TopiQAL

Topic Modeling based QA Model to answer topical COVID-19 questions

Hamsa Shwetha Venkataram
Data Scientist

Jet Propulsion Laboratory
California Institute of Technology

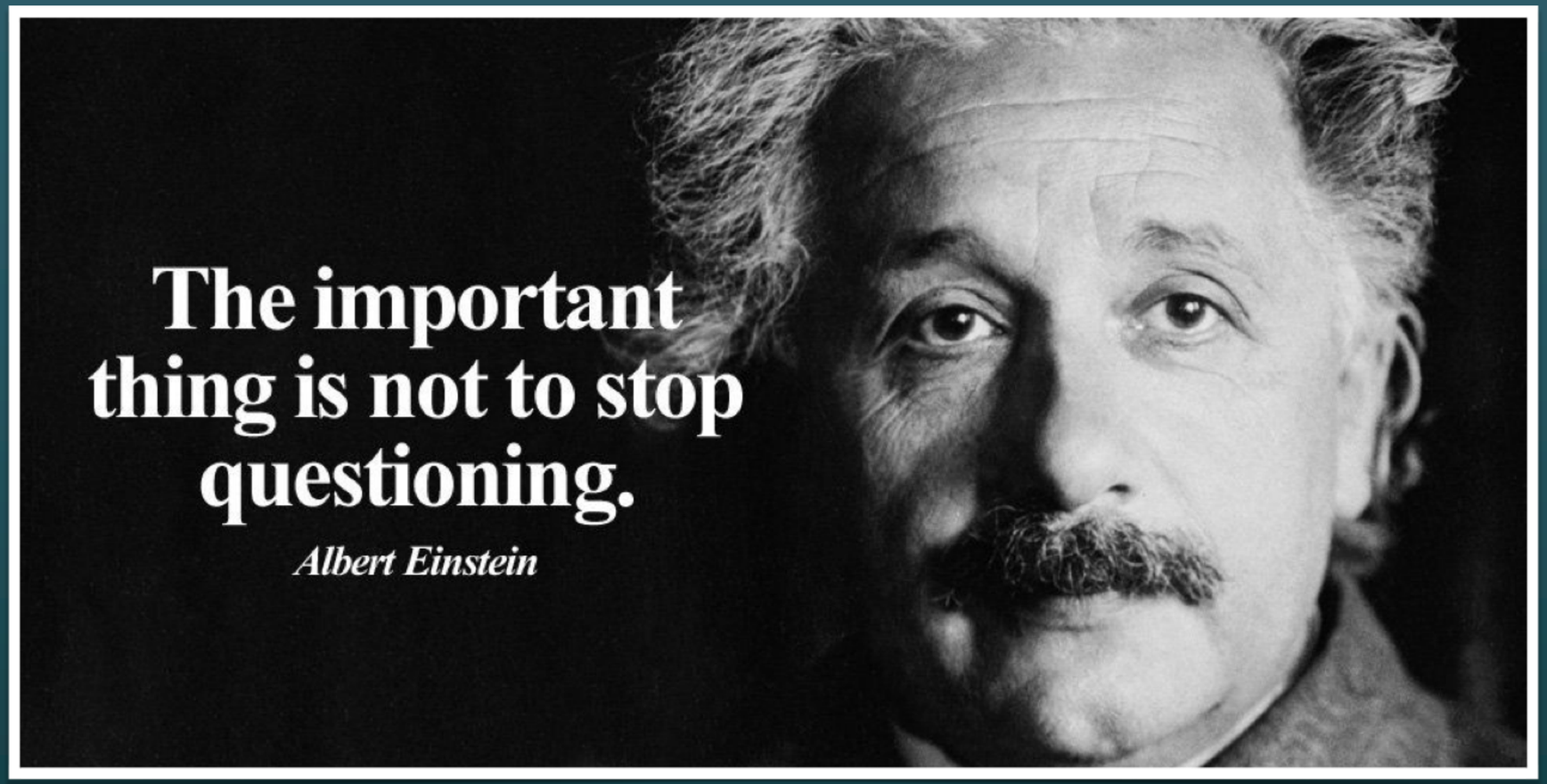
We all have questions.

Who am I ?

How big is the Universe ?

How many days in a Martian year ?

