

DIALEKTAL SO‘ZLAR VA NUTQ YOZUVLARINI SAQLASHNING DASTURIY YECHIMI

Hamroyeva Shahlo Mirdjonovna,
Filologiya fanlari doktori, professor v.b.
shaxlo.xamrayeva@navoiy-uni.uz
ToshDO‘TAU

Sanjarbek Baxodirov Rahmatali o‘g‘li,
Kompyuter lingvistikasi mutaxassisligi magistranti
sanjarbahodirov9901@gmail.com
ToshDO‘TAU

Ruxsora Muftillayeva Toshmuhammad qizi,
Kompyuter lingvistikasi mutaxassisligi magistranti
ruxsoramuftillayeva@gmail.com
ToshDO‘TAU

Annotatsiya. Ushbu maqolada Django veb-freymvorkidan foydalangan holda dialektal nutq yozuvlarini saqlash modeli tahlil qilinadi. “Recording” modeli audio fayllar, ularning transkripsiyasi, so‘zlovchining ma’lumotlari va yozib olingan joy kabi atributlarni o‘z ichiga oladi. Model dialektal tahlillar uchun qulay ma’lumotlar bazasini yaratishga imkon beradi. Maqolada modelning tarkibiy tuzilishi, maydonlarining funksional imkoniyatlari va ularning ahamiyati muhokama qilinadi. Django modellarining malumotlarni tartiblash, saqlash va qayta ishlashdagi afzalliklari ko‘rib chiqiladi. Ushbu yondashuv dialektologiya va tilshunoslik tadqiqotlarida amaliy ahamiyat kasb etadi. Modelni yanada takomillashtirish bo‘yicha tavsiyalar ham keltiriladi.

Abstract. This article analyzes a Django web framework-based model for storing dialectal speech recordings. The Recording model includes attributes such as audio files, their transcriptions, speaker information, and recording location. The model enables the creation of a structured database for dialectal analysis. The article discusses the structural composition of the model, the functionality of its fields, and their significance. It also examines the advantages of Django models in organizing, storing, and processing data. This approach has practical importance in dialectology and linguistic research. Additionally, recommendations for further improvement of the model are provided.

Аннотация. В данной статье анализируется модель сохранения диалектных речевых записей с использованием веб-фреймворка Django. Модель Recording включает в себя такие атрибуты, как аудиофайлы, их транскрипция, данные о говорящем и место записи. Модель позволяет создавать удобную базу данных для диалектных исследований. В статье рассматривается структурное строение модели, функциональные возможности ее полей и их значение. Обсуждаются преимущества использования моделей

Django для сортировки, хранения и обработки данных. Такой подход имеет практическое значение для диалектологических и лингвистических исследований. Также приводятся рекомендации по дальнейшему совершенствованию модели.

Kalit soʻzlar: *Django, dialektal nutq, maʼlumotlar bazasi, audio yozuvlar, transkripsiya, lingvistik tahlil, modellar, tilshunoslik.*

KIRISH

Nutqshunoslik, lingvistika va dialektologiya sohalarida dialektal nutqning tahlil qilinishi va hujjatlashtirilishi katta ahamiyatga ega. O'zbek tilida, mintaqaviy dialektlar va shevalar o'zining o'ziga xos fonetik, leksik va grammatik xususiyatlari bilan boydir. Ushbu xususiyatlarni to'plash va saqlash orqali tilshunoslar nafaqat tilning rivojlanishini, balki mintaqaviy o'zgarishlarni ham kuzatib borish imkoniyatiga ega bo'ladilar. Bugungi kunda bu jarayonlarni raqamlashtirish va tizimlashtirish uchun dasturiy yechimlar ishlab chiqish juda zarur.

Mazkur maqolada Django web-freymvorkidan foydalangan holda dialektal nutq yozuvlarini saqlash va tahlil qilish uchun ishlab chiqilgan “**Recording**” modeli va “**DialectWord**” modeli tahlil qilinadi. Ushbu modellar yordamida audio yozuvlar, soʻzlovchilarning ma'lumotlari, soʻzlarning mintaqaviy dialektlari, ularning transkripsiyasi, o'zbek tilining literatur turlariga mos keluvchi dialektal shakllari kabi ma'lumotlar saqlanadi.

