deepset

NLP SUMMIT HEALTHCARE

How to Build a QA Application with Haystack in 30 Minutes

Julian Risch
Senior Machine Learning Engineer
deepset

Question Answering

- Searching through large amounts of documents is slow

Question Answering

- Searching through large amounts of documents is slow

- Searching through one document is slow

Question Answering

- "How many people are affected by cancer-related fatigue?"

- Highlight the answer in the relevant document context

- Extractive QA, no hallucination

- Running on local machine

SPECIAL ARTICLE

Cancer-related fatigue: ESMO Clinical Practice Guidelines for diagnosis and treatment[†]

A. Fabi[1], R. Bhargava[2], S. Fatigoni[3], M. Guglielmo[4], M. Horneber[5], F. Roila[3], J. Weis[6], K. Jordan[7] & C. I. Ripamonti[4], on behalf of the ESMO Guidelines Committee

[1]Division of Medical Oncology, IRCCS Regina Elena National Cancer Institute, Rome, Italy; [2]William Osler Health System, Corporate Department of Research, Department of Oncology and Division of Palliative Care, Brampton, Canada; [3]Division of Medical Oncology, Ospedale Santa Maria della Misericordia, Perugia; [4]Oncology-Supportive Care Unit, Department Onco-Haematology, Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy; [5]Department of Internal Medicine, Division of Oncology and Hematology & Division of Pneumology, Paracelsus Medical University, Klinikum, Nuremberg; [6]Department of Self Help Research in Oncology, Comprehensive Cancer Center, University Medical Center, Freiburg; [7]Department of Medicine V, Hematology, Oncology and Rheumatology, University Hospital Heidelberg, Heidelberg, Germany

Available online 12 Mar

INTRODUCTION

Fatigue is the most common symptom experienced by patients during the cancer trajectory from diagnosis to the end of life and is defined as a distressing, persistent, subjective sense of physical, emotional and/or cognitive tiredness or exhaustion related to cancer or cancer treatment that is not proportional to recent physical activity and that interferes with usual functioning.[1,2] Cancer-related fatigue (CRF) is different from other types of fatigue by its severity and persistence and the inability to alleviate it through rest or sleep. CRF affects almost 65% of patients with cancer; over two-thirds of these patients describe CRF as severe for at least 6 months and one-third reports this as persistent fatigue for a number of years after treatment.[3–5] Up to 40% of patients report fatigue at cancer diagnosis, 80%—90% during chemotherapy (ChT) and/or radiotherapy (RT), in particular 17%—21% during ChT alone and 33%—53% during association of ChT and RT.[2,3] Moreover, hormonal therapy, targeted therapy and also immunotherapy can be responsible for fatigue. CRF related to immunotherapy has
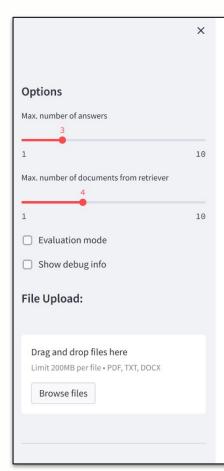
the type of tumour, stage of disease and treatment.[1,3,8]

CRF probably starts in the skeletal muscles due to a progressive reduction of physical activity (sometimes with physical interruption), but the brain is also critical as a central regulator of fatigue perception. Recently, tryptophan degradation and several cytokines and other pro-inflammatory mediators produced in response to cancer have been associated with fatigue; however, their direct role in pathogenesis of fatigue is controversial. Cytokines have been implicated in the pathophysiology of fatigue by possibly acting at multiple levels, including mood, muscle mass, strength and metabolic status. A recent review showed a positive correlation between fatigue and circulating levels of inflammatory markers; interleukin (IL)-6, IL-1 and neopterin values, in particular, were significantly associated with CRF.[1,3,8]

GENERAL PRINCIPLES OF CARE

Shared decision-making between the fatigued cancer pa-

(highlighted callout)

stence and the inability to alleviate it through rest
CRF affects almost 65% of patients with cancer;
thirds of these patients describe CRF as severe for

Question Answering

Healthcare Demo

Ask a question and see if Haystack can find the correct answer to your query!

