



XORAZM GIDRONIMLARI KORPUSINING NAZARIY ASOSLARI VA AMALIY AHAMIYATI

Aqmanova Shahnoza Alimboyevna,
UrDU O'zbek tilshunosligi kafedrasida dotsenti, PhD
shahnozaaqmanova.90@gmail.com
<https://orcid.org/0000-0002-9203-1722>

MAQOLA HAQIDA	ANNOTATSIYA
Qabul qilindi: 9-fevral 2026-yil Tasdiqlandi: 11-fevral 2026-yil Jurnal soni: 17-B Maqola raqami: 40 DOI: https://doi.org/10.54613/ku.v17i.1474 KALIT SO'ZLAR/ КЛЮЧЕВЫЕ СЛОВА/ KEYWORDS korpus lingvistikasi, gidronimiya, Xorazm gidronimlari, semantik munosabatlar, NLP, GIS-texnologiyalar, mikrotoponimika, ma'lumotlar bazasi.	Ushbu maqolada Xorazm vohasi gidronimlari elektron korpusini yaratishning nazariy-metodologik asoslari kompleks yondashuv asosida tahlil qilinadi. Tadqiqotda korpus lingvistikasi tamoyillari, onomastik materialni yig'ish va tasniflash mezonlari, gidrografik prinsip asosida obyektlarni tizimlashtirish masalalari yoritilgan. Matnlarni avtomatik qayta ishlash jarayonida tokenizatsiya, lemmatizatsiya, lingvistik teglash va Named Entity Recognition (NER) usullaridan foydalanish imkoniyatlari ko'rib chiqiladi. Gidronimlarning leksik-semantik, etimologik va strukturaviy xususiyatlari ochib berilib, ularning diaxronik qatlamlari hamda tarixiy variantlari korpus bazasida saqlash mexanizmi asoslanadi. Shuningdek, korpusni Geoinformatsion tizimlar (GIS) bilan integratsiyalash, obyektlarni elektron xaritalarga bog'lash va areal tahlillar o'tkazish imkoniyatlari tahlil etiladi. Maqolada mazkur korpusning lingvistik, tarixiy, arxeologik, etnografik hamda turizm sohalaridagi amaliy ahamiyati ilmiy jihatdan asoslab berilgan.

Kirish. XXI asr tilshunosligida til hodisalarini an'anaviy tavsiflash usulidan empirik ma'lumotlarga asoslangan korpus tahliliga o'tish jarayoni kuzatilmoqda. Korpus lingvistikasi elektron hisoblash texnikasi asosida so'ngi yillarda shakllangan va rivojlanib borayotgan sohadir. U lingvistik korpuslarni yaratish, ma'lumotlarni qayta ishlash usullari hamda ularni qo'llash texnologiyalarini o'rganadi. Korpus bir nechta tillarda elektron shaklda matnlar asosidagi qidiruv tizimi hisoblanadi [11].

Jahon tilshunosligida gidronimlar eng qadimiy va barqaror onomastik qatlam sifatida e'tirof etiladi va ularni raqamlashtirish tarixiy-lisoniy xotirani saqlashning samarali usuli hisoblanadi. Gidronimlar xalqning tarixi, tili va etnomadaniyatini o'zida aks ettiradi. Jahon onomastikasi bugungi kunda yangi – raqamli rivojlanish bosqichiga o'tmoqda. 2014-yil 25-29-avgust kunlari bo'lib o'tgan 25-xalqaro onomastika kongressida (Glazgo) onomastika tadqiqotlariga ilg'or texnologiyalarni joriy etish va ilmiy tadqiqot natijalarini targ'ib qilish muammosi ko'tarildi. S.Xamroqulovanning ta'kidlashicha, Kopengagen universiteti professori Peder Gammeltoft onomastikaning raqamli rivojlanishning yangi bosqichiga o'tish haqidagi ma'ruzasida ushbu masalaga alohida e'tibor qaratadi. Professorning fikricha, onomastikani yanada ommabop qilish uchun olimlar o'z tadqiqotlari natijalarini turli internet portallari va mobil ilovalar orqali tarqatishlari kerak [7.84].

