

Qurbanova Ə.M.

AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
afruz1961@gmail.com

TERMİNOQRAFİK FƏALİYYƏTİN AVTOMATLAŞDIRILMASI PROBLEMLƏRİ VƏ ONLARIN HƏLLİ YOLLARI

Məqalədə “leksikoqrafiya” və “terminoqrafiya” anlayışlarına, onların tarixi, nəzəri və praktiki xüsusiyyətlərinə baxılmışdır. Terminoloji lüğətlərin tərtibatının vahid prinsip və metodlarının işlənilməsinin zəruriliyi əsaslandırılmış, bunun üçün vacib şərtlər təsvir edilmişdir. Leksikoqrafiya və terminoqrafiyanın perspektiv istiqamətləri müəyyənləşdirilmiş, onlarda kompüterləşmənin üstünlükləri göstərilmişdir. Leksikoqrafik və terminoqrafik işlərin avtomatlaşdırılması prosesləri analiz edilmiş, bu istiqamətdə böyük həcmli terminoloji informasiyanın yığılması və sürətli emalına imkan verən terminoloji verilənlər bazasının yaradılmasının zəruriliyi əsaslandırılmışdır. Milli terminoloji informasiya sistemi çərçivəsində yaradılan terminoloji verilənlər bazasının imkanları göstərilmişdir.

Açar sözlər: *leksikoqrafiya, terminoqrafiya, terminoloji lüğət, leksikoqrafik informatika, terminoqrafik informatika, terminoloji verilənlər bankı.*

Giriş

İnformasiya texnologiyaları hazırkı dövrdə istənilən peşəkar fəaliyyət sahəsi kimi, linqvistikanın da ayrılmaz hissəsinə çevrilib. Belə ki, linqvistik tədqiqatlarda, tərcümədə və dilin öyrənilməsində kompüter və uyğun proqram təminatlarının istifadəsi bu gün böyük əhəmiyyət daşıyır. Linqvistikada informasiya texnologiyalarının istifadəsində birinci addım çap və şifahi formada olan dil materiallarının rəqəmsal formaya keçirilməsidir.

Linqvistikanın informasiya texnologiyalarının istifadəsilə əlaqəli bölmələrini müəyyən edərkən nəzəri və tətbiqi linqvistika arasındakı fərqlərə diqqət etmək lazımdır.

Nəzəri linqvistika dilçiliyin dövlət dilinin qurulması, onun tarixi və qanunauyğunluqlarını öyrənən bölməsidir. Tətbiqi linqvistika dilçiliyin keçən əsrin 20-ci illərindən başlayaraq inkişaf edən bölməsidir, dilin istifadəsilə praktiki məsələlərin həlli metodlarının işlənilməsi ilə əlaqəlidir.

Tətbiqi linqvistikanın bölməsi olan leksikoqrafik informatika kompüter lüğətlərinin yaradılması, linqvistik verilənlər bazası və leksikoqrafik işlərin dəstəklənməsi üçün proqram təminatlarının işlənilməsi məqsədini daşıyır. Ənənəvi və leksikoqrafik informatikanın əsas məsələləri lüğətlərin strukturu və lüğət məqalələrinin yerinin təyin olunması, eyni zamanda müxtəlif növ lüğətlərin tərtibatı prinsiplərinin işlənilməsidir [1].

Leksikoqrafik informatika və tətbiqi linqvistikanın perspektiv istiqamətlərindən biri elektron terminoloji lüğətlər və terminoloji bazalar üzərində işdən ibarətdir.

Xüsusi terminoloji lüğətlərin tərtib edilməsi ilə leksikoqrafiyanın bir bölməsi olan terminoqrafiya məşğul olur. Terminoqrafiya terminşünaslıqla sıx əlaqəlidir, terminşünaslıq dilin xüsusi leksik vahidi haqqında elmdir. Uyğun olaraq, terminoqrafik informatika – elektron terminoloji lüğətlərin tərtib edilməsi haqqında elmdir.

Terminoqrafik informatikanın prinsipləri ilə leksikoqrafik informatikanın prinsipləri ümumilikdə və tam olaraq üst-üstə düşür. Onların fərqi lüğət təsvirinin əsas obyektii ilə əlaqəlidir: leksikoqrafiyada bu adi söz və ya digər dil vahidləridir (söz birləşməsi, cümlə və s.), terminoqrafiyada isə bu termindir.

Termin – elmin metadilinin və insanın praktiki fəaliyyət sahəsinin dəqiq və birqiymətli tərfi olan, uyğun peşəkar sferadan xüsusi bilik tələb edən söz və ya söz birləşməsidir. “İnternet” sözü adi insan üçün ümumişlək sözdür, kompüter şəbəkəsinin mütəxəssisi üçün isə bu söz predmet sahəsinin nəhəng layıdır, uyğun olaraq, mütəxəssis üçün o termindir.

Beləliklə, termin onun terminlər sistemində reallaşan xassələri ilə birlikdə verilir. Terminlər sistemi bütövlükdə uyğun bilik sahəsini əks etdirir, onun hər komponenti (termin) isə bu bilik sahəsinin tərkibi adlandırılır.

Müasir dövrdə leksikoqrafiya nəzəriyyəsi və praktikasının ən çox tələb olunan istiqaməti terminoloji leksikoqrafiya və ya terminoqrafiyadır. Leksikoqrafiya, eləcə də terminoqrafiyada müasir kompüter texnologiyalarından istifadə məsələsi bu gün də çox aktualdır və detallı öyrənilmə tələb edir.

Bu işdə leksikoqrafiya və onun bir qolu olan terminoqrafiya işlərinin avtomatlaşdırılması ilə əlaqəli onun inkişaf perspektivlərinin müəyyənləşdirilməsi məsələsinə baxılır.

Leksikoqrafiya və terminoqrafiya

Lüğətlərin tərtib edilməsi sənəti uzun bir tarixə malikdir. Lakin leksikoqrafiya özünün anlayışlar aparatı və tədqiqat üsulları olan bir elmi fənn kimi XX əsrdə formalaşmışdır. Belə ki, keçən əsrin 70-ci illərindən başlayaraq leksikoqrafiya lüğətlərin hazırlanması “sənəti” olmaqla yanaşı, lüğət işi üzrə tətbiqi sahədən sərbəst elmi-tətbiqi fənnə çevrildi - lüğət işinin nəzəriyyəsi meydana çıxdı. Elmi ədəbiyyatda leksikoqrafiya dilçiliyin bir bölməsi kimi müəyyən olunur, lüğətlərin tərtib edilməsi nəzəriyyəsi və praktikası ilə məşğul olur, ictimai əhəmiyyətli funksiyaları yerinə yetirir: dilin öyrənilməsi; dilin təsviri və normallaşdırılması; dildaxili və dillərarası ünsiyyət; dilin elmi cəhətdən öyrənilməsi [2].

