



MATNLARNI PLAGIATGA TEKSHIRISH DASTURIY TA'MINOTINI YARATISH.

Elov Botir Boltayevich,

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti
Kompyuter lingvistikasi va raqamli texnologiyalar
kafedrası mudiri, t.f.f.d. (PhD), dots.

Muratbekova Shoira Dilshodbek qizi,

shoira.m96@gmail.com

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti
Kompyuter lingvistikasi mutaxassisligi
1-bosqich magistranti

Annotatsiya. Plagiat - bu kimningdir kontentidan ruxsatsiz foydalanishni o'z ichiga olgan intellektual o'g'irlik va firibgarlikdir. Bu asarning asl muallifiga (tasodifan yoki ataylab) hech qanday havola bermaydi. Plagiat, shuningdek, avvalgi tarkibdan ruxsatsiz qayta foydalanishni va uni yangi ish sifatida o'tkazishni o'z ichiga oladi. Shuning uchun bugungi kunga kelib ilmiy ishlarning orginalligini tekshirishga bo'lgan ehtiyoj nihoyatda ortdi. Ushbu maqolada plagiat tekshiruvini amalga oshiruvchi dasturlar va ularning ish faoliyati tahlil qilinadi.

Abstract. Plagiarism is an intellectual theft and fraud involving the unauthorized use of someone else's content. It makes no reference to the original author of the work (accidentally or intentionally). Plagiarism also involves the unauthorized use of old content and its transfer as new work. As a result, there is a growing need to verify the authenticity of scientific work. This article analyzes the programs that perform plagiarism testing and their performance.

Kalit so'zlar: *plagiat, dastur, internet, ma'lumotlar bazasi, veb-sayt, URL manzil, kontent, algoritm.*

Keywords: *plagiarism, software, internet, database, website, URL, content, algorithm.*

KIRISH. Tilda bir mazmundagi fikrni cheksiz ko'p usullarda ifodalash mumkin. Ma'lum fikrni so'zlarni joylarda qayta joylashtirish, ularni sinonimlar bilan almashtirish va boshqalar orqali o'zgacha ko'rinishda qayta yetkazish

mumkin. Sarlavha tabiiy tilda ma'noli jumla bo'lib, uning yordamida odam yangilik mohiyatini tushunishi mumkin. Sarlavha maksimal uzunligi 100 ta belgidan iborat. Bu har qanday belgi, jumladan raqamlar, tinish belgilari va maxsus belgilar bo'lishi mumkin.

Plagiat tekshiruvi - bu hujjatni yoki uning bir qismini plagiat yoki takroriy tarkibni tekshirishi mumkin bo'lgan onlayn ilova. Ushbu ilovalar kontent yoki hujjatlarning boshqa nusxalari uchun internetni skanerlaydi. Ko'pgina bunday vositalar veb-sayt URL manzilini tekshirish yoki hujjatlardan (.rtf, ex, .doc, .pdf, .txt, .docx va .odt) tarkib qo'shish imkonini beradi.

Quyida mashhur plagiat tekshiruvlari ro'yxati xususiyatlari bilan keltirilgan. Ro'yxatda ochiq kodli (bepul) va tijorat (pullik) dasturlar mavjud.

Tashqi plagiatni aniqlash

Tashqi plagiatni aniqlash - bu shubhali hujjatda plagiatga qarshi mos yozuvlar korpusi. Bu ma'lumotlar bazasida to'plangan mos yozuvlar korpusi, foydalanuvchi (yoki mijoz) sinovdan o'tkazmoqchi bo'lgan hujjatlar to'plami bo'lishi mumkin. Bu internet manbalaridan olingan hujjatlar to'plami ham bo'lishi mumkin

Tashqi plagiatni aniqlash bosqichlari[1]:

Plagiatni aniqlashda ikki asosiy bosqich odatda tashqi tomondan ishtirok etadi. Ular:

I. Qidiruv maydonini qisqartirish: bu qidirish uchun ma'lumotnoma korpusidagi mumkin bo'lgan manba hujjatlari va unga qarshi hujjat sinovdan o'tkaziladi. Bunda qidiruv maydonini kamaytirish uchun foydalaniladigan ma'lumotnoma korpusidagi hujjatlardan shubhali hujjatda plagiatni aniqlash kerak.

