Task website Go to landing page Log in Sign up

## ❤ Welcome to the submission website for The Mu-SHROOM shared task (SemEval 2025 Task 3)! ❤

Mu-SHROOM is a multilingual shared task to advance the SOTA in hallucination detection for content generated with LLMs.

Thanks for your interest in the task! You can download the data for your participation from <u>our shared task website</u>. Consider joining the <u>Google group</u> or the <u>shared task Slack</u> to stay informed!

AR			CA				CS				DE			
	Team	loU	Сог	Team	loU	Сог		Team	IoU	Сог		Team	loU	Сог
1	NotMSA	0.6700	0.648{ 1	UCSC	0.7211	0.7779	1	AILSNTUA	0.5429	0.5560	1	UCSC	0.6236	0.65
2	UCSC	0.6594	0.632{ 2	ccnu	0.6694	0.7479	2	UCSC	0.5393	0.5763	2	NotMSA	0.6133	0.61
3	smurfcat	0.6274	0.5864 3	smurfcat	0.6681	0.7127	3	NotMSA	0.5073	0.5516	3	ccnu	0.5917	0.60
4	Deloitte	0.6043	0.604( 4	AILSNTUA	0.6664	0.6986	4	HalluSearch	0.4911	0.4942	4	AILSNTUA	0.5820	0.63
5	ccnu	0.5995	0.658: 5	NCL-UoR	0.6602	0.7203	5	ccnu	0.4852	0.5541	5	ATLANTIS	0.5774	0.01
6	Cantharellus	0.5804	0.588( 6	NotMSA	0.6545	0.7126	6	smurfcat	0.4608	0.4676	6	Deloitte	0.5655	0.54
7	DeepPavlov	0.5628	0.5754 7	TUM-MiKaNi	0.5971	0.5551	7	Deloitte	0.4428	0.4808	7	Cantharellus	0.5639	0.53
8	BlueToad	0.5470	0.505{ 8	CL-UZH	0.5857	0.6420	8	NCL-UoR	0.4409	0.5285	8	LCTeam	0.5634	0.50
9	NCL-UoR	0.5390	0.571( 9	Deloitte	0.5295	0.5571	9	LCTeam	0.4051	0.4357	9	smurfcat	0.5608	0.57
10	HalluSearch	0.5362	0.525{ 10	Cantharellus	0.5231	0.5727	10	Cantharellus	0.3936	0.4239	10	TUM-MiKaNi	0.5569	0.50

-	Team	loU	Сог	Team	IoU	Сог		Team	IoU	Сог		Team	loU	Сог
11	LCTeam	0.5335	0.5537 11	HalluSearch	0.5215	0.5704	11	CL-UZH	0.3931	0.4098	11	NCL-UoR	0.5473	0.58
12	CL-UZH	0.5253	0.487 12	LCTeam	0.4924	0.4917	12	TUM-MiKaNi	0.3874	0.3738	12	BlueToad	0.5439	0.52
13	AILSNTUA	0.5140	0.575 <sup>-</sup> 13	nsu-ai	0.4682	0.5346	13	tsotsalab	0.3613	0.3668	13	HalluSearch	0.5187	0.50
14	TUM-MiKaNi	0.4778	0.511, 14	uir-cis	0.4644	0.5432	14	BlueToad	0.3514	0.3628	14	CL-UZH	0.5123	0.50
15 1	nsu-ai	0.4756	0.423( 15	tsotsalab	0.4607	0.5187	15	DeepPavlov	0.3422	0.3192	15	Swushroomsia	0.5093	0.49
16	tsotsalab	0.4673	0.476! 16	UMUTEAM	0.4301	0.4295	16	UMUTEAM	0.3380	0.3600	16	DeepPavlov	0.5040	0.61
17	REFIND	0.3743	0.1811 17	DeepPavlov	0.4179	0.6742	17	uir-cis	0.3060	0.2695	17	nsu-ai	0.4841	0.45
18 l	keepitsimple	0.3631	0.249! 18	keepitsimple	0.3161	0.3377	18	nsu-ai	0.3051	0.2948	18	UMUTEAM	0.4093	0.44
	Baseline	0.3614	0.0067 19	Go Bison	0.2731	0.3749	19	Go Bison	0.2978	0.3066	19	tsotsalab	0.3969	0.36
	(mark all)	0.2426	20	Baseline	0.2423	0.0600	20	keepitsimple	0.2895	0.2423	20	REFIND	0.3862	0.35
20	UMUTEAM	0.3436	0.421	(mark all)			21	REFIND	0.2761	0.0924	21	keepitsimple	0.3651	0.21
21	TrustAl	0.3428	0.238( 21	FENJI	0.1796	0.0600	22	Baseline	0.2632	0.1000	22	TU Munich	0.3476	-0.00
22	CUET_SSTM	0.3413	0.2247 22	CIMAT_NLP	0.1410	0.0690		(mark all)			23	Baseline	0.3451	0.01
23 9	Swushroomsia	0.3097	0.287، 23	Baseline	0.0800	0.0600	23	CIMAT_NLP	0.2201	0.1450		(mark all)	0.5 15 1	0.01
24 (	uir-cis	0.2722	0.447	(mark none)			24	Baseline	0.1300	0.1000	24	uir-cis	0.3400	0.40
25	TU Munich	0.2527	0.3200	Baseline (neural)	0.0524	0.0645		(mark none)			25	TrustAl	0.3323	0.51
26	Go Bison	0.2138	0.3844				25	FENJI	0.1073	0.1000		Go Bison	0.2522	0.27

