John Snow LABS

Applied Generative AI for Data Scientists

July 2024 Training & Certification

Training Workshop Outline

LLM & RAG with Spark NLP

Multi-Modal Language Models

Domain-Specific Language Models

Medical Chatbot

Generative AI Lab

Human-in-the-loop Workflows

John Snow Labs

is the team behind Spark NLP

100+ million

Downloads on PyPI.
“Most Widely Used NLP Library in the Enterprise.”

60% growth

In Spark NLP downloads since the 5.0 release for RAG & LLM pipelines

7 years

Straight with releases every two weeks

John Snow LABS

Spark NLP

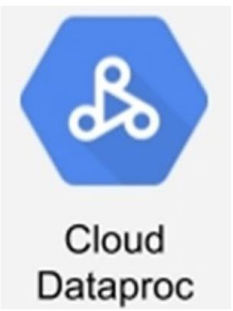
Apache 2.0 license

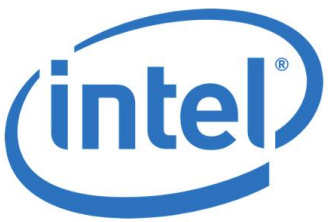
Documentation:
www.sparknlp.org

Community:
spark-nlp.slack.com

Own Curated Model Hub:
www.sparknlp.org/models

Optimized, Tested, Supported Integrations





udemy

IT & Software > IT Certifications > Natural Language Processing (NLP)

Spark NLP for Data Scientists

Unlock your NLP power with Spark NLP, the most popular NLP library in enterprises

Bestseller 4.3 ★★★★½ (57 ratings)

Spark NLP: Built for Scale

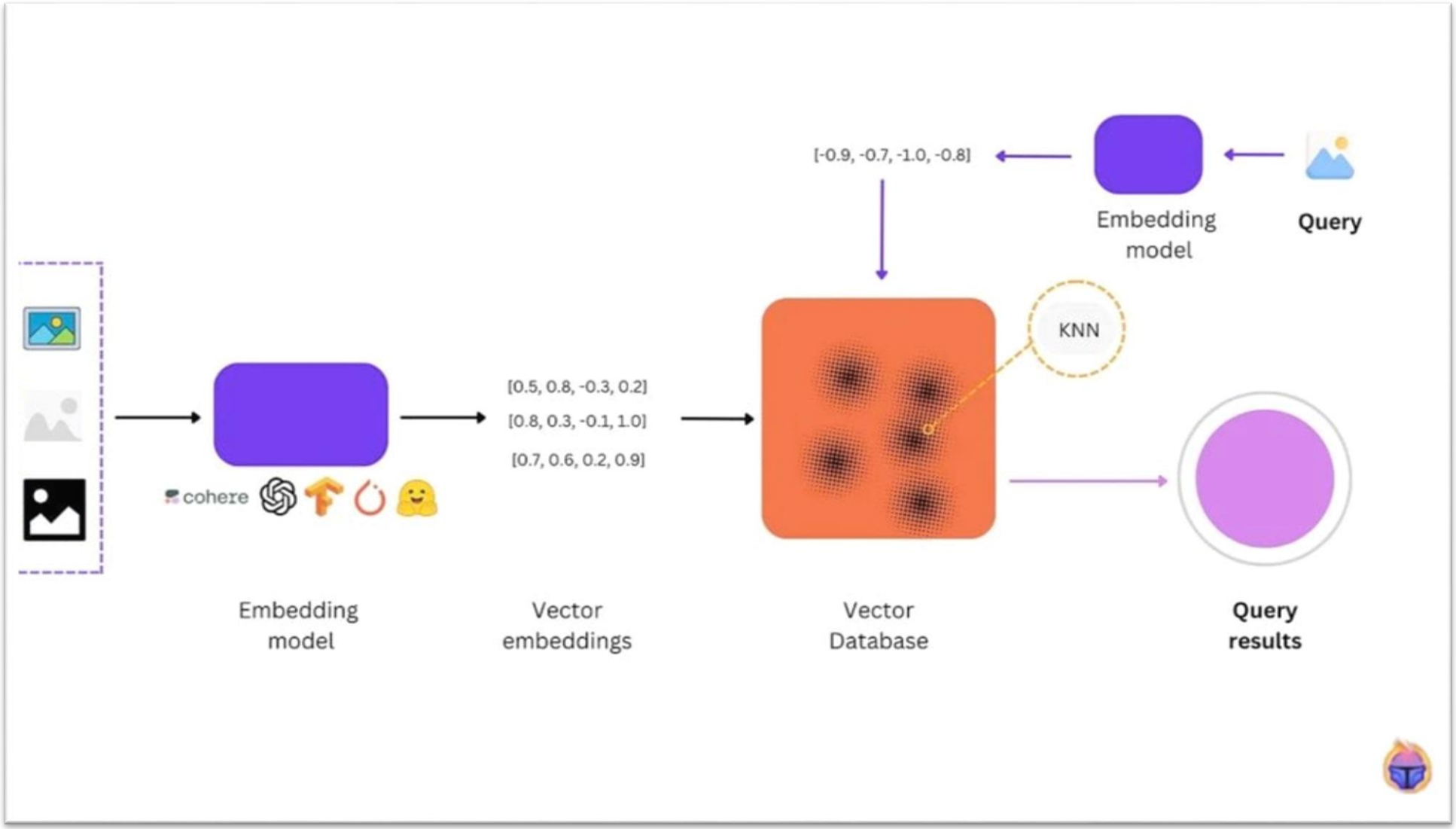
Compose multiple language models into custom pipelines