Dünya leksikoqrafiyası tarixində Vladimir Dal (Rusiya), Y.Qrimm, K.Qrimm və Q.Paul (Almaniya), E.Littre və P.Laruss (Fransa), İ.Malorini (İtaliya), V.Vartburq (İsveçrə), H.Vebster (ABŞ) kimi görkəmli leksikoqrafların adları tanınır. Onlar üçün leksikoqrafiya həyati məsələyə çevrilmişdir [3].

Müasir leksikoqrafiya – linqvistik biliyin müstəqil sahəsidir, ənənəvi olaraq iki hissəyə ayrılır [4]:

1. Nəzəri leksikoqrafiya, lüğətlərin tərtib edilməsi metodlarını və lüğətlərin konstruktiv elementlərini öyrənir;
2. Praktiki leksikoqrafiya, lüğətlərin tiplərini tədqiq edir və lüğətlərin müəyyən spesifik formaları (tarixi, frazeoloji, etimoloji və s.) üzrə klassifikasiyasını aparır.

Keçən əsrin 80-ci illərinin əvvəllərindən başlayaraq leksikoqrafiyanın bir bölməsi olan terminoqrafiya formalaşmağa başladı. Qeyd etmək lazımdır ki, leksikoqrafiya və terminoqrafiyanın problemləri ilə əlaqəli elmi ədəbiyyatda alimlərin işlərində “leksikoqrafiya” və “terminoqrafiya” terminlərinin birqiymətli olmayan interpretasiyası verilir.

Görkəmli ispan leksikoqrafı Julio Kasares özünün məşhur “Müasir leksikoqrafiyaya giriş” kitabında qeyd edir ki, leksikoqrafiya – lüğətlərin tərtib edilməsi üsulları və sənətidir [5].

Terminoqrafiya terminoloji lüğətlərin tarixi, nəzəriyyəsi və praktikası haqqında integrativ elmi-tətbiqi fəndir, müxtəlif elmi və praktiki məsələlərin həlli məqsədilə yaradılan terminoloji lüğətlərin optimal layihələndirilməsi, tərtibatı və istifadəsi üçün metodologiya, metod, metodika və texnologiyaları formalaşdırır [3].

Terminoloji lüğətlər müxtəlif fəaliyyət sahələrində ölkədaxili və beynəlxalq səviyyələrdə geniş ünsiyyət, həmçinin informasiya xidməti və sistemlərinin yaradılması və təmini üçün zəruridir [6].

A.S.Gerdin fikrincə, “terminoqrafiya” və “leksikoqrafiya” terminləri eyni bir anlayışı ifadə edir. Elmi-texniki leksikoqrafiyanın obyektı xüsusi (terminoloji) lüğətlərin yaradılması nəzəriyyəsi və praktikasından ibarətdir [7].

Terminoloji lüğətlərin yaradılması praktikası, başqa sözlə, terminoloji leksikoqrafiya, terminoqrafiya keçmiş Sovetlər Birliyində və Qərbdə eyni zamanda yaranıb. “Terminoqrafiya” termini “terminoloji leksikoqrafiya” söz birləşməsindən əmələ gəlmişdir. Terminoqrafiya tətbiqi terminşünaslığın prioritet istiqamətlərindən biridir [8].

Terminoloji (və ya elmi-texniki) leksikoqrafiya və ya terminoqrafiya xüsusi leksikanın lüğətlərinin layihələndirilməsi, tərtib edilməsi və istifadəsi haqqında elmdir [9].

K.Y.Averbuxun fikrincə, ənənəvi leksikoqrafiyanın içindən doğan terminoloji lüğətlərin yaradılması üzrə fəaliyyət hələ də onun əsas qaydalarını istifadə etməyə davam edir, lakin terminoloji lüğəti tərtib edənlərin istifadə etdiyi material, həll olunmalı məsələlər, onların həlli metodları yeni elmi istiqamətin – terminoqrafiyanın yaranmasına gətirib çıxartdı [10].

[11]-də leksikoqrafiyanın əsas məqsəd və vəzifəsinin hər hansı bir dilin lüğət tərkibinə daxil olan sözlərin, eyni zamanda frazeologiya vahidlərinin toplanıb sistemə salınmasından, onların törəməsi, mənsubiyyəti və mənalarının izah edilməsindən ibarət olduğu verilir.

Beləliklə, leksikoqrafiyanın predmeti lüğətlərin qurulması prinsipləri və metodlarının işlənilməsidir. Leksikoqrafiya nəzəriyyəsinin mərkəzində lüğətlərin tiplərinin öyrənilməsi durur. Analoji olaraq, terminoqrafiyanın predmeti xüsusi lüğətlərin qurulması metodologiyası və konkret nümunələrin işlənilməsidir. Terminoqrafiyanın mühüm məsələsi – terminoloji lüğətlərin təsnifatı və tipologiyası prinsiplərinin işlənilməsi, terminoloji lüğətlərin ayrı-ayrı tipləri üçün xüsusi leksikanın seçilməsinin elmi əsaslandırılmış prinsiplərinin çıxarılması, həmçinin xüsusi leksikadan istifadənin optimal şərtlərinin yaradılması üçün onlarda terminologiyanın vahidlərinin təqdimatı və təsvirinin vahid prinsiplərinin qurulmasıdır. Məlumdur ki, leksikoqrafiya leksikologiya ilə sıx əlaqəlidir, çox vaxt leksikoloji tədqiqatların nəticələrindən istifadə edir, öz növbəsində, leksikoloji problemlərin qoyuluşu və həlli üçün material təqdim edir. Bu cür sıx əlaqə terminoqrafiya və terminşünaslıq arasında da mövcuddur.

