

AnNER: Supporting Efficient Entity Annotation and Review Workflows for Knowledge Graph Construction

Umayar Reza, Nicholas Pease, Jinwu Wang, Torsten Hahmann

*Spatial Knowledge and Artificial Intelligence (SKAI) Lab
School of Computing and Information Science
University of Maine, Orono, Maine*

K-CAP 2025, Dayton, Ohio



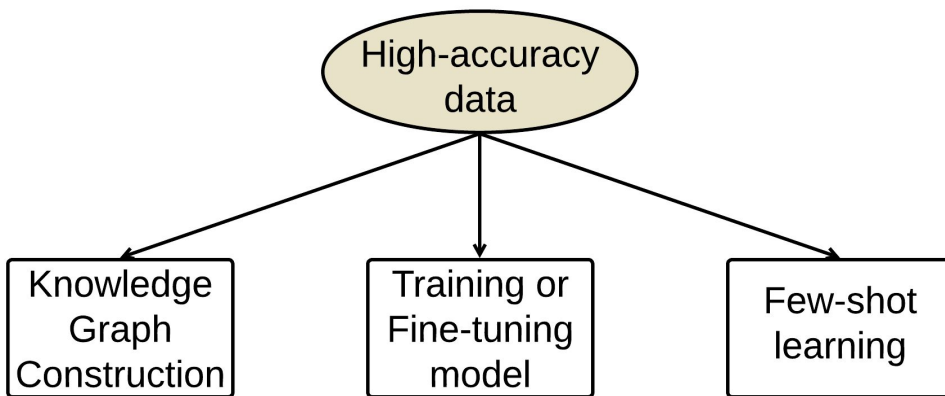
Motivation

- ❑ Models (including LLMs) make mistakes in entity recognition
 - ❑ *E.g., **mis**-classification, **mis**-identification, or **non**-recognition*
- ❑ Error increase in specialized domain entity recognition
 - ❑ *E.g., biomedical, chemical, material science*

Motivation

- ❑ Models (including LLMs) make mistakes in entity recognition
 - ❑ *E.g., **mis**-classification, **mis**-identification, or **non**-recognition*
- ❑ Error increase in specialized domain entity recognition
 - ❑ *E.g., biomedical, chemical, material science*

High-accuracy data is crucial for some downstream applications:



Human-in-the-Loop for Enhancing Data Accuracy

- ❑ Reviews, validates, and corrects model-generated entities
- ❑ Ensure higher accuracy and more consistency in data
- ❑ Creates high-quality inputs for domain-specific downstream tasks

Human-in-the-Loop for Enhancing Data Accuracy

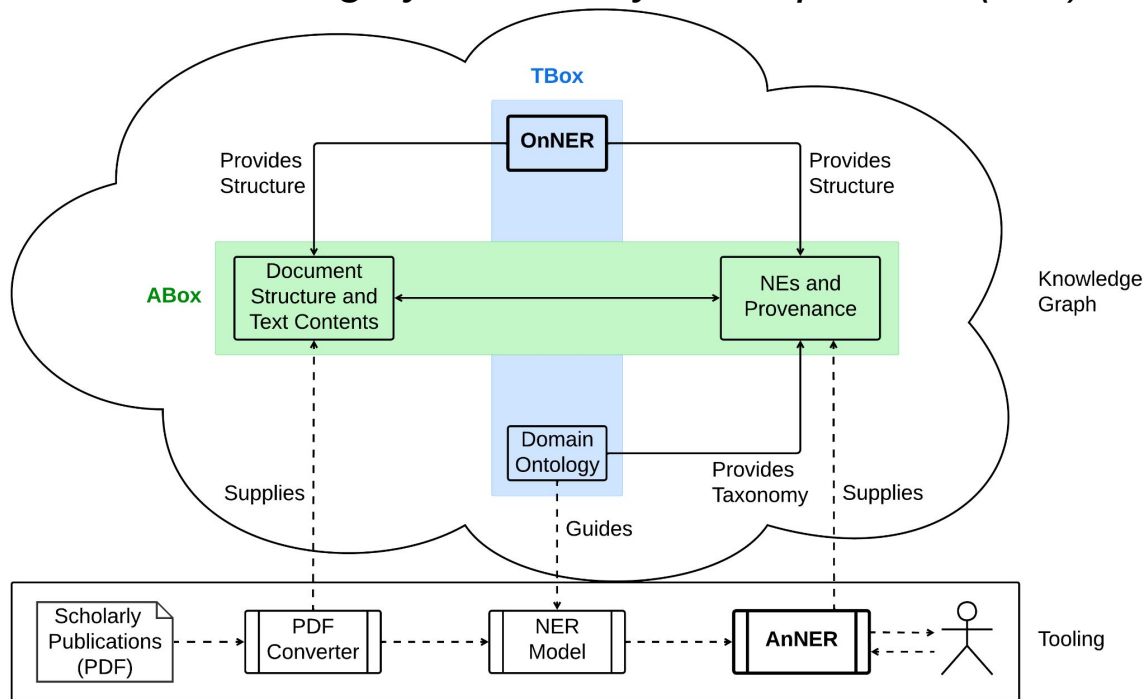
- ❑ Reviews, validates, and corrects model-generated entities
- ❑ Ensure higher accuracy and more consistency in data
- ❑ Creates high-quality inputs for domain-specific downstream tasks