Natively scalable:
Extends the Spark ML Pipeline

Speed Optimized:
Builds for Intel, Nvidia, Apple

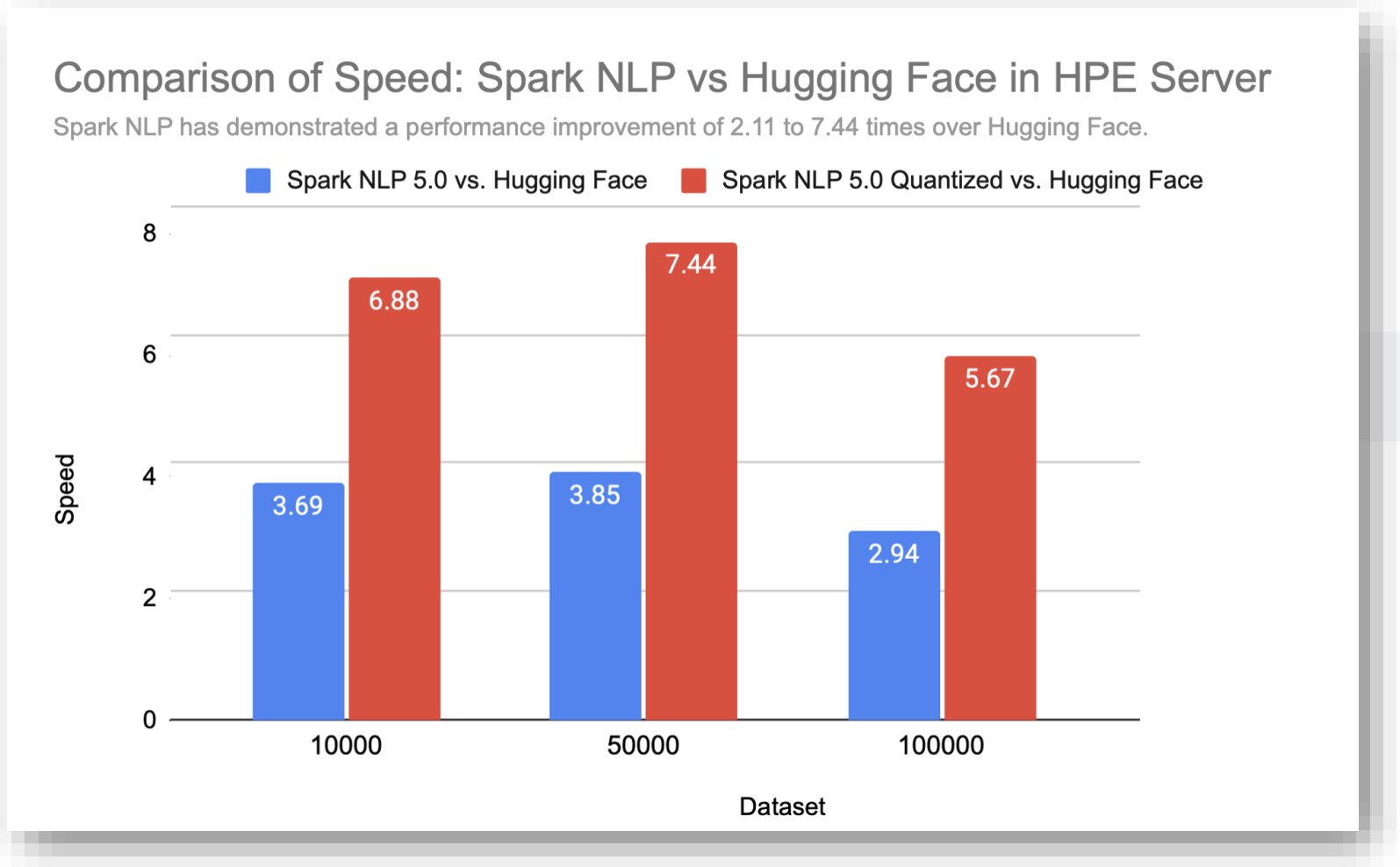
Own optimized codebase

RAG LLM Use Case: Populate a Vector Database



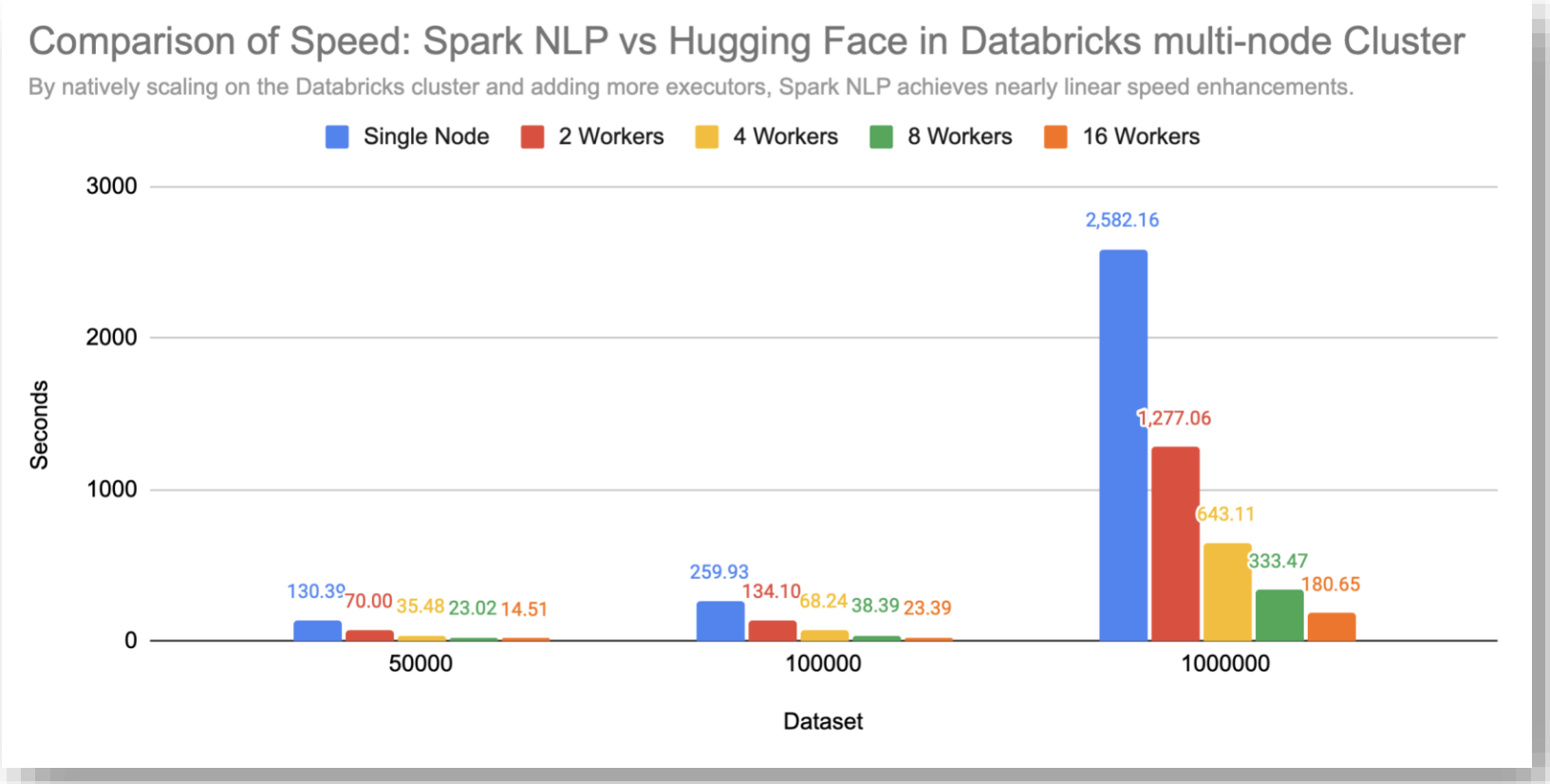
Calculate Embeddings 3x-7x Faster Than Hugging Face on a Single Server

- HPE Server:
  - AMD EPYC 7542 32-Core Processor
  - 80G memory
- Spark NLP based on Onyx Runtime vs. Hugging Face based on PyTorch



Scale Up With Zero Code Changes

- Databricks Single Node Cluster
  - 13.0 ML (includes Apache Spark 3.4.0, Scala 2.12)
  - c6i.8xlarge
  - 32 Cores
  - 64 GB Memory
- By natively scaling on the Databricks cluster and adding more executors, Spark NLP 5 achieves near-linear speedup.



Processing of 1,000,000 records was reduced
from 43 hours to 3 minutes with zero code changes

LLM Inference Use Case:
What if my text processing pipelines needs an LLM?

```
data = spark.createDataFrame([
    ("Is coffee consumption linked to improved cognitive function?",),
    ("What are the implications of quantum computing on data security?",),
    ("Describe the process of photosynthesis in plants.",),
    ("Can renewable energy effectively replace fossil fuels globally by 2040?",),
    ("What is the role of artificial intelligence in healthcare?",),
    ("How does the stock market influence the global economy?",),
    ("Explain the theory behind black holes and event horizons.",),
    ("What are the long-term effects of climate change on coastal ecosystems?",),
    ("Discuss the impact of virtual reality technology on education.",),
    ("What measures can be taken to reduce urban air pollution effectively?",)
], ["text"])
```

Question Answering, Text Generation, Summarization, Translation, Extraction

In Spark NLP, an LLM is just another pipeline step

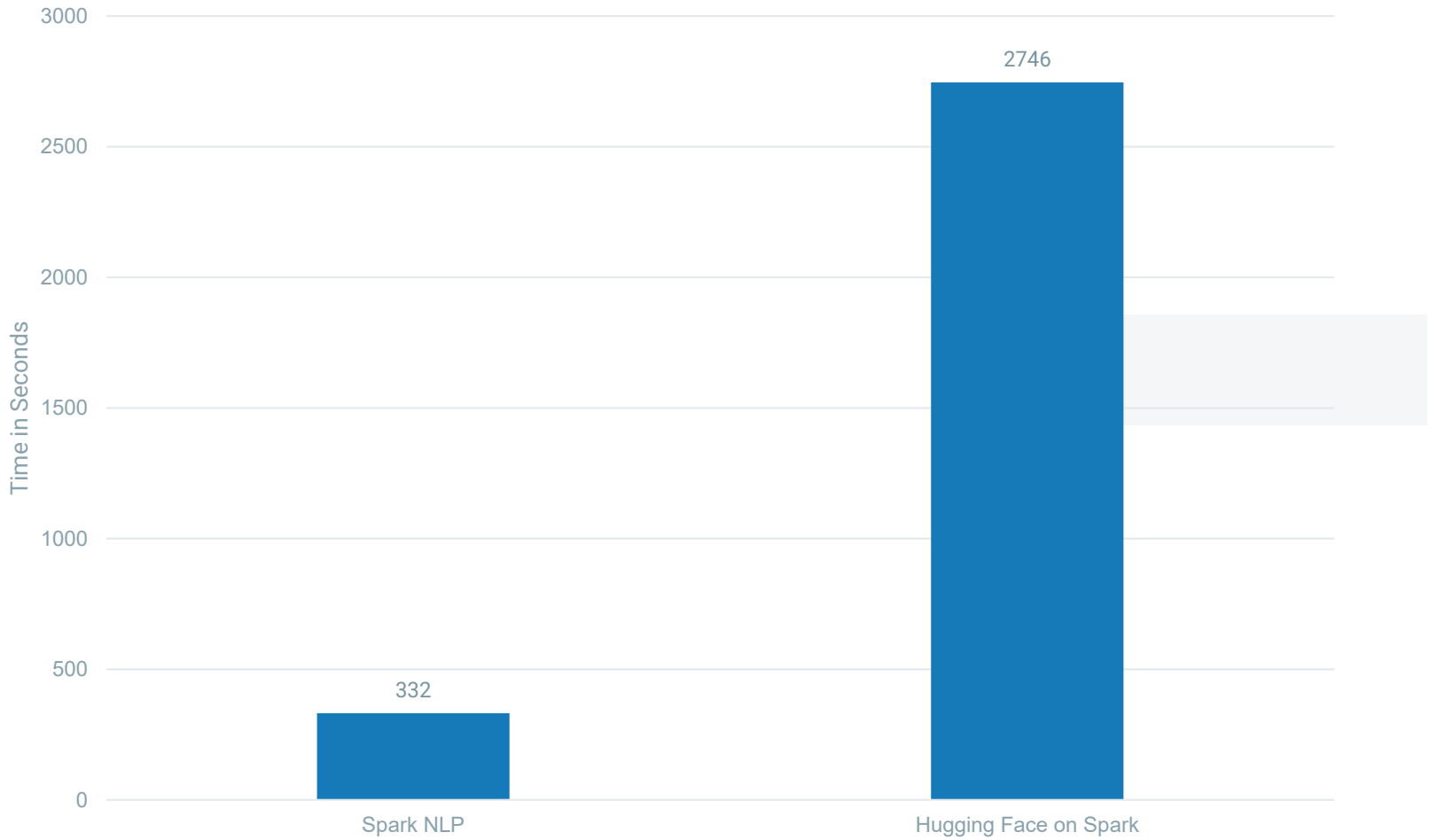
```
llama2 = LLAMA2Transformer.pretrained("llama2_openvino") \
    .setMaxOutputLength(150) \
    .setDoSample(False) \
    .setInputCols(["documents"]) \
    .setOutputCol("generation")

pipeline = Pipeline().setStages([document_assembler, llama2])
results = pipeline.fit(prompts_df).transform(prompts_df)
```

This makes it easy to distribute an LLM and process millions of questions efficiently.

Spark NLP 5: 8.2x Faster on 100 Prompts

- Server:
  - Databricks 13.3 LTS
  - Single m6id.4xlarge, 3rd Generation Intel Xeon (Ice Lake)
  - 64 GB memory
  - 16 Cores
- Dataset:
  - News summarization
    - Run on 100 rows
    - 38 tokens per row

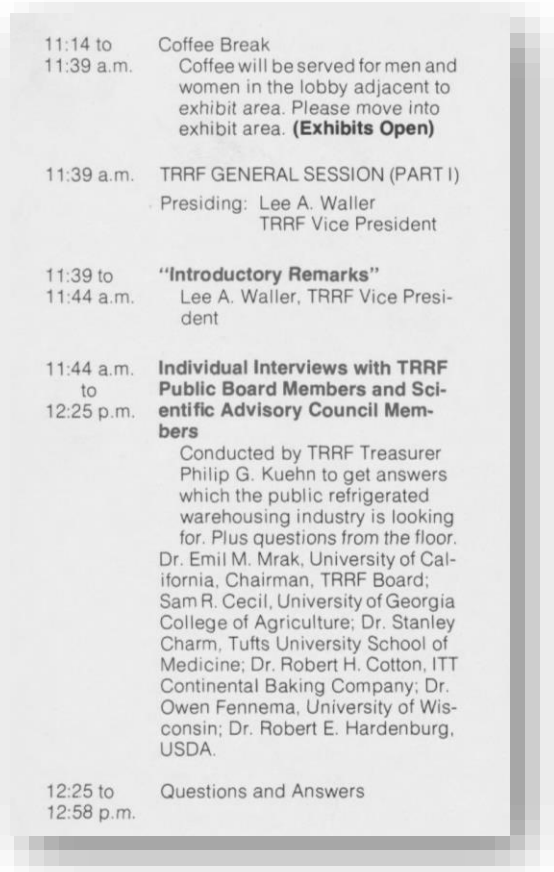


Efficient LLM Inference on CPU

- **On AWS the m6id.4xlarge server costs $0.9492 per hour on demand, or $436.54 per month on reserved pricing.**
- **This server delivers one summary in about 3.3 seconds on Spark NLP, and this scales without code changes.**
- **The software is free and open-source: Compute cost is your only cost here.**

Compute cost is the major cost driver for LLM inference.
Providing higher throughput on cheaper machines makes
small projects cheaper, and large projects viable.

