

Всячина:

Obsidian - <https://obsidian.md/> - хранение заметок и их связей
<https://yonote.ru/> — российский ноутбук

Трекер-менеджеры

YouTrack

https://www.jetbrains.com/ru-ru/youtrack/download/get_yourtrack.html?section=cloud —

Трекер от Jetbrains, работает только с VPN, бесплатно до 10 пользователей

Weeek <https://weeek.net/ru>

Kaiten <https://kaiten.ru/>

Текстовые редакторы

[Регулярные выражения](#) (в текстовых редакторах и др.)

[Notepad++](#) для Windows

[BBEdit](#) для macOS (бесплатная версия)

[Atom](#) для macOS и Linux

[Sublime](#) для macOS, Linux и Windows

Оформительство статей

Get glosses - <https://vantral.pythonanywhere.com/gloss> - сделать список глосс для статьи с примерами

LaTeX (<https://ru.overleaf.com/> для работы онлайн, у Антона есть платный, пишите)

Typst - <https://typst.app/> - моднее и современнее LaTeX

Программирование

Питон (<https://colab.google/> для работы онлайн)

R

<https://posit.cloud/> для работы онлайн

<https://openedu.ru/course/hse/RLING/> - курс на Открытом образовании (Г. А. Мороз)

https://agricolamz.github.io/2022_da4l/index.html - курс Анализ данных для лингвистов (Г. А. Мороз)

https://agricolamz.github.io/DS_for_DH/ - курс Наука о данных в R для программы

Цифровых гуманитарных исследований (Г. А. Мороз)

https://pozdniakov.github.io/tidy_stats/ - курс Анализ данных и статистика в R (И. Поздняков)

Клавиатуры своими руками

на мак: <https://ukelele.macupdate.com/>

на windows: [Microsoft Keyboard Layout Creator](#)

Нейросети

ChatGPT – только немного бесплатный

DeepSeek – полностью бесплатный, помог Стёпе перевести диссер с английского на русский

Claude – только немного бесплатный, решает лингвистические задачи лучше, чем ChatGPT

Phind - <https://www.phind.com/>

Gemini

Grok - <https://grok.com/>

Perplexity – академический вайб

Qwen

Яндекс API (платный?) в облаке
и др.

подробное перечисление и гайд по разным нейросетям:

<https://drive.google.com/file/d/18jd0g5BcVeysBq7SSjiONxrNFGTKMKi0/view?usp=sharing>

Можно подавать модели картинки с текстом и верный перевод и просить модель улучшить промпт, который позволит это сделать. А второй раз подавать результаты распознавания для улучшения распознавания.

Промпты для нейросетей - востребовано!

Литература

Список литературы хорошо делают нейросети, если дать им примеры и попросить не выдумывать. Просить проверять в интернете.

Хранить литературу, быстро оформлять библиографию

<https://www.zotero.org/>

Поискать литературу, которая плохо гуглится

Если не получается так:

google.com: “название” filetype:pdf

<https://www.academia.edu/>

<https://www.researchgate.net/>

Попробуйте так:

<https://annas-archive.org> (только через vpn) - включает libgen, scihub и

<https://ru.z-lib.fm/>

sci-hub чтобы скачивать по DOI

<https://i.twirpx.one>

Расширение Google Scholar PDF Reader -

<https://chromewebstore.google.com/detail/google-scholar-pdf-reader/dahenjhkooodjbpjheillca>

[dbppiidmhp?hl=en](#) - чтение академических текстов, составляет умное оглавление со ссылками, точно суммирует разделы, просто автоматический и удобный конспект

Поискать ответы на вопрос в статьях

<https://typeset.io/> - задаёте исследовательский вопрос, выдаёт статьи об этом с кратким ответом на вопрос по каждой

<https://elicit.com/> - аналогично

Узнать о статьях на схожую тематику

<https://inciteful.xyz/> - рисует граф статей, цитирующих друг друга, связанных ключевыми словами и т.д.

<https://www.connectedpapers.com/> - выдаёт похожие статьи, но платный

Разметка данных

Praat - <https://praat.softonic.ru/> - особенно для фонетики, интонации и т.д.

