

## O'ZBEK TILI UCHUN IMLO TAHRIRLOVCHI DASTUR YARATISHNING BOSQICHLARI

Sobirova Nazira G'anijon qizi

ToshDO'TAU tayanch doktoranti

+998901584815

Sobirova Zarnigor G'anijon qizi

+998903900196

<https://doi.org/10.5281/zenodo.13879682>

**Annotatsiya:** Hozirgi kunda odamlar o'zlarining haqiqiy hayotini turli xil biznes ilovalari: matn muharrirlari, kontentni boshqarish tizimlari (CMS), mijozlar bilan munosabatlarni boshqarish tizimlari (CRM) va boshqalardan foydalanmasdan, xabar almashish, qidirish, ko'rib chiqish va foydalanishsiz tasavvur qila olmaydi. Ularning barchasida odamlar matn terishda xatolarga yo'l qo'yishadi, so'zlarning to'g'ri yozilishini unutishadi, ba'zan grammatikadan noto'g'ri va hokazo. Yaxshi dasturiy ta'minot bu xatolarni eng oson va samarali tarzda tuzatishga yordam beradi. Ushbu maqolada aynan shu muammoga yechim bo'layotgan matnlarni tahrirlovchi dasturlar haqida so'z yuritiladi.

**Kalit so'zlar:** NLP, sun'iy intellekt, avtokorrektor, speller, tahrirlovchi dastur, lingvotexnologiya, lingvotizim

**Annotation:** Today, people can imagine their real life without using messaging, searching, browsing and using various business applications: text editors, content management systems (CMS), customer relationship management systems (CRM), etc. can't. In all of them, people make typos, forget the correct spelling of words, sometimes make mistakes in grammar, etc. A good software will help you fix these errors in the easiest and most effective way. This article talks about text editing programs that are the solution to this problem.

**Key words:** NLP, artificial intelligence, autocorrector, speller, editing program, language technology, language system

**Аннотация:** Сегодня люди не могут представить свою настоящую жизнь без использования обмена сообщениями, поиска, просмотра и использования различных бизнес-приложений: текстовых редакторов, систем управления контентом (CMS), систем управления взаимоотношениями с клиентами (CRM) и т. д. не могут. Во всех из них люди делают опечатки, забывают правильное написание слов, иногда допускают грамматические ошибки и т. д. Хорошее программное обеспечение поможет вам исправить эти ошибки самым простым и эффективным способом. В этой статье рассказывается о программах редактирования текста, которые являются решением этой проблемы.

**Ключевые слова:** НЛП, искусственный интеллект, автокорректор, орфография, программа редактирования, языковая технология, языковая система.

Dunyodagi og'zaki va yozma tillar xilma-xil bo'lib, o'ziga xos grammatik tuzilishga ega, bu esa ularni boshqarishni murakkablashtiradi, noqulayliklar kelib chiqishiga olib keladi. Grammatika, ayniqsa, samarali muloqot va ma'lumotlarni uzatish uchun juda muhimdir va yozma matnda yozuvchilar tomonidan turlicha uslub qoidasida qo'llaniladi. Ularning turli tillari va uslublari qoidalarini o'rganish algoritmlar va mashinalar uchun juda qiyin vazifadir.

Yaxshiyamki, bugungi kunda tabiiy tilni qayta ishlash (NLP) texnikasi mavjud bo'lib, bu muammoni hal qilish imkonini beradi.

NLP- sun'iy intellektning kichik bir sohasi bo'lib, kompyuterlar uchun inson tillarini tahlil qilish, tushunish va ma'no chiqarish jarayonidir. Kompyuterlar va inson tillari o'rtasidagi o'zaro aloqani ta'minlaydi. Ba'zi NLP ilovalari mavjud: avtomatik imloni tuzatish, hissiyotlarni tahlil qilish, soxta yangiliklarni aniqlash, mashina tarjimasini, savol va javoblar (savol-javob), chatbot va boshqalar ... Kompyuter va foydalanuvchi o'rtasida turli tillar uchun yaratilgan tizimlar sirasiga ***mashina tarjimasini, kompyuter yordamida tilga o'qitish hamda lingvistik vazifalarni bajaruvchi tizimlar korpusi, orfografik tekshiruvchi dasturlar*** kiradi.