Multi-Modal AI Use Case: Visual Document Understanding



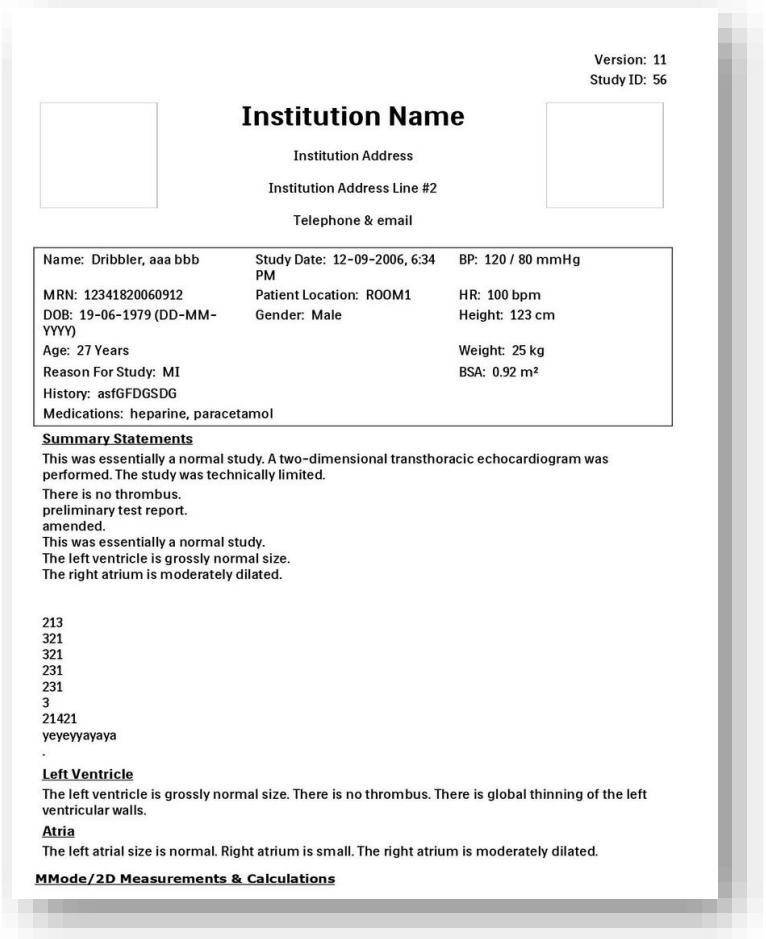
11:14 to 11:39 a.m.	Coffee Break Coffee will be served for men and women in the lobby adjacent to exhibit area. Please move into exhibit area. **(Exhibits Open)**
11:39 a.m.	TRRF GENERAL SESSION (PART I) Presiding: Lee A. Waller TRRF Vice President
11:39 to 11:44 a.m.	**"Introductory Remarks"** Lee A. Waller, TRRF Vice President
11:44 a.m. to 12:25 p.m.	**Individual Interviews with TRRF Public Board Members and Scientific Advisory Council Members** Conducted by TRRF Treasurer Philip G. Kuehn to get answers which the public refrigerated warehousing industry is looking for. Plus questions from the floor. Dr. Emil M. Mrak, University of California, Chairman, TRRF Board; Sam R. Cecil, University of Georgia College of Agriculture; Dr. Stanley Charm, Tufts University School of Medicine; Dr. Robert H. Cotton, ITT Continental Baking Company; Dr. Owen Fennema, University of Wisconsin; Dr. Robert E. Hardenburg, USDA.
12:25 to 12:58 p.m.	Questions and Answers

- Question : **When it finish the Coffee Break?**
- Answer : **11:44 to 11:39 a.m.**

- Question : **Who is giving the Introductory Remarks?**
- Answer : **lee a. waller, trrf vice presi- dent**

- Question : **Who is going to take part of the individual interviews?**
- Answer : **trrf treasurer**

Multi-Modal AI Use Case: Form Understanding



Version: 11
Study ID: 56

Institution Name

Institution Address

Institution Address Line #2

Telephone & email

Name: Dribbler, aaa bbb	Study Date: 12-09-2006, 6:34 PM	BP: 120 / 80 mmHg
MRN: 12341820060912	Patient Location: ROOM1	HR: 100 bpm
DOB: 19-06-1979 (DD-MM-YYYY)	Gender: Male	Height: 123 cm
Age: 27 Years		Weight: 25 kg
Reason For Study: MI		BSA: 0.92 m²
History: asfGFDGSDG		
Medications: heparine, paracetamol		

Summary Statements

This was essentially a normal study. A two-dimensional transthoracic echocardiogram was performed. The study was technically limited.
There is no thrombus.
preliminary test report.
amended.
This was essentially a normal study.
The left ventricle is grossly normal size.
The right atrium is moderately dilated.

213
321
321
231
231
3
21421
yeyeyyayaya
.

Left Ventricle

The left ventricle is grossly normal size. There is no thrombus. There is global thinning of the left ventricular walls.

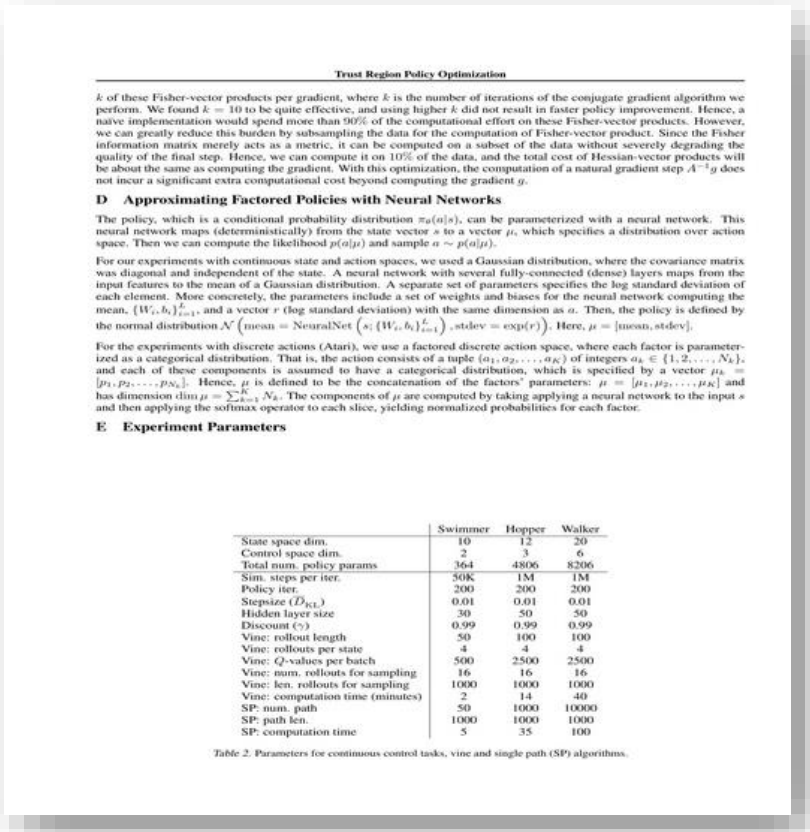
Atria

The left atrial size is normal. Right atrium is small. The right atrium is moderately dilated.

MMode/2D Measurements & Calculations

	key	value
0	Name:	Dribbler, bbb
1	Study Date:	12-09-2006, 6:34
2	BP:	120 80 mmHg
3	MRN:	12341820060912
4	Patient Location:	Location: ROOM1
5	HR:	100 bpm
6	DOB:	19-06-1979
7	Gender:	Male
8	Height:	123 cm
9	Age:	27 Years
10	Weight:	25 kg

Multi-Modal AI Use Case: Table Extraction



Trust Region Policy Optimization

k of these Fisher-vector products per gradient, where k is the number of iterations of the conjugate gradient algorithm we perform. We found $k = 10$ to be quite effective, and using higher k did not result in faster policy improvement. Hence, a naive implementation would spend more than 90% of the computational effort on these Fisher-vector products. However, we can greatly reduce this burden by subsampling the data for the computation of Fisher-vector product. Since the Fisher information matrix merely acts as a metric, it can be computed on a subset of the data without severely degrading the quality of the final step. Hence, we can compute it on 10% of the data, and the total cost of Hessian-vector products will be about the same as computing the gradient. With this optimization, the computation of a natural gradient step $A^{-1}g$ does not incur a significant extra computational cost beyond computing the gradient g.

D Approximating Factored Policies with Neural Networks

The policy, which is a conditional probability distribution $\pi_\theta(a|s)$, can be parameterized with a neural network. This neural network maps (deterministically) from the state vector s to a vector μ, which specifies a distribution over action space. Then we can compute the likelihood $p(a|\mu)$ and sample $a \sim p(a|\mu)$.

For our experiments with continuous state and action spaces, we used a Gaussian distribution, where the covariance matrix was diagonal and independent of the state. A neural network with several fully-connected (dense) layers maps from the input features to the mean of a Gaussian distribution. A separate set of parameters specifies the log standard deviation of each element. More concretely, the parameters include a set of weights and biases for the neural network computing the mean, $\{W_i, b_i\}_{i=1}^L$, and a vector r (log standard deviation) with the same dimension as a. Then, the policy is defined by the normal distribution $\mathcal{N}\left(\text{mean} = \text{NeuralNet}\left(s; \{W_i, b_i\}_{i=1}^L\right), \text{stdev} = \exp(r)\right)$. Here, $\mu = [\text{mean}, \text{stdev}]$.

For the experiments with discrete actions (Atari), we use a factored discrete action space, where each factor is parameterized as a categorical distribution. That is, the action consists of a tuple $(a_1, a_2, \ldots, a_K)$ of integers $a_k \in \{1, 2, \ldots, N_k\}$, and each of these components is assumed to have a categorical distribution, which is specified by a vector $\mu_k = [p_1, p_2, \ldots, p_{N_k}]$. Hence, μ is defined to be the concatenation of the factors' parameters: $\mu = [\mu_1, \mu_2, \ldots, \mu_K]$ and has dimension $\dim \mu = \sum_{k=1}^K N_k$. The components of μ are computed by taking applying a neural network to the input s and then applying the softmax operator to each slice, yielding normalized probabilities for each factor.

