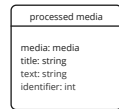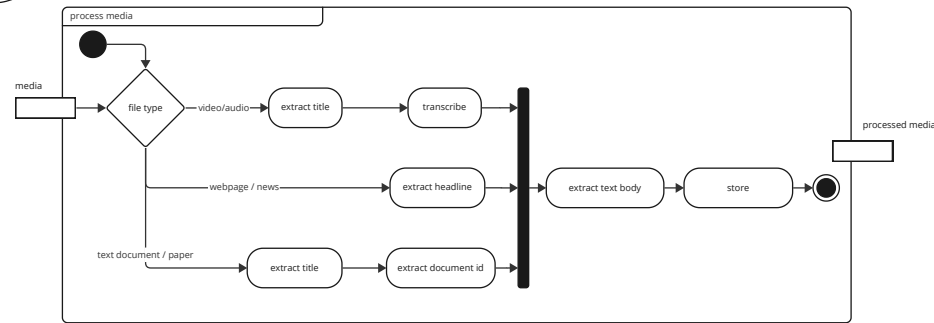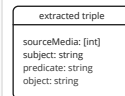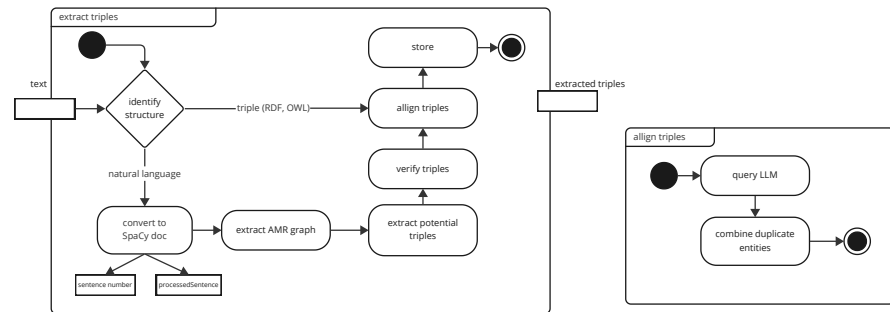# Scientific veracity scoring of news reports

by Constantin Tremel
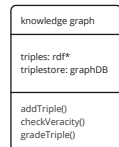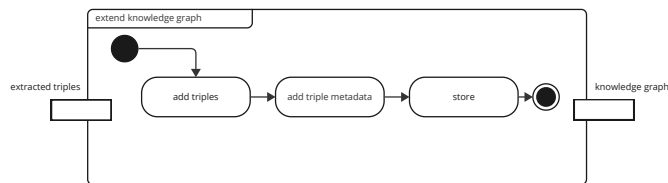(constantin.tremel@stud.uni-hannover.de)

With the vast amount of available media, it is impossible for humans to process it all. Some media receives significant attention, but its alignment with trusted scientific knowledge remains unclear.

To address this issue, we propose calculating a scientific veracity score by testing popular media, such as news articles and videos, against trusted sources, such as papers and journals.
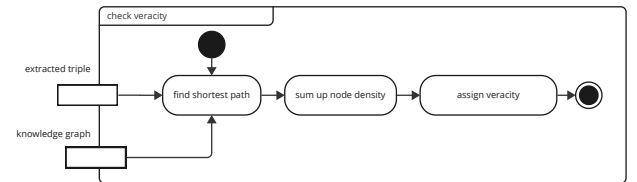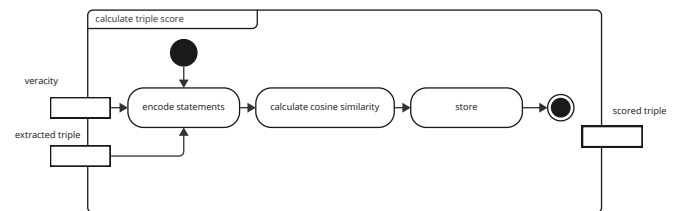
1

process media

media · file type · video/audio → extract title → transcribe · webpage / news → extract headline → extract text body → store → processed media · text document / paper → extract title → extract document id

processed media

media: media
title: string
text: string
identifier: int

2

extract triples

text · identify structure · triple (RDF, OWL) → allign triples · natural language → convert to SpaCy doc → extract AMR graph → verify triples → extract potential triples → store → extracted triples
sentence number · processedSentence

allign triples

query LLM → combine duplicate entities

3

extend knowledge graph

extracted triples → add triples → add triple metadata → store → knowledge graph

4

check veracity

extracted triple → find shortest path → sum up node density → assign veracity
knowledge graph

5

calculate triple score

veracity · extracted triple → encode statements → calculate cosine similarity → store → scored triple

6

calculate media score

scored triples → average triple scores → store → media score

media

fileLocation: string
fileType: string
downloadUrl: string
accessDate: string

process trusted media

1 process media → 2 extract triples → 3 extend knowledge graph

extracted triple

sourceMedia: [int]
subject: string
predicate: string
object: string

process popular media

1 process media
2 extract triples → list of triples → <<iterative>> check veracity (4) → calculate triple score (5) → list of triples → calculate media score (6)

knowledge graph

triples: rdf*
triplestore: graphDB

addTriple()
checkVeracity()
gradeTriple()

scored triple

triple: extractedTriple
veracity: float
clearness: float
confidence: float

media score

veracity: float
clearness: float
confidence: float