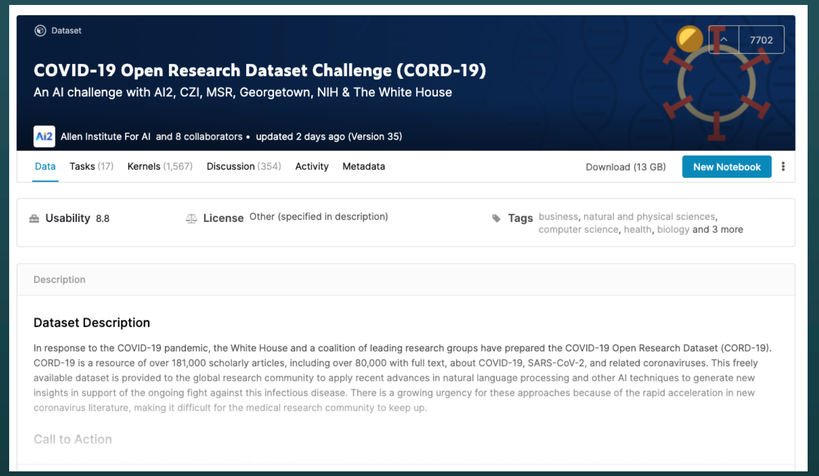
Is there a vaccine to COVID-19 ?



https://pbs.twimg.com/media/D1lyBC7X0AAYn4d.jpg

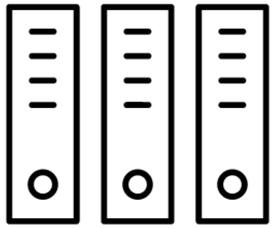
Let’s step back a bit ..

The White House and a coalition of leading research groups have prepared the COVID-19 Open Research Dataset (CORD-19). CORD-19 is a resource of over 29,000 scholarly articles, including over 13,000 with full text, about COVID-19, SARS-CoV-2, and related coronaviruses.



Challenge Submission: https://www.kaggle.com/venkat1949/topic-modeling-based-biobert-qa-system

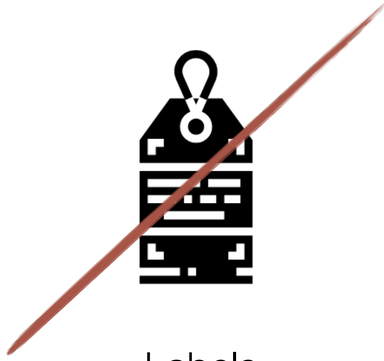
Articles



CORD-19 Dataset

Answers

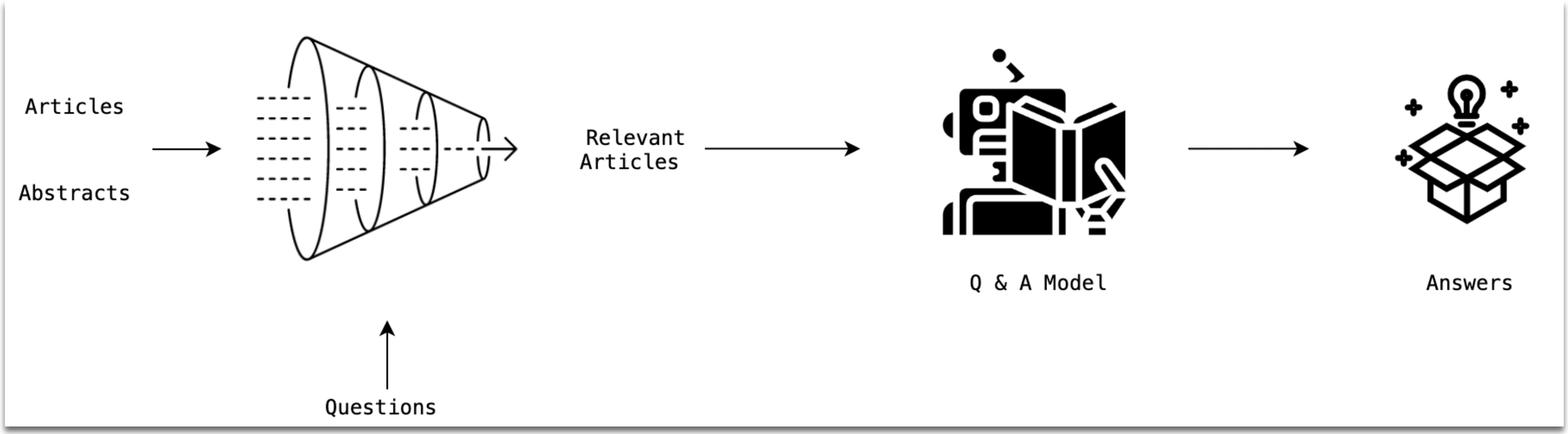
Subject Matter Experts



Labels

Tasks - Questions

Why TopiQAL ?