A Tool for ***Named Entity Annotation and Review (AnNER)***:

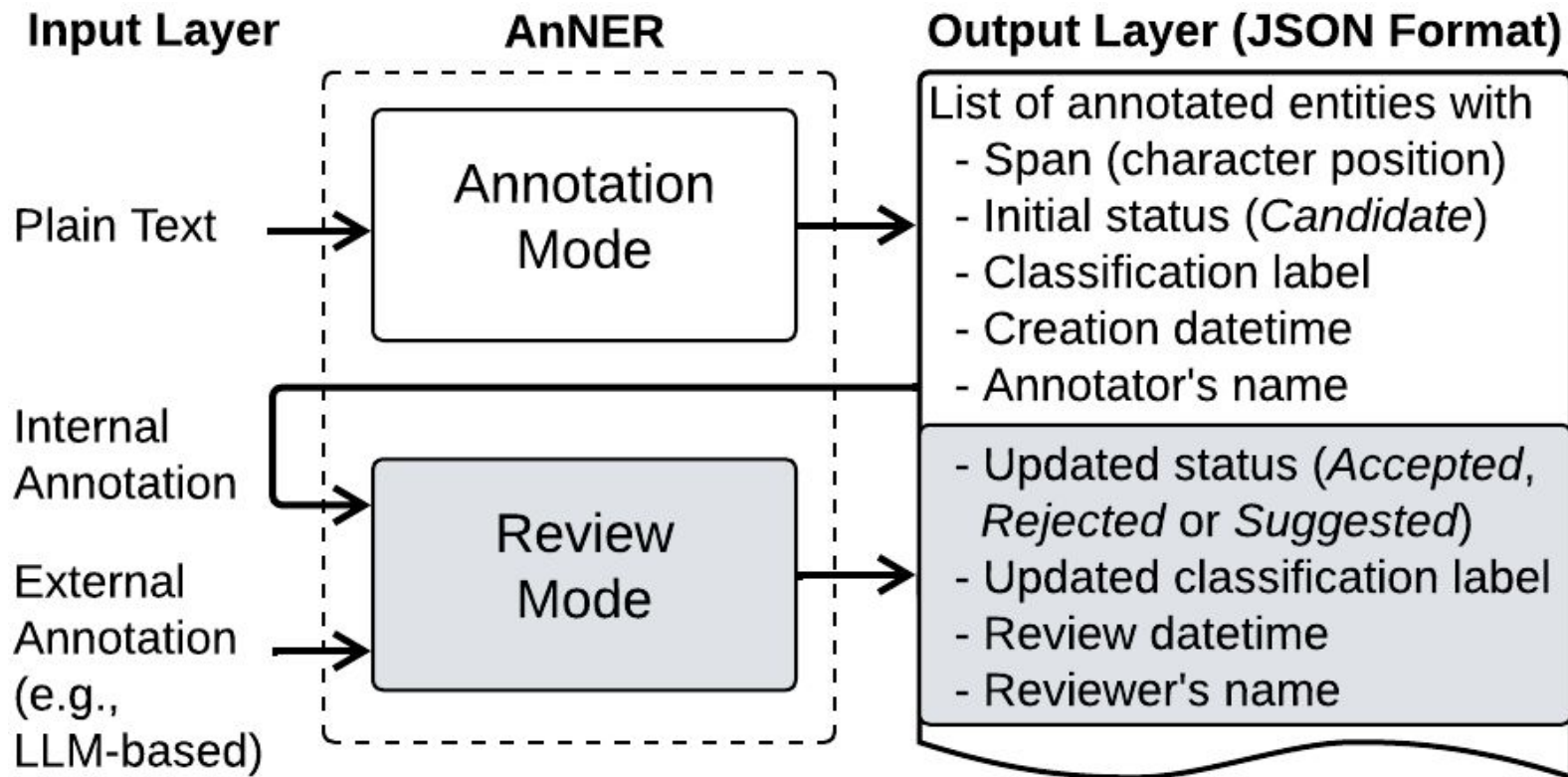
- Facilitates validation of model predictions or manual annotations
- Keeps record of provenance for both review and annotation workflows