Note: do not use keywords, but full-fledged questions. The demo is not optimized to deal with keyword queries and might misunderstand you.

| How can we prevent a central nervous system relapse? |
| 52/100 |

| Run | Random question |

Results:

...nts with aggressive B-cell lymphoma. Blood (ASH Annual Meeting Abstracts) 2014; 124: 394. 54. Tilly H, Lepage E, Coiffier B et al. Intensive conventional chemotherapy (ACVBP regimen) compared with standard CHOP for poor-prognosis aggressive non- Hodgkin lymphoma. Blood 2003; 102: 4284–4289. 55. Abramson JS, Hellmann M, Barnes JA et al. Intravenous methotrexate ANSWER as central nervous system (CNS) prophylaxis is associated with a low risk of CNS recurrence in high-risk patients with diffuse large B-cell lymphoma. Cancer 2010; 116: 4283–4290. 56. Cheah CY, Herbert KE, O'Rourke K et al. A multicentre retrospective comparison of central nervous system prophylaxis strategies among patients with high-r...

Relevance: 80.74 - **Source:** Diffuse Large B-Cell Lymphoma.pdf

Options

Max. number of answers

3

1   10

Max. number of documents from retriever

4

1   10

☐ Evaluation mode

☐ Show debug info

File Upload:

Drag and drop files here

Limit 200MB per file • PDF, TXT, DOCX

Browse files

Building a QA Application

- Three docker images are all you need!

elasticsearch	haystack by deepset	Streamlit
Data	**NLP**	**UI**

Building a QA Application

- Install Docker
  - https://docs.docker.com/get-docker/

- Clone GitHub repository
  - `git clone` https://github.com/deepset-ai/haystack-demos.git

- Build and start containers
  - `docker-compose up`

Building a QA Application

Data

NLP

UI

Building a QA Application

elasticsearch

Data

haystack
by deepset

NLP

Streamlit

UI
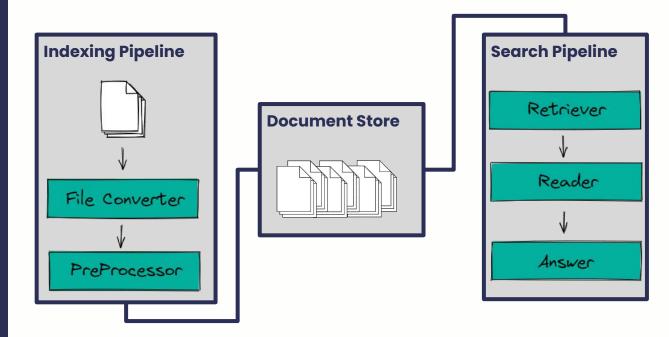
What is Haystack?
- Fully open source NLP framework
- The core NLP tasks covered
- Production focused

Integration with HF model hub
- Easy to download latest models
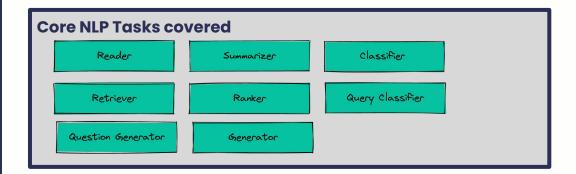
It's end-to-end, not just models
- Connector between components
- Efficient storage options
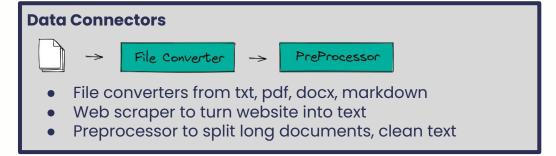
Indexing Pipeline

File Converter

PreProcessor

Document Store

Search Pipeline

Retriever

Reader

Answer

haystack
by deepset

Features of Haystack

Core NLP Tasks covered

Reader	Summarizer	Classifier
Retriever	Ranker	Query Classifier
Question Generator	Generator	

Data Connectors

📄 → File Converter → PreProcessor

- File converters from txt, pdf, docx, markdown
- Web scraper to turn website into text
- Preprocessor to split long documents, clean text

