

# CheckThat! 2025

8<sup>th</sup> edition

**Subjectivity, Fact-Checking, Claim Extraction & Normalization, and Retrieval**

<https://checkthat.gitlab.io>



# Our Organization Team



**Firoj Alam**



**Julia Maria  
Struß**



**Tanmoy  
Chakraborty**



**Stefan Dietze**



**Salim Hafid**



**Katerina  
Korre**



**Arianna  
Muti**



**Preslav Nakov**



**Federico  
Ruggeri**



**Sebastian  
Schellhammer**



**Vinay Setty**



**Megha  
Sundriyal**

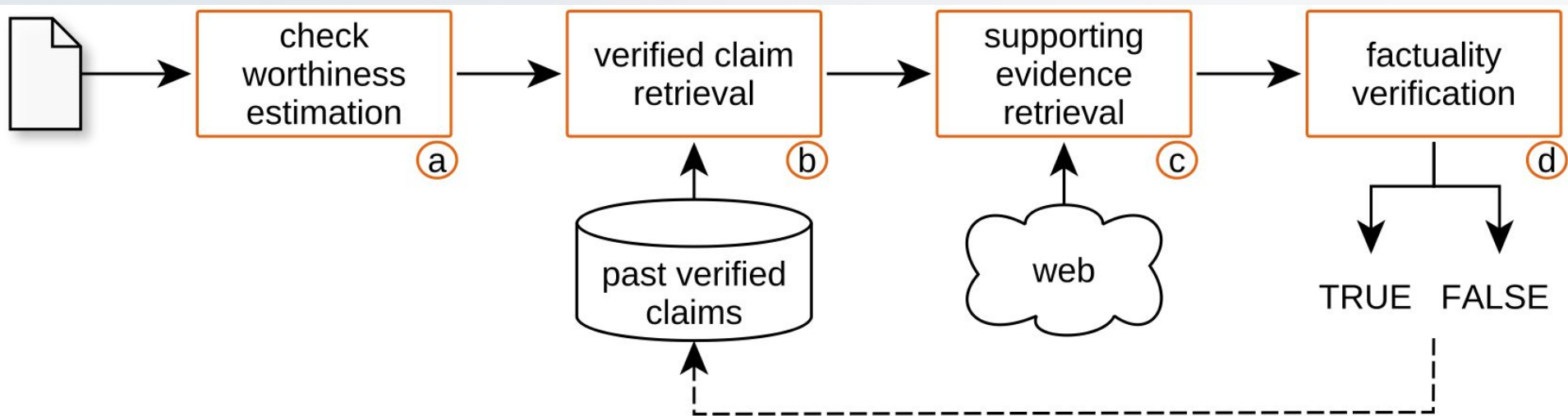


**Konstantin  
Todorov**

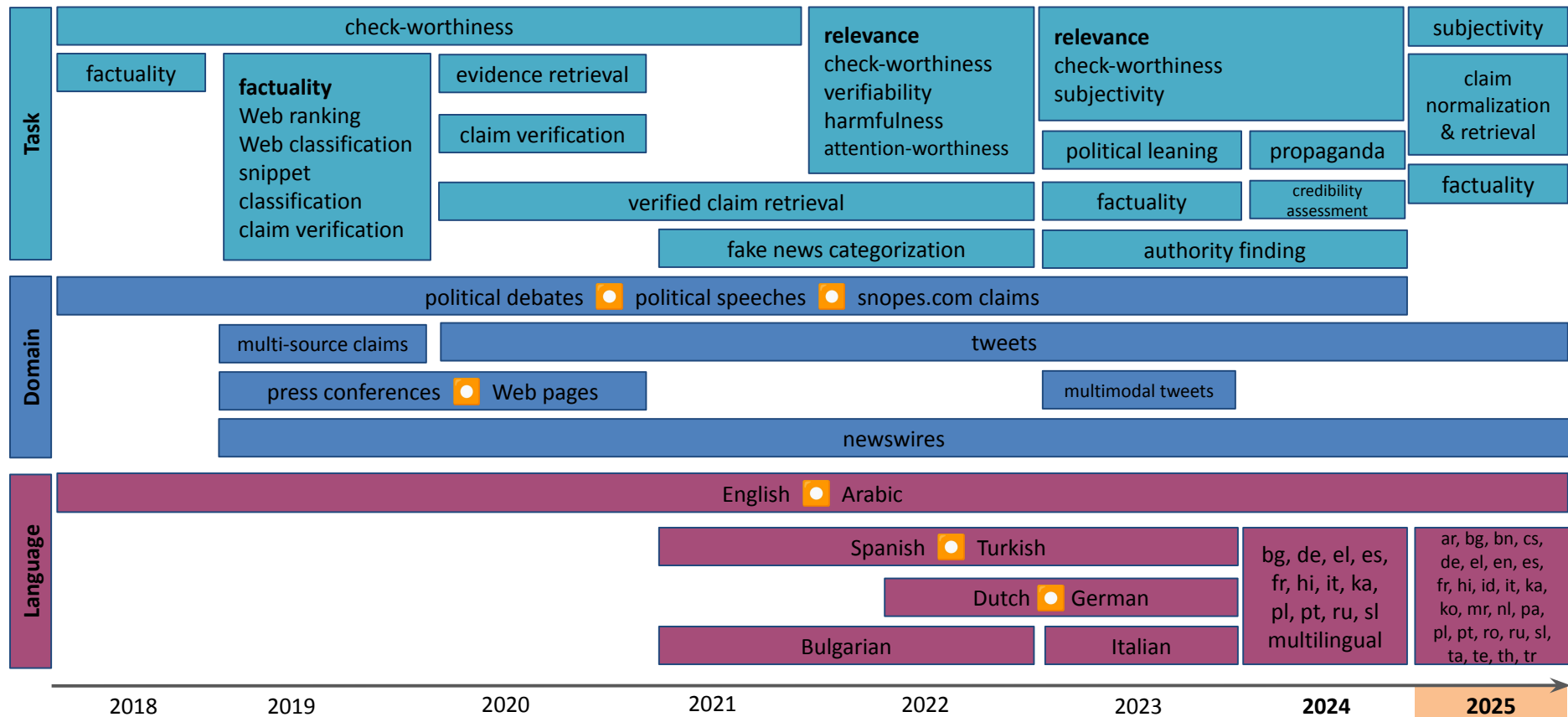


**Venkatesh  
Viswanathan**

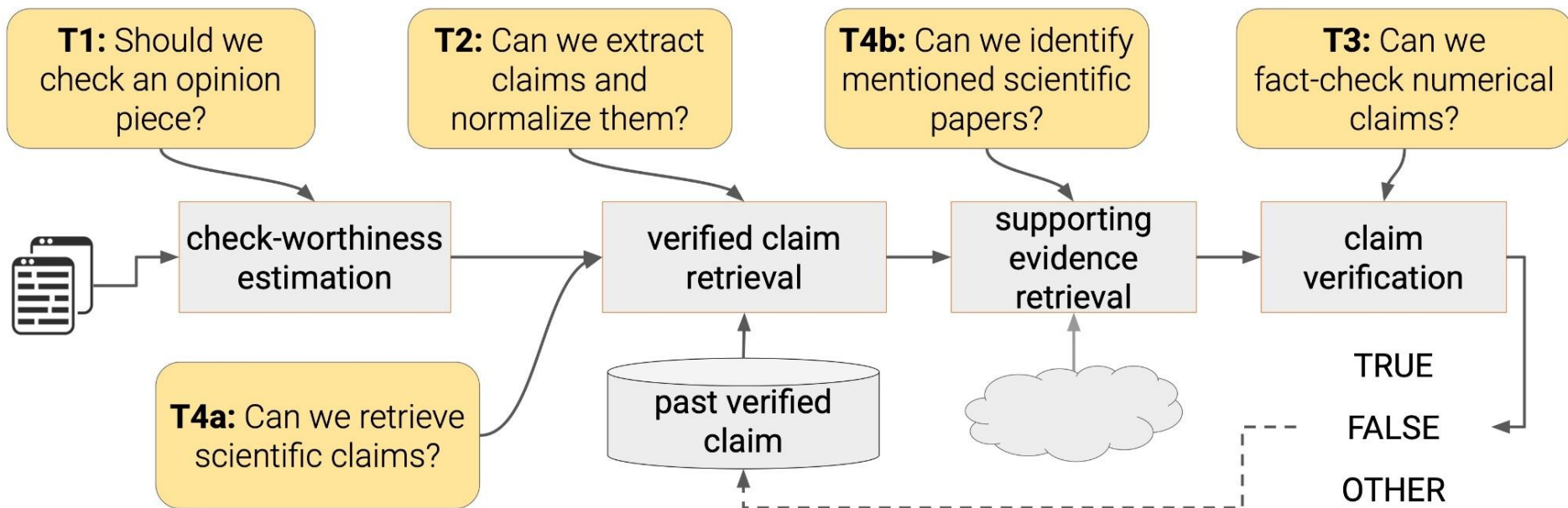
# The CheckThat! Verification Pipeline



# The CLEF CheckThat! Lab:Tasks, Lang & Data



# The Verification Pipeline and 2025 Tasks



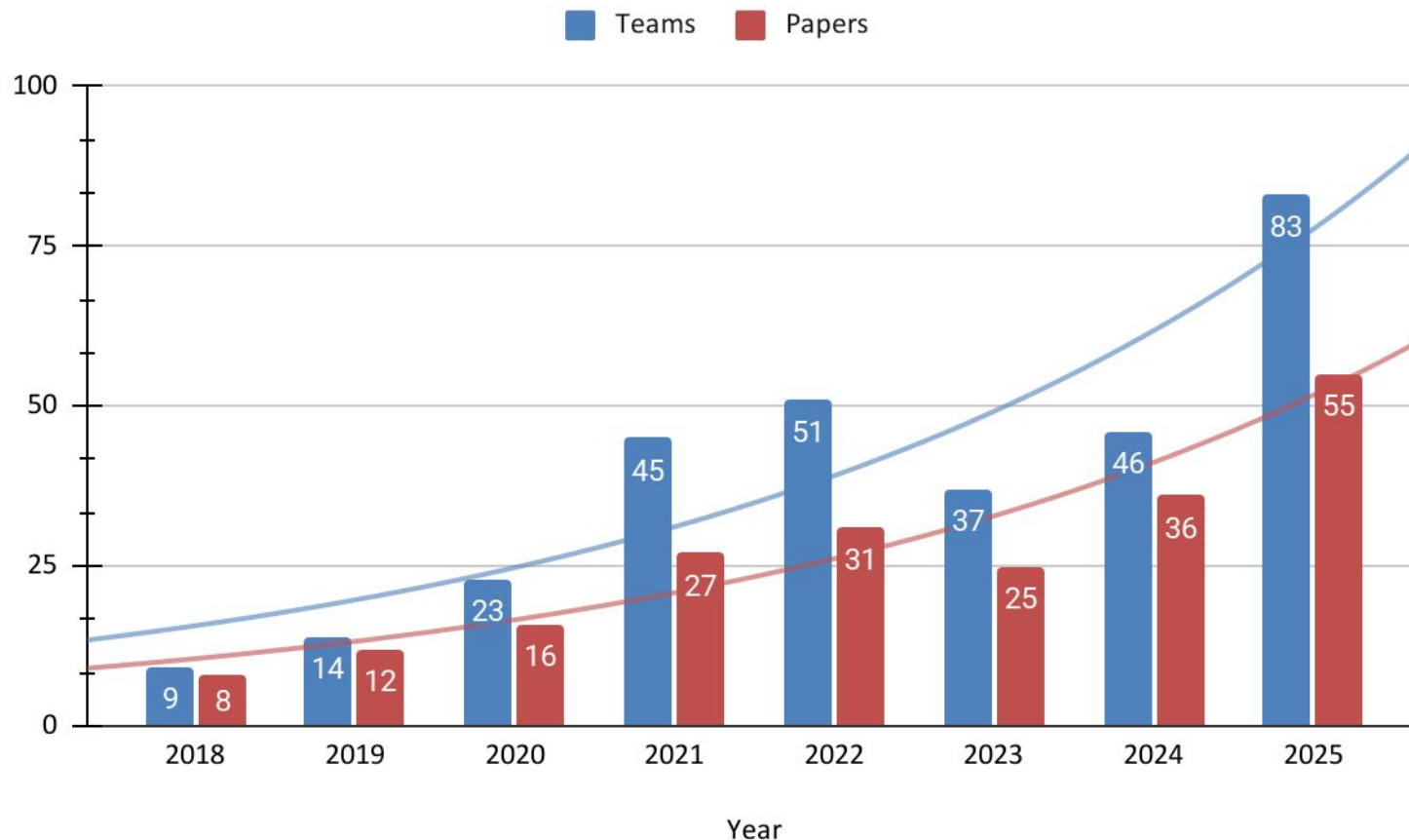