Imloni tekshiruvchi programmlar *orfokorrektorlar* yoki avtokorrektorlar, shuningdek, spellerlar yoki *spel-checkerlar* deb ataladi. Kompyuter tekshiruvchi paytida matndagi so'zlar tizimli lug'atlarda allaqachon mavjud bo'lgan yoki maxsus tuzilgan ro'yxatlardan yaratilgan so'z shakllari bilan taqqoslanadi. Agar so'z topilmasa - bu qandaydir sabablarga ko'ra matn muharririning lug'atlariga kiritilmaganligini yoki xatoga yo'l qo'yganligini anglatadi. Ikkala holatda ham speller dastur foydalanuvchiga bunday "yangi" so'zga e'tiborini qaratishni taklif qiladi (masalan, qizil to'lqinsimon chiziq bilan ta'kidlaydi).

Matn protsessorlariga singib ketgan imlo tekshiruvchi har qanday matnni qayta ishlash dasturi yoki vositasi uchun zarur bo'lgan asosiy vazifalar sifatida xatolarni aniqlash va tuzatishda yordam beradi.

Matnni avtomatik tahrir qilishga bo'lgan ehtiyojning o'sib borishi natijasida imloni avtomatik tekshirish jarayoni shaxsiy EHM dasturiy ta'minotidan veb-brauzerlarga ko'chirilgan bo'lib, Internet tizimida ishlaydi.[1]

Biz har kuni noutbuklar, uyali telefonlar, kompyuterlar va boshqalarda foydalanish mumkin bo'lgan avtotuzatish funksiyasini amalga oshirish uchun imlo tekshiruvchi algoritmlaridan foydalanishimiz mumkin. Imlo tekshiruvchi matndagi imloviy xatolarni tahlil qilish va tasdiqlash uchun ishlatiladigan vosita sifatida so'nggi yillarda uning roli diversifikatsiya qilindi va u aniqlangan imloviy xatolarni ehtimoliy tuzatishni, avtomatik tuzatish funksiyasini, hatto kontekst va grammatik takliflar asosida umuman gapni tuzatishni taklif qilish uchun ham qo'llaniladigan bo'ldi.

Avtomatik tahrirlash yo'nalishida avtomatik tahrir qiluvchi dasturlar ishlab chiqilgan, ular *ruscha, inglizcha, nemischa, fransuzcha, italyancha* matnlarni tahrirlashga yo'naltirilgan.

Matnlarni tahrirlovchi dastur *turk, tatar, qozoq, qirg'iz* tillari uchun yaratilgan. O'zbek tili uchun ham tahrirlovchi dasturning lingvistik ta'minotini yaratish bo'yicha bir qator ishlar olib borildi, ushbu jarayon davom etmoqda. Xususan, Sh.Xamroyeva, N.Abdurahmonova, D.O'rinboyeva[2],[3],[4]lar izlanishlarida matnlarni tahlil qilish bo'yicha muayyan fikrlar berilgan.

Ixtiyoriy matn tahrirlovchi dasturning algoritmi quyidagi bosqichlardan iborat bo'ladi:

Matn tarkibidagi so'zlarni kompyuter o'qib oladi;

Harbir so'z lug'atdagi imlosi bilan solishtiriladi, leksik va grammatik atributlariga ko'ra so'zlarning lug'ati tuziladi;

Ajratilgan segment birliklar morfologik va sintaktik lingvistik me'yorlarga ko'ra tekshiriladi.

### Imlo dasturlari

**Matn.uz** saytida o'zbek tilidagi matnni bepul tekshirish mumkin. Dastur xato so'zlarning to'g'ri variantlarini taklif qiladi.

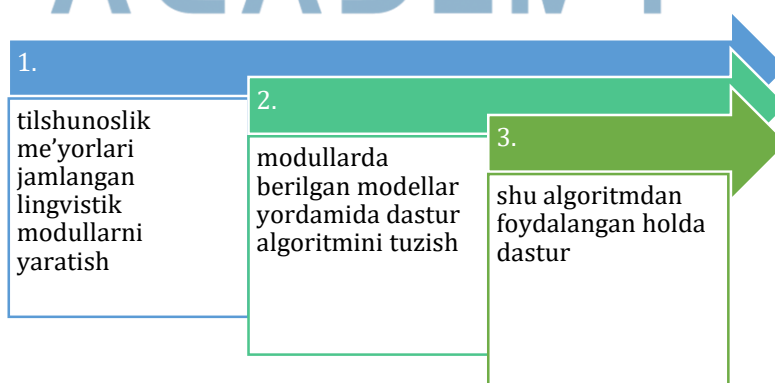
**Savodxon.uz** pullik dasturi ayni damda o'zbek tilidagi eng mukammal imlo dasturidir.

**Tahrirchi:tilmoch** - Turkiy tillar uchun maxsus yaratilgan sun'iy idrok o'zbek va qoraqalpoq tillariga tabiiy va sifatli tarjima qilishni biladi.