II. To'liq qidiruv: bu har tomonlama qidiruvni o'z ichiga oladi. Plagiatni aniqlash uchun turli usullar qo'llanilgan.

Tilga bog'liq metodologiya N-gramlardan foydalanadi.

N-gramm tavsifi

“n-gramm n ta elementdan iborat ulashgan matn yoki nutqning berilgan ketma-ketligidan iborat. Elementlar ilovaga ko‘ra juftlashadigan fonema, bo‘g‘in, harf, so‘z yoki asos bo‘lishi mumkin. N-grammlar odatda matn yoki nutq korpusidan yig‘iladi.

“N-gramm kattaroq satrning n-belgili bo‘lagidir”. Ba‘zan bo‘shliqlar birlashtirilishi mumkin. Satrning N-grammlari turlicha yo‘llar va turli o‘lchamlar uchun qabul qilinishi mumkin. Masalan, “u odatda maktabga taksida boradi.”

1-gram: “u” , “boradi” , “odatda” , “maktabga” , “ taksida” [2]

2-gram: “u boradi” , “taksida boradi” , “maktabda boradi” , “odatda boradi”.

X plagiat tekshiruvi

Plagiat tekshiruvi X plagiat skaneridan foydalanish oson. Bu tadqiqot ishlari uchun eng yaxshi plagiat tekshiruvlaridan biri bo‘lib, mazmuni ommaviy ravishda osongina solishtirish imkonini beradi.

Xususiyatlari: bu bir nechta qidiruv tizimlari orqali plagiat ma’lumotlarini aniqlay oladigan tadqiqot ishlari uchun eng yaxshi plagiat tekshiruvlaridan biri.

Ilova kontentni hech qanday qiyinchiliksiz yonma-yon solishtirish imkonini beradi.

Plagiat tekshiruvi X ma’lumotlaringizni maxfiy saqlaydi.

Bu sizga DOCX, DOC, PDF, HTML va RTF fayllarini import qilish imkonini beradi. Ushbu ilova ingliz, fransuz, ispan, italyan va boshqalarni o‘z ichiga olgan ko‘p tillarni qo‘llab-quvvatlaydi. Siz rang kodli zichlikka asoslangan natijalarni olishingiz mumkin. 24/7 qo‘llab-quvvatlaydi.

ProWritingAid

ProWritingAid - bu sizga o‘z kontentingizni ko‘plab veb-sahifalar, akademik maqolalar, o‘ziga xoslikni ta’minlash uchun nashr etilgan ishlarni tekshirish imkonini beruvchi vositadir. Bu sizga plagiat hisobotini onlayn yoki plagiat orqali yuborish imkonini beradi.

Xususiyatlari: Bu o‘z-o‘zini tahrirlashda yordam beradi. ProWritingAid 20 dan ortiq chuqurlashtirilgan yozish hisobotlarini taqdim etadi. Siz uni word,

Chrome, Google Docs, Gmail va boshqalar bilan birlashtira olasiz. Uskunadan foydalanish oson.

Keng doiradagi fayllarni qo'llab-quvvatlaydi. Bu sizga ma'lumotlar, savollar, g'oyalar, so'zlar va boshqalarni o'z ichiga olgan ishingiz misollarini topishga yordam beradi. ProWritingAid sizga so'zlaringiz boshqaruvini saqlab qolish imkonini beradi.

Grammarly

Grammarly grammatika va plagiatni tekshirish uchun keng qo'llaniladigan vositadir. Ushbu ilova ko'plab veb-sahifalardagi plagiatni aniqlay oladi. Bu sizning matningizda takroriy tarkib bor yoki yo'qligini bilishga yordam beradi.