E Experiment Parameters

	Swimmer	Hopper	Walker
State space dim.	10	12	20
Control space dim.	2	3	6
Total num. policy params	364	4806	8206
Sim. steps per iter.	50K	1M	1M
Policy iter.	200	200	200
Stepsize (D_{KL})	0.01	0.01	0.01
Hidden layer size	30	50	50
Discount (γ)	0.99	0.99	0.99
Vine: rollout length	50	100	100
Vine: rollouts per state	4	4	4
Vine: Q-values per batch	500	2500	2500
Vine: num. rollouts for sampling	16	16	16
Vine: len. rollouts for sampling	1000	1000	1000
Vine: computation time (minutes)	2	14	40
SP: num. path	50	1000	10000
SP: path len.	1000	1000	1000
SP: computation time	5	35	100

Table 2. Parameters for continuous control tasks, vine and single path (SP) algorithms.

Unnamed: 0	Swimmer	Hopper	Walker
State space dim.	10	12	20
Control space dim.	2	3	6
Total num. policy params	364	4806	8206
Sim. steps per iter.	50K	1M	1M
Policy iter.	200	200	200
Stepsize (DKL)	0.01	0.01	0.01
Hidden layer size	30	50	50
Discount (γ)	0.99	0.99	0.99
Vine: rollout length	50	100	100
Vine: rollouts per state	4	4	4

Visual NLP

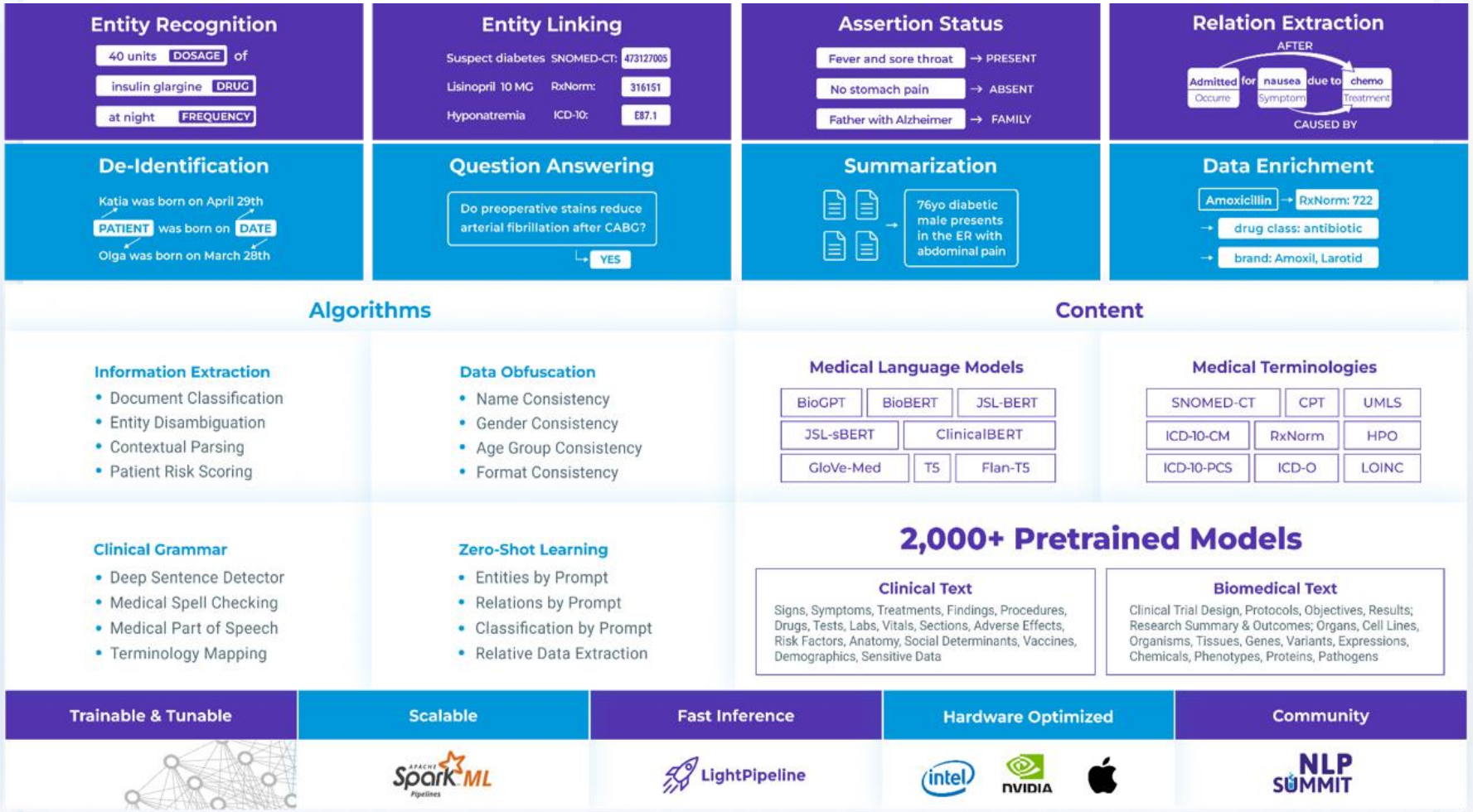
Form understanding	Object Character Recognition	Table detection & extraction	Noisy image enhancement
Visual document classification	Visual entity recognition	Signature detection	Image de-identification

Read	Enhance		Recognize	Extract	Generate
Text or PDF	Binarizer	Adaptive scaler	Regions	Text from Images (OCR)	Extracted text & layout (HOCR)
Scanned PDF	Adaptive Tresholding	Noise Scorer	Tables	Data from Tables (CV)	Structured data tables
DOCX	Erosion	Remove objects	Signatures	Named Entities from Forms	Highlighted named entities
Image	Skew corrections	Morphology opening	Dates	De-identification	De-identified text, PDF, or DICOM
DICOM	Scaler	Cropper	Brands	Visual Document Classification	Bounding boxes & coordinates

300+ live demos & Python notebooks:

`nlp.johnsnowlabs.com/demos`

Healthcare NLP & LLM

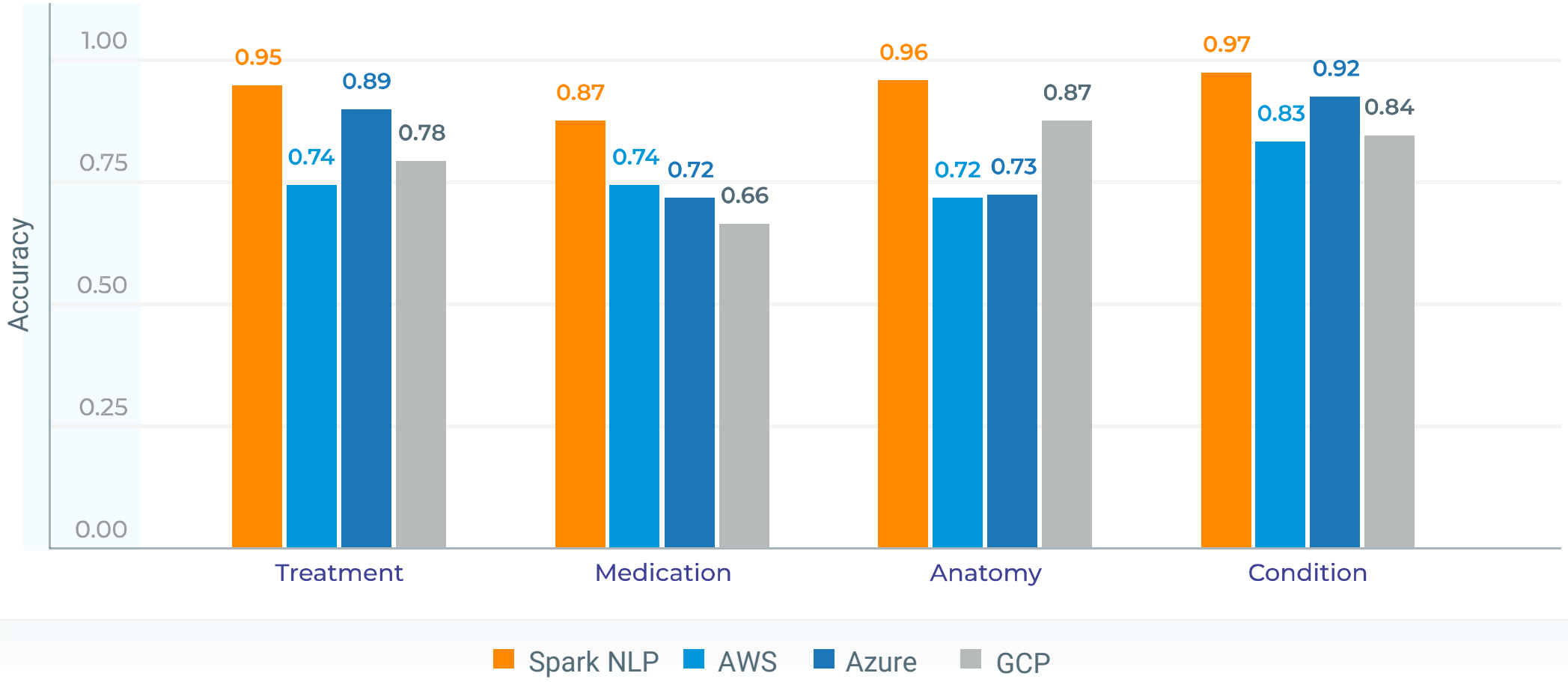


Peer-reviewed papers:

johnsnowlabs.com/peer-reviewed-papers/

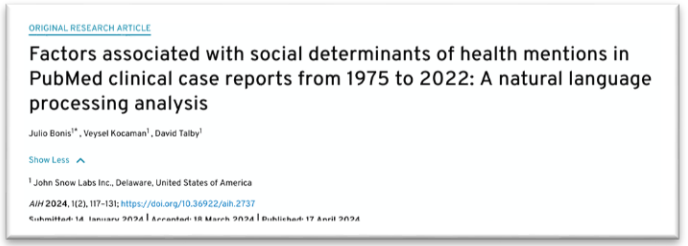
4-6x Fewer Errors than AWS, Azure, & GCP

www.johnsnowlabs.com/comparison-of-key-medical-nlp-benchmarks-spark-nlp-vs-aws-google-cloud-and-azure/



Domain-Specific Language Models Use Case: Extracting Social Determinants of Health

The patient reported a history of substance use SUBSTANCE_USE, specifically alcohol ALCOHOL and marijuana SUBSTANCE_USE, which began during their college EDUCATION years. They also disclosed a history of childhood trauma CHIDHOOD_EVENT related to emotional abuse VIOLENCE_OR_ABUSE by a family member. The patient is currently experiencing financial difficulties FINANCIAL_STATUS and is unemployed EMPLOYMENT, which has caused significant stress MENTAL_HEALTH and impacted their access to healthcare services ACCESS_TO_CARE. Additionally, the patient has been diagnosed with hypertension HYPERTENSION and hyperlipidemia HYPERLIPIDEMIA, and struggles with maintaining a healthy diet DIET due to limited access to healthy food FOOD_INSECURITY options and a lack of social support SOCIAL_SUPPORT. The patient has no current legal issues and identifies as bisexual SEXUAL_ORIENTATION. They report limited