**Yurida-ORFO** dasturi esa pullik va faqat [Microsoft Word](#) uchun mo'ljallangan.

2014-yildan buyon YURIDA firmasi o'zbek tilining tahrirlovchi dasturi bo'yicha amaliy ishlar olib bormoqda. Ham krill, ham lotin alifbosiga asoslangan usbu dastur so'z shakllarini va imlosini tekshirish imkoniyatiga ega. Zamonaviy tahrirlovchi dasturlar singari sintaktik va semantik jihatdan tekshirish imkoniyatiga erishish uchun soha doirasida ko'plab ilmiy-amaliy ishlar amalga oshirilmogda.

Avtomatik tahrirlovchi dasturni yaratish uchun :



ni yaratishimiz kerak bo'ladi.

Dunyo tillari tahlili uchun yaratilgan dasturiy vositalar, qidiruv tizimlari (Google, Yandex, Google translate) o'zbek tili uchun mos kelmaydi. Chunki ular hind-yvropa, slavyan tillar oilasiga kiruvchi tabiiy tillar uchun ishlab chiqilgan. Har bir til o'zining ma'lum nozikliklariga ega, bir tilning morfoanalizatori, parseri, spelleri boshqa tilning qoidalariga to'g'ri kelmaydi.

O'zbek tili uchun ham Google, Yandex kabi axborot-qidiruv tizimlaridagi qidiruv, Google translatedagi tarjima sifati talabga javob bermaydi, xatoliklar judayam ko'p.

Bugun o'zbek tilidagi matnlarni tahrir qilish inson mehnati bilan amalga oshirilmoqda. Bu esa o'z navbatida matnlarni tahrirlashga ko'p vaqt sarflanishi, iqtisodiy xarajatlarning oshishi, inson tomonidan yo'l qo'yiladigan xatoliklarga sabab bo'moqda. Shu bois o'zbekcha matnlarni tahrir va tahlil qiluvchi dasturning lingvistik ta'minotini yaratish muhim ilmiy-amaliy ahamiyat kasb etadi.[5]

Xulosa qilib aytganda, zamonaviy internet davrida yuqori sifatli kontent ma'lumotlarni tarqatish va odamlarni atrofida sodir bo'layotgan muhim voqealardan xabardor qilishda yozuv muhim boylikdir. Tarkibning sifati, asosan, matn mualliflari tomonidan yo'l qo'yilgan imlo xatolari, noto'g'ri yozilgan so'zlar va grammatik xatolar bilan belgilanadi. Noto'g'ri imlo qidiruvlari tegishli ma'lumotni topishda qiyinchiliklarga olib keladi. Shu sababli ham hozirda matnlarni tahrir qiluvchi dasturlarga talab katta.

Shuning uchun ham jahondagi har bir til uchun o'zining qonun-qoidalari asosida yaratilgan lingvotizim va texnologiyalarni ishlab chiqilishi kerak. Bu esa o'z navbatida tilshunos hamda dasturchi soha vakillari oldidagi muhim masaladir. O'zbek tilining lingvistik ta'minoti qanchalik mukammal qilib yaratilsa, ishlab chiqiladigan avtomatik tahrir dasturlari ham shunchalik aniq va samarali natija beradi.

### References:

1. Сокирко А. В. Морфологические модули на сайте [www.aot.ru](http://www.aot.ru) (Электронный ресурс) / А. В. Сокирко. Режим доступа: <http://www.aot.ru/docs/sokirko/Dialog2004.htm>.; Russian Link
2. Grammar Parser [Электронный ресурс] / С. Протасов. Режим доступа: <http://sz.ru/parser/>.; Русский морфологический словарь Дмитрия Григорьева [Электронный ресурс]. Режим доступа: <http://www.aot.ru/download/LexGroup.rar>.; Парсер Mystem [Электронный ресурс] /Yandex. <http://company.yandex.ru/technology/products/mystem/mystem.xm>!
3. Xamroyeva Sh. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari: Filol.fan.bo'yicha falsafa doktori (PhD)...dis. avtoref. – Qarshi, 2018. – 53 b.
4. Abdurahmonova N. Inglizcha matnlarni o'zbek tiliga tarjima qilish dasturining lingvistik ta'minoti: Filol.fan.bo'yicha falsafa doktori (PhD)...dis. avtoref. – Toshkent, 2018. – 47 b.
5. O'rinboyeva D. Xalq og'zaki ijodi: janriylisoniy va lingvostatistik tadqiq muammolari: Filol. fan. bo'yicha dokt. (DSc) ... diss. avtoref. – Samarqand, 2019. – 74 b.
6. Abjalova M.A. Tahrir va tahlil dasturlarining lingvistik modullari /monografiya / Toshkent: Nodirabegim, 2020. – 176 b.