	Team	loU	Сог
27	CIMAT_NLP	0.2044	0.0775
28	HalluciSeekers	0.1180	0.0572
29	Hallucination Detectives	0.0760	0.0275
30	FENJI	0.0467	0.0067
30	FENJI  Baseline (mark none)	0.0467	
	Baseline (mark none)		

	Team	loU	OU Cor		Team	IoU	Сог
26	Baseline			27	FENJI	0.1624	0.01
	(neural)			28	HalluciSeekers	0.0573	0.04
				29	Baseline (neural)	0.0318	0.10
				30	Baseline (mark none)	0.0267	0.01
				31	S1mT5v-FMI	0.0267	0.01

EN		ES			EU				FA				
	Team	IoU	Сог	Team	loU	Сог	Team	loU	Сог		Team	loU	Сог
1	iai_MSU	0.6509	0.6294 1	ATLANTIS	0.5311	0.0137 1	NotMSA	0.6129	0.6202	1	AILSNTUA	0.7110	0.69
2	UCSC	0.6146	0.546 <sup>-</sup> 2	CIMAT_NLP	0.5209	0.5237 2	UCSC	0.5894	0.5826	2	UCSC	0.6949	0.69
3	ATLANTIS	0.5698	0.000( 3	NCL-UoR	0.5146	0.546, 3	ccnu	0.5784	0.6121	3	NotMSA	0.6693	0.67
4	HalluSearch	0.5656	0.536( 4	ccnu	0.5125	0.541! 4	AILSNTUA	0.5550	0.5805	4	ccnu	0.6600	0.67
5	ccnu	0.5394	0.5509 5	AILSNTUA	0.5004	0.564{ 5	Cantharellus	0.5339	0.5038	5	NCL-UoR	0.6586	0.67
6	NotMSA	0.5314	0.520( 6	UCSC	0.4794	0.6021 6	HalluSearch	0.5251	0.4789	6	Cantharellus	0.6551	0.68
7	AILSNTUA	0.5308	0.638 7	LCTeam	0.4434	0.433! 7	TUM-MiKaNi	0.5237	0.4709	7	smurfcat	0.6375	0.62