Context: The CelloGraph Project

- ❑ Goal: Accelerating knowledge discovery for novel uses of cellulosic materials
- ❑ Funded by US Dept. of Agriculture (USDA NIFA & ARS) and US Forest Service
- ❑ Cellulose is a renewable, highly versatile yet inexpensive (raw) material



AnNER: Two Independent User Modes



AnNER: A Simple Annotation Interface

File Edit Annotator HelpSample.txt - **Annotate Mode**

✓

CHEMICAL

MATERIAL

PROPERTY

APPLICATION

ADD TAG

EDIT TAGS

Cellulose is mainly used to produce paperboard APPLICATION ⇌ ✕ and paper MATERIAL ⇌ ✕ . Conversion of cellulose from energy crops MATERIAL ⇌ ✕ into biofuels such as cellulosic ethanol CHEMICAL ⇌ ✕ is under development as a renewable fuel source .

29 Words 189 Characters Paragraph 1/2 4 Annotations

BACK

NEXT

AnNER: A Dedicated Review Interface

File Edit Annotator Help Umayer-annotations.json - Review Mode

1 CHEMICAL 2 MATERIAL 3 PROPERTY 4 APPLICATION

Accepted

Suggested

Rejected

Candidate

Cellulose is mainly used to produce paperboard and paper . Conversion of cellulose from energy crops into biofuels such as cellulosic ethanol is under development as a renewable fuel source .

29 Words 189 Characters Paragraph 1/2 5 Annotations

BACK NEXT

AnNER: Annotation Lifecycle

Example paragraph:

... the role of the ***nanoparticle shape*** ... on the modulation of the mass transfer properties in nanocomposites ...

Classification:

Property

Position in paragraph:

214

Length:

18

AnNER: Annotation Lifecycle

Example paragraph:

... the role of the **nanoparticle shape** ... on the modulation of the mass transfer properties in nanocomposites ...

Classification:

Property

Position in paragraph:

214

Length:

18

“... the role of the **nanoparticle shape** ... on the..”

└─ 227, 233 (shape)

└─ Candidate

└─ Structure

└─ 2025-12-05T14:05:06Z

└─ GPT-4o

{ entity span
status
classification
datetime
annotator

AnNER: Annotation Lifecycle

Example paragraph:

... the role of the **nanoparticle shape** ... on the modulation of the mass transfer properties in nanocomposites ...

Classification:

Property

Position in paragraph:

214

Length:

18

“... the role of the **nanoparticle shape** ... on the..”

└─ 227, 233 (shape)

└─ Candidate

└─ Structure

└─ 2025-12-05T14:05:06Z

└─ GPT-4o

└─ Rejected

└─ Structure

└─ 2025-12-07T15:21:18Z

└─ Jinwu Wang

{ entity span
status
classification
datetime
annotator

reviewer

AnNER: Annotation Lifecycle

Example paragraph:

... the role of the **nanoparticle shape** ... on the modulation of the mass transfer properties in nanocomposites ...

Classification:

Property

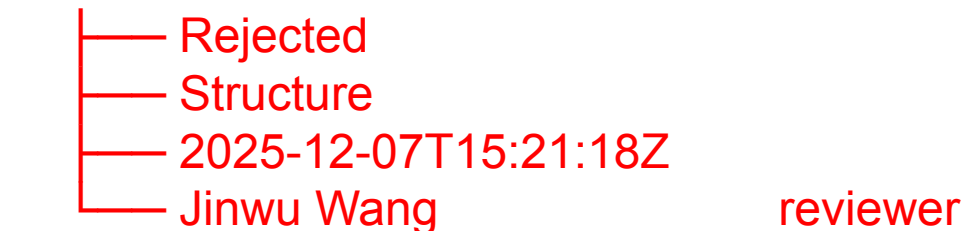
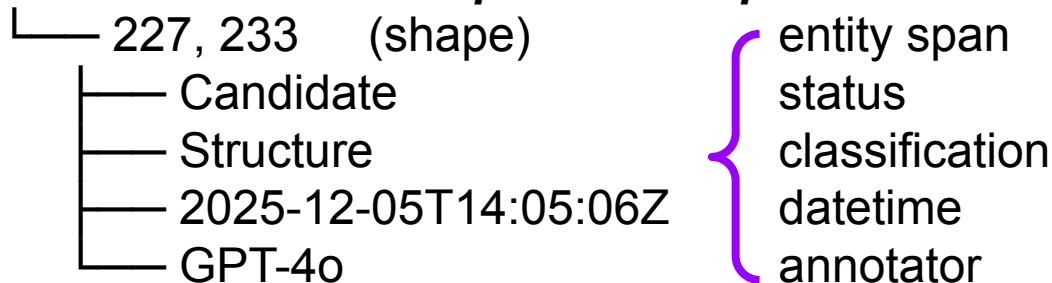
Position in paragraph:

214

Length:

18

“... the role of the **nanoparticle shape** ... on the..”

