



PERSONAL NAMES SPELL-CHECKING – A STUDY RELATED TO UZBEK

Isroilov Jasur
Namangan state university



Erasmus+



2017 yilning fevral oyidan iyul oyigacha Erasmus+ ning almashinuv dasturi doirasida Polsha davlatining “Silesian University of Technology” universitetida ilmiy stajirovka oshirib qaytdim.

U yerda NLP(Natural language processing) bo'yicha prof. Przemyslaw Szmajda rahbarligida **“O‘zbek tilini avtomatik sintaktik taxlili”** mavzusida ilmiy izlanish olib bordim.

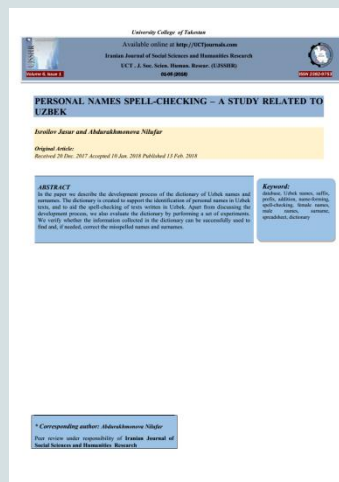
Ushbu almashinuv dasturi menga har tomonlama foydali bo'ldi. Ilmiy taqdiqot uchun kerakli materiallarni oldim, ilmiy raxbarimdan yo'nalish oldim, yuqori impact faktorli jurnallar talabiga javob beruvchi MAQOLA yozishni o'rgandim...



Qanday natijalarga erishdim?

2 ta ilmiy maqola tayyorladim va nufuzli jurnallarda chop etildi:

- ✓ **Personal names spell-checking –a study related to uzbek** (Iranian Journal of Social Sciences and Humanities Research, Iran, 01-05.2018, Impact factor=2.85)
- ✓ **O‘zbek tili lotin alifbosi uchun klaviatura yoki “unicode” masalasi** (Til va adabiyot ta’limi, 2018 yil, 6-son, OAK ro‘yhatida)





Qanday natijalarga erishdim?

Lotin alifbosiga asoslangan o'zbek alifbosidagi muammolarni yechimi uchun innovatsion g'oya tayyorladim:

Alifbodagi 5 ta qo'sh belgilar uchun UNICODE dan kod olgan holda, yangi klaviatura dasturini ishlab chiqish orqali alifboni texnik muammolarni hal etish masalasini Alisher Navoiy nomidagi o'zbek tili va adabiyoti universiteti qoshidagi ishchi guruhga taqdim etdim.



O' yoki O'
G' yoki G'



Qanday natijalarga erishdim?

O'zbek tilidagi matnlarni avtomatik sintaktik taxlil qiluvchi dastur uchun boshlang'ich algoritm yaratishga muvaffaq bo'ldim:

Ushbu dastur yordamida lotin alifbosida yozilgan matnlarni nafaqat grammatik taxlil qilib bera oladi, balki, matndagi grammatik xatolar haqida ma'lumot ham bera oladi.

O'zbek tilini kompetensiyali avtomatik o'qitish tizimi mavzusidagi loyixani tayyorladim:

Innovatsion rivojlanish vazirligi tomonidan e'lon qilingan amaliy loyixalar tanloviga O'zbek tilini kompetensiyali avtomatik o'qitish tizimi mavzusidagi 180 mln. so'mlik loyixani taqdim etdim. Hozirgi kunda ushbu loyixa ekspertlar tomonidan ko'rib chiqilmoqda.





O'zbek tilini kompyuter lingvistikasidagi muammolari



matn
taxrirllovchi
dasturlar

mashina
tarjimasi

elektron
lug'atlar

O'zbek tilini
korpusini
yaratish

avtomatik
o'qitish
tizimini
ishlab
chiqish

nutqni
avtomatik
tushunish
tizimini
ishlab
chiqish
(Automatic
Speech
Recognition)

matndan
faktlarni
avtomatik
ajratib olish
(text mining)



O'zbek tilidagi matnlarni avtomatik taxlil qilish dasturi uchun birinchi navbatda ma'lumotlar bazasini shakllantirish lozim!!!

Izlanishlar shuni ko'rsatmoqdaki, o'zbek tilida 1 000 000 dan ortiq go'shimchalar kombinatsiyasi mavjud ekan. Shuning uchun statik ma'lumotlar bazasida foydalana olmaymiz. Dasturiy ta'minotimiz uchun dinamik massivdan foydalanishimiz zarur bo'ladi. Quyidagi maqolam orqali o'zbek ismlarini ma'lumotlar bazasini tayyorlash bo'yicha yozilgan.

[Personal names spell-checking –a study related to Uzbek](#)



THANK YOUR ATTENTION!