

Jurnalistikada mətnlərin intellektual linqvistik təminatı problemləri

Sünbül Zalova

İnformasiya Texnologiyaları İnstitutu,
Bakı Dövlət Universiteti, Bakı, Azərbaycan
sunbulzalova@gmail.com

Xülasə — Tezisdə linqvistik xüsusiyyətlərin jurnalistikada rolu, jurnalist mətnlərində informasiya qəbulunu asanlaşdırmaq üçün mətnlərin keyfiyyətinin yaxşılaşdırılması, süni intellekt texnologiyalarının jurnalistikaya tətbiqinin yaratdığı yeni istiqamətlər haqqında məlumat verilmiş. linqvistik təminat vasitələri, vətəndaş jurnalistikası ilə peşəkar jurnalistikanın dil-üslub fərqləri ilə bağlı tədqiqatlar araşdırılmış, müasir jurnalistikada yazı işlərinin keyfiyyətinin artırılması üçün informasiya kommunikasiya texnologiyalarının imkanlarından istifadənin zəruriliyi əsaslandırılmışdır.

Açar sözlər — linqvistik təminat vasitələri; vətəndaş jurnalistikası; süni intellekt

I. GİRİŞ

Jurnalistika sahəsi cəmiyyətin problemlərinin araşdırılması, hadisələrin düzgün təqdim edilməsi kimi məsələlərin həllində vacib rol oynayır. Burada yazı dili və onun xüsusiyyəti çox önəmlidir. Jurnalistikanın hər bir janrında dilin incəliklərinə və imkanlarına yaxşı bələd olmaq və onlardan məharətlə istifadə etmək tələb olunur. Jurnalistlər hadisələri, faktları, problemləri söz və cümlələrin düzgün

seçilməsi ilə müxtəlif bucaqlardan təqdim edə bilirlər [1]. Burada hər hansı bir məsələnin neqativ və ya pozitiv təqdim olunması bir çox hallarda jurnalistin seçdiyi ifadələrdən asılıdır.

II. İNTELLEKTUAL LİNGVİSTİK TƏMİNAT VASİTƏLƏRİ

Mətn axtarış sistemlərində tammətli sənədlərlə iş prosesində təbii dilin emalı vasitələrindən geniş istifadə olunur. Bu vasitələr mətn axtarış sistemlərinin mürəkkəb komponentləridir. Təbii dilin emalı vasitələri mətn sənədlərindəki, sorğulardakı terminləri, sözləri, söz birləşmələrini və s. təhlil edir. Təbii dilin emalı vasitələri mətnlərdəki məzmununu aydınlaşdırır, terminlər arasındakı münasibətləri aşkar edir və səmərəli axtarışı təmin edir. Mətnlərin keyfiyyətinin yaxşılaşdırılması məqsədi ilə mətni axtarış sistemlərində linqvistik təminat vasitələri kompleksindən istifadə olunur. Bunlara müxtəlif lüğətlər, tezauruslar, sistemin predmet sahəsinin ontoloji xüsusiyyətlərini əks etdirən materiallar daxildir [3].

Sistem lüğətləri dedikdə təbii dilin emalı ilə məşğul olan mətn axtarış sistemlərində ümumdil leksikası, və predmet sahəsi

leksikası üzrə lüğətlərdən istifadə edilir [3]. Bu lüğətlər mətni morfoloji təhlil və axtarış edən zaman müxtəlif qrammatik formalardakı sözləri tanımağa imkan verir.

Tezaurus – dilin əsas anlayışları lüğətidir [3]. Tezaursları xüsusi lüğətlər kimi də adlandırmaq olar. Əsasən mətn tipli sənədləri təqdim etmək və sorğunu genişləndirmək üçün tezaurus adlanan xüsusi lüğətlərdən istifadə olunur.