Document Stores

- Elasticsearch, OpenSearch, SQL, FAISS, Milvus, Weaviate, Pinecone, Qdrant, …

Haystack Pipelines

Document Store

Question Answering

Query
Retriever
Reader
Answer

- Clinical practice guidelines about oncology as .pdf files

- Retrieve the most relevant documents

- Extract answers from documents

Haystack Pipelines

Document Store

Question Answering

Query

Retriever

Reader

Answer

Summarization
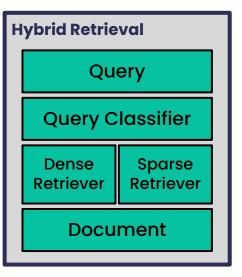
Query

Retriever

Summarizer

Document

Hybrid Retrieval

Query

Query Classifier

Dense Retriever | Sparse Retriever

Document

Haystack Pipelines
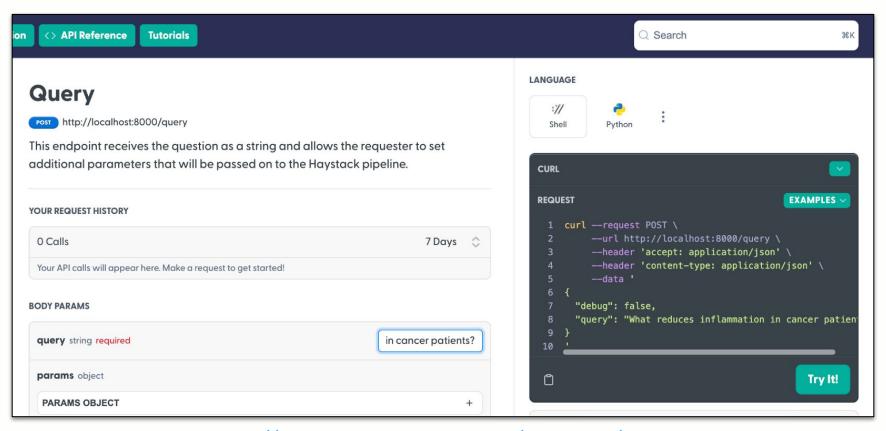
```yaml
components:
  - name: DocumentStore
    type: ElasticsearchDocumentStore
    params:
      host: localhost
  - name: Retriever
    type: EmbeddingRetriever
    params:
      document_store: DocumentStore
      embedding_model: sentence-transformers/multi-qa-mpnet-base-dot-v1
      top_k: 10
  - name: Reader
    type: FARMReader
    params:
      model_name_or_path: dmis-lab/biobert-large-cased-v1.1-squad
      context_window_size: 700
  - name: PDFConverter
    type: PDFToTextConverter
  - name: Preprocessor
    type: PreProcessor
    params:
      split_by: word
      split_length: 250
```
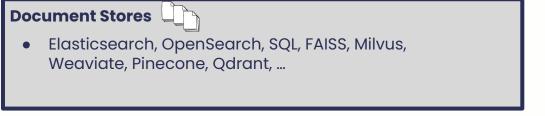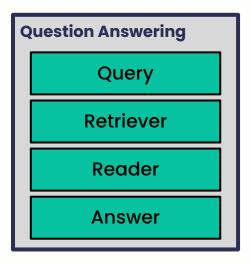
Haystack Pipelines

```yaml
pipelines:
  - name: query
    nodes:
      - name: Retriever
        inputs: [Query]
      - name: Reader
        inputs: [Retriever]
  - name: indexing
    nodes:
      - name: PDFConverter
        inputs: [File]
      - name: Preprocessor
        inputs: [PDFConverter]
      - name: Retriever
        inputs: [Preprocessor]
      - name: DocumentStore
        inputs: [Retriever]
```

Haystack API
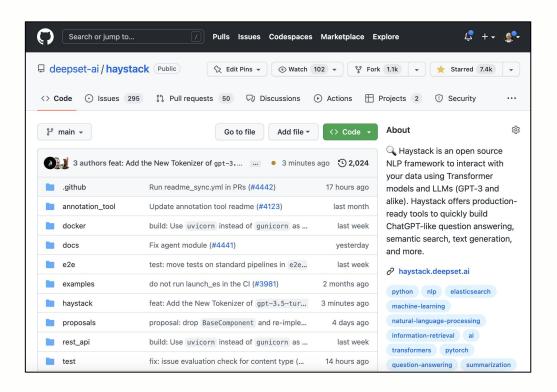

https://docs.haystack.deepset.ai/reference/query

What's Next?

- Try out different pipelines
- Collect user feedback
- Create evaluation data
- Tune pipeline parameters

- Fine-tune models
- Reduce model size
- Calibrate confidence scores
- Try QA on tabular data

deepset

Thank you!

www.deepset.ai

https://github.com/deepset-ai

Julian Risch

julian.risch@deepset.ai