Recording modeli, audio fayllar va ularning transkripsiyalarini saqlash imkoniyatini yaratadi, shu bilan birga, soʻzlovchining shaxsiy ma'lumotlari, yoshi, jinsi va kasbi kabi atributlarni ham kiritadi. **DialectWord** modeli esa soʻzlarning turkumlarini, mintaqaviy o'zgarishlarni va ularning o'zbek tilidagi literatur ekvivalentlarini saqlash uchun ishlatiladi. Bu modellarning birgalikda ishlashi dialektal nutqni tizimlashtirish va ilmiy tahlil qilishda katta qulaylik yaratadi.

DialectWord modelidagi **WORD_CATEGORIES** maydoni, soʻz turkumlarini belgilashda yordam beradi va har bir dialektal soʻzning grammatik xususiyatlarini tasniflash imkonini beradi. Ushbu tasniflash tizimi tilshunoslik tadqiqotlarida tilning semantik va sintaktik strukturasi haqida muhim ma'lumotlarni olishda foydalaniladi. Modeldagi boshqa maydonlar, masalan, **region**, **meaning**, va **example_sentence**, soʻzlarning mintaqaviy tafovutlari va ularning amaliy qo'llanilishiga oid ma'lumotlarni to'plashda yordam beradi.

Ushbu yondashuv, nafaqat dialektal soʻzlarni, balki ularning fonetik va semantik xususiyatlarini ham tizimlashtirish orqali o'zbek tilining boyligini va xilma-xilligini saqlashga xizmat qiladi. Maqolada keltirilgan kod va modellar, tilshunoslar va ilmiy tadqiqotchilarga mintaqaviy soʻzlar va nutq shakllarini tahlil

qilishda yordam berishi mumkin. Django modellarining qulayliklari va ularning ma'lumotlarni saqlashdagi ahamiyati chuqur tahlil qilinadi.

ASOSIY QISM

Modelning Tuzilishi. Recording modeli va uning maqsadi. Recording modeli dialektal audio yozuvlarni saqlash, tahlil qilish va foydalanuvchi ma'lumotlarini jamlash uchun ishlab chiqilgan.

```
class Recording(models.Model):
    dialect_word = models.ForeignKey(
        DialectWord,
        on_delete=models.CASCADE,
        related_name='recordings',
        null=True,
        blank=True
    )
    audio_file = models.FileField(upload_to='recordings/')
    speaker_name = models.CharField(max_length=100, default='Unknown')
    speaker_age = models.IntegerField(default=0)
    speaker_gender = models.CharField(
        max_length=10,
        choices=[
            ('male', 'Erkak'),
            ('female', 'Ayol')
        ],
        default='male'
    )
    speaker_occupation = models.CharField(max_length=100, default='Unknown')
    transcription_text = models.TextField(
        help_text="Audio matni transkripsiyasi",
        default=""
    )
    recorded_date = models.DateField(default=date.today)
    location = models.CharField(
        max_length=200,
        help_text="Yozib olingan joy",
        default='Nurota'
    )
    notes = models.TextField(
        blank=True,
        help_text="Qo'shimcha ma'lumotlar"
    )

    def __str__(self):
        if self.dialect_word:
            return f'Recording of {self.dialect_word.word} by {self.speaker_name}'
        return f'Recording by {self.speaker_name}'

class Meta:
    ordering = ['-recorded_date']
```

Ushbu model har bir yozuv haqida quyidagi ma'lumotlarni o'z ichiga oladi:

- Audio fayl
- So'zlovchi ma'lumotlari (ismi, yoshi, jinsi, kasbi)
- Transkripsiya (yozib olingan nutqning matnli shakli)
- Yozib olingan sana va joy
- Qo'shimcha izohlar

Recording modeli dialektal tadqiqotlar uchun to'plangan audio ma'lumotlarni tartiblash, saqlash va keyinchalik analiz qilish imkoniyatini beradi.

Modeldagi maydonlarning tavsifi va ularning roli. Quydagi jadvalda **Recording** modelining har bir maydoni va uning roli ko'rsatilgan: (Qarang: 1-jadval)

1-jadval. Modeldagi maydonlarning tavsifi

Maydon nomi	Ma'lumot turi	Tavsifi
dialect_word	ForeignKey	Yozuvga tegishli dialektal so'zni aniqlash uchun ishlatiladi
audio_file	FileField	Audio yozuvni saqlash va yuklash imkonini beradi
speaker_name	CharField	So'zlovchining ismini saqlaydi
speaker_age	IntegerField	So'zlovchining yoshini belgilaydi
speaker_gender	CharField (choices)	So'zlovchining jinsini (erkak yoki ayol) belgilaydi
speaker_occupation	CharField	So'zlovchining kasbi haqida ma'lumot
transcription_text	TextField	Audio yozuvning matnli shakli (transkripsiya)
recorded_date	DateField (default=date.today)	Yozib olingan sana
location	CharField	Yozib olingan joy haqida ma'lumot
notes	TextField (blank=True)	Qo'shimcha izohlar uchun maydon

Recording modeli **dialect_word** maydoni orqali **DialectWord** modeliga bog'langan. Bu bog'lanish **ForeignKey** orqali amalga oshirilgan bo'lib, har bir audio yozuv muayyan dialektal so'zga tegishli bo'lishi mumkin.