Articles
Abstracts
Relevant
Articles
Q & A Model
Answers
Questions

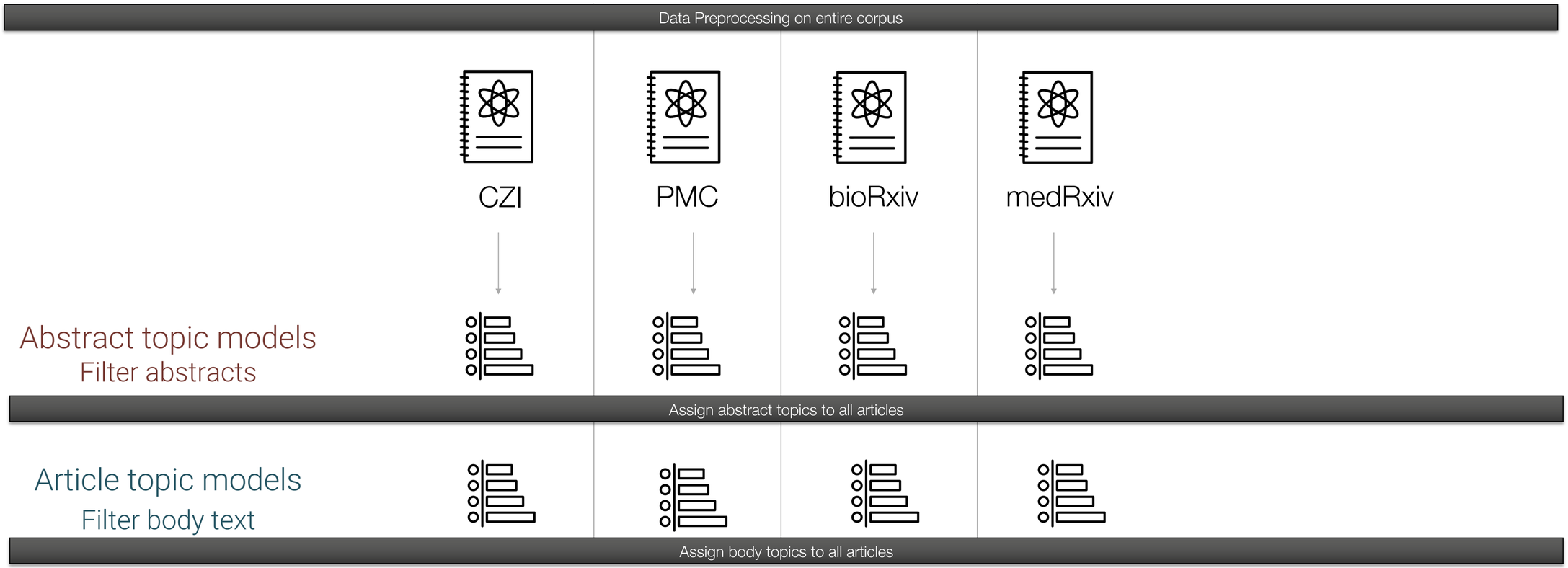
Sums up all that we do!

Topic Models

Topic Models

- Discover hidden thematic structure
- Latent Dirichlet Allocation
- Interpretability of models
- Hierarchical modeling
  - Journals
  - Abstracts
  - Articles
- Setting priors for word-topic combination
  - GuidedLDA
  - Gensim (eta)
- Evaluation
- Visualization
  - pyLDAVis

Hierarchical Topic Modeling



Each topic is seeded with words from 10 Kaggle Tasks

(Bio)BERT- SQuaD

BERT

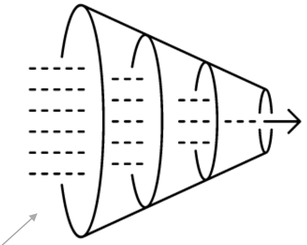
- What is BERT ?
- How can BERT answer questions ?
  - SQuaD Dataset
- BioBERT
- Fine-tuning and Inference on TPUs
- Google Colab notebooks and gsutil
- Evaluation