Terminşünaslar tərəfindən tədqiq olunan bir çox problemlər praktikada terminoloji lüğətlərin işlənilməsi zamanı ortaya çıxır və bu problemlərin həlli lüğətlərin tərtib edilməsi metodlarına təsir edir. Eyni zamanda xüsusi leksikanın istənilən sahəsinin öyrənilməsi terminoqrafiya ilə əlaqəlidir, belə ki, terminologiyanın aşkarlanması, tədqiqatı və nizamlanması üzrə işlərin nəticələri adətən lüğət formasında tərtib olunur. Terminoqrafiya terminşünaslıqla sıx əlaqəlidir, belə ki, terminlərin çoxmənalılıq, sinonimi və omoniminin qəbul edilməsi və onların digər dillərdə ekvivalentlərinin seçilməsi kimi problemlər terminşünaslıq çərçivəsində həll edilir. Ona görə də bir çox terminşünas hesab edir ki, terminşünaslıq terminoqrafiya üçün nəzəri bazadır və ya terminoqrafiya terminşünaslığın bir bölməsidir [12].

Son dövrlərdə terminoqrafiyaya aid problemlər dairəsinin müəyyənlişməsi ona terminşünaslıq və leksikoqrafiyanın kəşf etməsində sərbəst bilik sahəsi kimi baxmağa imkan verir.

Terminoqrafiya nəzəriyyəsinin problemlərinin həlli üçün ilkin şərtlər lüğətlərin tipologiyasının işlənilməsi ilə əlaqəli elmi-tədqiqat prosesində leksikoqrafik parametrlərin konsepsiyasının ayrılması oldu: müxtəlif tipli (izahlı, tərcümə, informasiya və s.) terminoloji lüğətlərin tərtib edilməsi üzrə iş təcrübəsinin ümumiləşdirilməsinə tələbat; terminoloji lüğətə xüsusi işgüzar nəsr janrı kimi baxılması; terminoqrafik işlərin avtomatlaşdırılması; avtomatik lüğətlərin və terminoloji verilənlər bazasının yaradılması.

Terminoloji lüğətlərin tərtibatı prinsipləri

Məlum olduğu kimi, lüğətlər insan həyatında böyük rol oynayır. Lüğətlər müxtəlif dillərdə danışan insanların öz aralarında ünsiyyət qurmasına imkan yaradır. Lüğətlər, həmçinin dogma və xarici dilin öyrənilməsi vasitəsidir, bilik ehtiyatını genişləndirməyə kömək edir, mövcud bilik səviyyəsinin möhkəmlənməsinə xidmət edir. Müasir dövrdə terminoloji lüğətlər elmin və texnikanın sürətli inkişafında xüsusilə böyük rol oynayır. Son dövrlər terminoloji lüğətlərin sayı sürətlə çoxalır və digər lüğət formalarının sayının da artmasına səbəb olur.

Azərbaycanda 1920-1952-ci illər ərzində terminoloji lüğətlərin sayı digər lüğət formalarına nisbətən on dəfələrlə çox olmuşdur. Keçən əsrin 50-ci illərindən indiyə kimi 300-dək ikidilli, çoxdilli və izahlı lüğətlər nəşr olunmuşdur [13].

Yeni bilik sahələri və yeni tətbiq sahələrinin lüğətləri meydana çıxır: informasiya texnologiyaları, idarəetmə sistemləri və s. Nəticə olaraq xüsusi lüğətlərin yeni formaları meydana

çıxır. Məhz terminoloji lüğətlərin tərtib edilməsinə praktiki tələbat terminoqrafiyanın konkret metodlarına əsaslanan nəzəri müddəaların işlənməsini tələb edir.

Hazırkı dövrdə hər il müxtəlif sahə mütəxəssisləri tərəfindən çoxlu sayda terminoloji lüğətlər hazırlanır. Bu mütəxəssislər leksikoqrafiya sahəsində heç bir hazırlığa malik deyillər. Eyni zamanda mövcud metodiki materiallar yalnız lüğətlərin konkret formalarına tətbiq olunur və öz aralarında kifayət qədər qarşılıqlı əlaqəli deyillər. Xüsusi lüğətlərin ayrı-ayrı formalarında ümumi qəbul edilmiş tərtibat qaydaları mövcud deyil. Ona görə də lüğət tərtib edənlər çətin vəziyyətdə qalırlar, yalnız öz təcrübələrinə və istifadəçilərin istəyinə əsaslanaraq bu cür mürəkkəb məsələlərin öhdəsindən gəlirlər. Bu da lüğətlərin hazırlanması və tərtibatı metodlarında vahid qaydaların olmamasına gətirir, onları hazırlayanların işini çətinləşdirir, iş vaxtını artırır, lüğətin keyfiyyətini azaldır [14].

[13]-də lüğətlərin tərtibatı üçün aşağıdakı şərtlərin yerinə yetirilməsi verilmişdir:

1. Sözlüyə daxil ediləcək anlayışların toplanması;
2. Toplanmış terminlərin sistemləşdirilməsi;
3. Sistemləşdirilmə zamanı terminlərdən anlayışları dəqiq ifadə edənlərin seçilməsi;
4. Lüğətdə terminin verilməsi (izahlı, ikidilli, çoxdilli) yollarının müəyyənləşdirilməsi;
5. Lüğətdə sözlüyün verilməsi prinsiplərinin hazırlanması;
6. Lüğətə aid statistik məlumatların verilməsi;
7. Lüğətdən istifadə qaydalarının hazırlanması;

Bilik sahəsi obyektivdir, termin və terminlər sistemi isə konkret bir dilə bağlıdır, ona görə də terminoqrafiyanın mühüm məsələsi terminlərin standartlaşdırılması və unifikasiyası, eyni zamanda dünyanın müxtəlif dillərinə birmənalı tərcüməsidir.

Termin sistemlərinin unifikasiyası terminoloji standartlara əsaslanır. Hazırda dünyada termin sistemlərinin təşkili üzrə 20 mindən çox standart var. Bundan başqa müxtəlif səviyyələrdə - beynəlxalq, dövlət və hətta ayrı-ayrı firma və şirkətlər səviyyəsində terminoloji standartlar mövcuddur. Bununla əlaqədar olaraq termin və termin sistemlərinin unifikasiyası məsələsi də dövlət və yerli dil siyasətinin tərkib hissəsi olmalıdır [15]. Belə ki, terminlərin çoxmənalılığı və omonimlik xüsusiyyətləri, yaxın termin sistemləri arasında uzlaşmanın olmaması, çətin tələffüz olunan və yaxşı səslənməyən ixtisarlarla terminoloji uyğunluğun yaradılması elmi-texniki inkişaf üçün maneədir.

