

## BACKGROUND

1

High cost + time requirement

## OBJECTIVE

2

Generate idiomatic instances with LLMs

## METHOD

3

Prompt GPT4 for context, then generate samples

# IDIOM CORPORA CONSTRUCTION VIA LARGE LANGUAGE MODELS



## EVALUATION

4

Sequence labelling task on idiom detection

## RESULTS

5

Not in human-level yet promising

## FUTURE WORK

6

Additional languages and prompting techniques