ELAN - <https://archive.mpi.nl/tla/elan> - элан - аудио, видео, диаризация

Elan-to-Word - <https://vantral.pythonanywhere.com/> - из разметки elan в красивый шаблон word (текст курсивом, глоссы small caps, перевод с одинарными кавычками, нумерация примеров)

ELAN-CorpA

https://llacan.cnrs.fr/res_ELAN-CorpA_en.php - мифическое улучшение элана, позволяет гlosсировать автоматически как в FieldWorks (но у Тани Казаковой не работает)

[ВидеоИнструкция](#) от коллеги Оли Кузнецовой

Saymore - <https://software.sil.org/saymore/> - прослушать, перевести, записать ещё произнесение (для носителей)

FieldWorks

<https://software.sil.org/fieldworks/> - разметка текстов, морфонологические правила, метаинформация, словарь, корпус и всё унифицировано в одном месте (но нет для мака, но нет привязки к аудио)

<https://languagedepot.org/register> - для коллективной работы в FieldWorks

Морфологический анализатор своими руками

Uniparser - <https://uniparser-morph.readthedocs.io/en/latest/> - написать правила

Трансдьюсеры (lexd и twol)

https://agricolamz.github.io/2025_morphological_transducers/ - миникурс-инструкция

<https://morfessor.readthedocs.io/en/latest/> - Тане неочевидно, как с этим работать

Корпус

Tsakorpus

<https://tsakorpus.readthedocs.io/en/latest/> - из файлов с разметкой в классический корпус

Пример тсакорпса: https://adyghe.web-corpora.net/adyghe_corpus/search

(не для хорошего онлайн корпуса)

Antconc - <https://www.laurenceanthony.net/software/antconc/> - быстрые конкордансы, коллокации и ключевые слова для коллекции текстов

Voyant - <https://voyant-tools.org/> - сайт - быстрые конкордансы, коллокации и ключевые слова для коллекции текстов

Создание словарей

GoldenDict – <https://goldendict.org/>

Lexonomy – <https://www.lexonomy.eu/>

TLex - <https://tshwanedje.com/tshwanelex/> (за пиратской версией писать Марусе Пупыниной)

в чем лучше размечать аудиофайлы и тексты для словаря?

в чём лучше делать сам словарь?

на каких платформах размещать онлайн-словари?

какие есть прям хорошие примеры онлайн-словарей?

<https://www.dargwa.ru/dictionary> - пример от Эла

Выравнивание текстов с переводами

Euclid - [ссылка](#) - табличка для выравнивания удобнее, чем два столбика в excel

Машинное обучение - <https://github.com/averkij/lingtrain-aligner>

OCR - optical character recognition

ABBYY - <https://pdf.abbyy.com/> - платно, но есть пиратские версии

Tesseract — <https://github.com/tesseract-ocr/tesseract>

Транскрибус — <https://www.transkribus.org/> (каждый месяц начисляется 50 кредитов, на которые можно распознать 50 рукописных страниц текста или 100 печатных) - гораздо лучше других для рукописных и печатных (с нестандартной орфографией) текстов

Программы, которые не используют скрытый алгоритм сжатия, который может молча изменять некоторые знаки:

<https://www.pdf-xchange.com/>

ASR - automatic speech recognition

Распознать русский

Whisper - <https://huggingface.co/bond005/whisper-podlodka-turbo> - бесплатно

Телеграм-боты для транскрибации (с разделением говорящих - оплата поминутно):

@VoiceTextAI_bot

@Voicee_AI_Bot

@GoTranscribe_BOT

@MyMeetAIBot

@GeckoVoiceBot

@charla_ru_bot

Распознать популярный язык

<https://huggingface.co/models>

Обучить ASR на своём датасете

w2v2-bert -

<https://huggingface.co/blog/fine-tune-w2v2-bert#fine-tune-w2v2-bert-for-low-resource-asr-with-%F0%A4%97-transformers>

mms - https://huggingface.co/blog/mms_adapters

whisper - <https://huggingface.co/blog/fine-tune-whisper>

и др.