Yuxarıda deyilənlərə uyğun olaraq müasir leksikoqrafiya və terminoqrafiyanın perspektiv istiqamətləri aşağıdakı kimi ifadə edilir:

- Müxtəlif məqsədlər üçün müxtəlif tip lüğətlərin layihələndirilməsi, tərtibatı və effektiv istifadəsi metodologiyası, metodikası və texnologiyasının həyata keçirilməsi;
- Parametrik leksikoqrafiya və terminoqrafiyanın inkişafı;
- Leksikoqrafiya və terminoqrafiya nəzəriyyəsinin problemlərinin dərinləşməsi;
- Leksikoqrafiya və terminoqrafiya ilə inteqrasiya olunmuş yeni elmi paradigmalardan təsdiqlənməsi;
- Lüğətlərin mövcud növləri və janrlarının təkmilləşdirilməsi;
- Yeni prinsip və məsələlər əsasında yeni tip lüğətlərin yaradılması;
- Lüğətlərin yaradılmasının sürətləndirilməsi və leksikoqraf/terminoqrafın işinin yüngülləşdirilməsi üçün lüğət işinin informasiyalaşdırılması və kompüterləşdirilməsi və s.

Məlumdur ki, müəyyən predmet sahəsinin terminləri xüsusi terminlər lüğətində yığılır və təsvir olunur. Bu lüğətlər deskriptiv və normativ, ümumi və xüsusi, izahlı və tərcümə, əlifba sırası ilə və tezaurus ola bilər.

Müasir kompüter texnologiyaları müxtəlif bilik sahələri üzrə böyük sayda terminlər massivinin emalı və saxlanılmasına imkan verir. Bu cür terminlər massivi terminoloji verilənlər bazası (TVB) adlandırılır.

Beləliklə, hazırkı dövrdə terminoloji bilik biliyin təqdimatının yeni formalarını axtarır: verilənlər bazası, biliklər bazası və s.

Leksikoqrafiya və terminoqrafiyanın avtomatlaşdırılması

Leksikoqrafiya və onun bir bölməsi olan terminoqrafiyada kompüterləşmənin üstünlükləri məlumdur: kompüter geniş sözlüyə, sitat materialına, çoxmənalı söz haqqında informasiyaya əlyətərliyi tez təqdim edə bilir.

Mətnin emalını həyata keçirən xüsusi kompüter proqramı böyük həcmdə lüğət və mətn informasiyasının saxlanması və emalına imkan verir, həmçinin bir və çoxdilli lüğətlərin yaradılmasında istifadə oluna bilər.

Lüğətlərin sosial əhəmiyyəti günü-gündən artır: onlar yalnız bilik sahəsini qeyd etmir, həmçinin elmi idrakın etibarlı alətinə xidmət edir. Bu səbəbdən leksikoqrafiya və terminoqrafiyanın avtomatlaşdırılması hazırda çox aktual məsələlərdən biridir.

Müasir linqvistikanın qarşısında duran məsələlər mətnin avtomatik emalı üçün kompüter texnikasından istifadəni ön plana çəkir.

Burada aşağıdakı istiqamətləri ayırmaq olar:

- Müxtəlif tip informasiya sistemlərinin linqvistik təmini;
- Maşın tərcüməsi;
- Təbii dili anlayan sistemlərin işlənilməsi;
- Səsli nitq signalından ibarət informasiyanın istifadəsi sisteminin işlənilməsi və s.

Hələ keçən əsrin 70-ci illərində Y.N.Marçuk leksikoqrafiya işlərinə kompüterin tətbiqi imkanları haqqında yazmışdı [16]. Belə ki, kompüter ona daxil edilən leksik vahidlərin əlifba sırası ilə nizamlanması, eyni zamanda leksikoqrafların sərf etdikləri çox böyük zaman və əmək tələb edən daha çətin işləri yerinə yetirə bilir.

Leksikoqrafiya, eləcə də terminoqrafiyada müasir kompüter texnologiyalarından istifadə məsələsi bu gün də çox aktualdır və detallı öyrənilmə tələb edir.

Müasir hesablama vasitələri leksikoqrafik işləri praktiki olaraq hər mərhələdə - sitatların seçilməsindən başlayaraq lüğətin redaktəsi və çapına qədər avtomatlaşdırılmasına imkan verir. Bu cür işlərin avtomatlaşması, kompüter proqramlarının geniş yayılması ayrı-ayrı leksikoqrafların (terminoqrafların) əmək məhsuldarlığını yüksəldir. Bunun nəticəsində də yeni istiqamət – leksikoqrafik informatika formalaşır. Leksikoqrafik informatika avtomatik lüğətlərin yaradılması, həmçinin leksikoqrafik işlərin dəstəklənməsi proqramlarının işlənilməsi kimi məsələləri özünə daxil edir.

Qeyd edək ki, elmi ədəbiyyatda maşın daşıyıcılarında leksikoqrafiyanın tərif, ifadəsi üçün bir sıra terminə rast gəlmək olur: leksikoqrafik informatika, hesablama leksikoqrafiyası, maşın leksikoqrafiyası, avtomatik leksikoqrafiya və avtomatlaşdırılmış leksikoqrafiya. “Leksikoqrafik informatika” termini hal-hazırda daha çox istifadə olunandır.

Leksikoqrafik informatika bu gün özündə ilk olaraq sürətlə inkişaf edən kompüter sənayesini təqdim edir, onun yaranması və yayılmasının əsas üsullarından biri elmi biliyin lüğət şəklində salınmasıdır. Leksikoqrafik informatika, həmçinin dilçilikdə tətbiqi elmi fənni təqdim edir, bu istiqamət lüğətlərin tərtib edilməsi nəzəriyyəsi və praktikasında kompüter texnologiyalarından istifadəni öyrənir.

Leksikoqrafik informatika leksikoqrafiya kimi, nəzəri və praktiki hissələrə bölünür. Birinci bölmə leksikoqrafik informatika nəzəriyyəsi öyrənir və kompüter lüğətlərinin tipləri, kompüter leksikoqrafiyalaşması qaydalarının seçilməsi və s. kimi məsələlərə baxır.

Leksikoqrafik informatikanın əsas istiqamətlərinə aşağıdakılar daxildir:

- Müxtəlif lüğətlərin (tezlik, terminoloji və s.) kompüter vasitələrinin köməyi ilə mətnlərdən avtomatik alınması;
- Təbii dilin emalı sistemi üçün kompüter lüğətlərinin tərtib edilməsinin nəzəri və praktiki aspektləri;
- Ənənəvi lüğətlərin maşın versiyasının yaradılması və həyata keçirilməsi.

