

😊 Datasets, итоги!

[Ask a question](#)

Что ж, это было настоящее путешествие по библиотеке 😊 Datasets — поздравляем, вы зашли так далеко! Со знаниями, которые вы получили из этой главы, вы сможете:

- Загружать наборы данных из любого места, будь то Hugging Face Hub, ваш ноутбук или удаленный сервер в вашей компании.
- Обработать свои данные, используя сочетание функций `Dataset.map()` и `Dataset.filter()`.
- Быстро переключаться между форматами данных, такими как Pandas и NumPy, с помощью `Dataset.set_format()`.
- Создавать свой собственный набор данных и отправлять его в Hugging Face Hub.
- Строить свои эмбединги документов с помощью модели Transformer и создавать семантический поисковик с помощью FAISS.

В Главе 7 мы будем использовать все это с пользой, поскольку мы углубимся в основные задачи NLP, для которых отлично подходят модели Transformer. Однако, прежде чем идти вперед, проверьте свои знания о 😊 Datasets с помощью быстрого теста!

<> [Update on GitHub](#)