Adabiyotlar tahlili. Dunyoda onomastika bo'yicha tadqiqotlarni muvofiqlashtiruvchi yirik markazlar mavjud bo'lib, ular nomlar bazasini ya'ni korpusini shakllantirish bilan shug'ullanadi: ICOS (International Council of Onomastic Sciences) [15] onomastik fanlar bo'yicha xalqaro kengash bo'lib, u dunyo miqyosidagi tadqiqotlarni birlashtiradi va *Onoma* jurnalini nashr etadi; *English Place-Name Society* (EPNS) [17] Buyuk Britaniyadagi eng qadimgi markazlardan biri bo'lib, ular bir necha o'n yilliklardan buyon toponimik korpuslar ustida ishlab kelmoqda. An'anaviy korpuslardan farqli o'laroq, maxsus onomastik korpuslar yaratish bo'yicha quyidagi yirik loyihalar ham bor: *Trismegistos* (Names in the ancient world) [18] qadimgi dunyo nomlari (misr, yunon, lotin) bo'yicha eng yirik onomastik bazalardan biri. Unda 40000 dan ortiq nomlar va ularning variantlari tizimlashtirilgan; *Namenforschung.net* [14] Germaniyadagi Mainz akademiyasi tomonidan yuritiladigan loyiha bo'lib, nemis familiyalari va joy nomlarining raqamli korpusini (Digital Dictionary of Surnames in Germany) yaratishga ixtisoslashgan; *Lexicon of Greek Personal Names* (LGPN) [16] Oksford universiteti loyihasi bo'lib, unda qadimgi yunon antroponimlarining to'liq raqamli bazasi jamlangan.

O'zbek tilshunosligida N.Abdurahmonova, B. Elov, Sh.Hamroyeva [3] kabi olimlar tomonidan "O'zbek tili milliy korpusi" ning yaratilishi sohada katta burilish yasadi. Biroq, umumiy milliy korpus tarkibida hududiy va onomastik xususiyatlarni to'liq aks ettiruvchi ixtisoslashgan kichik korpuslarga ehtiyoj mavjud. Shu nuqtayi nazardan mintaqaviy onomastik korpuslarni yaratish va ularni bosqichma-bosqich Milliy korpus bazasiga havola qilish zaruratini

ta'kidlash joiz. Xorazm vohasi o'zining qadimiy irrigatsiya madaniyati va o'ziga xos dialektal leksikasi (*arna, yop, soqa*) bilan ajralib turadi. Ushbu gidronimlarning maxsus korpusini yaratish nafaqat lingvistik, balki tarixiy va kognitiv ahamiyatga ega deyish mumkin.

Tadqiqot metodologiyasi. Mazkur tadqiqotda korpus lingvistikasi tamoyillari asosida Xorazm vohasi gidronimlari jamlanib, matnlar tokenizatsiya va lemmatizatsiya jarayonlari orqali avtomatik qayta ishlandi hamda lingvistik teglash amalga oshirildi. Gidronimlar gidrografik tasnif asosida (obyekt turi va hududiy joylashuviga ko'ra) tizimlashtirilib, ularning leksik-semantik va strukturaviy xususiyatlari qiyosiy-tavsifiy hamda statistik tahlil usullarida o'rganildi. Shuningdek, korpus ma'lumotlari GIS texnologiyalari bilan integratsiyalanib, obyektlarning fazoviy vizualizatsiyasi va areal tahlili bajarildi.

Natijalar. Til korpusi – ma'lum maqsadda yig'ilgan matnlar majmuini tashkil etuvchi til birliklarining elektron shakldagi maxsus dasturiy ta'minot asosidagi yig'indisidir. Hozirgi kunda lingvistik tadqiqot va amaliy topshiriqlar yechimi uchun til korpuslari zamonaviy tilshunoslikning inkor etib bo'lmaz ish quroliga aylandi. Regional toponimika, jumladan, suv havzalari nomlarini o'rganuvchi gidronimika sohasida an'anaviy qog'oz kartotekalardan voz kechilib, ochiq informatsion tizimlar va raqamli ma'lumotlar bazalariga o'tilmoqda. Xorazm vohasi – O'rta Osiyoning qadimiy va eng yirik sun'iy sug'orish tizimlariga ega hududlaridan biri sifatida ulkan gidronimik merosga ega. Ushbu boy madaniy-tarixiy merosni elektron korpus shaklida tizimlashtirish bugungi kun filologiyasi va kiberleksikografiyasi oldidagi muhim vazifalardanidir. A.V.Superanskaya ta'kidlaganidek, mintaqaviy tadqiqotlar muayyan yopiq hududda qo'llanilgan barcha onimlarni ularning dastlabki qayd etilishidan to bugungi kungacha o'rganishni maqsad qiladi [13.217]. Onomastik birliklar korpusi va ularning raqamli bazasini yaratish bo'yicha dunyo miqyosida juda keng ko'lamli tadqiqotlar olib borilmoqda. Bugungi kunda bu soha nafaqat tilshunoslik, balki kompyuter lingvistikasi va sun'iy intellekt (Named Entity Recognition – NER) yo'nalishlarining ajralmas qismiga aylangan.