Birinci və ikinci istiqamətlər leksikoqrafik işlərin dəstəklənməsi məqsədilə proqram təminatının işlənilməsi ilə məşğul olur.

Leksik materialın emalı üçün kompüter texnologiyalarından geniş istifadə olunur. Bu cür texnologiyalar leksikoqrafın öz fəaliyyəti prosesində istifadə etdiyi verilənlərin emalı əməliyyatını həyata keçirən alqoritmi təqdim edir.

Mətnin kompüter emalının əsas komponenti mətn verilənlərinin daxil edilməsi və yadda saxlanılmasına, leksikoqrafın iş prosesində onları dəyişməyə və nəticələri ənənəvi kağız daşıyıcılarında olduğu kimi, “elektron” formada maqnit, maqnit-optik və optik informasiya daşıyıcılarında çap etməyə imkan verən kompüter vasitələridir.

Hazırkı dövrdə məşhur lüğət və sorğuların maşın daşıyıcılarına daxil edilməsi və onların bazasında yeni lüğətlərin yaradılması aktual məsələlərdən biridir. Əvvəllər çap edilmiş kitab rəflərində olan lüğətlərin kompüterə köçürülməsi onların hər birinin əlçatanlığına imkan yaradır.

[17]-də lüğətlərin yaradılmasına tətbiq edilən linqvistik verilənlərin yaradılması və toplanmasının üç müxtəlif metodu təsvir olunmuşdur: istifadəçi kontenti; viki modeli; kraudsorsinq modeli. Nəticə ondan ibarətdir ki, hər üç metod leksikoqrafiya və terminoqrafiya üçün əhəmiyyətli potensialdır.

Terminoqrafiyanın əsas məsələlərindən biri də kompleks avtomatlaşdırılmış terminoloji lüğətlərin yaradılmasıdır. Bu mətn üzrə müxtəlif lüğətlərin vahid proqram kompleksində birləşməsi və bu informasiyanın əsas lüğətə ötürülməsi ola bilər. Bütün bunlar kompüter lüğətindən müxtəlif aspektli informasiyanın alınmasına imkan verir.

Beləliklə, terminoqrafik informatika iki əsas istiqamətdə inkişaf edəcək:

–kompüter texnologiyası əsasında çap formasında lüğətlərin yaradılması;

–yalnız kompüterin yaddaşında və ya maqnit daşıyıcılarında fəaliyyət göstərən elektron lüğətlərin yaradılması.

Məlumdur ki, hazırkı dövr çap formasında lüğətlərdən elektron lüğətlərə keçid dövrüdür. Məhz elektron lüğətlərin hesabına leksikoqrafiya və terminoqrafiya nəzəriyyəsi və praktikasının sürətli genişlənməsi baş verdi və XXI əsrin birinci onilliyində dünyanın 40 dili üçün elektron lüğətlərin 600-dən çox tipi var idi [18]. Elektron lüğətlərin əsas tipləri aşağıdakılardır [19]:

1. İnformasiya-axtarış tezaurusları - informasiya-axtarış dillərindən istifadə etməklə informasiya-axtarış sisteminə xidmət edir;
2. Verilənlər bazası;
3. Terminoloji verilənlər bazası;
4. Biliklər bazası;
5. Terminoloji biliklər bazası;
6. Maşın korpusları və s.

Axtarışın effektivliyinin yüksəldilməsi informasiya-axtarış tezauruslarının keyfiyyətinin yüksəldilməsindən asılıdır. Bu səbəbdən informasiya-axtarış tezauruslarının optimallaşdırılması metodları işlənilmişdir ki, bu metodlar tezaurus üçün daha optimal leksikanın seçilməsinə imkan yaradır [20].

Qeyd edək ki, XXI əsrin əvvəllərindən başlayaraq elektron terminoqrafiya sahəsində İngiltərə liderlik edir. Bu, müasir elektron lüğətlərin prototipləri – “İnternetdə sığu resursları” adlanan və kompakt-disklərdə olan lüğətlər meydana çıxan vaxtdan başlayıb [21].

Bəzi tədqiqatçılar güman edirlər ki, “məhz lüğət əməyinin kompüterləşməsi leksikoqrafiyanın gələcək sifətini təyin edir. Leksikoqrafik verilənlərin kompüter emalı sayəsində lüğət yaradıcılığının sürətli inkişafı davam edəcək, ola bilsin ki, XXI əsr leksikoqrafiyanın Qızıl Əsri adlandırılacaq” [18].

Hazırkı dövrdə analiz üçün böyük sayda etibarlı verilənlər mövcuddur. Bu səbəbdən leksikoqraflar elə dil texnologiyaları axtarırlar ki, o verilənlərin analizini sadələşdirən və sürətləndirən alət və metodlar təqdim etsin.

[22]-də iki müxtəlif leksikoqrafik və terminoqrafik tədqiqat layihəsi təqdim olunmuşdur. Hər iki layihədə korpusun verilənlərinin avtomatik çıxarılması üçün eyni metod istifadə olunmuş və əldə edilmiş verilənlərin keyfiyyətinin oxşar və fərqli cəhətləri təqdim olunmuşdur. Leksikoqrafiyada və terminoqrafiyada avtomatlaşdırmanın gələcək perspektivləri verilmişdir.

Leksikoqrafik informatika və terminoqrafik informatika keçid dövrünün fənləridir - bir çox onilliklər ərzində ənənəvi leksikoqrafik (terminoqrafik) əl yazısı praktikasından yeni kağızsız informasiya texnologiyalarına keçiddir.

Leksikoqrafik işlərin avtomatlaşdırılması ilə nəticələnən müxtəlif növ terminoloji lüğətlərin yaradılması təcrübəsinin ümumiləşdirilməsinə tələbat, lüğətlər üçün xüsusi leksikanın seçilməsində artan çətinlik, ümumi leksikoqrafiyada lüğətlərin tipologiyası üzərində iş və lüğətlərin ən effektiv kompozit həllərinin seçilməsi zərurəti terminoqrafiya nəzəriyyəsinin inkişafını stimullaşdırdı.

Terminoloji verilənlər bazası

Yuxarıda qeyd edildiyi kimi, terminoloji lüğətlərin yaradılmasının böyük əmək tutumu ilə əlaqədar olaraq son zamanlar lüğətlərin hazırlanması üçün avtomatlaşdırılmış metodlar işlənir. Leksikoqrafik işlərin avtomatlaşdırılması böyük həcmli terminoloji informasiyanın yığılması və sürətli emalına imkan verən terminoloji verilənlər bazasının (TVB) yaradılmasına gətirib çıxartdı.