ORIGINAL RESEARCH ARTICLE

Factors associated with social determinants of health mentions in PubMed clinical case reports from 1975 to 2022: A natural language processing analysis

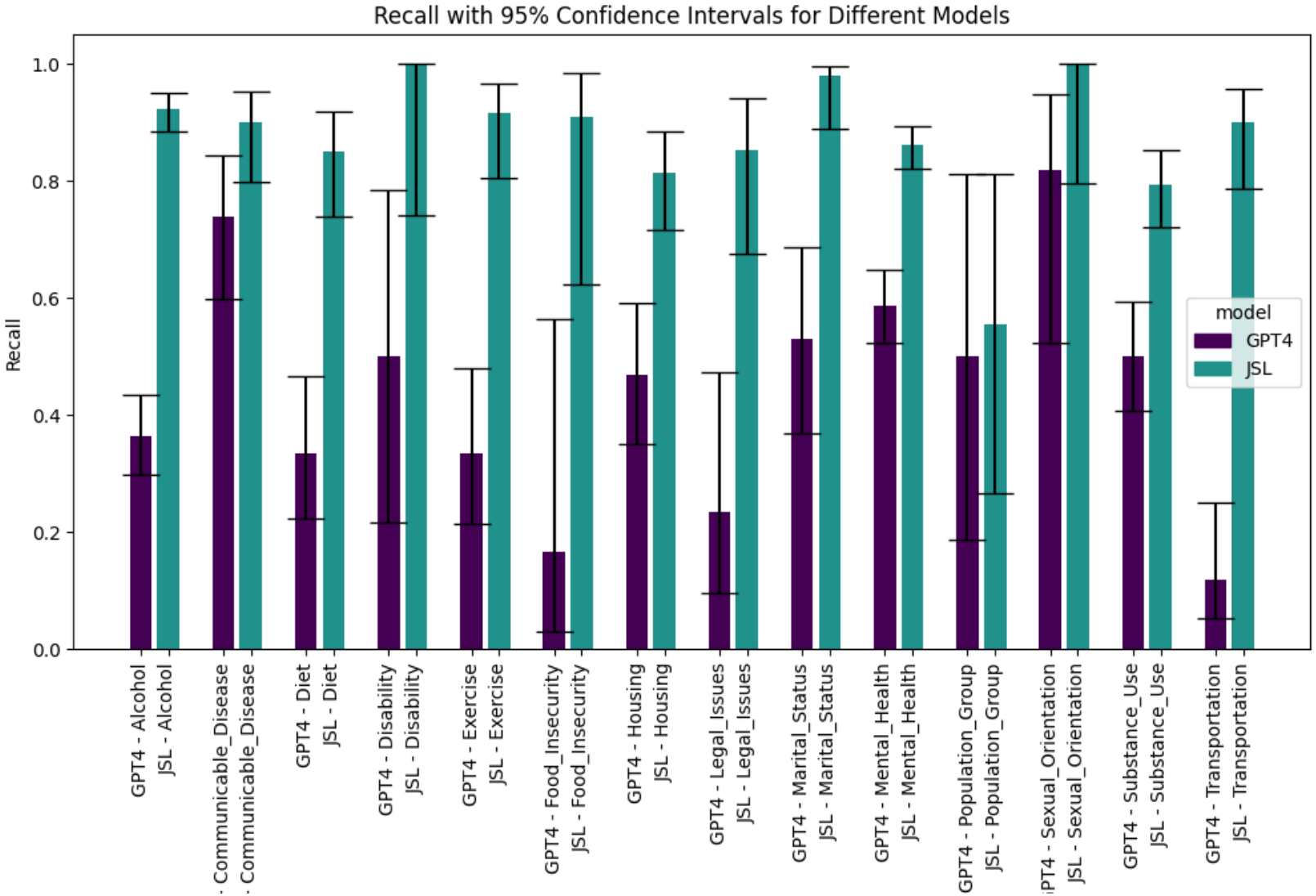
Julio Bonis[1*], Veysel Kocaman[1], David Talby[1]

Show Less

[1] John Snow Labs Inc., Delaware, United States of America

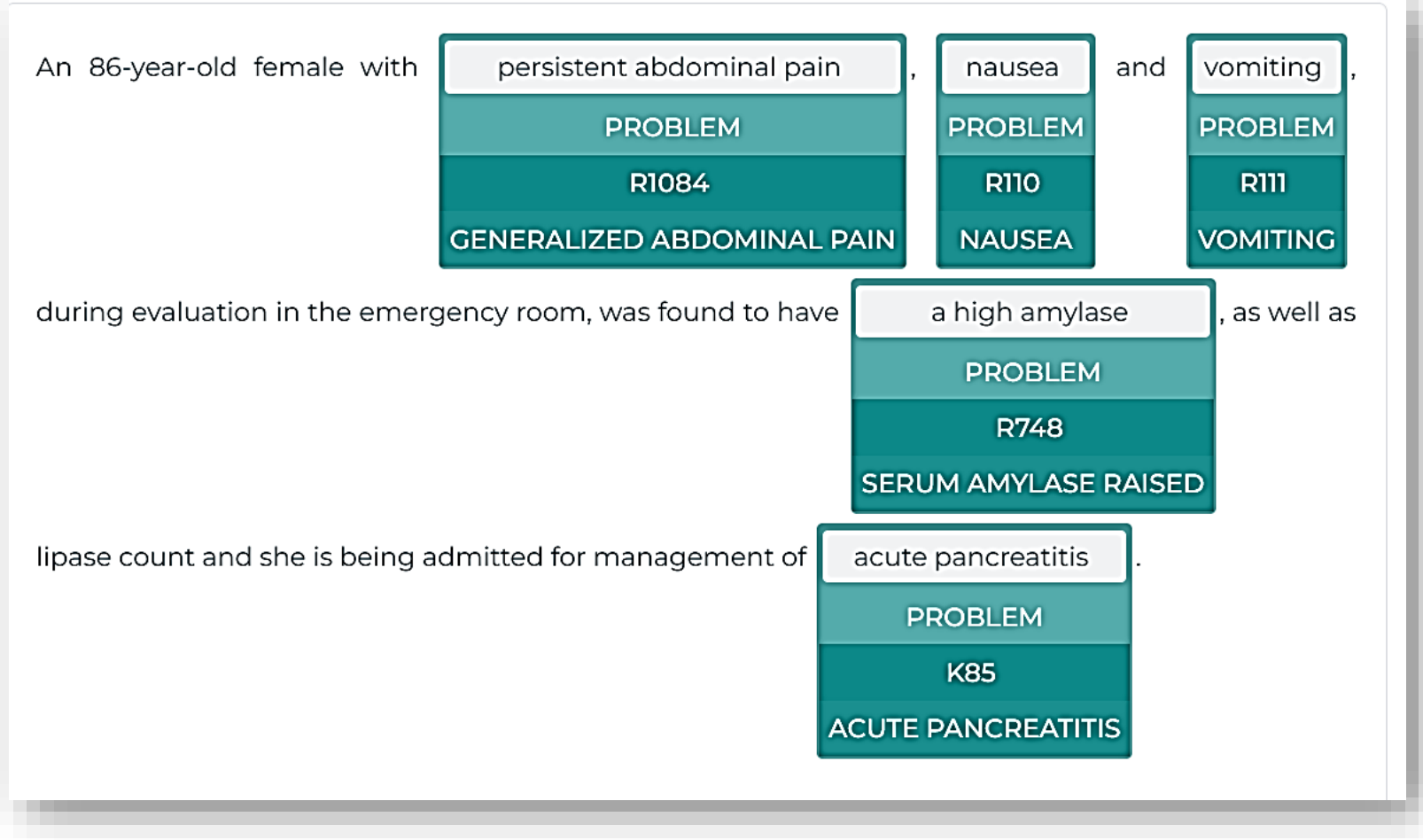
AIH 2024, 1(2), 117–131; https://doi.org/10.36922/aih.2737

Benchmark: Extracting Social Determinants of Health



GPT-4 makes **3 times more mistakes** than John Snow Labs' current SDoH models.

Benchmark: Extracting Diagnosis Codes



Extracting ICD-10-CM codes is done with a **76% success rate vs. 26% for GPT-3.5 and 36% for GPT-4.**

https://www.johnsnowlabs.com/comparing-spark-nlp-for-healthcare-and-chatgpt-in-extracting-icd10-cm-codes-from-clinical-notes/

Benchmark: Medical Test Summarization

- Preferred twice as often as GPT-4 summaries in a blind test by medical doctors
- Small 3B model, designed for fast inference on commodity hardware
- Runs behind your firewall, without sending your data to an external API

Text:

Medical Specialty: Pediatrics - Neonatal, Sample Name: Chest Closure

Description: Delayed primary chest closure. Open chest status post modified stage 1 Norwood operation. The patient is a newborn with diagnosis of hypoplastic left heart syndrome who 48 hours prior to the current procedure has undergone a modified stage 1 Norwood operation. (Medical Transcription Sample Report)

PROCEDURE: Delayed primary chest closure.

INDICATIONS: The patient is a newborn with diagnosis of hypoplastic left heart syndrome who 48 hours prior to the current procedure has undergone a modified stage 1 Norwood operation. Given the magnitude of the operation and the size of the patient (2.5 kg), we have elected to leave the chest open to facilitate postoperative management. He is now taken back to the operative room for delayed primary chest closure.

PREOP DX: Open chest status post modified stage 1 Norwood operation.

POSTOP DX: Open chest status post modified stage 1 Norwood operation.

ANESTHESIA: General endotracheal.

COMPLICATIONS: None.

FINDINGS: No evidence of intramediastinal purulence or hematoma. He tolerated the procedure well.

DETAILS OF PROCEDURE: The patient was brought to the operating room and placed on the operating table in the supine position. Following general endotracheal anesthesia, the chest was prepped and draped in the usual sterile fashion. The previously placed AlloDerm membrane was removed. Mediastinal cultures were obtained, and the mediastinum was then profusely irrigated and suctioned. Both cavities were also irrigated and suctioned. The drains were flushed and repositioned. Approximately 30 cubic centimeters of blood were drawn slowly from the right atrial line. The sternum was then smeared with a vancomycin paste. The proximal aspect of the 5 mm RV-PA conduit was marked with a small titanium clip at its inferior most aspect and with an additional one on its rightward inferior side. The sternum was then closed with stainless steel wires followed by closure of subcutaneous tissues with interrupted monofilament stitches. The skin was closed with interrupted nylon sutures and a sterile dressing was placed. The peritoneal dialysis catheter, atrial and ventricular pacing wires were removed. The patient was transferred to the pediatric intensive unit shortly thereafter in very stable condition. I was the surgical attending present in the operating room and in charge of the surgical procedure throughout the entire length of the case.