ForeignKey orqali Recording modeli **DialectWord** modeli bilan “ko‘p-birga” (many-to-one) bog'lanadi.

Ma'lumotlarni tartiblash (ordering) va `__str__` metodi

- `ordering = ['-recorded_date']` — Ushbu parametr orqali yozuvlar oxirgi yozib olingan sanadan boshlab tartiblanadi.

- `__str__` metodi — Model obyektini matn sifatida aks ettirish uchun ishlatiladi:

```
def str (self):
    if self.dialect_word:
        return f'Recording of {self.dialect_word.word} by {self.speaker_name}'
```

```
return f"Recording by {self.speaker_name}"
```

Bu funksiya natijasida ma'lumotlar ko'rinishi yanada tushunarli bo'ladi.

Kodning Tahlili. Django model maydonlarining to'g'ri tanlanganligi.

Django model maydonlari to'g'ri tanlangan va ma'lumotlarning turiga qarab mos keluvchi model maydonlari ishlatilgan. Masalan(Qarang:2-jadval)

2-jadval. Django model maydonlari(fields)

Django maydoni	Ma'lumot turi
CharField	Matn (100-200 belgi)
IntegerField	Butun son (yosh)
FileField	Fayllarni yuklash (audio)
TextField	Katta hajmdagi matn (transkripsiya)
DateField	Sana (recorded_date)

null=True, blank=True ishlatilishi va ularning farqi(Qarang:3-jadval)

3-jadval. *null=True, blank=True* ishlatilishi va ularning farqi

Parametr	Vazifasi
null=True	Baza darajasida maydon NULL qiymat olishiga ruxsat beradi
blank=True	Forma orqali to'ldirish shart emasligini bildiradi

Masalan, `dialect_word` maydonida `null=True, blank=True` belgilangan bo'lib, bu foydalanuvchi ba'zi yozuvlarda dialektal so'zni kiritmasligi mumkinligini bildiradi.

choices argumenti orqali jinsni tanlash. Django'da "choices" argumenti orqali foydalanuvchi faqat ma'lum qiymatlar ichidan tanlash imkoniga ega bo'ladi. Masalan:

```
speaker_gender = models.CharField(
    max_length=10,
    choices=[
        ('male', 'Erkak'),
        ('female', 'Ayol')
    ],
    default='male'
)
```

Bu yondashuv foydalanuvchining noto'g'ri qiymat kiritish ehtimolini kamaytiradi.

AMALIY TATBIQ

Ushbu modeldan qanday foydalanish mumkin?

- Ma'lumotlar bazasida dialektal so'zlar va ularning audio yozuvlarini saqlash
- O'zbek tilining mintaqaviy dialektlarini ilmiy tahlil qilish
- Audio yozuvlar orqali lingvistik tadqiqotlar olib borish

Ushbu model lingvistik tadqiqotlar va dialektal tahlillar uchun muhim ahamiyat kasb etadi. Uning imkoniyatlari keng bo'lib, turli sohalarda samarali qo'llanilishi mumkin. Ayniqsa, O'zbek tilining mintaqaviy xususiyatlarini chuqur o'rganish, audio materiallarni to'plash va ilmiy ma'lumotlarni tizimli tarzda saqlash imkonini beradi.

Ma'lumotlar bazasida dialektal so'zlar va ularning audio yozuvlarini saqlash. O'zbek tilining dialektal so'zlarini saqlash va ularni mintaqalar bo'yicha tasniflash tilshunoslik tadqiqotlari uchun muhim ahamiyatga ega. ***Recording*** modeli har bir dialektal so'zga tegishli ***audio yozuvni*** saqlash imkonini beradi. Bu esa tadqiqotchilarga ***so'zlarning talaffuzi, fonetik o'zgarishlari va kontekstda ishlatilishini tahlil qilish*** imkoniyatini taqdim etadi.