# CLEF CheckThat! Lab 2025 Tasks



1. **Subjectivity detection** (3rd)
2. **Claim Normalization** (1st)
3. **Fact-Checking Numerical Claims** (1st)
4. **Scientific Web Discourse** (1st)

										ar	bg	en	de	it	el	pl	ro	uk	
ar	bn	cs	de	el	en	fr	hi	ko	mr										
id	nl	pa	pl	pt	rm	es	ta	te	th										
																	ar	en	es
																			en

# Evolution in Terms of Participation





# CheckThat! 2025 Task 1

## Subjectivity in News Articles

Determine whether a sentence from a news article is subjective or objective.

Language	Sentence	Class
Arabic	وجدت بوحيد نفسها بين يدي ضباط المستعمَر الفرنسي فريسة ينهش لحمها بكل الطرق.	SUBJ
	كما تدخل نترات الأمونيوم في صناعة المتفجرات خاصة في مجال التعدين والمناجم.	OBJ
Bulgarian	Думите на Тръмп са просто думи, докато тези на Обама означават война.	SUBJ
	Аз се почувствах се глупаво, когато разбрах фактите.	OBJ
English	<i>But the state's budget is nothing like a credit card.</i>	SUBJ
	<i>The plan incorporates cash payments supplemented by contingent contributions.</i>	OBJ
German	<i>Den Grünen bleibt nur, immer wieder darauf hinzuweisen, dass sie selbst gerne ein bisschen großzügiger wären -sich damit aber leider nicht durchsetzen können.</i>	SUBJ
	<i>Mitte November kündigte die Ampel-Koalition an, das zu ändern.</i>	OBJ
Italian	<i>Inoltre paragonare immagini di attori paparazzati per strada a foto di studio photo-shoppate non ha senso.</i>	SUBJ
	<i>Il presidente russo, Vladimir Putin, ha visitato Kaliningrad per incontrare gli studenti dell'Università Kant e tenere un incontro sullo sviluppo della regione.</i>	OBJ



# CheckThat! 2025 Task 1

## Model Highlights

---

- Transformers were most popular, both monolingual and multilingual.
- Many performed feature selection for improvement
- Few approaches relied on LLM-based translation and data augmentation

[illegible]



# CheckThat! 2025 Task 2

## Claim Normalization

*Given noisy social media posts the task is to transform them into clear, concise, and verifiable statements known as normalized claims, which capture the core factual assertion of a post.*

**Task 2 was offered in 20 languages:** English, German, French, Spanish, Portuguese, Hindi, Marathi, Punjabi, Tamil, Arabic, etc.