Inference

Inference Cycle

Question

Abstract Topic Models

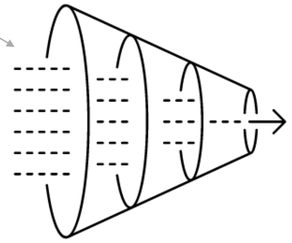


Articles

Topics

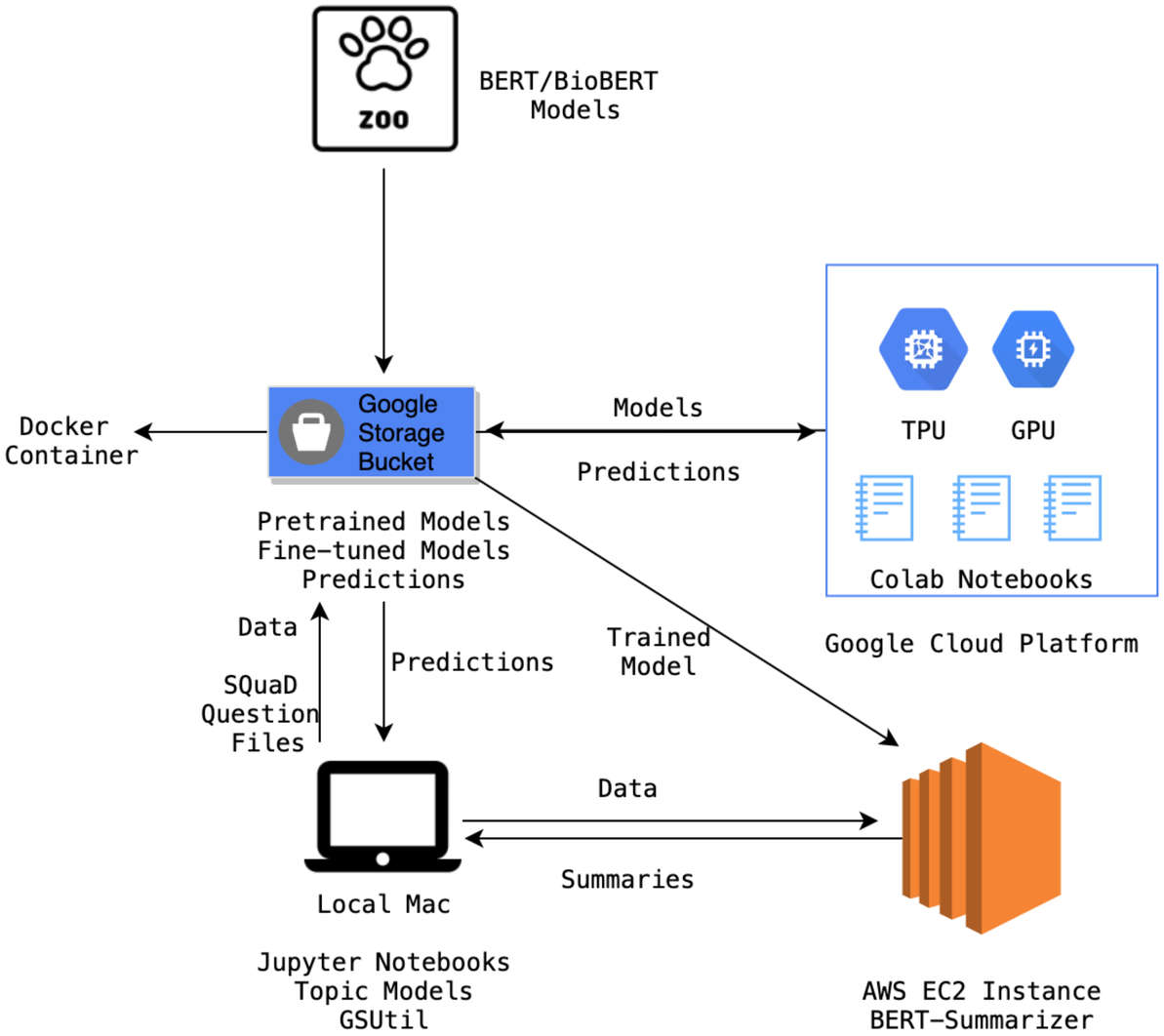
Filtered article sections

Body text topic models



Context

Architecture & BioBERT Answers



Question: Methods evaluating potential complication of Antibody-Dependent Enhancement (ADE) in vaccine recipients.
Answer: Nsp3b and E-channel

Question: Exploration of use of best animal models and their predictive value for a human vaccine.
Answer: no target has been predicted

Question: Capabilities to discover a therapeutic (not vaccine) for the disease, and clinical effectiveness studies to discover therapeutics, to include antiviral agents.
Answer: scientists have come up with three strategies for developing new drugs

Question: Exploration of use of best animal models and their predictive value for a human vaccine.

Team

Hamsa Shwetha

Annie Didier

Asitang Mishra

Anastasia Mensikova

Dr. Chris Mattmann

Thank You!

References

- Icons from Noun Project
- https://github.com/vi3k6i5/GuidedLDA
- https://github.com/nasa-jpl-cord-19/topiQAL