



ALANIA DIGITAL



Цифровое будущее осетинского языка

ООО «Цифровая Алания»





Проблема

Справочник «Ethnologue» отмечает, что в мире насчитывается **около шестисот тысяч носителей осетинского языка**, тогда как общая численность осетин превышает отметку в **750 тысяч**.

Данная разница объясняется недостаточной сохранностью языка, который на сегодняшний день в атласе языков мира, опубликованном **ЮНЕСКО**, имеет статус «находящийся под угрозой исчезновения» (endangered/unsafe), хотя ещё в 2021 году именовался просто **«уязвимым» (vulnerable)**.



Критерии жизнеспособности языка

- 1 абсолютное число говорящих на языке;
- 2 процент говорящих на языке от общего числа населения;
- 3 трансгенерационная передача языка;
- 4 отношение членов сообщества к родному языку;
- 5 перемещение групп пользователей языка;
- 6 отношение правительства и институтов к языку;
- 7 качество и доступность материалов для языкового образования;
- 8 реакция на новые сферы действия и средства коммуникации.



«Смерть языка – смерть народа?» *

«Когда исчезает народный язык, – народа нет более!»

(К. Д. Ушинский)


«Для познания нравов какого ни есть народа старайся прежде всего
изучить его язык»

(Пифагор)

«Величайшее богатство народа – его язык»

(М. А. Шолохов)

* Название большой коллективной монографии, вышедшей в 2019 году при Институте этнологии и антропологии Российской Академии Наук им. Н. Н. Миклухо-Маклая.



Причины снижения значимости осетинского языка *

- 1 Ограниченная передача родного языка и традиционной культуры младшим поколениям.
- 2 Недостаточное уважение к осетинскому языку, общение на котором многие считают непрестижным.
- 3 Дефицит контента на осетинском языке, особенно детского.

* Бесолова Е. Б. и др. Вестник РУДН. Серия: Теория языка. Семиотика. Семантика. 2021. Т. 12. № 4. С. 1247–1259

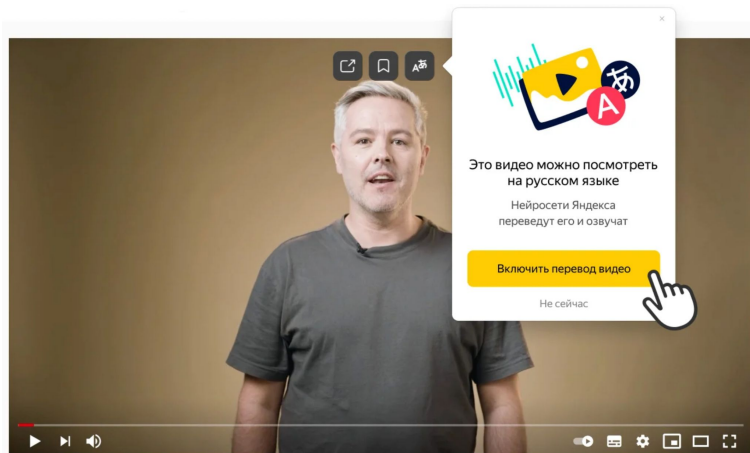


Дефицит детского контента

Просмотр мультфильмов и видеоблогов на любом языке, отличном от родного, влияет на формирование у детей определённых речевых паттернов, потому как они начинают использовать для общения язык любимых персонажей и блогеров.

Таким образом, **одной из основных проблем сохранения осетинского языка является недостаточное количество контента**, особенно детского.

Автоматический дубляж видео



Современные технологии позволяют создавать потоковый дубляж на основе искусственного интеллекта.

Так, два года назад в **Яндекс-браузер** был внедрён автоматический закадровый перевод видео.

В этом году **Youtube** введёт функцию автоматического дубляжа на нескольких языках.

Этапы цифрового развития

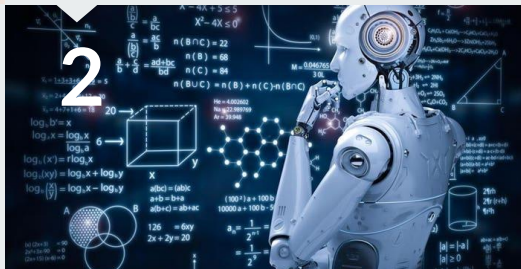


Большой электронный словарь

Формирование корпуса языка и обучение нейронной сети для машинного перевода текстов