	Team	IoU	Сог	Team	loU	Сог	Team	loU	Сог		Team	loU	Сог
8	TUM-MiKaNi	0.5249	0.5363 8	smurfcat	0.4342	0.440( 8	Deloitte	0.5218	0.5157	8	LCTeam	0.6018	0.45
9	smurfcat	0.5241	0.5963 9	NotMSA	0.4162	0.545( 9	smurfcat	0.5195	0.4697	9	Deloitte	0.5754	0.51
10	Deloitte	0.5234	0.560{ 10	Deloitte	0.4065	0.585: 10	NCL-UoR	0.5105	0.5974	10	BlueToad	0.5711	0.57
11	NCL-UoR	0.5195	0.5477 11	CL-UZH	0.4051	0.508! 11	CL-UZH	0.5071	0.5108	11	TUM-MiKaNi	0.5465	0.42
12	Swushroomsia	0.5030	0.463; 12	HalluSearch	0.3883	0.445( 12	BlueToad	0.5061	0.4571	12	CL-UZH	0.5108	0.49
13	DeepPavlov	0.4989	0.602 <sup>-</sup> 13	Cantharellus	0.3869	0.423( 13	LCTeam	0.4804	0.5499	13	UMUTEAM	0.4677	0.39
14	CL-UZH	0.4850	0.4824 14	TUM-MiKaNi	0.3739	0.5027 14	nsu-ai	0.4368	0.4210	14	HalluSearch	0.4443	0.47
15	YNU-HPCC	0.4807	0.407! 15	uir-cis	0.3447	0.3104 15	keepitsimple	0.4193	0.3525	15	nsu-ai	0.3729	0.38
16	LCTeam	0.4725	0.553{ 16	UMUTEAM	0.2980	0.415; 16	REFIND	0.4074	0.2713	16	keepitsimple	0.3132	0.35
17	Cantharellus	0.4721	0.461: 17	nsu-ai	0.2854	0.396( 17	DeepPavlov	0.3872	0.3214	17	DeepPavlov	0.2405	0.18
18	BlueToad	0.4688	0.450! 18	GIL-IIMAS	0.2807	0.324: 18	Baseline	0.3671	0.0000	18	Baseline	0.2028	0.01
19	GIL-IIMAS	0.4607	0.501!	UNAM			(mark all)				(mark all)		
	UNAM		19	BlueToad	0.2787	0.4267 19	tsotsalab	0.3524	0.0000	19	tsotsalab	0.2028	0.01
20	CIMAT_NLP	0.4577	0.3707 20	TrustAl	0.2683	0.498: 20	UMUTEAM	0.3272	0.3925	20	uir-cis	0.1661	0.39
21	tsotsalab	0.4454	0.394( 21	DeepPavlov	0.2614	0.398! 21	uir-cis	0.2916	0.3989	21	Go Bison	0.1190	0.06
22	advacheck	0.4443	0.3432 22	TU Munich	0.2578	0.3229 22	Go Bison	0.2461	0.1707	22	HalluciSeekers	0.1126	0.07
23	nsu-ai	0.4436	0.4571 23	Swushroomsia	0.2466	0.248( 23	CIMAT_NLP	0.1755	0.0522	23	CIMAT_NLP	0.0316	0.39

Team

**FENJI** 

Baseline

(neural)

Baseline

(mark none)

loU

Сог

0.0208 0.1004 25

0.0101 0.0000 26

0.1326 0.0000 24 FENJI

Team

Baseline

(neural)

Baseline

(mark none)

loU

0.0028

0.0001 0.10

0.0000 0.01

Сог

0.01

	Team	IoU	Cor	Team	loU	Сог
24	uir-cis	0.4025	0.478 24	REFIND	0.2348	0.130{ 24
25	VerbaNexAl	0.3810	0.3642 25	keepitsimple	0.2131	0.233! 25
26	UMUTEAM	0.3667	0.496( 26	Baseline	0.1853	0.0137
27	keepitsimple	0.3660	0.2104	(mark all)		26
28	TU Munich	0.3646	0.216 <sup>2</sup>	tsotsalab	0.1853	0.0132
29	REFIND	0.3525	0.1082	FunghiFunghi	0.1616	-0.0986
30	Baseline	0.3489	0.0000	Go Bison	0.1341	0.3643
50	(mark all)	0.5 105	30	FENJI	0.1325	0.0132
31	MALTO	0.3269	0.3104 31	COGUMELO	0.1321	0.1013
32	RaggedyFive	0.3151	0.303{ 32	Baseline	0.0855	0.0132
33	COGUMELO	0.3107	0.227	(mark none)		
34	HalluRAG-	0.3093	0.083	S1mT5v-FMI	0.0855	0.0132
	RUG		34	Baseline	0.0724	0.0359
35	TrustAl	0.2980	0.5642	(neural)		
36	FunghiFunghi	0.2943	0.011L	HalluciSeekers	0.0519	0.0266
37	Hallucination Detectives	0.2142	0.1682			
38	FENJI	0.1856	0.0000			