Summary:

A newborn with hypoplastic left heart syndrome underwent a delayed primary chest closure under general endotracheal anesthesia. The chest was prepped and draped in a sterile fashion, and mediastinal cultures were obtained. The mediastinum and cavities were irrigated and suctioned, and the sternum was closed with stainless steel wires and subcutaneous tissues with interrupted monofilament stitches. The patient tolerated the procedure well and was transferred to the pediatric intensive unit in stable condition.

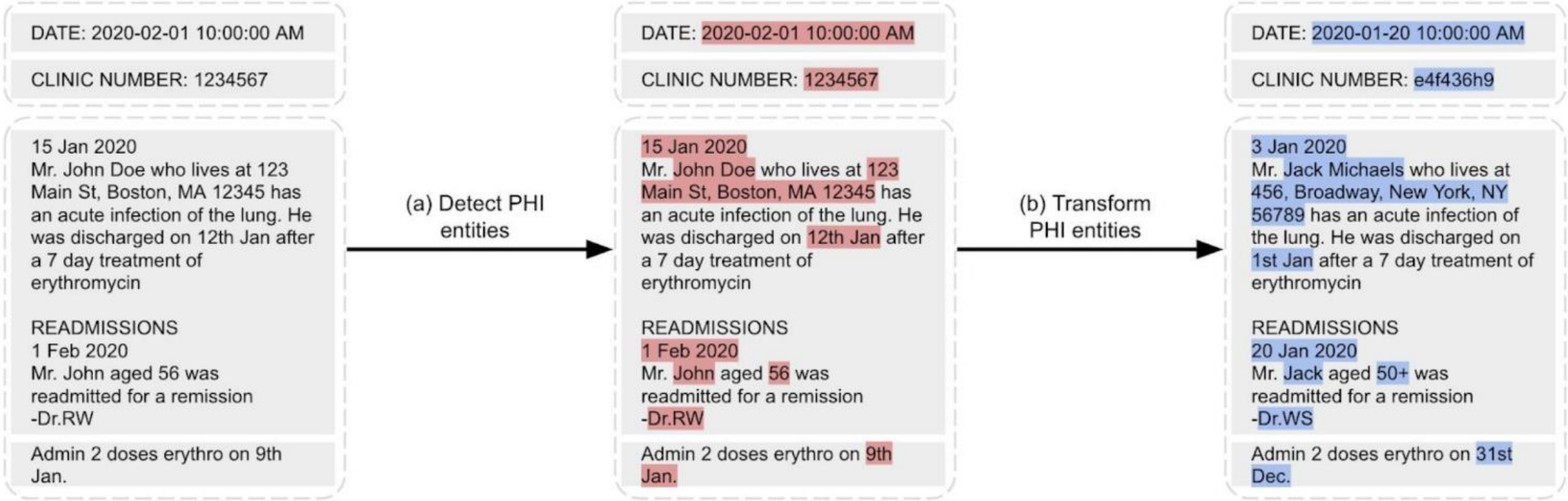
Benchmark: De-Identification

"The proposed system makes 50%, 475%, and 575% fewer errors than the comparable AWS, Azure, and GCP services respectively while also outperforming ChatGPT by 33%. It exceeds 98% coverage of sensitive data across 7 European languages, without a need for fine tuning."

[Submitted on 13 Dec 2023]

Beyond Accuracy: Automated De-Identification of Large Real-World Clinical Text Datasets

Veysel Kocaman, Hasham Ul Haq, David Talby



https://www.johnsnowlabs.com/a-comprehensive-comparison-of-chatgpt-and-spark-nlp-for-healthcare-in-de-identification-of-sensitive-data-phi/

Domain-Specific LLM's Usually Outperform General-Purpose LLM's

John Snow LABS

[Submitted on 30 May 2023 (v1), last revised 29 Mar 2024 (this version, v7)]

Domain Specialization as the Key to Make Large Language Models Disruptive: A Comprehensive Survey

Chen Ling, Xujiang Zhao, Jiaying Lu, Chengyuan Deng, Can Zheng, Junxiang Wang, Tanmoy Chowdhury, Yun Li, Hejie Cui, Xuchao Zhang, Tianjiao Zhao, Amit Panalkar, Dhagash Mehta, Stefano Pasquali, Wei Cheng, Haoyu Wang, Yanchi Liu, Zhengzhang Chen, Haifeng Chen, Chris White, Quanquan Gu, Jian Pei, Carl Yang, Liang Zhao

Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing

YU GU*, ROBERT TINN*, HAO CHENG*, MICHAEL LUCAS, NAOTO USUYAMA, XIAODONG LIU, TRISTAN NAUMANN, JIANFENG GAO, and HOIFUNG POON, Microsoft Research

Research Letter | Cardiothoracic Imaging | January 17, 2024

General-Purpose Large Language Models Versus a Domain-Specific Natural Language Processing Tool for Label Extraction From Chest Radiograph Reports

Authors: Cody H. Savage, MD, Hyoungsun Park, MS, Kijung Kwak, BA, Andrew D. Smith, MD, PhD, Steven A. Rothenberg, MD, Vishwa S. Parekh, MD, Florence X. Doo, MD, and Paul H. Yi, MD | AUTHOR INFO & AFFILIATIONS

[Submitted on 4 Feb 2024]

A Survey of Large Language Models in Finance (FinLLMs)

Jean Lee, Nicholas Stevens, Soyeon Caren Han, Minseok Song

Pretrained Domain-Specific Language Model for Natural Language Processing Tasks in the AEC Domain

Zhe Zheng[1], Xin-Zheng Lu[1], Ke-Yin Chen[1], Yu-Cheng Zhou[1], Jia-Rui Lin[1,*]

*Corresponding author, E-mail: lin611@tsinghua.edu.cn; jiarui_lin@foxmail.com

(1. Department of Civil Engineering, Tsinghua University, Beijing, 100084, China)

Training Workshop Outline

LLM & RAG with Spark NLP

Multi-Modal Language Models

Domain-Specific Language Models

Medical Chatbot

Generative AI Lab

Human-in-the-loop Workflows

State-of-the-Art Medical LLM

JSL-MedMX	91.82
Med-PaLM2	84.09
GPT-4	82.97
Llama3-FT-Med	77.71

* on the Open Medical LLM Leaderboard Benchmark

MedQA (USMLE)

- 1,273 real-world questions from the US Medical License Exams (USMLE) to test general medical knowledge

PubMedQA

- 500 questions constructed to test reasoning over biomedical research texts, especially their quantitative contents

MedMCQA

- 4,183 questions from Indian medical entrance exams (AIIMS & NEET PG) spanning 2.4k healthcare topics, designed to address real-world medical entrance exam questions

MMLU

- College-level questions on Clinical knowledge (265), Medical genetics (100), Anatomy (135), Professional medicine (272), College biology (144), and College medicine (173)

First 7B LLM to Beat GPT-4 on PubMedQA

JSL-MedLX-7B	78.4
GPT-4	75.1
Single Human Performance	78.0

* Reproducible using the LLM evaluation harness

Question:
Do preoperative statins reduce atrial fibrillation after coronary artery bypass grafting?

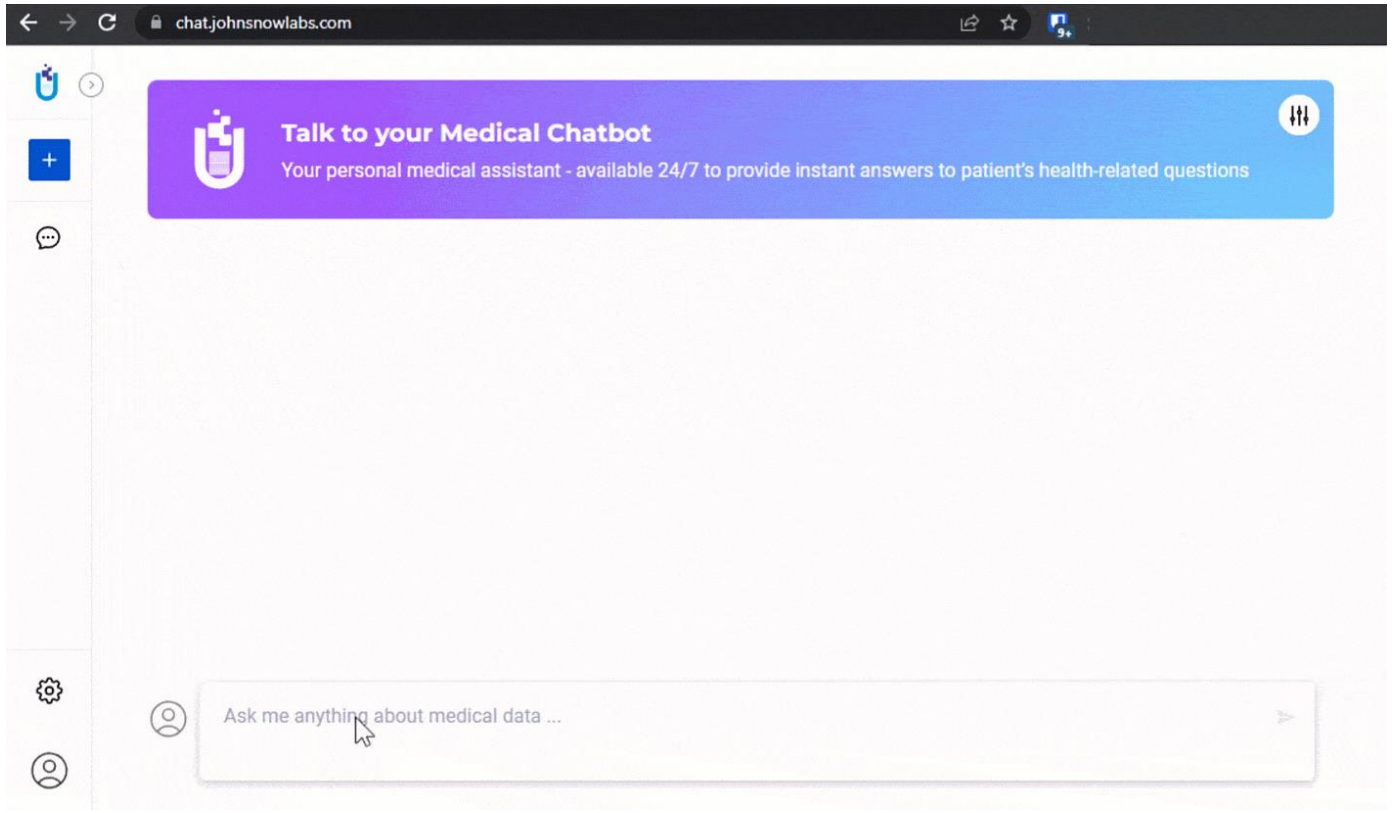
Context:
(Objective) Recent studies have demonstrated that statins have pleiotropic effects, including anti-inflammatory effects and atrial fibrillation (AF) preventive effects [...]
(Methods) 221 patients underwent CABG in our hospital from 2004 to 2007. 14 patients with preoperative AF and 4 patients with concomitant valve surgery [...]
(Results) The overall incidence of postoperative AF was 26%. Postoperative AF was significantly lower in the Statin group compared with the Non-statin group (16% versus 33%, p=0.005). Multivariate analysis demonstrated that independent predictors of AF [...]

