikdir. Kataloqlar saytların axtarışı üçün, axtarış sistemləri isə səhifələrin axtarışı üçün daha münasib hesab olunurlar. Axtarış sistemlərinin sayı da kifayət qədərdir. Onlardan bəziləri yaxşı axtarış aparır, digərləri tapılan sənədləri yaxşı nizamlayır, digərləri rahat interfeysə malikdirlər və s. Odur ki, İnternet istifadəçiləri çox vaxt bir neçə axtarış sistemindən istifadə etməklə, axtardıqları sənədləri tapmağa çalışırlar. Lakin bu zaman istifadəçi əksər halda çox böyük sayda sənədlər alır və onların da çoxu bir-birini təkrarlayır.

Bu problemin həlli üçün *metaaxtarış sistemlərindən* (onlara həmçinin multiaxtarış sistemləri də deyilir) istifadə olunur. Öz-özlüyündə bu sistemlər axtarış aparmırlar. Onlar axtarış üçün sorğunu eyni vaxtda bir neçə axtarış sisteminə və ya kataloqa ünvanlayır, sonra isə alınmış nəticələri birləşdirir və bu zaman təkrarlanmaları aradan qaldırırlar. Metaaxtarış sistemləri az rast gəlen nadir sözlərə görə axtarış üçün daha səmərəli olurlar. Tanınmış metaaxtarış sistemlərinə qısa nəzər yetirək.

Metacrawler (www.metacrawler.com) - İnternet-in qərb seqmentində metaaxtarışın liderlərindən biridir. Axtarış üçün sorğunu aşağıda sadalanan 15-ə qədər ən güclü axtarış sistemlərinə və kataloqlara göndərir: Yahoo, Altavista, Google, Lycos, Directhit, Webcrawler, excite, Findwhat, Goto.com, İnternet Keywords, Kanoodle, Metacatalog, Looksmart, Sprinksbyabout. Bu saytlardan bəziləri o qədər də tanınmayıblar. Lakin axtarışın tamlığı üçün onların da rolu olur. «Geniş axtarış» (*advanced search*) rejiminə keçməklə göstərilən sistemlərdən ən yaxşılarını seçmək olar. Metacrawler hər bir axtarış sisteminin nəticələrindən ilk 30 ünvan götürür, təkrarlanmaları aradan qaldırır, alınmış ümumi siyahını istinad reytinginə görə nizamlayıb istifadəçiyə təqdim edir.

Sorğuda bir neçə açar sözdən istifadə etdikdə sistem sözlərin hamısına (all), istənilən birinə (any) və ya bütövlükdə söz birləşməsinə (phrase) görə axtarış apara bilir.

ixquick (www.ixquick.com) ən yaxşı metaaxtarış sistemlərindən biri hesab olunur. Bu sistemə verilən sorğu 14-ə qədər müxtəlif axtarış sistemlərinə istiqamətləndirilə bilər: Aol, Live Directory, Altavista, Look Smart, Excite, Lycos, Fast Search, Msn, Go To, Snap, Hotbot, Webcrawler, Infoseek, Yahoo. İstifadəçi İxquick-in pəncərəsində bu sistemlərin adlarından sol tərəfdə yerləşdirilmiş işarəni qeyd etməklə, onlardan yalnız məqsədəuyğun olanlarını işə sala bilər.

Sorğuya görə tapılmış sənəd (ünvanı və qısa xarakteristikası) müəyyən sayda ulduz işarələri ilə müşayiət edilir. Ulduzların sayı axtarış nəticələrinin alınmasında payı olan axtarış sistemlərinin sayını göstərir. Pəncərənin aşağı hissəsində axtarış payı olan hər bir sistemin adı və tapdığı sənədlərin sayı verilir.

Metaaxtarış sistemlərindən ilkin axtarış üçün də istifadə etmək olar.

Baxılan metaaxtarış sistemlərindən başqa digər sistemlər, o cümlədən, rus metaaxtarış sistemləri də var. Rusiyanın metaaxtarış sistemlərindən ən çox tanınanlarına Poisk (www.informika.ru/Windows/intern/poisk/main.html) və Search (<http://search.da.ru>) sistemlərini misal göstərmək olar.