P. Beyker ta'kidlaganidek, korpuslar tildagi yashirin qonuniyatlarini ochishga va intuitiv qarashlarni statistik dalillar bilan isbotlashga xizmat qiladi [5]. Xorazm gidronimlari korpusida ham nomlarning chastotasi va so'zlararo birikuv orqali hududning lisoniy manzarasini yaratish mumkin. Korpus hududning onomastik stratigrafiyasini, ya'ni onomastik tizimlar, jarayonlar va hodisalarining nisbiy xronologiyasini aks ettirishi kerak. Gidronimik qatlamlar turli davrlarga oid til va madaniyat qoldiqlarini o'zida tashiydi [10.44].

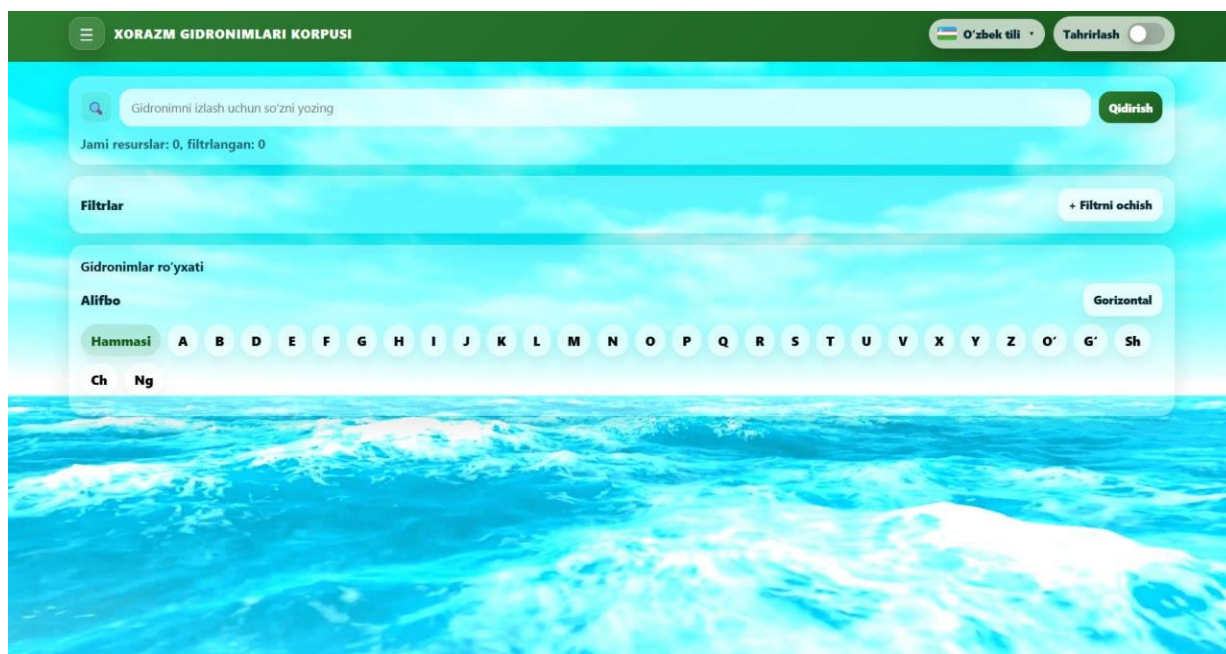
Korpusni loyihalashtirish uning reprezentativligini, ya'ni zaruriy hajm, miqdor, janr hamda materiallarning xilma-xilligini ta'minlashdan boshlanadi. Xorazm gidronimlari korpusi reprezentativ jihatdan talablarga mos bo'lishi uchun dastavval ushbu korpusni yaratishning nazariy va texnologik yondashuvlariga to'xtalamiz.