TVB-nin yaradılması böyük resurslar tələb edir, onun tərkibinin dəyişdirilməsi isə böyük həcmli işlə əlaqəlidir. Ona görə də TVB-də terminoloji informasiyaya daxil edilmiş tərkibin öncədən tədqiq olunması çox vacibdir.

Terminoqrafiya nəzəriyyəsinin təqdimatı nöqteyi nəzərindən TVB-nin işlənməsi üçün problem olan onun təyinatı, funksiyaları, əsas məsələləri, daxil edilən informasiyanın tərkibi, lüğət informasiyasının mənbələrinin tiplərinin öyrənilməsi kimi problemlərin analiz edilməsi zəruridir.

TVB-nin inkişaf perspektivləri müasir ekspert sistemlərinin meydana çıxması və intellektual sistemlərin işlənməsi və yeni fəaliyyət sahələrinin – bilik mühəndisliyinin yaranması ilə əlaqədardır. TVB biliyin təqdimatı sisteminin yeni formalarını və biliklər bazasında xüsusi informasiyanın təşkili vasitələrini təqdim edir.

Terminoloji işin elmi və texniki ədəbiyyat və sənədlərin mətnlərinin emalının ənənəvi metodları ilə birlikdə avtomatlaşdırılması TVB-nin yaradılmasını təmin etdi. TVB – terminoloji verilənlərin yığılması, saxlanması, emalı və ötürülməsinin kompüterin köməyiylə yerinə yetirildiyi sistemdir. Böyük massivdə terminoloji informasiyanın struktur təşkili və avtomatik emalının zəruriliyi terminologiya sahəsində bu istiqaməti aktual etmişdir.

Yaradılma məqsədinə görə TVB-ni iki qrupa bölmək olar:

- elmi-texniki ədəbiyyatın və sənədlərin tərcüməsi işinə yönəlmiş TVB;
- standartlaşdırılmış və tövsiyə edilmiş terminologiya haqqında informasiya ilə təminat üçün təyin edilmiş TVB.

TVB-nin geniş istifadəsi sahələrinə aid etmək olar:

- müxtəlif bilik sahələrinin mütəxəssislərinə sorğu-informasiya xidmətinin göstərilməsi;
- elmi-texniki ədəbiyyatın ənənəvi tərcüməsinin təmin olunması;
- maşın tərcümə sisteminin təmin olunması;
- avtomatlaşdırılmış informasiya sistemlərinin linqvistik təmini;
- terminologiyanın nizamlanması üzrə işin təmini;
- terminoloji lüğətlərin hazırlanması və çap edilməsi;
- müəyyən terminlərin unifikasiyası;
- dilin (yaradıldığı dil) tərkibi haqqında elmi hesabatların hazırlanması.

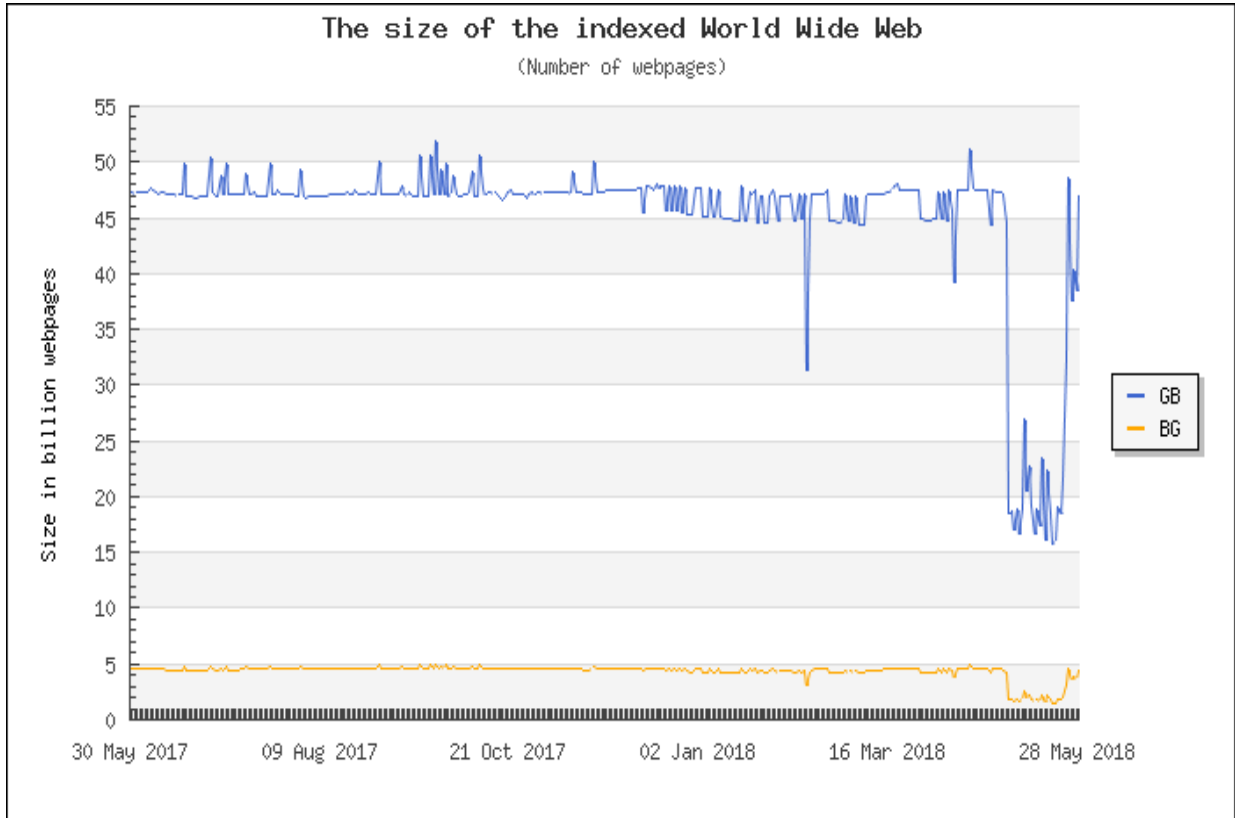
Hazırda müxtəlif bilik sahələri və praktiki fəaliyyət sferaları üzrə terminoloji verilənlər bazası yaradılır. Müasir dövrdə işlənən çoxdillli terminoloji verilənlər bazası terminlərin semantikasının nizamlanması və dillərarası terminoloji uyğunluğun qurulması üzrə beynəlxalq qərarların qeydə alınması vasitəsi olmalıdır. Bu cür verilənlər bazası terminlərin linqvistik və

məntiqi özəllikləri, onların istifadəsi, çoxdilli ekvivalentləri və nizamlanma dərəcəsi haqqında informasiyanın yığılması və saxlanmasına imkan verir.