Təbii dildə təqdim olunan mətn sənədlərinin və axtarış sorğularının adekvat dərki üçün predmet sahəsinin əsas anlayışları və onların arasındakı müxtəlif semantik əlaqələri təyin edən vasitələrə *ontoloji materiallar* deyilir. Bura predmet sahəsinin tezaurusu və birinci dərəcəli məntiq dillərini (*Birinci dərəcəli məntiq dedikdə predikat məntiqi, kəmiyyət məntiqi və birinci dərəcəli predikat hesabı nəzərdə tutulur ki, bu da riyaziyyat, fəlsəfə, dilçilik və kompüter elmlərində istifadə olunan formal sistemlərin məcmusudur*) və s. aid etmək olar [4].

İnformasiya kommunikasiya texnologiyalarının (İKT) sürətli inkişafı jurnalist yazılarında mətlərin oxunaqlığının yoxlanmasına, mətnlərdə ölkədə qəbul edilməmiş terminlərin aşkarlanmasına, vulqar sözlərin avtomatik redaktəsinə şərait yarada bilər. İKT-nin köməyi ilə, əsasən də, süni intellekt vasitəsilə bu gün mətnlərin keyfiyyəti avtomatik yoxlanılır və yaxşılaşdırılır.

III. SÜNİ İNTELLEKTİN JURNALİSTİKA SAHƏSİNDƏ YARATDIĞI İMKANLAR

Struktur linqvistik xüsusiyyətlər yazılı ünsiyyətin mühüm aspektidir. Araşdırmalar göstərir ki, linqvistik mürəkkəblik insanların informasiyanı emal etməsində mühüm rol oynayır. Tədqiqatçılar avtomatlaşdırılmış məzmun təhlili metodlarından istifadə edərək müxtəlif vətəndaş və peşəkar jurnalist təşkilatları arasında dil mürəkkəbliyindəki fərqləri araşdırıb. Təhlil nəticəsində müxtəlif kütləvi informasiya vasitələrində siyasi informasiyanın təqdim edilməsi nümunələrinin fərqli olduğu göstərilmişdir. Araşdırmada jurnalistikada mətn mürəkkəbliyinin rolu və ənənəvi, tabloid qəzetlərinin və vətəndaş jurnalistikası məqalələrinin mətnin mürəkkəbliyi və ölçüləri baxımından necə fərqləndiyi barədə dəyərli fikirlər təqdim olunmuşdur [2]. Bu araşdırmada üç müxtəlif media növündən olan siyasi jurnalist mətnlərinin struktur xüsusiyyətləri, xüsusən də peşəkar qəzet məqalələri, tabloid qəzet məqalələri və siyasi veb-bloq məqalələri müqayisə edilmişdir. Tədqiqatın məqsədi mətn mürəkkəbliyinin bu medialar arasında fərqinin olub-olmadığını, xüsusən də vətəndaş jurnalistikası mətnlərinin digər iki peşəkar mətn növündən fərqlənib-fərqlənmədiyini araşdırmaq idi. Dilçilikdən tutmuş sosial elmlərə qədər bir sıra sahələrdən mətn mürəkkəbliyi ölçüləri seçilmiş və ~930.000 məqaləyə tətbiq edilmişdir. Müəyyən edilmişdir ki, bu ölçülər faktiki olaraq mətn mürəkkəbliyinin üç müxtəlif ölçüsünü ölçür: sintaktik, semantik və məzmun mürəkkəbliyi.

Maşın tərcüməsi, avtomatik referatlaşma, mətnin mövzuya uyğun təsnifatlaşdırılması süni intellektin jurnalistika sahəsində yaratdığı yeni imkanlardır. Süni intellektin jurnalistikaya verdiyi töhfələri aşağıdakı kimi təsnif etmək olar:

Mətnlərin oxunaqlığının və mürəkkəbliyin səviyyəsinin qiymətləndirilməsi dedikdə, mətnlərin mürəkkəbliyinin qiymətləndirilməsi üçün riyazi modellərin qurulmasının zəruriliyinin əsaslandırılması nəzərdə tutulur. Həmçinin mövcud qiymətləndirmə düsturları Azərbaycan dilinin özəllikləri nəzərə alınmaqla uyğunlaşdırılmalıdır. Bu prosesə Azərbaycan dilindəki mətnlərin mürəkkəbliyin səviyyəsinin qiymətləndirilməsi üçün avtomatlaşdırılmış texnologiyaların işlənilib hazırlanması ilə yanaşı bəzi dünya ədəbiyyatı nümunələrinin orijinalının Azərbaycan dilinə tərcümələrinin oxunaqlığının müqayisəsi və hazırlanmış metodika əsasında məktəb dərslərinin mürəkkəbliyinin şagirdlərin yaş səviyyəsinə uyğun olub-olmamasının qiymətləndirilməsi kimi məsələlər də daxildir [5].