Muhokama. Ma'lumki, matmlarga avtomatik ishlov berish jarayonida tokenizatsiya (matnni so'zlarga ajratish) va lemmatizatsiya

(soʻzshakllarni uning lugʻatdagi oddiy shakli – lemmaga keltirish) usullari qoʻllaniladi. Xorazm vohasi gidronimlari korpusida ham dastlabki bosqich matn tarkibidagi token va lemmalarni aniqlash hisoblanadi. Hozirgi holatda tarkibida voha gidronimlari mavjud boʻlgan 1000 dan ziyod gaplarni yigʻishga erishdik. Ular tarixiy, badiiy, ilmiy manbalardan toʻplandi. Toponimlarni toʻplash va korpusini shakllantirgan olimlar anʼanaviy alifbo tartibidan voz kechib, nomlarni hududiy-geografik bogʻliqlik asosida – daryoning quyilishidan manbasiga qarab yoxud obyektlarning suv tizimidagi joylashuviga koʻra tizimlashtirishni koʻzda tutuvchi gidrografik tamoyil korpus arxitekturasini uchun eng maqbuli hisoblanishini nazarda tutib ish yuritganlar. Buni Y.Gordovanning fikrlari ham tasdiqlaydi: “Kataloglar nafaqat koʻplab toponimik tadqiqotlar uchun muhim manbalarga aylandi, balki toponimikadagi geografik nomlar alifbo tartibida emas, balki tizimli roʻyxat gʻoyasini ishlab chiqish uchun asos boʻlib xizmat qildi [10.44]”. Shunga koʻra biz ham Xorazm gidronimlari korpusida umumiy elektron katalogni alifbo tartibida emas, geografik obyekt joylashuvi va obyekt turiga qarab tasniflashni maʼqul koʻrdik. Maʼlumki, gidronimlar suv obyekti turiga qarab alohida umumiy birlik sifatida ham turlarga ajratiladi. Xususan, potamonimlar – daryolar, soylar; limnonimlar – koʻllar va hovuzlar, suv omborlari; pelagonimlar – dengiz va uning qismlari, okeanonimlar – okeanlar va helonimlar – botqoqliklarni tashkil qiladi.

Oʻrganishlar natijasida Xorazm gidronimlarining korpus bazasi uchun qator leksik birliklar ajratildi. Matnlarda Amudaryo va uning qadimgi oʻzanlari bir necha xil nomlar bilan keladi. Masalan, manbalarda “*Moziyda, tarix silsilalarida hozirgi Amudaryo nomi, atamasi bir necha marta oʻzgartirilgandir...*” Oq daryo, Oʻgʻiz, Panj, Okus Oks, Akes, Jayhun kabi. Uning oʻzanlari Oqcha daryo, Suvyorgan, Davdon, Daryoliq nomlari bilan yuritilgan” [6.84]. Dastlabki yozma manbalarda, xususan, “Avesto”da Amudaryo bilan bogʻliq “*Men – Ahura Mazda bunyod etgan oʻn oltinchi sarzamin va goʻzal yurt Rangha daryosi sarchashmasi atrofidagi besar xalqlar yashovchi makondir*”, “*Unga qurbonlik keltirdi yaratguvchi Ahura Mazda Oriy Vaja sahrosida, Mushfiq Doityo qirgʻogʻida...*”, “*Mujda beraman va olqishlayman pokiza suv, rad ashavan Ardvini*”, “*Farogʻ Kart daryosi kengliklarida tugʻyon boshlandi*” [4] kabi ifodalar uchraydi va bu gidronimlar Amudaryoning qadimgi nomlari ekanligi ayon boʻladi. Bunday leksik xilma-xillik korpusning xronologik va diaxronik oʻzgarishlar modulida saqlanadi.

Voha gidronimlari korpusida leksikografik natijalar ham nazarda tutilgan boʻlib, albatta, elektron lugʻat yaratish qoidalari va imkoniyatlaridan kelib chiqib, Xorazm gidronimlari elektron lugʻatini ham korpus tarkibiga kiritish maqsad qilingan. Unda alifbo tartibida materiallar joylanadi va qidiruvni osonlashtirish uchun lotin yozuviga asoslangan oʻzbek alifbosi asosida gidronimni bildiruvchi birlik filtrlab olinishi mumkin.



Korpus leksikografiyasi – maʼlumotlarni saqlash, uzatish va analiz qilish kabi vazifasini bajarishga yoʻnaltirilgan kompyuterining muayyan fazasida koʻp qoʻllaniladigan matn yoki akustik korpusi asosida tilni oʻrganish sohasi hisoblanadi [9.46]. Gidronimlarni turli lingvistik va ekstralingvistik jihatdan oʻrganishning eng maqbul usuli sifatida mazkur gidronimlar lugʻatidan foydalanish hamma uchun manfaatli boʻlishi kutilmoqda. Lugʻatda gidronimlar kelib chiqishiga koʻra etimologik, morfologik va semantik xususiyatlariga koʻra teglanadi.