Kanadanın Ottava Universitetinin süni intellekt laboratoriyasında terminoloji biliklər bazası (TBB) konsepsiyası işlənmişdir. TBB konsepsiyasının reallaşdırılması və inkişafı biliyin əldə edilməsi prosesinin asanlaşdırılmasına yönəlmişdir [23].

Hazırda müxtəlif ölkələrdə onlarla iri TVB mövcuddur. Rusiyada iri TVB Ümumrusiya standartlaşma və keyfiyyət üzrə kompleks informasiya elmi-tədqiqat institutunda [24], eləcə də, Ümumrusiya elmi və texniki informasiya institutunda mövcuddur. Dünyada ən iri TVB Almaniyanın "Siemens" firmasıdır, 8 dildə 2,5 milyon terminoloji yazını özündə saxlayır. Almaniya Standartlaşdırma İnstitutunda və digər idarəetmə orqanlarında TVB fəaliyyət göstərir. Lüksemburqda Avropa Komissiyasının terminoloji bürosunda iri beynəlxalq TVB yaradılmışdır [25]. "Eurodicautom" adı altında yaradılan TVB bir neçə dildə fəaliyyət göstərir. Fransa, Kanada, İcveç, Belçika və digər ölkələrdə də TVB yaradılmış və fəaliyyət göstərir. Beynəlxalq Standartlaşdırma Təşkilatının nəhəng TVB-si vardır [26].

İnformasiya cəmiyyəti çoxlu sayda verilənlər istehsal edir. Bu məlumatların böyük hissəsi ixtisaslaşmışdır: xüsusi lüğətlərdə kodlaşdırılmış predmetyönlü biliyin məzmunudur; ekspert məqsəd qrupu tərəfindən başa düşülür; xüsusi biliklərin ötürülməsində əsas maneə olan terminologiyanın tərcüməsidir.



Şəkil 1. İndekslənmiş Ümumdünya Hörümçək Torunun (İnternet) ölçüsü

GB = Google və Bing üzrə çeşidlənib.

BG = Bing və Google üzrə çeşidlənib.

İnsan biliyinin böyük hissəsi strukturlaşmamış formada mətnlərdə olur, çünki bilik aydın şəkildə göstərilməyib. Biliyə sürətli çıxış əldə etmək üçün anlayışlar və onlar arasında münasibətləri çıxarmaq lazımdır.

Ümumdünya hörümçək torunun (World Wide Web) ölçüsü üzrə aparılan araşdırmaların nəticələrinə görə, hazırda indeksləşmiş şəbəkə 4,48 milyard səhifəni ehtiva edir [27]. İndeksləşmiş World Wide Web-in minimal ölçüsü Google, Bing və Yahoo Search-də indeksləşmiş səhifələrin

sayının qiymətləndirilməsinə əsaslanır. Şəkil 1-də Bing-lə (BG) və Google-la (GB) başlayan iki ümumi qiymətləndirmə təqdim olunur. Axtarış sisteminin ölçüsü böyük mətn korpusunda alınan sözlərin tezliyi ilə sorğuların sayını birləşdirən metoda əsasən qiymətləndirilir (bax. Zipf Qanunu).

Qeyd edək ki, Azərbaycanda Milli Terminoloji İnformasiya Sisteminin (MTİS) konsepsiyası işlənmişdir [28]. Müxtəlif sahələr üzrə hazırlanmış terminoloji lüğətlərin vahid informasiya sistemində toplanması məqsədilə MTİS çərçivəsində Milli terminoloji veb-portal yaradılmışdır [29]. Milli terminoloji veb-portalın vasitəsilə elm və texnikanın müxtəlif sahələrinə aid terminləri, onların mümkün dillərə tərcüməsini əldə etmək mümkün olacaqdır.

Beləliklə, TVB böyük inkişaf potensialına malikdir, onlar avtomatlaşdırılmış anbardan ciddi elmi araşdırmaların aparılması vasitəsinə çevrilirlər, böyük praktiki əhəmiyyətə malikdirlər.

Nəticə

Tədqiqat işində terminoqrafiya işlərinin avtomatlaşdırılması ilə əlaqəli onun inkişaf perspektivlərinin müəyyənləşdirilməsi məsələsinə baxıldı və aşağıdakı nəticələr əldə olundu:

- Hazırda terminoqrafiyaya müstəqil elm sahəsi kimi baxmağa bütün əsaslar var: özünün tədqiqat obyektı var – terminoloji lüğətlər; özünün predmeti var – xüsusi lüğətlərin yaradılması və qiymətləndirilməsi prinsiplərinin müəyyənləşdirilməsi, həmçinin termin sistemlərinin müxtəlif növ lüğətlərdə təsviri problemlərinin həlli.
- Məlum oldu ki, müasir terminoqrafiyanın əsas problemləri xüsusi lüğətlərin etibarlı klassifikasiyasının yaradılması, ayrı-ayrı lüğət formalarına tələbatın və onların keyfiyyətinin yüksəldilməsi üsullarının analizi, lüğətdə terminoloji informasiyanın seçilməsi və təşkili metodlarının, həmçinin terminoqrafik işlərin avtomatlaşdırılması üsullarının öyrənilməsi və tətbiqindən ibarətdir.
- Terminoqrafiya işlərinin avtomatlaşdırılması aşağıdakı məsələlərin həllinə imkan yaradacaqdır:
 - terminoloji lüğətlərin tərtibinə vahid yanaşmanı təmin edəcək;
 - lüğətləri tərtib edərkən onların əsas xüsusiyyətlərini təyin etmək üçün optimal ardıcılıığı seçməyə imkan verəcək;
 - lüğətlərin tərtibi ilə məşğul olan çoxlu sayda mütəxəssisin öz işini təşkil etməsi və nəticələri obyektiv qiymətləndirməsi üçün imkanlar yaranacaqdır.
- Azərbaycanda yaradılan TVB termin yaradıcılığı işlərinin səmərəliliyinin yüksəldilməsinə xidmət edir və aşağıdakı elmi məsələlərin həllinə imkan verəcəkdir:
 - Azərbaycan dilinin terminoloji sisteminin modelləşdirilməsi;
 - ümumelmi və ümumnəzəri tezaurların qurulması;
 - Azərbaycan terminologiyasının tədqiqatı.