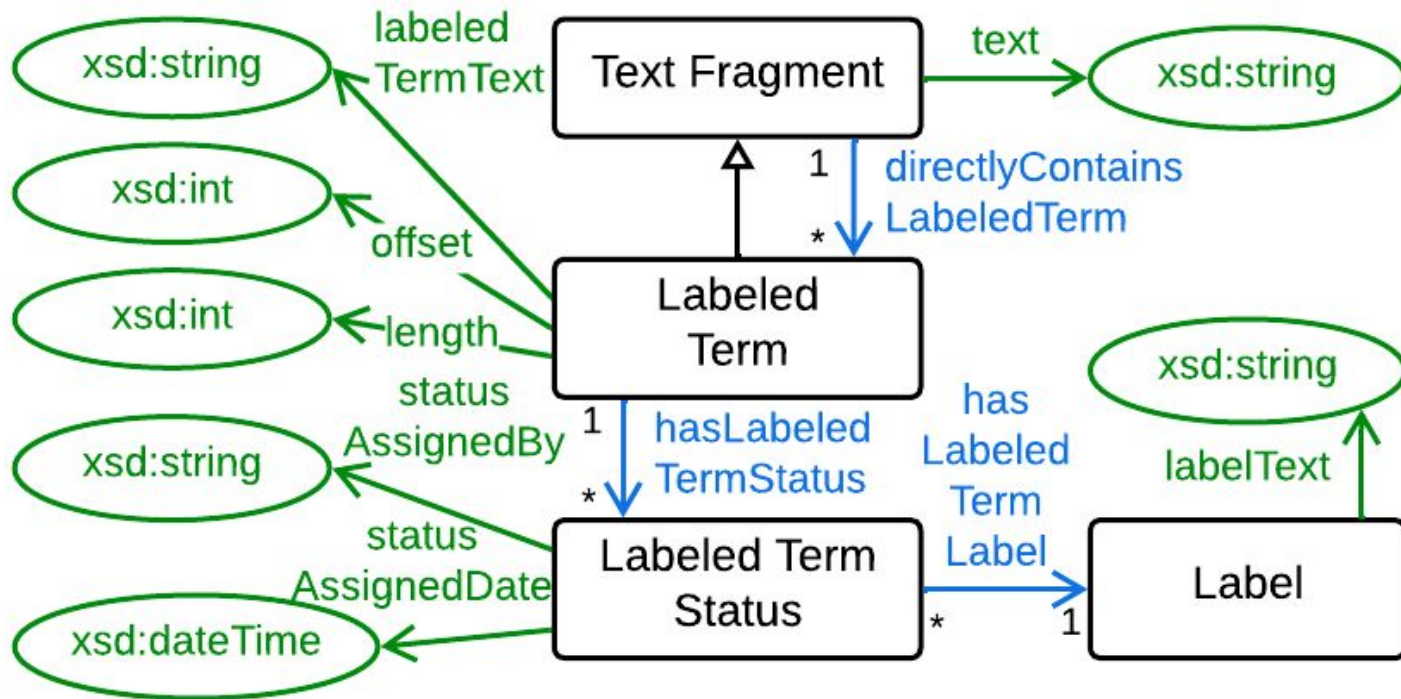


214, 233 (nanoparticle shape)



reviewer

AnNER: OnNER-Inspired Annotation Model



Reza, Zhang, Hahmann (2024). ***OnNER: An Ontology for Semantic Representation of Named Entities in Scholarly Publications.*** FOIS '24.

GitHub: <https://github.com/theSKAILab/OnNER>

Documentation: <https://theskailab.github.io/OnNER>



AnNER: Named Entity Annotation and Review Tool



Open a Text or JSON file to begin

Hint: You can also drag and drop files into this window!



GITHUB



DOCUMENTATION

This tool was designed and developed in the Spatial Knowledge and Artificial Intelligence (SKAI) lab at University of Maine.

The research was supported in part by the U.S. Department of Agriculture through a National Institute of Food and Agriculture (NIFA) grant

Award #[2021-67022-34366](#)

Thank You!

Contact:

a.reza@maine.edu