Django yordamida foydalanuvchilar dialektal so'zlar bazasiga **yangi yozuvlar qo'shishi**, mavjudlarini **tahrirlashi** va ularni **mintaqalar bo'yicha filtrlashi** mumkin. Har bir so'z bilan bog'liq **talaffuz namunasi (audio fayl)** ham saqlanadi, bu esa tadqiqotchilarga **real nutq namunalarini to'plash va tahlil qilishga** yordam beradi.

Bunday tizimni yaratish orqali:

- **Dialektal lug'atlar shakllanadi**
- **So'zlarning talaffuzi saqlanib qoladi**
- **Mintaqaviy shevalarning fonetik jihatlari o'rganiladi**

Misol uchun, **Recording** modelidan foydalanib, quyidagi ma'lumotlar saqlanishi mumkin:

- Toshkent shevasiga oid **“pachakilamoq”** so'zi
- Uning adabiy o'zbek tilidagi muqobili **“cho'pchak aytmoq”**
- **Audio yozuv** orqali bu so'zning talaffuzi
- **Transkripsiya va lug'aviy izoh**

O'zbek tilining mintaqaviy dialektlarini ilmiy tahlil qilish. O'zbekistonning turli hududlarida turlicha talaffuz qilinadigan va ishlatiladigan so'zlarni tahlil qilish orqali ***tilshunoslik va dialektologiya*** yo'nalishida ilmiy tadqiqotlar olib borish mumkin. Ushbu model ***so'zlarning semantik va fonetik xususiyatlarini*** o'rganish imkonini beradi.

Dialektal tadqiqotlar uchun bu model quyidagi jihatlarida qo'llanilishi mumkin:

- **Mintaqalar bo'yicha lug'aviy farqlarni aniqlash**
- **Sheva va adabiy til o'rtasidagi tafovutlarni o'rganish**
- **O'zbek tilining tarixiy rivojlanish jarayonlarini kuzatish**

Masalan, Toshkent shevasida keng tarqalgan “havo” so'zi Xorazm mintaqasida “avo” tarzida ishlatiladi. Ushbu model yordamida bu kabi farqlar **audio materiallar** orqali saqlanib, o'rganilishi mumkin.

Bundan tashqari, tadqiqotchilar **mintaqalar bo'yicha dialektal so'zlarni taqqoslab, ularning kelib chiqishiga oid xulosalar** chiqarishlari mumkin.

Audio yozuvlar orqali lingvistik tadqiqotlar olib borish. Nutqning audio shaklda saqlanishi tadqiqotchilarga *real muloqot jarayonlarini tahlil qilish, fonetik o'zgarishlarni o'rganish va talaffuz farqlarini aniqlash* imkonini beradi. Ovozli materiallardan foydalanish orqali quyidagi tadqiqotlarni amalga oshirish mumkin:

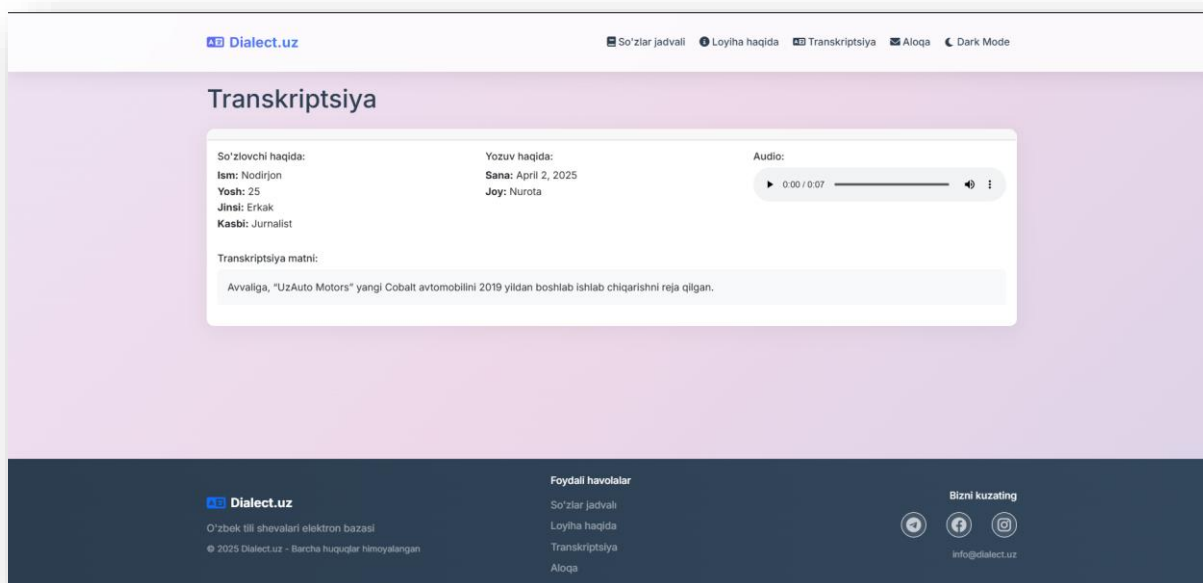
- Sheva va lahjalarning talaffuz xususiyatlarini o'rganish
- Turli yosh guruhlari va kasblarga mansub shaxslarning nutqini tahlil qilish
- Davrlar o'tishi bilan o'zbek tilining o'zgarishlarini kuzatish

