

# Multimodal Idiomaticity Representation

## Extending the AdMIRe shared task

Thomas Pickard & Aline Villavicencio

# Motivation

- A number of **idiomaticity detection** tasks have been introduced
  - FLUTE ([Chakrabarty et al 2022](#))
  - SemEval 2022 Task 2a ([Tayyar Madabushi et al 2022](#))
  - MAGPIE ([Haagsma et al 2020](#))
- These predominantly focus on **English**
- Datasets contain ‘construction artifacts’ ([Boisson et al 2023](#))
  - Tuned models perform well even when context or target expression hidden
- Instruction-prompted LLMs perform moderately well ([Phelps et al 2024](#))
  - ... in English
  - ... but not as well as fine-tuned models or human speakers
- Still room to improve quality of idiom representations
- Can we **avoid artifacts** and require better ‘understanding’ of **idiomatic meaning**?

## AdMIRe (SemEval 2025)

- [SemEval 2025](#) Task 1: [Advancing Multimodal Idiomaticity Representation](#)
  - Training phase **open now**
  - Evaluation starts **10th January**
  - [Google group](#)
- Key idea: **visual** and **visual-temporal** modalities should require **better representations** of idiomatic & literal meanings of idioms
- To enable participation with **text only**, tracks using descriptive image captions
- English training data available now
  - PT-BR available soon

## Subtask A – Static Images

- Given a particular **noun compound** and a **context sentence** in which it appears:
- Select which of 5 images **best** represents the target NC in the sense (literal/idiomatic) in which it is used in the context sentence
- E.g. for **bad apple**



We have to recognize that this is not the occasional bad apple but a structural, sector-wide problem

## Subtask A – Static Images

- Given a particular **noun compound** and a **context sentence** in which it appears:
- Select which of 5 images **best** represents the target NC in the sense (literal/idiomatic) in which it is used in the context sentence
- E.g. for **bad apple**



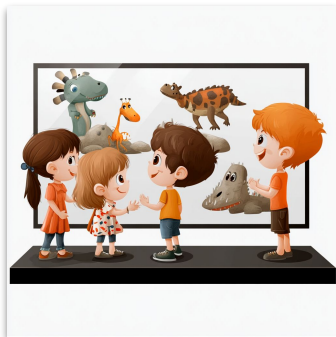
However, if ethylene happens to be around (say from a bad apple), these fruits do ripen more quickly.

## Subtask B – Image Sequences

- Many idiomatic meanings are difficult to capture in a static image
  - E.g. **kangaroo court** – literal sense is highly imageable, but the idiomatic is complex
- Given an idiomatic **noun compound** and 2 images forming **part of a sequence** representing either the idiomatic or literal sense of the NC:
- **Complete the sequence** by selecting the third image from 4 candidates
- Additionally, identify whether the sequence represents the **idiomatic** or **literal** sense

## Subtask B – Image Sequences

- Given an idiomatic **noun compound** and 2 images forming **part of a sequence** representing either the idiomatic or literal sense of the NC:
- Complete the sequence** by selecting the third image from 4 candidates
- E.g. for **bad apple**



## Extension

- Our proposal is to extend the dataset for AdMiRe:
  - More idiom types (e.g. verbal idioms)
  - More examples
  - **Additional languages**
- We're also planning to conduct **human evaluation**
- Would be good to re-run the expanded task in SemEval 2026



# Organizational Plan

## Current Timeline

- Evaluation start 10 January 2025
- Evaluation end by 31 January 2025
- Paper submission due 28 February 2025
- Notification to authors 31 March 2025
- Camera ready due 21 April 2025
- SemEval workshop Summer 2025  
(co-located with a major NLP conference)

## Tentative Timeline

- Call for data contribution 31 March 2025
- Sample data available 31 March 2025
- End of data contribution period 1 September 2025
- Training data available 27 October 2025
- Evaluation start 12 January 2026
- Evaluation end by 1 February 2026
- Paper submission due 1 March 2026
- Notification to authors 29 March 2026
- Camera ready due 19 April 2026
- SemEval workshop Summer 2026

## Organizers

- Tom Pickard, University of Sheffield, UK
- Wei He, University of Sheffield, UK
- Maggie Mi, University of Sheffield, UK
- Dylan Phelps, University of Sheffield, UK
- Carolina Scarton, University of Sheffield, UK
- Marco Idiart, Federal University of Rio Grande do Sul, Brazil
- Aline Villavicencio, University of Exeter, UK
- Doğukan Arslan, Istanbul Technical University, Türkiye
- Gülşen Eryiğit, Istanbul Technical University, Türkiye