	Team	IoU	Сог
39	Go Bison	0.1325	0.2752
40	DUTJBD	0.0571	-0.1883
41	HalluciSeekers	0.0542	0.1530
42	HausaNLP	0.0325	0.4226
43	Baseline (mark none)	0.0325	0.0000
44	Baseline (neural)	0.0310	0.1190

FI	FR					HI		IT				
	Team	loU	Сог	Team	loU	Сог	Team	loU	Сог	Team	loU	Сог
1	UCSC	0.6483	0.6498 1	Deloitte	0.6469	0.6187 1	ccnu	0.7466	0.7847 1	UCSC	0.7872	0.78
2	NotMSA	0.6422	0.5467 2	TUM-MiKaNi	0.6314	0.5157 2	UCSC	0.7441	0.762! 2	AILSNTUA	0.7660	0.81
3	smurfcat	0.6310	0.553! 3	NotMSA	0.6195	0.5553	AILSNTUA	0.7259	0.7602 3	smurfcat	0.7478	0.62
4	Deloitte	0.6307	0.6356 4	Swushroomsia	0.5937	0.542! 4	smurfcat	0.7164	0.5964 4	NotMSA	0.7369	0.75
5	TUM-MiKaNi	0.6267	0.575′ 5	UCSC	0.5868	0.5597 5	NotMSA	0.6842	0.7252 5	Swushroomsia	0.7274	0.72
6	AILSNTUA	0.6235	0.6204 6	smurfcat	0.5838	0.515! 6	LCTeam	0.6601	0.5122 6	NCL-UoR	0.7123	0.76
7	CL-UZH	0.6014	0.4736 7	DeepPavlov	0.5831	0.544( 7	Cantharellus	0.6572	0.6909 7	ccnu	0.7060	0.74
8	nsu-ai	0.5874	0.4922 8	AILSNTUA	0.5812	0.610: 8	BlueToad	0.6447	0.6844 8	Deloitte	0.7059	0.61

	Team	loU	Сог	Team	loU	Сог	Team	IoU	Сог	Team	IoU	Сог
9	DeepPavlov	0.5845	0.482´9	CL-UZH	0.5765	0.441 9	CL-UZH	0.6377	0.6687 9	LCTeam	0.7013	0.54
10	Cantharellus	0.5714	0.564€ 10	LCTeam	0.5634	0.488: 10	Deloitte	0.6322	0.639 10	Cantharellus	0.6907	0.71
11	BlueToad	0.5694	0.490€ 11	ATLANTIS	0.5190	0.4117 11	NCL-UoR	0.6286	0.683( 11	CL-UZH	0.6833	0.70
12	HalluSearch	0.5681	0.5297 12	nsu-ai	0.5181	0.433! 12	TUM-MiKaNi	0.5835	0.4964 12	TUM-MiKaNi	0.6787	0.53
13	ccnu	0.5117	0.563′ 13	Cantharellus	0.5147	0.5317 13	HalluSearch	0.5265	0.519! 13	BlueToad	0.6388	0.66
14	NCL-UoR	0.5096	0.496! 14	tsotsalab	0.4896	0.457! 14	DeepPavlov	0.5117	0.732( 14	HalluSearch	0.5484	0.56
15	REFIND	0.5061	0.196! 15	ccnu	0.4823	0.5724 15	nsu-ai	0.4771	0.4438 15	DeepPavlov	0.5280	0.55
16	Swushroomsia	0.4955	0.4298 16	REFIND	0.4734	0.075; 16	Swushroomsia	0.4534	0.4789 16	UMUTEAM	0.4413	0.46
17	Baseline	0.4857	0.000( 17	keepitsimple	0.4651	0.275( 17	UMUTEAM	0.4510	0.438€ 17	nsu-ai	0.4396	0.44
	(mark all)		18	TU Munich	0.4547	0.009( 18	keepitsimple	0.3598	0.3508 18	keepitsimple	0.4009	0.38
18	tsotsalab	0.4857	0.000(	Baseline	0.4543	0.000( 19	TrustAl	0.3144	0.505( 19	uir-cis	0.3967	0.49
19	TU Munich	0.4857	0.0032	(mark all)		20	TU Munich	0.2807	0.3297 20	TrustAl	0.3441	0.28
20	UMUTEAM	0.4563	0.5126 20	BlueToad	0.4385	0.3797	Baseline	0.2711	0.0000 21	TU Munich	0.3319	0.42
21	keepitsimple	0.4554	0.3323 21	HalluSearch	0.4366		(mark all)	0.2	22		0.3255	
22	LCTeam	0.4221	0.5300 22	Go Bison	0.4164	0.399( 22	tsotsalab	0.2711	0.0000			
23	Go Bison	0.3996	0.3433 23	NCL-UoR	0.4058	0.4187 23	Go Bison	0.2586	0.3217	Baseline (mark all)	0.2826	0.00
24	CIMAT_NLP	0.3742	0.031( 24	TrustAl	0.3799	0.4997 24	uir-cis	0.0613	0.5586 24	tsotsalab	0.2826	0.00