	Social Media Post	Normalized Claim
English	Something to #consider don't you #think ? Something to #consider, don't you #think? Something to #consider, don't you #think? 40 years worth of research...*no vaccine for HIV *At least 100 years of research...no vaccine for cancer Ongoing research... no vaccine for the common cold Less than a year for a Covid vaccine?	Vaccines for HIV, cold, and cancer should deter you from getting the Covid-19 vaccine.
German	Das reiche Deutschland, wir haben das geringste Durchschnittseinkommen, die geringsten Renten und die dümmsten Wähler. ( <i>Translation: Rich Germany, we have the lowest average income, the lowest pensions and the stupidest voters.</i> )	Deutschland hat geringste Durchschnittseinkommen und Renten. ( <i>Translation: Germany has the lowest average incomes and pensions.</i> )

# CheckThat! 2025 Task 2

## Model Highlights



Team	English	Arabic	German	French	Hindi	Marathi	Indonesian	Punjabi	Polish	Portuguese	Spanish	Tamil	Thai	Bengali	Telugu	Dutch	Czech	Greek	Romanian	Korean
dfkinit2b [39]	1	1	2	2	1	1	2	1	2	2	2	1	3	1	1	1	1	1	1	1
DS@GT [40]	2	2	1	1	2	4	1	5	1	1	1	3	1	4	5	5	3	4	4	3
TIFIN [41]	3	5	5	6	7	6		4	5		5	5		5	6	4				
AKCIT-FN [42]	4	6	3	3	5	5	3	2	3	3	3	2	2	3	2	2	4	2	2	2
Factiverse and IAI [43]	5	7	4	4	8	9	4	7	6	6	6	8	4	6	7		5	5	5	
rohan_shankar <sup>+</sup>	6																			
manan-tifin <sup>+</sup>	7			7	9	7			5					5	6					
MMA [44]	8	3	7	8	6	3	5	6	7	4	4	6								
UNH [45]	9																			
Investigators [46]	10									8										5
OpenFact [47]	11	4	6	5	4	2	6	3	4	5	7	4	5	2	3	3	2	3	3	4
Nikhil_Kadapala <sup>+</sup>	12																			
aryasuneesh <sup>+</sup>	13		5	6	7	6		4			5	5	6							
JU_NLP@M&S [48]	14																			
uhh_dem4ai <sup>+</sup>	15																			
UmuTeam [49]	16	8	8	9	10	8	7	8	8	7	9	7	7	7	8	6	6	6	6	6
VSE <sup>+</sup>	17																			
saivineetha [50]					3									4						