Jahon tajribasi shuni koʻrsatadiki, zamonaviy toponimik va gidronimik bazalar ochiq kodli Geoinformatsion tizimlar (GIS) orqali elektron xaritalarga birlashtirilishi shart. Ushbu yondashuv Y.Zaxarova tadqiqotlarida ham taʼkidlangan [12]. Xorazm gidronimlari korpusi obyektlarni elektron xaritalarda (masalan, OpenStreetMap tizimida) vizual koʻrish, areal makon tahlillarini oʻtkazish imkoniyatini taqdim etadi. Har bir obyekt “Hudud”, “Obyekt turi”, “Rasmiy nomlanishi”, “Tarixiy manbalarda uchrashi” va “Xalq talqini” kabi metamaʼlumotlar orqali xaritaga bogʻlanadi. Bu xususiyat foydalanuvchilarga Xorazmdagi suv inshootlari, jumladan, qadimgi oʻzanlar evolyutsiyasini, mavjud koʻllar yoki suvi qurib boshqa xil yer maydonlariga aylantirilgan hududlar haqida muayyan tasavvurga ega boʻlishga, tahlil qilishga koʻmak beradi. Korpus lingvistikasi boʻyicha salmoqli tadqiqotlarni olib borgan olimlar fikricha “Korpusdan tashqari yana qaysi axborot banki savollarga tez, oson javob berishi mumkinligini tasavvur etib boʻlmaydi” [8.17]. Demak, Xorazm vohasi gidronimlari bazasi shunday shakllanishi kerakki, ushbu sohada barcha masalalar korpusda oʻz aksini topgan boʻlishi zarur.

Onomastik materialni xaritalash va shular asosida mintaqaviy toponimik atlasni yaratish ham bir xil darajada muhim va dolzarb vazifadir. Onomastik xarita onomastik faktlar va hodisalarni fazoviy tahlil qilish, ularning tarqalish zonalarini va mintaqalar ichida va oʻrtasidagi aloqalari haqida maʼlumot berish, umuman olganda, katta hajmdagi yangi va muhim maʼlumotlarni taqdim etish imkonini beradi [10]. Gidronimik korpus qadimgi xalqlarning migratsiyasi hamda iqlim oʻzgarishlarini oʻrganishda ulkan ahamiyatga ega. S.P.Tolstov, B.V.Andrianov kabi olimlarning tadqiqotlariga koʻra, mil.avv. II ming yillik oxirlarida Janubiy Dovdon, Qangadaryo qurib qolganligi hamda Oqchadaryo oʻzanida qadimgi Xorazm sivilizatsiyasi shakllanganligi tarixiy-toponimik faktlar hisoblanadi. Korpus orqali bunday joy nomlarini arxeologik madaniyatlar xaritalari bilan tahlil qilish usuli muvaffaqiyatli amalga oshiriladi.

Gidronimlar, ayniqsa kichik daryolar, irmoqlar, botqoqliklar va koʻllar nomlari (mikrotoponimlar) eng tez oʻzgaruvchan guruhdir. O.Vintseler taʼkidlaganidek, “mikrotoponimik material u yoki bu hudud tarixi uchun eng qimmatli kapitaldir” va uni kech boʻlmasdan yigʻish kerak. Korpus anʼanaviy ogʻzaki nomlarni yoʻqolib ketishdan saqlab qoladi. Barcha kichik ariqlar, koʻllar (masalan, Ulli Shoʻrkoʻl, Devonkoʻl) va daryo boʻylaridagi hududlar nomlarini korpuslashtirish milliy merosni yoʻqolishdan saqlaydi. Rasmiy nom sifatida saqlanib qolmagan, mahalliy xalq tomonidan nom berilgan koʻpgina suv obyektlari nomlari ham korpus orqali yuzaga chiqadi.