Цифровые образовательные платформы



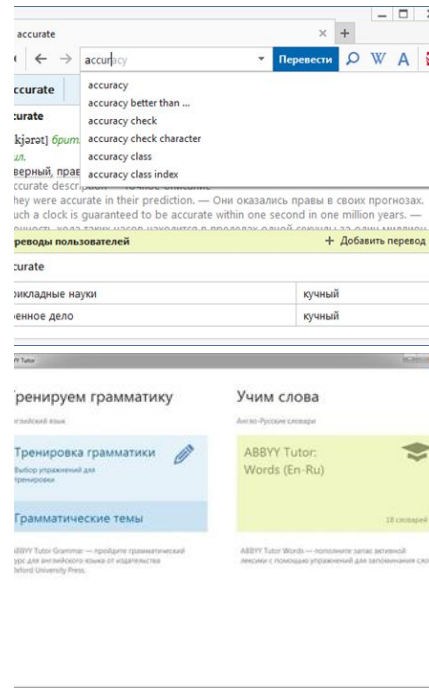


Электронный словарь

Создание максимально полного электронного словаря русско-осетинской языковой пары, представленного в формате веб и мобильного приложений с функцией внесения изменений.

Ключевой особенностью электронного словаря станет реализация сервисов API (программный интерфейс приложения) для сторонних разработчиков, что будет способствовать появлению образовательных проектов, создаваемых энтузиастами.

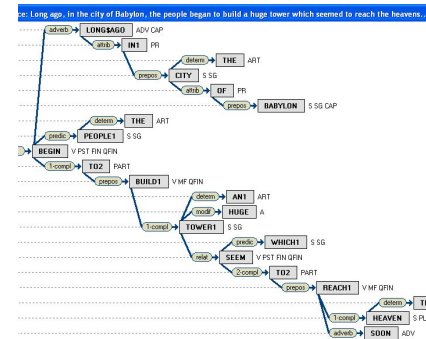
Благодаря API пользователи будут иметь доступ к единому словарю, утверждённому научным сообществом, что исключит возможность неверного перевода.



Корпус осетинского языка

Создание и развитие электронного текстово-речевого корпуса осетинского языка будет включать в себя:

- 1 | разработку концептуальной модели корпуса осетинских текстов
- 2 | формирование текстовой коллекции и ее нормализацию
- 3 | озвучивание текстовой коллекции профессиональными дикторами
- 4 | разработку методологии лингвистической аннотации
- 5 | поиск в корпусе и визуализацию результатов поиска





Источники текстов для языкового корпуса

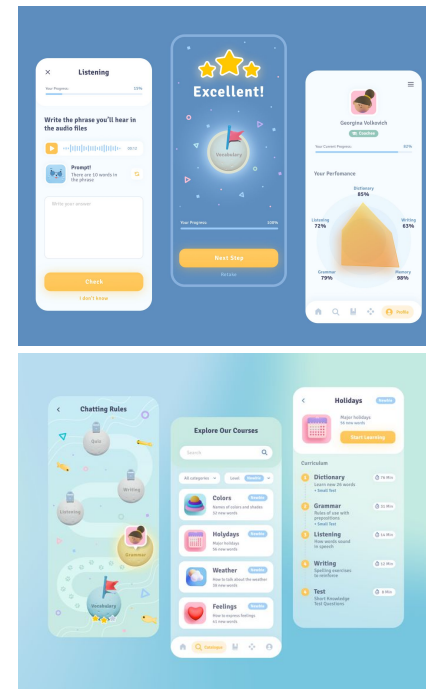
- 1 художественная литература
- 2 научно-исследовательские работы
- 3 средства массовой информации
- 4 социальные сети
- 5 кинематограф
- 6 техническая документация
- 7 публицистика
- 8 нормативные акты

Цифровые образовательные платформы

Нацелены в игровой форме способствовать развитию базовых знаний языка у детей дошкольного и школьного возраста.

Помогут выстроить обучение для взрослых, начинающих изучать язык.

Станут вспомогательным ресурсом для педагогов, учащихся дошкольных и общеобразовательных учреждений и их родителей, если будут разработаны в соответствии с тематическим планом утверждённых учебных пособий.





Дорожная карта





Команда проекта

- 1 ООО «Цифровая Алания» (аккредитованная IT-компания)
- 2 научно-исследовательская группа (СОГУ, СОИГСИ, ЛИТ ОИЯИ)
- 3 эксперты в области машинного обучения и отраслевые эксперты
- 4 разработчики и дизайнеры
- 5 лингвисты, переводчики (не менее 100 человек), методисты
- 6 профессиональные дикторы



ALANIA DIGITAL