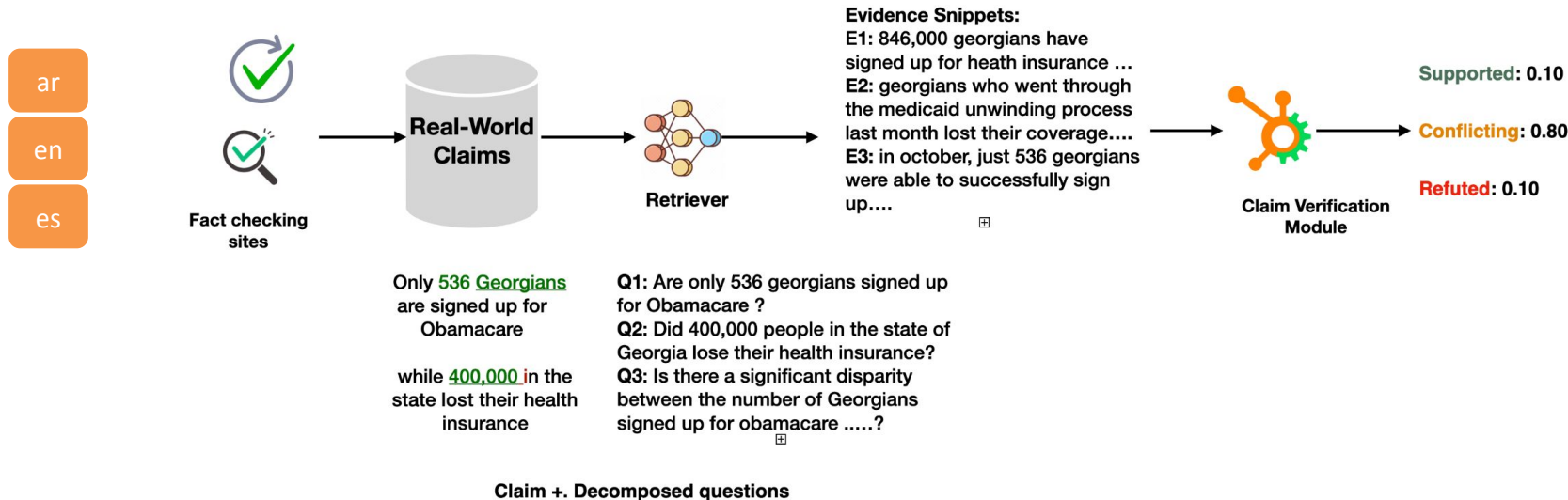
Team	Setting	Data	Approach	Model Family
	Monolingual			
	Zero-Shot			
	Content Filtering			
	Deduplication			
	Data Augmentation			
	Fine-Tuning (Full/PEFT)			
	Zero-Shot Prompting			
	Retrieval-Augmented ICL			
	Self-Reflection/Reasoning			
	Ensemble/Hybrid System			
	T5-family (T5, Flan-T5)			
	BART			
	GPT-family			
	Llama-family			
	Qwen			
	Gemma			
dfkinit2b [39]	✓	✓	✓	✓
DS@GT [40]	✓	✓	✓	✓
TIFIN [41]	✓	✓	✓	✓
AKCIT-FN [42]	✓	✓	✓	✓
Factiverse and IAI [43]	✓	✓	✓	✓
MMA [44]	✓	✓	✓	✓
UNH [45]	✓	✓	✓	✓
Investigators [46]	✓	✓	✓	✓
OpenFact [47]	✓	✓	✓	✓
JU_NLP@M&S [48]	✓	✓	✓	✓
Saivineetha [50]	✓	✓	✓	✓
UmuTeam [49]	✓	✓	✓	✓

# CheckThat! 2025 Task 3

## Fact-checking numerical claims



**Task :** *Given a numerical claim and the retrieved evidence snippets, the goal is to predict if the evidence supports, refutes, conflicting or is unrelated to the numerical claim.*



# CheckThat! 2024 Task 3: Model Highlights



Dataset statistics for different languages, such as English, Spanish and Arabic.

Split	English				Spanish				Arabic			
	T	F	C	Total	T	F	C	Total	T	F	C	Total
Train	1,824	5,770	2,341	9,935	127	1,200	179	1,506	975	1,216	-	2,191
Dev	617	1,795	672	3,084	30	299	48	377	274	313	-	587
Test	717	2,275	664	3,656	115	1,539	152	1,806	206	276	-	482