Axtarış sistemlərinin kataloqları. Dünya miqyasında axtarış sistemlərinin sayı çoxdur. Onların içərisində ən çox istifadə olunalar (məsələn, Altavista, Google, Yandex) və an az istifadə olunanlar var. Axtarış sisteminin ən çox və ya ən az istifadə olunması onun xarakteristikaları və imkanları ilə təyin olunur. Odur ki, axtarış sistemlərinin xarakteristikaları haqqında müqayisəli məlumatın olması çox vacibdir. Bu cür məlumat xüsusi kataloqlarda verilir. Bu baxımdan Rusiyanın İnternet fəzasında təşkil edilən axtarış sistemlərinin kataloqları diqqətə layiqdir.

Rusiyanın İnternet fəzasında (Runet) reallaşdırılan ən rahat və tam kataloqlardan biri Search KİT (www.alf.ru/search/) adlı kataloqdur. Bu kataloqda 150-dən çox dünya və 50-yə qədər rusdilli axtarış sistemləri, reytinglər və axtarış kataloqları haqqında məlumat və onların ünvanları verilir. Kataloqdan istifadə edərkən orada göstərilmiş hər bir axtarış sistemine kataloqun təqdim etdiyi formada sorğu vermək olar. Hər bir axtarış sisteminin rus dilində işləyə bilməsi imkanı xüsusi işarə ilə qeyd edilir. Kataloq informasiya axtarışının bütün istiqamətlərini əhatə edir. Həmin kataloq vasitəsilə ənənəvi axtarış sistemləri ilə yanaşı, xüsusi axtarış sistemlərindən (faylların axtarışı, MP3 formatında musiqinin axtarışı və s.) də istifadə etmək olar.

Axtarış sistemlərindən əlavə, SEARCH KİT kataloqu İnternet-in digər faydalı resurslarına (elektron kitabxanalarına, saytlara, pulsuz xidmətlərə, tərcüməçilərə və s.) istinadları da özündə saxlayır.

Yalnız rus axtarış sistemləri haqqında informasiyanı əhatə edən BUKİ kataloqunda 100-dən artıq rus axtarış serverləri haqqında məlumat toplanıb.

Portallar. «Portal» sözünün hərfi mənası «darvaza», «binaya giriş» deməkdir. İnternet-də isə portal dedikdə özündə bir neçə (bəzən onlarla) saytı: axtarış sistemlərini, İnternet səhifələrinin kataloqunu, xəbərlər xidmətini, elektron poçt xidmətini, elektron ensiklopediyasını, elektron mağazanı, elektron elanlar lövhəsini və s. birləşdirən güclü informasiya sistemi başa düşülür. Daha böyük portallar isə sadalanan İnternet resurslarının hər birindən bir neçəsini özündə cəmləyir, məsələn, bir neçə elektron ensiklopediyanı, bir neçə müxtəlif kataloqları, müxtəlif mövzulara aid xəbərlər qruplarını və s. Portalın bu cür müxtəlif informasiya resursları həm ümumi interfeysə və istinadlara görə, həm də ümumi axtarış sisteminə görə bir-biri ilə əlaqəli olurlar. Portalların əsas üstünlüyü də buradan irəli gəlir: portalın axtarış sistemində verilən sorğuya cavab kimi portalın bütün resurslarına istinadlar almaq olar.

Portal sistemləri əsasən ayrı-aynı mövzu sahələrinə görə təşkil edirlər. Tematik portallardan əlavə, hər hansı ərazinin, ölkənin və ya şəhərin informasiya resurslarını birləşdirən portallara da rast gəlinir. Portallar ölçülərinə görə də müxtəlif olurlar: *ən* nəhəng portallardan tutmuş, adi istinadlar yığımından ibarət olan kiçik portallara qədər. Portalın adı İnternet axtarış sistemindən əsas fərqi ondan ibarətdir ki, burada lazımi cari informasiyanı: hava haqqında məlumatı, son xəbərləri, valyuta məzənnəsini, telekanalların proqramlarını və s. almağa çox az vaxt sərf olunur.

