

Final proj: 1st meeting

Todo:

- ☐ meeting with tutor @12.12.2023 afternoon tbd
- ☐ QA:
 - data: abstract or full content of the paper?
 - how: data analysis?
 - discuss the pipeline
 - the standard of a good answer generation
 - benchmark of evaluation

Group Project

- Milestone description (10%) ddl: 11.01.2024
- Project report + evaluation (40%)
- Code + running system (50%) ddl: 04.03.2024

Timeline

- 11.01.2024 - First milestone: **retrieval and data analysis?**
 - Contact your mentor and setup a meeting
 - Mentor will / need to talk to all team members
 - Describe what you have done so far
 - GitHub should reflect what you present
- 04.03.2024, 2pm, project deadline
 - Any commit after this time is ignored.

Topic

Question Answering (QA) system: Use the dataset provided to create a system to answer questions about our data in natural language.

eg:

Q: What are the requirements for refrigerated storage cabinets?

A: The key requirements include: Energy efficiency, Product labeling, and Technical documentation [Regulation (EU) 2015/1095]

How to Build such System?

(Step 1) (@Qiaowen Hu)

- Start with data acquisition. how? e-search, e-direct
- Collect the document subset needed.
- Find a way to properly store the data. opensearch?
- Find a way to search through the documents. 数据结构设计?
pmid, title, abstract, keywords, (author, publish_date)

(Step 2)

- Create a system that considers the entire document set or a small relevant subset to generate the correct answer.
- Mainly two stages
 - **Retrieval:** find documents that are relevant for the question. embedding+ ranking relevance to the input question? how to evaluate performance of retrieval? @Qiaowen Hu
 - **Answer generation:** generate an answer based on the selected subset. Annotated question-answer pairs, model selection, @Xiaoqing Cai

(Step 3) @Binwu Wang

- Build an interface for the user to connect with the system. 技术选择:js+flask? svelte+fastapi? restful api
- Command line in simplest form
- Website, mobile app.... (be creative!)

(Step 4) *no idea*

- Come up with an evaluation strategy
 - Make use of manual annotation. annotated in test set and **Precision, Recall, F1 Score:**
 - Find an automated annotation strategy. BLEU, ROUGE, or METEOR,
 - Use correct metrics.
- You learn about evaluation metrics in the course
- Take care of edge cases and know the limitations of your system.

<https://pubmed.ncbi.nlm.nih.gov/?term=intelligence+%5BTitle%2Fabstract%5D&filter=simsearch1.fha&filter=years.2014-2024&sort=date>

PubMed Example

Line-field confocal optical coherence tomography: new insights for psoriasis treatment monitoring

C Orsini ¹, E Trovato ¹, G Cortonesi ¹, M Pedrazzani ², M Suppa ³, P Rubegni ¹, L Tognetti ¹, E Cinotti ¹

Affiliations + expand
PMID: 37823360 DOI: 10.1111/jdv.19568

Abstract

Background: Line-field confocal optical coherence tomography (LC-OCT) is a new, valid means for a rapid and non-invasive in vivo examination of the epidermis and upper dermis, allowing digital interpretation and measurement of high-resolution images on a cellular level. Given these properties, it may represent a valid tool for monitoring psoriasis during treatment, allowing a new method to set a precise objective severity of the disease.

Objectives: We aimed to investigate the potentialities of LC-OCT in the non-invasive monitoring of microscopical changes associated with moderate-severe plaque psoriasis (PP) during the treatment with the most common biological drugs.

Materials and methods: We performed LC-OCT imaging of PP lesions from 17 patients before and after 8 weeks of treatment. The clinical severity of the single lesions was evaluated using a lesion score (LS), designed considering three parameters: erythema, desquamation, and infiltration. LC-OCT images were segmented by artificial intelligence and evaluated based on three microscopic criteria: the thickness of the stratum corneum, the thickness of the living epidermis, and the undulation of the dermo-epidermal junction.

Results: LC-OCT digital analysis allowed recognition and quantification of the three microscopic criteria, showing a reduction of all these during the follow-up. Furthermore, a high correlation between change in LS and the thickness of the stratum corneum and the thickness of the living epidermis was found.

Conclusion: LC-OCT can non-invasively monitor the response of PP to different treatments. Morphometric changes occurring in the psoriatic lesion during the 8-week treatment period were identified by in vivo LC-OCT and measured by using artificial intelligence. Although future studies are required, based on these preliminary results, LC-OCT may represent a valid potential tool for precise monitoring of therapeutic response.

- What metadata is needed?
- Only abstracts are required
- Document has multiple sections
 - Do you keep the structure?

https://eur-lex.europa.eu/search.html?name=browse-by%3Alegislation-in-force&type=named&displayProfile=allRelAllConsDocProfile&qid=1696858573178&CC_1_CODED=12

Eur-lex Example

Document 32022R0869

Regulation (EU) 2022/869 of the European Parliament and of the Council of 30 May 2022 on guidelines for trans-European energy infrastructure, amending Regulations (EC) No 715/2009, (EU) 2019/942 and (EU) 2019/943 and Directives 2009/73/EC and (EU) 2019/944, and repealing Regulation (EU) No 347/2013

PE/2/2022/REV/1

OJ L 152, 3.6.2022, p. 45–102 (BG, ES, CS, DA, DE, ET, EL, EN, FR, GA, HR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV)

in force

ELI: <http://data.europa.eu/eli/reg/2022/869/oj>

Expand all Collapse all

Languages, formats and link to OJ

HTML	BG	ES	CS	DA	DE	ET	EL	EN	FR	GA	HR	IT	LV	LT	HU	MT	NL	PL	PT	RO	SK	SL	FI	SV
PDF																								
Official Journal																								

Multilingual display

English (en) Please choose Please choose [Display](#)

Text

3.6.2022 Official Journal of the European Union L 152/45

REGULATION (EU) 2022/869 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 30 May 2022

on guidelines for trans-European energy infrastructure, amending Regulations (EC) No 715/2009, (EU) 2019/942 and (EU) 2019/943 and Directives 2009/73/EC and (EU) 2019/944, and repealing Regulation (EU) No 347/2013

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Article 172 thereof,

Having regard to the proposal from the European Commission,

After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee ⁽¹⁾,

Having regard to the opinion of the Committee of the Regions ⁽²⁾,

Acting in accordance with the ordinary legislative procedure ⁽³⁾,

Whereas:

- What metadata is needed?
- Only english language
- Structure of the document matters
 - Entire text
 - Paragraph
 - Sentences
- Choose the database according to the model used in later stages.