Korpusning matn kontenti va uning tuzilishini loyihalash auditoriya foydalanuvchilarining maqsadiga moslashtiriladi. Shu oʻrinda, tilning imkoniyati hisobga olinsa, eng katta hajmdagi korpus

ham tilni mukammal darajada har jihatdan o'rganish uchun asos bo'lmashligi mumkinligini ta'kidlash o'rinli [2:159]. Darhaqiqat, yaratiladigan lingvistik korpuslar butunlay til hodisalarini va tildagi barcha so'z va so'zshakllarini qamrab ololmaydi. Qachonki, tilshunoslikning turli sohalari bo'yicha maxsus yoki subkorpuslar yaratilsa, bunday korpuslar umumiy bazaga birlashtirilsa, shundagina muayyan natijaga erishish mumkin.

Xulosa. Gidronimik korpus suv obyektlari nomlarining matndagi qo'llanilishi, grammatik shakllari va semantik xususiyatlarini o'zida jamlagan axborot tizimidir. "Xorazm gidronimlari korpusi" loyihasi o'zbek tilshunosligida axborot-qidiruv imkoniyatlarini kengaytiruvchi

FOYDALANILGAN ADABIYOTLAR

1. Abjalova M. va boshqalar. Kompyuter lingvistikasi: O'quv qo'llanma. – Toshkent, 2021.
2. Abdurahmonova N. O'zbek tili electron korpusining kompyuter modellari. – Toshkent, 2021. – 213 b.
3. N. Abdurahmonova, B. Elov, Sh. Hamroyeva, Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti "KOMPYUTER LINGVISTIKASI: MUAMMOLAR, YECHIM, ISTIQBOLLAR" V Xalqaro ilmiy-amaliy konferensiya Vol. 1 №. 01 (2025) 209.
4. Авесто. Аскар Маҳкам таржимаси. – Тошкент: Шарқ, 2001. - 383 б.
5. Olga N. Prokhorova . Corpus analysis in the study of the genitive constructions formed with hydronyms as the target component in English and Russian. XLinguae, Volume 16 Issue 1, January 2023, ISSN 1337-8384, eISSN 2453-711X 224.
6. Umid Bekmuhammad. Esabiy va Gurlan tarixnomasi. – Xorazm, 2025. – 351 b.
7. Xamroqulova S. O'zbekistonning milliy toponimik geoaxborot bazasini yaratish. "Zamonaviy dunyoda tabiiy fanlar. Nazariy va amaliy izlanishlar" nomli onlayn konferensiya materiallari. – B.84.
8. Zaxarov V., Mengliyev B. Korpus lingvistikasi. O'quv qo'llanma – Toshkent, 2021. – 167 b.

muhim leksik va hududiy ma'lumotlar bazasiga aylanadi. Hidrografik tamoyillarga, maxsus lingvistik teglash hamda NLP dasturlari va Geoinformatsion tizimlarga (GIS) asoslangan bunday loyiha mutaxassis tilshunolar, tarixchilar, geograflar hamda keng jamoatchilik uchun Xorazmning boy madaniy-tarixiy, suv inshootlari bilan bog'liq sivilizatsiyasi haqidagi bilimlarni raqamli asrda mustahkamlashga xizmat qiladi. Shuningdek, Korpusdagi gidronimik afsonalar va haqiqatlar, xalq rivoyatlari va boshqa ma'lumotlar turizm sohasida, hududning o'ziga xos obrazini yaratishda va ta'lim sohasida keng qo'llanilishi mumkin.

9. Бархударов С.Г., Новиков Л.А. Каким должен быть учебный словарь? // Русский язык за рубежом. — 2001. — № 3. — С. 46-50.
10. Гордова Ю. Ю. Лингвистические проблемы региональной ономастики в исторической перспективе. Диссертация... доктора филологических наук. – Москва, 2018.– С.44.
11. Карпович С.Н. Русскоязычный корпус текстов scdm-ru для построения тематических моделей / Труды СПИИРАН. 2015. Вып. 2(39). ISSN 2078-9181 (печ.), ISSN 2078-9599.
12. Захарова Е. В. Разработка геоинформационного ресурса по топонимии Карелии: историко-культурный аспект // Финно-угорский мир. 2022.
13. Суперанская А.В., Сталтмане В.Э., Подольская Н.В., Султанов А.Х. Теория и методика ономастических исследований / Отв. ред. А.П. Непокупный. Изд. 3-е. – М.: Книжный дом «ЛИБРОКОМ», 2009. с. 217]
14. <http://xn--namenforschung-uf3h.net/>
15. <https://www.iso.org/organization/9981232.html>
16. <https://www.lgpn.ox.ac.uk/>
17. <https://www.nottingham.ac.uk/research/groups/epns/>
18. https://www.trismegistos.org/ref/about_naw.php