Yuxarıda baxılan güclü axtarış sistemlərindən bəzilərinə (məsələn, Altavista, Yahoo, Rambler) yerinə yetirdikləri xidmətlərə görə portal kimi baxmaq olar. Hazırda İnternet-də portalların sayı get-gedə artır və güman edilir ki, yaxın gələcəkdə onlar İnternet-in əsas informasiya resurslarını təşkil edəcək.

Portallara misal olaraq rus portallarından «Kirill və Mefodiy» (www.km.ru), «Kulichki» (www.kulichki.com), azəri portallarından «Azeri_info» (www.azeri_info.com), «Science» (www.science.az) və s. göstərmək olar.

NƏTİCƏ

İnformasiya cəmiyyətinin inkişafı vətəndaşların biliklər və məlumat tələbatının ödənilməsi, informasiya və kommunikasiya texnologiyalarının köməkliyi ilə insanların təhsil səviyyəsinin yüksəldilməsi, ümumilikdə ölkənin intellektual potensialının artırılması ilə sıx bağlıdır.

İnformasiya cəmiyyətinin əsasını şəbəkə informasiya texnologiyaları təşkil edir və istənilən ölkədə bu cəmiyyətinin formalaşması əhalinin fəal hissəsinin informasiya və kommunikasiya texnologiyalarından kütləvi şəkildə istifadəsindən asılıdır. Bu texnologiyaların insanlar tərəfindən qəbul edilməsi, onların kompüter biliklərinin artırılması, əhalinin beynəlxalq İnternet şəbəkəsinin informasiya mənbəyi və rabitə vasitəsi kimi xidmət etmək imkanları barədə məlumatlandırılması istiqamətlərində mühüm addımlar atılmalıdır.

Müasir cəmiyyətin maraq və prioritetlərini dolğun və çevik əks etdirən dövlət informasiya siyasəti idarəetmə sisteminin ayrılmaz hissəsi olaraq, ölkənin inkişafın səviyyə və dinamikasını müəyyən edir, onun dünya birliyində layiqli yer tutmasına əsas verir.

Yeni elmi istiqamət kimi XX əsrin son onilliyində formalaşmış kommunikasiya texnologiyaları qısa müddət ərzində cəmiyyətin informasiya resursları ilə təmin olunmasında mühüm rol oynamışdır.

İnternetin informasiya-axtarış sistemləri – nəhəng informasiya fəzasında zəruri məlumatların axtarılıb tapılması üçün nəzərdə tutulmuş intellektual alətdir. İndiyədək informasiya axtarışının modelləşdirilməsi, səmərəliliyinin və intellektuallığının artırılması istiqamətində xeyli elmi-tədqiqat və təcrübi işlərin görülməsinə baxmayaraq, bu sahədə hələ də arzu olunan nəticələr əldə edilməmişdir.

Bütün axtarış xidmətləri istifadəçilərin sorğularına cavab olaraq onların tələbatlarının ödənilməsi üçün daha münasib olan informasiya mənbələrinə istinadların siyahısını təqdim edir. Belə siyahıda lazımsız, istifadəçinin axtarış sorğusuna və tələbatına uyğun olmayan mənbələrə daha çox rast gəlinir.

Axtarış sorğusuna uyğunluq (relevantlıq) dərəcəsinin müəyyən edilməsi zamanı web-saytlarda reallaşdırılan naviqasiya qraflarının (hiperistinadların və hörümçək torunun) quruluşu, web-səhifələrdə aparılan dəyişikliklərin və onların həyat dövrünün (yaşama müddətinin) vaxtında nəzərə alınması böyük əhəmiyyət kəsb edir.

Mühazirə mətnində İnternetin informasiya massivində, yəni web-fəzada informasiya axtarışının təşkili və həyata keçirilməsi mexanizmləri, eləcə də sənədlərin indeksləşdirilməsi və axtarılması üsul və vasitələri, informasiya-axtarış sistemlərinin iş prinsipləri ətraflı şəkildə öz əksini tapmışdır.

Eyni zamanda informasiyanın indeksləşdirilməsi və axtarışı üzrə mövcud alqoritmlər, İnternet üçün nəzərdə tutulmuş informasiya-axtarış sistemlərinin yaranma və inkişaf tarixi, arxitekturası və iş prinsipləri, eləcə də axtarış sistemləri barədə məlumat təqdim edilmişdir.