

ИСПОЛЬЗУЕМЫЕ ТЕХНОЛОГИИ

CORE:

- Python 3.10+
- Gensim (Doc2Vec)
- SpaCy (NLP)
- PyMuPDF, python-docx

GUI:

- PyQt6
- Matplotlib/Seaborn

ОПТИМИЗАЦИЯ:

- Многопоточность
- Кэширование результатов
- Поточковая обработка

ОЦЕНКА:

- TF-IDF baseline
- BM25 baseline
- MAP, MRR, P@k, R@k