Misol uchun, Recording modeli yordamida **xalq orasida keng tarqalgan iboralar va so'z birikmalarining talaffuzi** saqlanib, ularning **zamonaviy til bilan bog'liqligi** tahlil qilinishi mumkin.

Ushbu model nafaqat **tilshunoslar**, balki **sun'iy intellekt** va **nutqni qayta ishlash** (speech processing) sohasida ishlovchilar uchun ham muhim ahamiyat kasb etadi. AI-modellar uchun **dialektal nutqni tushunish** va **turli talaffuzlarni analiz qilish** muhim vazifalardan biri hisoblanadi.

Audio fayllarni yuklash va ularni qayta ishlash imkoniyatlari. Recording modelidagi `audio_file` maydoni audio fayllarni saqlash va ularni analiz qilish imkonini beradi. Django yordamida media fayllarni yuklash va ularni qayta ishlash quyidagicha amalga oshiriladi (Qarang: 1-rasm)

```
recording = Recording.objects.create(  
    speaker_name="Ali",  
    speaker_age=30,  
    speaker_gender="male",  
    audio_file="recordings/sample.mp3"  
)
```



1-rasm. Dasturdagi ko'rinishi

Ushbu model dialektal nutqni raqamlashtirish, saqlash va ilmiy tahlil qilish imkoniyatini beradi. Django modellaridan foydalanish orqali samarali va tizimli ma'lumotlar bazasini shakllantirish mumkin.

XULOSA

Mazkur maqolada **Recording** va **DialectWord** modellarining tuzilishi, ularning ma'lumotlar bazasida qanday ishlashi va amaliy qo'llanilishi haqida batafsil tahlil qilindi. Ushbu modellar O'zbek tilining dialektal xususiyatlarini saqlash, o'rganish va tahlil qilish imkonini beradi. **Recording** modeli audio yozuvlarni saqlash va ularga tegishli til birliklarini bog'lashga xizmat qilsa, **DialectWord** modeli dialektal so'zlarni adabiy til shakllari bilan solishtirish, ularning ma'nosini va mintaqaviy xususiyatlarini aniqlashga yordam beradi.

Modeldagi maydonlarning to'g'ri tanlanishi, **“ForeignKey”** bog'lanishlarining mavjudligi va **“ordering”** parametrining qo'llanilishi ma'lumotlarni tartibli saqlashga va samarali ishlov berishga imkon yaratadi. Shuningdek, audio fayllarni yuklash va ularni qayta ishlash funksiyasi lingvistik tadqiqotlar uchun yangi imkoniyatlar ochadi.

Ushbu modellar dialektologiya, lingvistika va sun'iy intellekt sohalarida amaliy qo'llanishga ega bo'lib, tilshunoslik tadqiqotlari uchun muhim ma'lumotlar bazasini shakllantirishga xizmat qiladi. Shu bilan birga, O'zbek tilining mintaqaviy xususiyatlarini saqlash va o'rganish orqali milliy til merosimizni kelajak avlodlarga yetkazishga ko'maklashadi.



Foydalanilgan adabiyotlar:

1. Холова М. ўзбек миллий шевалари корпусини тузишнинг лингвистик асослари (Бойсун тумани “ж”ловчи шевалари мисолида):Фил. фан. бўйича фалсафа доктори (PhD)...дисс. Автореф. – Термиз, 2022.
2. http://rusling.narod.ru/qqq_corp_nonslav_other.htm
3. <http://www.corpusrural.es>
4. <https://freidok.uni-freiburg.de/proj/>
5. [Nordic Dialect Corpus \(uio.no\)](http://NordicDialectCorpus.uio.no)
6. <https://www.helsinki.fi/varieng/CoRD/corpora/Dialects/fieldwork>
7. [Региональная этнолингвистика \(ethnolex.ru\)](http://ethnolex.ru)