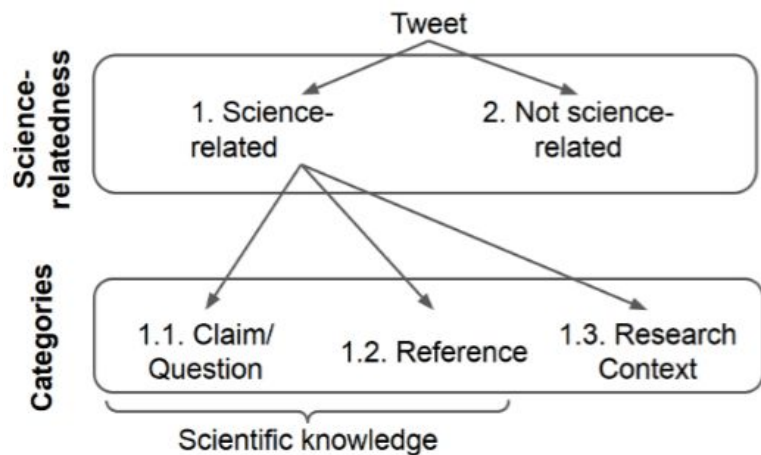
Team	Language			Model														Macro-F1		
	Arabic	Spanish	English	BM25	cross-encoder	gpt-4o-mini	Qwen	Llama	DeepSeek	ModernBERT	Math-Roberta	RoBERTa-base	QWQ-32B	Qwen-8B	Deberta-Large-MNLI	mxbai-rerank-large-v1	granite-3.3-8b-instruct	Arabic	Spanish	English
LIS [45]	☐	☐	☐										☐					50.34	96.15	59.54
DS@GT-CheckThat! [41]			☐															-	-	52.10
TIFIN [39]	☐		☐	☐											☐		☐	55.36	-	55.70
ClaimIQ [11]			☐															-	-	42.43
FraunhoferSIT [59]			☐						☐									-	-	51.00
NGU_Research [1]	☐	☐	☐	☐	☐	☐			☐			☐						63.52	24.41	-
JU_NLP [23]	☐		☐	☐	☐	☐												36.38	-	48.83
CornellNLP [26]			☐	☐	☐	☐		☐										-	-	48.57
UGLPN [80]			☐	☐	☐													-	-	45.53
UCOM_UNAM_PLN [2]	☐	☐		☐	☐	☐							☐					-	35.95	-
News-polygraph*	☐			☐	☐	☐		☐										-	-	42.86

# CheckThat! 2025 Task 4

## Scientific Web Discourse

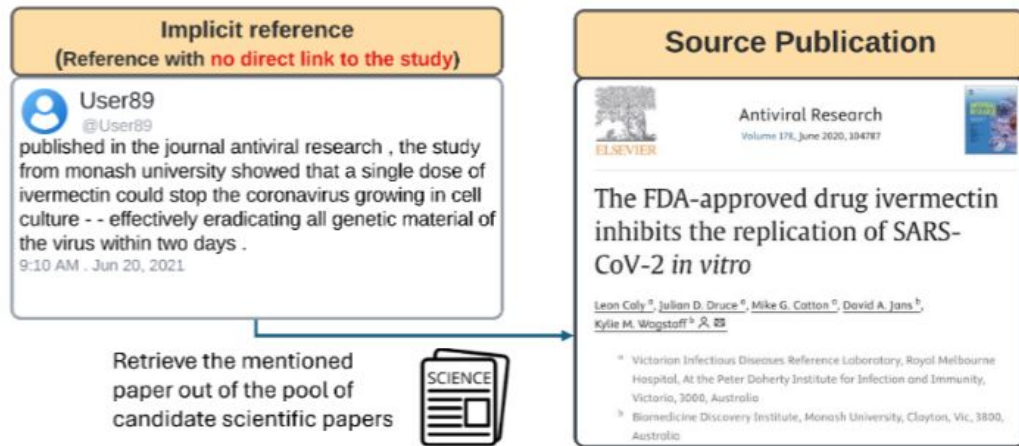


### Task 4a: Scientific Web Discourse Detection



**Objective:** Detecting different forms of scientific web discourse (e.g., **claims**, **references**)

### Task 4b: Scientific Claim Source Retrieval



**Objective:** Retrieving source publications from which claims and references originate

# CheckThat! 2025 Task 4

## Model Highlights



### Task 4a: Scientific Web Discourse Detection

Team	Models					Misc.	Perf.			
	DeBERTa-v3	SciBERT	Twitter-RoBERTa	LLMs	Others	Data Augmentation	Ensemble	Other Optimizations	Macro-avg. F1-Score	Rank
ClimateSense [18]			■		■		■		0.7998	1
VerbaNexAI [20]	■						■	■	0.7983	2
SBU-SCIRE [21]	■					■		■	0.7917	4
DS@GT [23]	■			■	■		■		0.7685	6
DeBERTa-v3 Baseline	■								0.7668	7
TurQUaz [24]				■			■		0.7615	8
JU_NLP [25]		■	■				■		0.7347	9