Mətnlərdə istifadə olunan terminlərin müqayisəli analizi dedikdə, mətnlərdə ölkədə qəbul edilməmiş terminlərin aşkarlanması (məsələn: “Gəncə Destinasiya Mərkəzi” tipli adlara qarşı terminaloji filtrasiyadan istifadə) və Milli Terminoloji İnformasiya Sistemi ilə müqayisəsi nəzərdə tutulur. Müasir dövrdə bir çox elm sahələrində də (leksikoqrafiya, terminşünaslıq, informasiya axtarışı və s.) emal olunmalı informasiyanın həcmünün

artım dinamikası termin və açar sözlərin avtomatik seçilməsi məsələsini aktual edir. Hazırda təbii dilin emalı sahəsində qoyulan qaydaları təqdim edən mətnlərdən terminlərin avtomatik çıxarılması üçün bir çox fərqli yanaşma və sistem işlənmişdir [6].

Mətnlərdə vulqar sözlərin aşkarlanması İnternet mühiti şifahi nitqdən fərqlidir və burada müəyyən texniki əsaslar çərçivəsində vulqarizmlərə qarşı mübarizə aparmaq mümkündür. Milli dillərdə yaradılan vulqarizm bazalarının inteqrasiyası ilə qlobal vulqarizm bazasının yaradılması və UNESCO-nun koordinatorluğunun bu baza üzrə AntiVulgar sisteminin tərtibi İnternet mühitində vulqarizmlərə qarşı uğurlu mübarizə platforması kimi nəzərdən keçirilməlidir.

Mətnlərdə müxtəlif növ sirr daşıyan məlumatların müəyyənəşdirilməsi. Müasir dövrimdə texnologiyanın sürətli inkişafı ilə əlaqədar olaraq informasiyanın toplanması, saxlanması, emalı və ötürülməsi prosesləri müasir İKT vasitələrindən istifadə etməklə həyata keçirilir. Bu isə konfidensial informasiyanın asanlıqla başqalarının əlinə keçməsi təhlükəsini artırır. İnformasiya təhlükəsizliyinin təmin edilməsinin mühüm şərtlərindən biri informasiyanın konfidensiallığı məsələsidir. İnformasiya resursları ilə işləyən insanlar ayrı-ayrı vətəndaşlara və ya təşkilatlara aid olan informasiyanın konfidensiallığını qorumaq üçün hüquqi məsuliyyətlərini dərk etməli və informasiya təhlükəsizliyini təmin etməlidirlər.

Müxtəlif növ sirr daşıyan məlumatların olub-olmamasının aşkarlanması müvafiq qanunvericilik aktları ilə tənzimlənməlidir.

Mətnə plagiatlığın aşkarlanması. Plagiatlıq dedikdə başqa müəllifə məxsus fikir və ideyaların öz adından təqdim olunması nəzərdə tutulur [7]. Bu hadisəni oğurluq və piratçılıqlada sıx əlaqələndirmək olar. Plagiatlıq həmçinin kompüter virusları, spamlar, fişinq və s. kimi kibercinayətkarlığın bir növü hesab olunmalıdır. Plagiatlıq İnternetdən, kitabdan, jurnaldan və s. müxtəlif mənbələrdən birbaşa mətnin köçürülməsi və yaxud az sayda dəyişiklik etməklə baş verir.

Plagiatlığı avtomatik aşkarlama kompüter proqramları vasitəsilə həyata keçirilir və hazırda plagiatlığı aşkarlayan çoxlu sayda proqram təminatları mövcuddur: PlagAware, PlagScan, Check for Plagiarism, iThenticate, PlagiarismDetection, Academic Plagiarism, The Plagiarism Checker, Urkund, Docoloc və s. [7].