Ədəbiyyat

1. Щипицина Л.Ю. Информационные технологии в лингвистике : учеб. пособие, М., ФЛИНТА: Наука, 2013, 128 с.
2. Бобунова М. А. Русская лексикография XXI века , Учеб. Пособие, М., Флинта : Наука, 2009, 200 с.
3. Комарова З. И. Когда и почему возникла терминография как научно-прикладная дисциплина? // Актуальные проблемы германистики, романистики и русистики : сб. науч. работ, Екатеринбург, Изд-во УрГПУ, 2016, ч. 3, с. 90—107.
4. Дамбуев И.А. Современная лексикография: статус и направления развития, Бурятия, Россия // Вестник Бурятского Государственного Университета, 2011, стр. 16-21.
5. Касарес Х. Введение в современную лексикографию, Изд-во иностранной лит-ры, 1958, 354 с.
6. Герд А. С. Прикладная лингвистика. СПб., СПбГУ, 2005, 268 с.

7. Герд А. С. Научно-техническая лексикография, Прикладное языкознание. СПб., Изд-во СПб, 1996, 287 с.
8. Лейчик В. М. Терминоведение (Предмет, методы, структура), М., КомКнига, 2006, 256 с.
9. Гринёв-Гриневиц С. В. Введение в терминоведение, М., Москва, Лицей, 1993, 230 с.
10. Авербух, К. Я. Общая теория термина, Иваново, 2004, 252 с.
11. Qurbanov A. Müasir Azərbaycan ədəbi dili. I cild. Bakı, “Nurlan”, 2003, 450 s.
12. Гринев С.В. Введение в терминографию, М., Изд-во МПУ, 1995, 158 с.
13. Sadıqova S. Azərbaycan dilinin terminologiyası. Bakı, “Elm”, 2011, 380 s.
14. Комарова З.И. Терминография: предмет, задачи, проблематика. Гуманитарные термины в специальной литературе, Межвуз.сб. Горький, 1990, с. 104-111.
15. Баранов, А Н. Введение в прикладную лингвистику, Учебник; МГУ, Филол. фак., Изд. 3-е., Москва: Изд-во ЛКИ, 2007, 358 с.
16. Марчук Ю.Н. Компьютерная лингвистика: учеб. Пособие, М., АСТ: Восток-Запад, 2007, 317 с.
17. Rundell M., Hanks P., Schryver G.M. Crowdsourcing, wikis, and user-generated content, and their potential value for dictionaries // International Handbook of Modern Lexis and Lexicography, 2015, pp. 1-16.
18. Дубичинский В. В. Терминография и стандарты, Лексикография русского языка : учеб. Пособие, М., Флинта: Наука, 2009. 432 с.
19. Гринёв-Гриневиц С. В. Терминография, Терминоведение, Учеб. Пособие, М., Академия, 2008, 304 с.
20. Мамедова М.Г., Скороходько Э.Ф. Метод оптимизации информационно-поискового тезауруса // НТИ, сер. 2 – Информационные процессы и системы, 1985, стр. 12–17 (Москва).
21. Терминология и знание: материалы, 5-го Международного симпозиума, М., Принт Про, 2017, 367 с.
22. Kosem I., Gantar P., Logar N., Krek S. Automation of Lexicographic Work Using General and Specialized Corpora: Two Case Studies, Proceedings of the XVI EURALEX International Congress: The User in Focus, Lexicography and Language Technologies , 2014, pp. 355-364.
23. Meyer I., Skuce D., Bowker L., Eck K. Towards a New Generation of Terminological Resources: An Experiment in Building a Terminological Knowledge Base, <http://acl.ldc.upenn.edu/C/C92/C92-3146.pdf>
24. Всероссийский научно-исследовательский институт классификации, терминологии и информации по стандартизации и качеству, <http://www.vniiki.ru>
25. IATE - European Terminology Database, <http://www.iate.europa.eu>
26. International Organization for Standardization, Online Browsing Platform, <http://www.iso.org/obp/ui>
27. The size of the World Wide Web. <http://www.worldwidewebsite.com>
28. Əliquliyev R.M., Qurbanova Ə.M. Azərbaycanda terminoloji informasiya sisteminin yaradılmasının konseptual əsasları // İnformasiya cəmiyyəti problemləri, Bakı, 2011, №1, s. 3– 8.
29. Azərbaycan Respublikasının Nazirlər Kabineti yanında Terminologiya Komissiyası, <http://www.terminologiya.az>

УДК 004.65:06:80

Гурбанова Афруз М.

Институт Информационных Технологий НАНА, Баку, Азербайджан

afruz1961@gmail.com

Проблемы автоматизации терминографической деятельности и пути их решения

В статье рассматриваются понятия «лексикография» и «терминография», их история, теоретические и практические особенности. Отмечена необходимость разработки единых принципов и методов составления словарей, описаны необходимые для этого условия. Определены перспективные направления лексикографии и терминографии, отмечены преимущества их компьютеризации. Проанализированы процессы автоматизации работ по лексикографии и терминографии и отмечена необходимость создания терминологического банка данных, позволяющего накапливать и быстро обрабатывать большие объемы терминологической информации. Показаны возможности терминологического банка данных, созданного в рамках Национальной терминологической информационной системы Азербайджана.

Ключевые слова: лексикография, терминография, терминологический словарь, компьютерная лексикография, компьютерная терминография, терминологический банк данных.

Afruz M. Gurbanova

Institute of Information Technology of ANAS, Baku, Azerbaijan

afruz1961@gmail.com

The problems of automation of terminographic activity and their solution ways

The article reviews the concepts of "lexicography" and "terminography", their history, theoretical and practical features. The need to develop common principles and methods for compiling dictionaries is noted, and the necessary conditions for this are described. Prospective directions of lexicography and terminography are determined, and the advantages of their computerization are mentioned. The processes of automation of works on lexicography and terminography are analyzed. The necessity of building terminological data bank enabling the accumulation and rapid processing of large volumes of terminological information is noted. The opportunities of terminological data bank established within the framework of the National Terminological Information System of Azerbaijan are presented.

Keywords: lexicography, terminography, terminological dictionary, computer lexicography, computer terminography, terminological data bank