Long Answer:
(Conclusion) Our study indicated that preoperative statin therapy seems to reduce AF development after CABG.

Short Answer:
yes

John Snow Labs' Medical Chatbot

- Delivers Superior Accuracy for Healthcare-Specific Tasks
- Runs privately behind your firewall: No third-party APIs or data sharing
- Updates all medical knowledge daily
- Can consistently reproduce results
- Does not hallucinate, cites its sources
- Explains its answers
- Adds private knowledge bases
- Tunable guardrails and brand voice



John Snow Labs' Medical Chatbot

- ✓ Delivers Superior Accuracy for Healthcare-Specific Tasks
- ✓ Runs privately behind your firewall: No third-party APIs or data sharing
- ✓ Updates all medical knowledge daily
- ✓ Can consistently reproduce results
- ✓ Does not hallucinate, cites its sources
- ✓ Explains its answers
- ✓ Adds private knowledge bases
- ✓ Tunable guardrails and brand voice

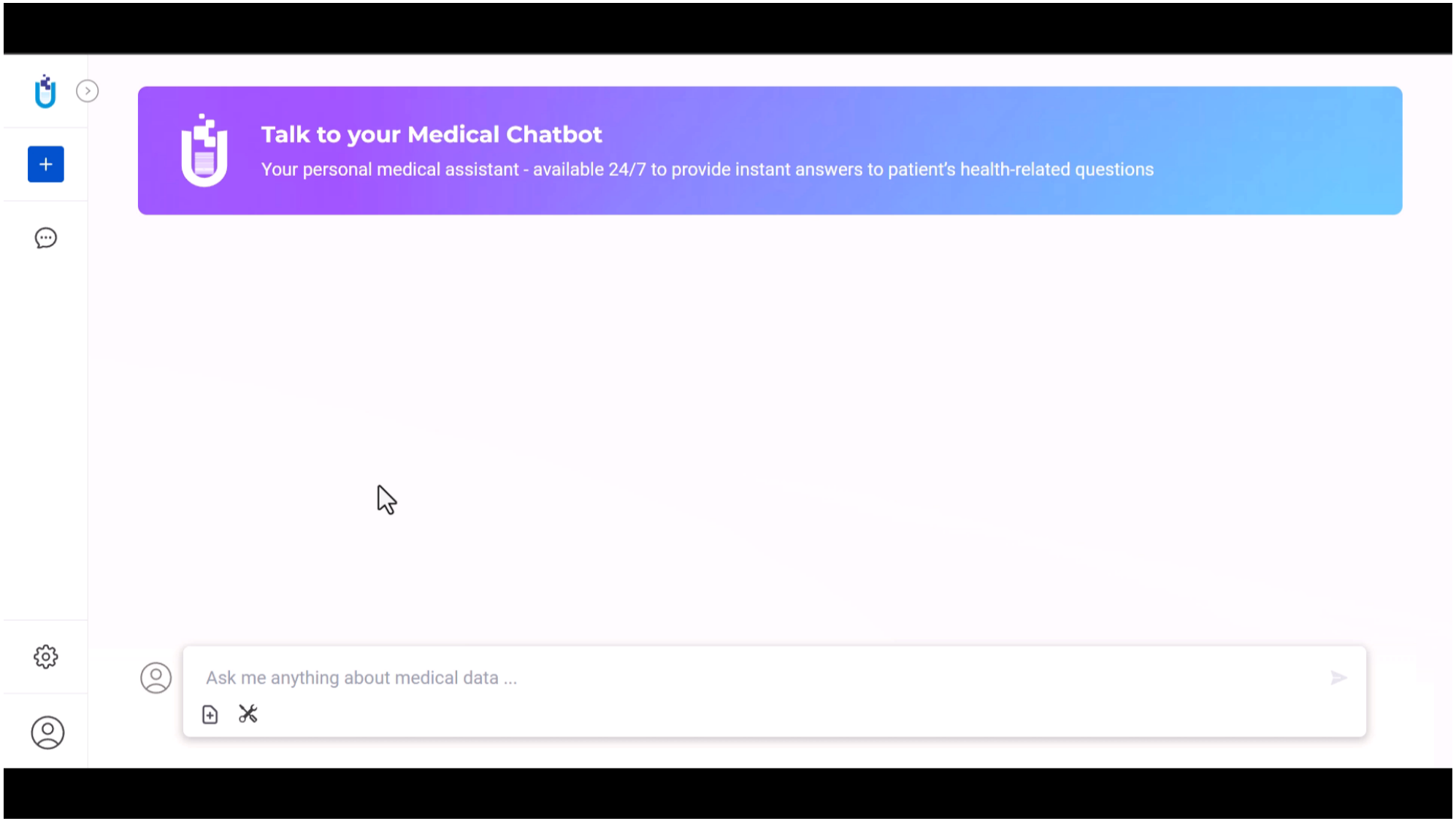
→ Multi-LLM, Multi-Agent Architecture

→ Comes with Medical Knowledge Bases

→ Built-in RAG Backend

→ Enterprise Grade Security & Scalability

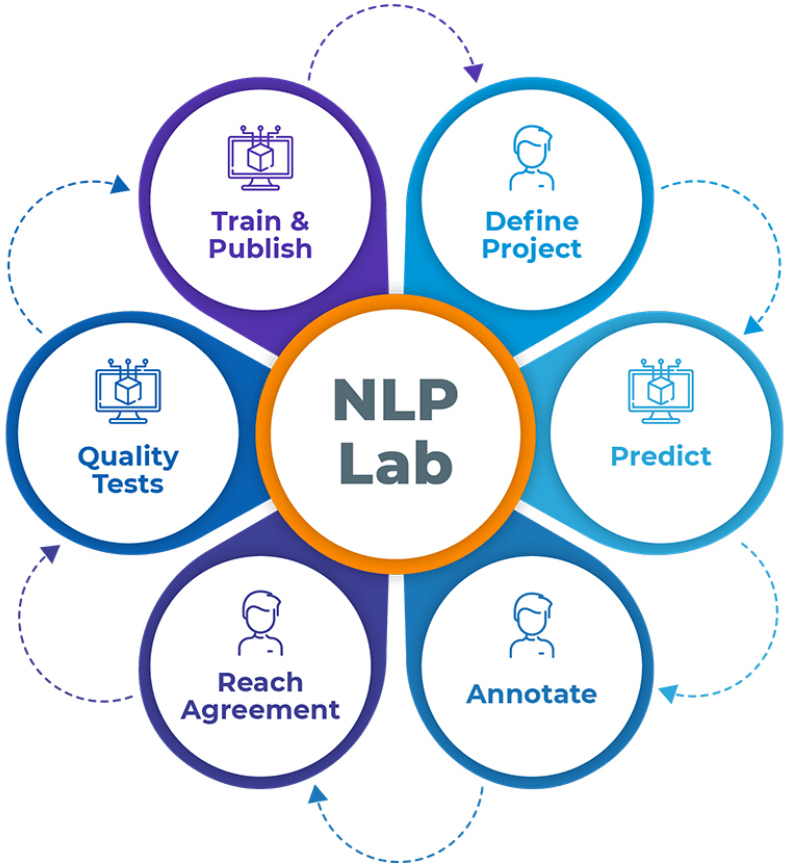
Document Q&A


Talk to your Medical Chatbot
Your personal medical assistant - available 24/7 to provide instant answers to patient's health-related questions
Ask me anything about medical data ...

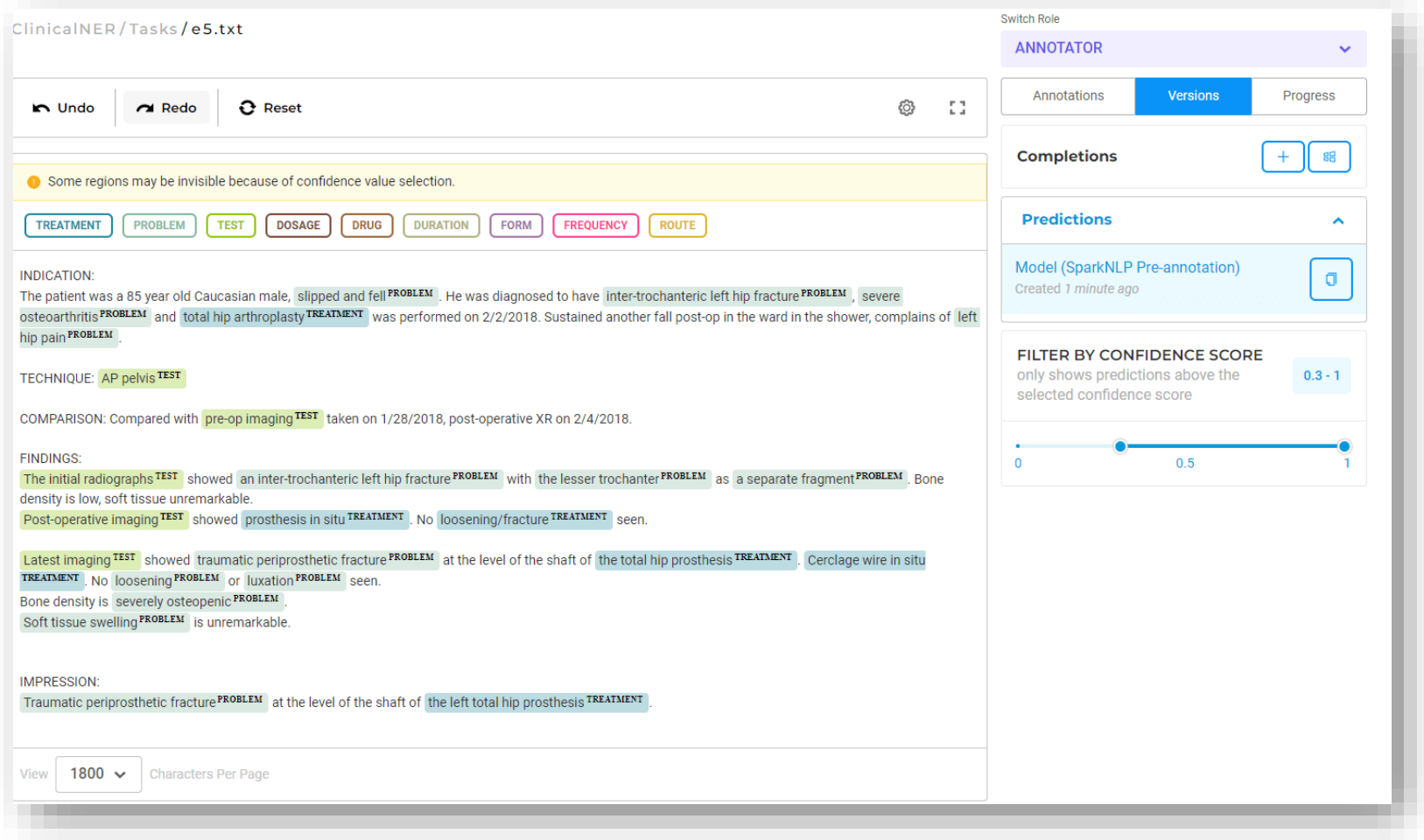
The Generative AI Lab

The No-Code NLP Platform:

- Annotate Text & Images
- AI Assisted Annotation
- Train & Tune NLP Models
- Models, Rules, and Prompts Hub
- Manage Projects & Teams
- Enterprise Security & Privacy



The Generative AI Lab: Building Small, Task-Specific Language Models



The "Old School" Use Case:

Define the task for a new model

Import documents to Annotate

AI-Assisted Pre-Annotation

Manual work by domain experts

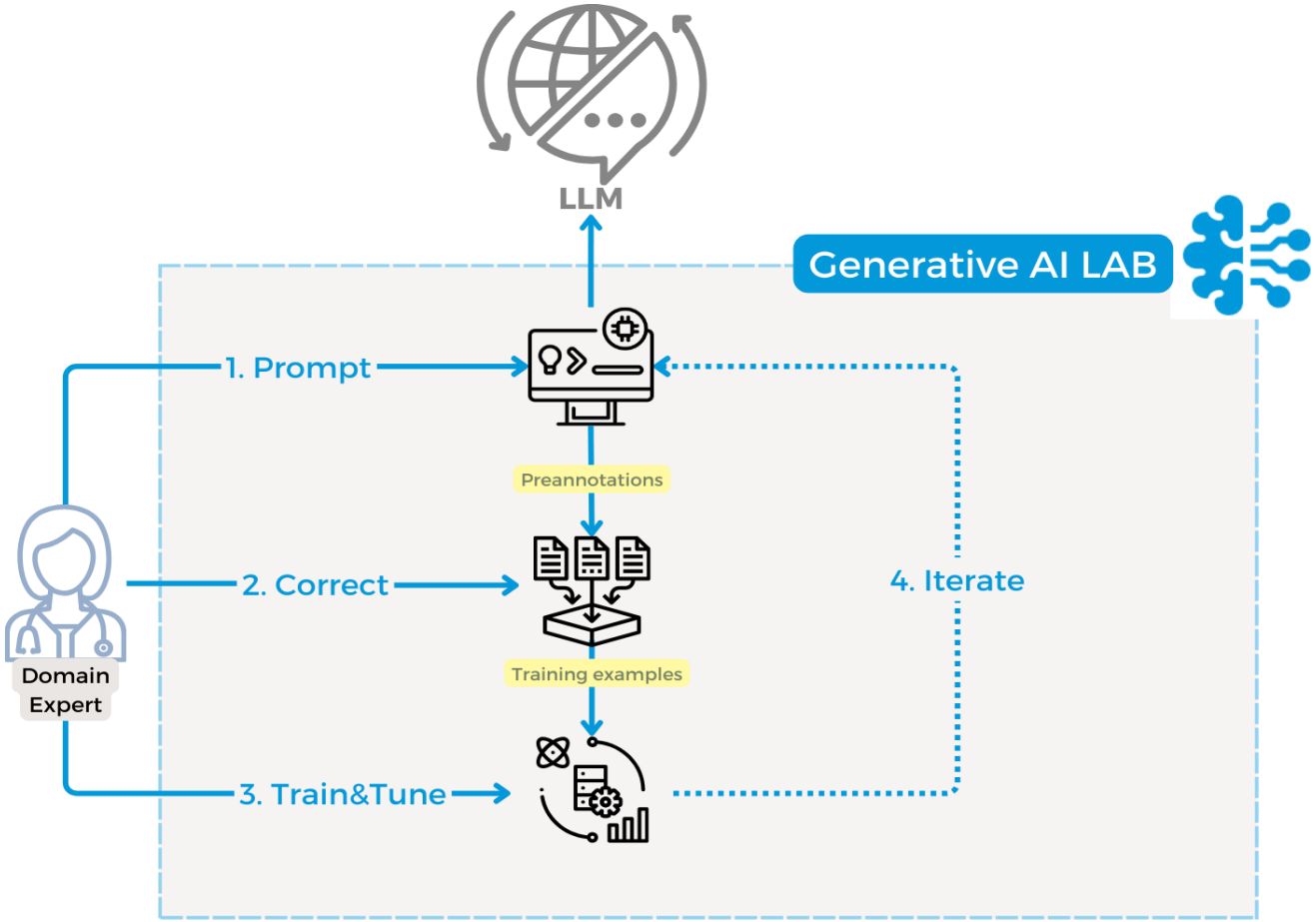
Train or tune a language model

Export the model or the data

Generative AI Lab Use Cases

1. Use LLMs to bootstrap task-specific models
2. Human-in-the-loop workflows to validate LLMs when regulatory-grade accuracy is required
3. Organize and share models, prompts, and rules within one private enterprise hub

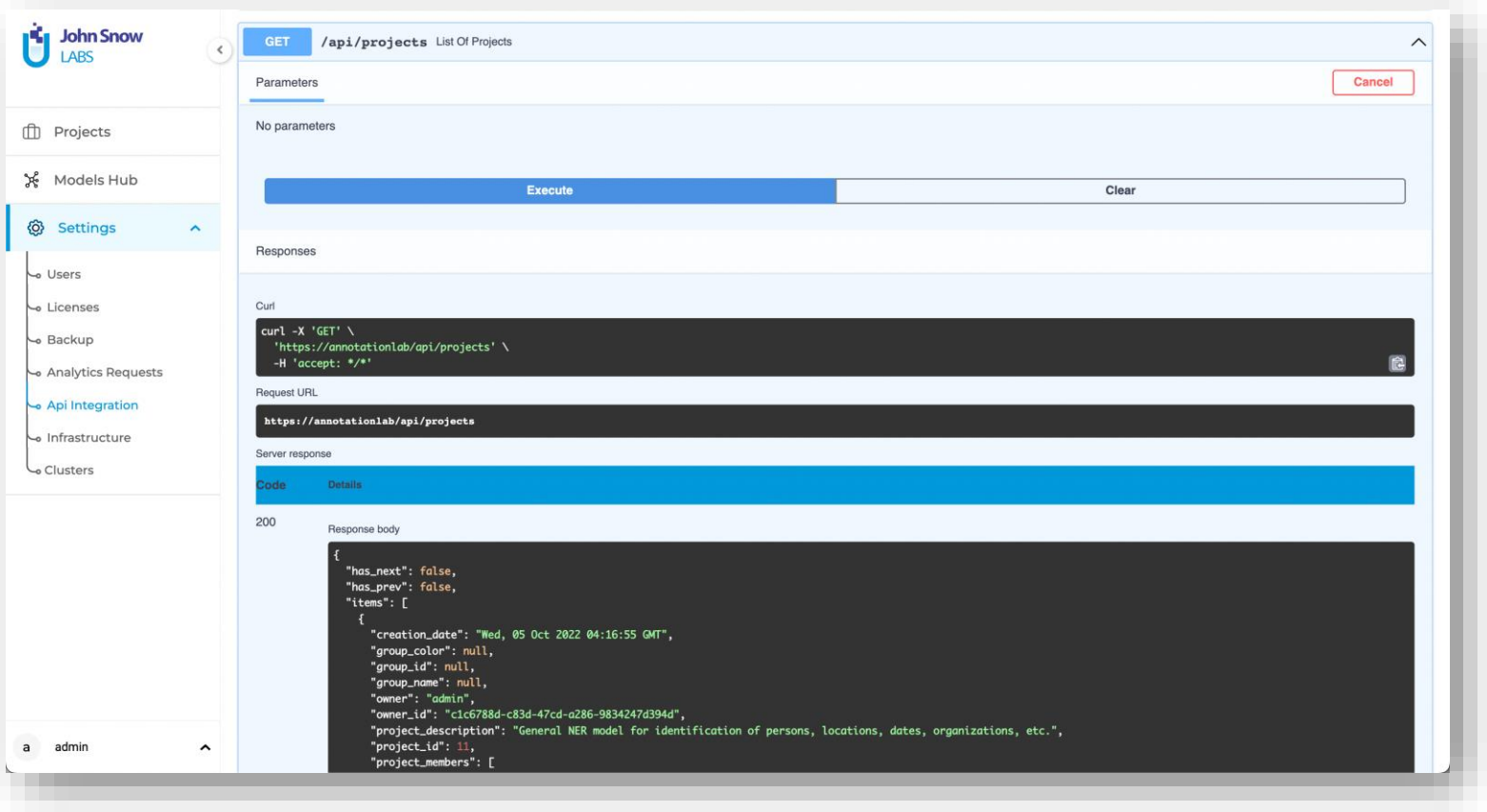
From LLMs to task specific AI Models



From LLMs to task specific AI Models

1. **Prompt**
   - ✓ Entity Extraction, Classification
   - ✓ ChatGPT or Private Zero-Shot Models
2. **Correct**
   - ✓ Intuitive UI for checking and correcting results
   - ✓ Collaboration and task distribution across teams of domain experts
3. **Train & Tune**
   - ✓ No-code model training
   - ✓ Support for experiments
4. **Iterate**
   - ✓ Check model performance metrics and iterate as needed

Building Enterprise-Grade Human-in-the-Loop Pipelines



- Versioning Human & AI Work
- Full Audit Trails
- Custom Approval Workflows
- API Integration
- Authentication & Authorization
- Private On-Premise Deployment
- Scale to Large Teams & Projects

How to Get The Most from This Training

1. **Try the libraries and tools hands-on**
2. Ask questions
3. Learn how the tools apply to use cases: www.johnsnowlabs.com/customers

Thank you!