Maşın tərcüməsi prosesi sözlərin, cümlələrin və ya böyük həcmli mətnin tərcüməsi olan kompüter sistemi vasitəsilə təbii dilin başqa təbii dilə avtomatik çevrilməsidir [8]. Tərcümə proqramlarının özləri də iki kateqoriyaya bölünür: bütün mətni tərcümə edən və hər bir sözün ayrı-ayrılıqda mənasına uyğun sözlər siyahısını təqdim edən proqramlar.

Avtomatik referatlaşdırma mətndən zəruri məlumatların çıxarılması və bu məlumatın əsas məzmununu saxlamaqla yığcam formada istifadəçiyə təqdim edilməsi prosesidir. Mətnlərin avtomatik referatlaşdırılması üçün klasterləşmə və

optimallaşdırma metodlarına əsaslanan iki mərhələli cümlə seçimi modeli təklif olunub [9].

Mətnin təsnifatlandırılması prosesi göstərilən təsnifatlandırma sistemi çərçivəsində məzmunu uyğun olaraq mətnin hansı kateqoriyaya aid olmasını avtomatik müəyyən edir [10]. Lakin cümlələr oxşar olduqda bu tip məsələlərin həlli çətinləşir, qeyri-səlis klasterləşmə tələb olunur.

Yazı üslubunu analiz etməklə, müəllifin müəyyənlişməsi. Hər bir müəllif öz dəsti-xətti, üslubu və araşdırması ilə fərqlənir. Müasir İKT vasitələrinin köməyi ilə mətnlərdə müəllifin dəsti-xəttinin avtomatik olaraq aşkarlanması mümkündür.

NƏTİCƏ

Nəticə olaraq demək olar ki, İKT-nin köməyi ilə mətnləri daha anlaşılıqlı formada təqdim etmək, tanınmayan terminlərdən, vulqar sözlərdən təmizləmək mümkündür. Həmçinin plagiatlığın yoxlanılması və referatlaşma, mətnlərin təsnifatlandırılması və xarici dillərdən avtomatik tərcüməsi mətnlərin keyfiyyətinin yaxşılaşdırılmasında texnologiyanın necə vacib rol oynadığının əyani sübutudur. Müasir İKT vasitələrinin tətbiqi ilə jurnalistikada da yazı işlərinin keyfiyyəti artırıla bilər.

İSTİNADLAR

- [1] John Richardson “Language and journalism: An expanding research agenda”, *Critical Discourse Studies* 2015, vol 12, pp 370-373
- [2] Petro Tolochko, Hajo G. Boomgaarden, *Journalism Studies*, 2018, vol 19, pp 1786-1803
- [3] İsa Musayev, Mətləb Əlizadə, *İnformatikanın əsaları (dərs vəsaiti)*, 2014

- [4] https://en.wikipedia.org/wiki/First-order_logic
- [5] İ.Sadıgov, "Mathematical and Information Models for Evaluating Readability of Texts in Azerbaijani Language", 2018, El-Cezeri, vol. 5, no. 3, pp. 888-903, Sep.
- [6] Qurbanova Ə.M, Mətnlərdən terminlərin avtomatik çıxarılması metodlarının araşdırılması və müqayisəli analizi, İnformasiya cəmiyyəti problemləri, 2021, №1, səh. 55–69
- [7] R.S. Hasanova, F.Ş.Əsgərov, Plagiatlıq və onunla mübarizə vasitələri: antiplagiat sistemləri, İnformasiya cəmiyyəti problemləri, 2019, №2, səh. 94–101
- [8] Ramiz M. Aliguliyev, Makrufa Sh. Hajirahimova “An optimization model for automatic text summarization”, Problems of information technology, 2015, №2, pp 84–90
- [9] Rasim M. Alguliyev, Ramiz M. Aliguliyev, Nijat R. Isazade, Asad Abdi, Norisma Idris, COSUM: Text summarization based on clustering and optimization, “Expert Systems”, 2019, vol 36
- [10] Алекперова И.Я., Анализ тональности текста как эффективный подход к анализу поведения // Информационные технологии, 2022, том 28, № 11, стр. 590–596