### Task 4b: Scientific Claim Source Retrieval

Team	Models				Misc.	Perf.		
	Dense Retrieval	Sparse Retrieval	Re-ranking	LLMs	Data Augmentation	Style transfer	MRR@5	Rank
AIRwaves [26]	■		■				0.67	2
Deep Retrieval [27]	■	■	■	■			0.66	3
ATOM [28]	■		■				0.66	4
SBU-SCIRE [21]	■		■				0.65	5
SeRRa [29]	■		■				0.61	8
Claim2Source [30]	■	■		■		■	0.59	12
DS@GT [31]		■	■	■	■	■	0.58	16
BM25 Baseline		■					0.43	28





# Programme

## CT! oral session 1: Thursday 11th September, 14:15 to 15:45

14:15 - Introduction to the CheckThat! Lab

15:00 - **Task 1 & 2:** Three talks on Subjectivity and Claim Normalization

## CLEF poster session 3: Thursday 11th September, 15:45 to 16:30

## CT! oral session 2: Thursday 11th September, 16:30 to 18:00

16:30 - **Task 2:** One talk on Claim Normalization

16:45 - **Task 3:** Three talks on Numerical Claim Verification

17:30 - **Task 4:** Two talks on Numerical Claim Verification

## CT! oral session 3: Friday 12th September, 11:30 to 13:00

11:30 - **Invited talk.** Rubén Míguez Pérez

Details on the CheckThat! website:

<http://checkthat.gitlab.io/clef2025/#lab-programme>

## **Title:** Automated detection of disinformation campaigns targeting brands. Turning automated fact-checking into a profitable business

**Speaker:** Rubén Míguez



Rubén Míguez holds a PhD in Telecommunications Engineering from the University of Vigo and an MBA from the School of Industrial Organization. He started his career in research at the University of Vigo, specializing in intelligent systems. In 2018, he joined Newtral as a product leader and head of technology, where he focused on leveraging AI to combat misinformation. He is also the founder of a tech startup and has received several accolades, including the National University Entrepreneur Award and the Antonio Palacios Award for Innovation.

Rubén has served as a mentor on platforms such as Startup Pirates and has presented at various national and international events. Rubén Míguez is the CTO and co-founder of TrueFlag.ai, a Newtral spin-off dedicated to applying automated fact-checking for brand protection. TrueFlag is the first multilingual and multimodal SaaS platform designed to detect and prevent disinformation campaigns in real-time.

**Title: Automated detection of disinformation campaigns targeting brands. Turning automated fact-checking into a profitable business**

### **Abstract:**

In this talk, Rubén Míguez will share his work at Trueflag.ai, a cutting-edge startup focused on leveraging automated fact-checking technologies to detect targeted disinformation campaigns on social media directed at major brands. He will discuss how the original technology developed for fact-checkers at Newtral has evolved into a new AI framework—combining open-source and proprietary models—designed to identify harmful narratives affecting brands across industries such as energy and banking. The presentation will outline the key challenges that lie ahead, highlight the solutions already deployed, and trace the transition from a monolithic, BERT-based architecture to a multi-agent framework for disinformation detection. In addition, we'll explore the core business use cases and scientific objectives driving the development of this technology, offering the audience a broader perspective on what it takes to transform automated fact-checking into a viable and impactful business model—while addressing the real-world needs of fact-checkers and researchers alike.

# CheckThat! 2025

<https://checkthat.gitlab.io>



MOHAMED BIN ZAYED  
UNIVERSITY OF  
ARTIFICIAL INTELLIGENCE



GESIS Leibniz-Institut  
für Sozialwissenschaften



Heinrich Heine  
Universität  
Düsseldorf



University  
of Stavanger



INDRAPRASTHA INSTITUTE of  
INFORMATION TECHNOLOGY  
DELHI



UNIVERSITÉ DE  
MONTPELLIER

SciencesPo



Università  
Bocconi  
MILANO

# Acknowledgements

