

# Accurate Cervical Spine Fracture Detection from Radiology Report: Harnessing Privacy-preserving Open-Source Large Language Models (LLMs)

Sanaz Vahdati, M.D.<sup>1</sup>, Bardia Khosravi, M.D.<sup>1</sup>, Bradley J. Erickson, M.D.<sup>1</sup>

<sup>1</sup>Mayo Clinic Artificial Intelligence Laboratory, Department of Radiology, Rochester MN

## BACKGROUND

Open-source Language Models (LLMs) have emerged as powerful tools for data abstraction from radiologic reports. These models leverage their natural language processing capabilities to extract critical information from complex medical texts, enhancing the efficiency of healthcare professionals. The open-source approach provides a more secure and compliant option for medical data abstraction, as it eliminates the need to transmit sensitive patient information through a commercial API, ensuring privacy and HIPAA compliance while offering greater control and flexibility for healthcare organizations. Here, we use an open-source LLM to determine the presence and level of acute cervical spine fractures in radiologic reports.

## METHODS

We acquired radiology reports from 700 non-contrast cervical spine CT scans conducted between January and February 2022. Of these reports, 110 were utilized for initial model prompting and fine-tuning. The remaining 590 served as a test dataset to assess the model's performance. The primary aim of the model was to generate a JSON file that could indicate the presence or absence of an acute cervical vertebral fracture and specify its associated location. This specific challenge was chosen because the model needed to distinguish among various anatomical regions and discern between chronic and acute fractures. To accomplish this, we employed an open-source Large Language Model (LLM) known as the Zephyr-7B-Alpha and deployed it locally. We applied a combination of chain-of-thought (CoT) and few-shot prompting techniques to instruct the model in producing the final JSON output containing the results (Figure 1).

## RESULTS

The model attained a 98% accuracy rate, correctly identifying acute cervical spine fractures in 583 out of 590 cases. In cases where reports indicated the "possibility" of fracture, the model reported the presence of fracture. Particularly noteworthy is that the model accurately predicted multiple fracture sites where multiple locations of fracture were present in one report.

## CONCLUSIONS

Despite the widespread recognition of ChatGPT, its limitation lies in offering solely a commercially hosted API, which mandates the transmission of patient data beyond the confines of the healthcare institution. By employing proper prompts, open-source LLMs can excel in their performance without the necessity of sharing patient data.

Prompt

You are a radiologist tasked with analyzing a CT cervical spine report and classifying the presence and absence of cervical vertebrae fracture(C1-C7). Consider the medical terms below:

- The cervical spine is made of 7 vertebrae(C1-C7)
- Atlas is another name for C1.
- Axis is another name for C2.
- C2 has a structure called odontoid, also known as dens.
- After C7 comes the thoracic spine with 12 vertebrae(T1-T12).

CoT:

- 1- Summarize the impression in one or two sentences (at most 30 words) and say if there is an acute cervical vertebrae fracture.\
- 2- Estimate which vertebrae are fractured. If you find any mention of acute cervical vertebrae fracture (not healing or chronic), print out comma-separated list of vertebrae that you found to have fracture (C1, C2, .., C7). If you do not find any fracture level, print out "no fracture found"..\
- 3- There might be several fractured vertebrae in one report, and you should mention all of them.\
- 4- Do not consider Anterolisthesis and retrolisthesis as fractures.\
- 5- Only consider vertebra fractures that start with C in your report. Do not report fractures in the thoracic T1-T12 or lumbar (L1-L5) spine.\
- 6- You can not print the final result "absent" and then give a "level of fracture"..\
- 7- We do not want findings such as rib fracture or thyroid nodule evaluation.\
- 8- Do not report post-operation reports as fracture.\

FIGURE 1: Demonstration of education and chain of thoughts for refining of the prompt.