Xususiyatlari: Ushbu ilova kontentni ommaviy ravishda qidirish imkonini beradi. Grammarly matn onlayn matnga yoki ma'lumotlar bazasiga mos kelganda plagiat haqida ogohlantirish beradi. Bu sizga sakkiz milliarddan ortiq veb-sahifalar bilan matningizni tekshirish orqali plagiatni aniqlashga yordam beradi. U ko'plab yozish uslublari, grammatika, lug'at va sintaksis bo'yicha takliflarni taklif etadi. Grammatik tezkor hisobot kontentda qancha takrorlash muammosi borligini bilish imkonini beradi. Uskuna hujjatingiz uchun umumiy originallik ballini hisoblab chiqadi. Ushbu plagiat vositasi sizning yozishingizni maxfiy saqlaydi.

CopyScape

CopyScape bulutga asoslangan vosita bo'lib, plagiatni qidirish va kontentni o'g'irlashning turli holatlarini topish imkonini beradi. Asl kontentning URL manzilini kiritish orqali uni osongina ishlatish mumkin. Bu veb-sahifa nusxalari uchun cheksiz skanerlash imkonini beruvchi eng yaxshi bepul plagiat tekshiruvlaridan biridir.

Xususiyatlari: Kontentni to'plamda qidirish. Bu ikkita tarkibni yonma-yon taqqoslash imkonini beradi. Ushbu plagiat vositasi kontentingizdagi takroriylikni tekshirish uchun shaxsiy indeksni taqdim etadi.[3] Ishni boshqarish yoki kuzatish imkonini beradi. Plagiatni tekshirish uchun ushbu veb-sayt sizga allaqachon bilgan



boshqa saytdan natijani filtrlash imkonini beradi. U foydalanuvchilarga qulay muhitni taklif etadi. Offlayn tarkibni indekslashni qo'llab-quvvatlaydi.

Plagiat tekshiruvi plagiat kontentni aniqlaydi va ko'rsatadi. Ushbu vositalar hujjatingizni kichik bo'laklarga va iboralarga bo'lib, keyin qidiruv tizimlarida har bir iborani qidiradi. Bunday ilovalar shunga o'xshash tarkibni topish uchun kontent yoki hujjatlardan boshqa foydalanish uchun internetni skanerlaydi.

XULOSA. Xulosa qilib shuni aytish joizki, Plagiat tekshiruvi paragraflar, jumlar yoki g'oyalardagi o'xshash tarkibni aniqlash uchun murakkab xususiy algoritmlardan foydalanadi. Agar biron bir sahifada bir xil matn bo'lsa, jumla plagiat bo'ladi. Plagiatni aniqlash dasturi hujjatlarni real vaqt rejimida skanerlaydi, aniq natija beradi. Siz yaratgan kontentda oldingi mualliflarning ilhomlari bo'lishi ehtimoli yuqori. Agar siz o'z kontentingizni plagiat tekshiruvchisida tekshirmasangiz, bu muallif huquqining buzilishi bo'ladi. Bunday holda, qoidabuzarga nisbatan jinoiy ish qo'zg'atilishi mumkin. Agar kontentingiz mualliflik huquqi qonunlarini buzsa, plagiat qilishda aybdor bo'lgan shaxs sifatida sudga murojaat qilishingiz ehtimoli katta. Ko'pchilik plagiat tekshiruvi ma'lumotlar maxfiyligi qonunlariga mos keladi. Ular qat'iy siyosatga ega va xodimlar siz yaratgan tarkibga kirish huquqiga ega emasligini tekshirishadi. Ularning serverlarida saqlangan tarkibni istalgan vaqtda o'chirishingiz mumkin.

Plagiatning aniqligi siz foydalanadigan Plagiat dasturiga bog'liq. Plagiatning aniqligi ikki omil bilan belgilanadi: algoritmi va Plagiat dasturi tomonidan foydalaniladigan ma'lumotlar bazasi hajmi. Ba'zi plagiat tekshiruvchilari faqat to'g'ridan-to'g'ri plagiatni aniqlay oladi, yaxshi plagiat detektorlari esa shunga o'xshash jumlar yoki so'zlar qo'llaniladigan plagiatni ham aniqlaydi.

FOYDALANILGAN ADABIYOTLAR RO'YXATI:

- [1] <https://habr.com/ru/post/341148/>
- [2] <https://habr.com/ru/post/447068/>
- [3] <https://habr.com/ru/post/142589/>