	Team	loU	Сог		Team	IoU	Сог		Team	IoU	Сог		Team	IoU	Сог
25	TrustAl	0.2955	0.1777	25	CIMAT_NLP	0.3533	0.071 2	25	Baseline	0.0029	0.1429	25	FENJI	0.2765	0.00
26	uir-cis	0.2459	0.3366	26	UMUTEAM	0.3200			(neural)				Go Bison	0.2675	0.40
27	FENJI	0.0941	0.0000	27	FunghiFunghi	0.3095	-0.152		FENJI	0.0000	0.0000	27	FunghiFunghi	0.2111	-0.21
28		0.0042	0.0924	28	uir-cis	0.2286		27	Baseline (mark none)	0.0000	0.0000	28	CIMAT_NLP	0.1899	0.04
	(neural)			29	FENJI	0.0844	0.0000					29	HalluciSeekers	0.0350	0.02
29	S1mT5v-FMI	0.0000	0.0014	30	HalluciSeekers	0.0500	0.0447					30	Baseline	0.0104	0.08
30	Baseline (mark none)	0.0000	0.0000	31	Baseline	0.0022	0.0208						(neural)		
	(mark none)				(neural)							31		0.0000	0.00
				32	Baseline (mark none)	0.0000	0.0000						(mark none)		

0.0000 0.0000

SV			ZH						
	Team	IoU	Сог	Team	IoU	Сог			
1	UCSC	0.6423	0.5204 1	YNU-HPCC	0.5540	0.3518			
2	NotMSA	0.6364	0.4224 2	LCTeam	0.5232	0.5171			
3	Deloitte	0.6220	0.537, 3	nsu-ai	0.4937	0.3813			
4	smurfcat	0.6174	0.5007 4	DeepPavlov	0.4900	0.2529			
5	AILSNTUA	0.6009	0.5627 5	smurfcat	0.4842	0.2529			

33 S1mT5v-FMI

	Team	loU	Сог	Team	loU	Cor
6	TUM-MiKaNi	0.5886	0.393( 6	CL-UZH	0.4790	0.1783
7	BlueToad	0.5854	0.4267	Baseline	0.4772	0.0000
8	HalluSearch	0.5622	0.4290	(mark all)	0.4772	0.0000
9	CL-UZH	0.5612	0.412!	tsotsalab	0.4772	0.0000
10	NCL-UoR	0.5547	0.458	TUM-MiKaNi	0.4735	0.4095
11	nsu-ai	0.5478	0.344;	UCSC	0.4707	0.3966
12	DeepPavlov	0.5380	0.414	keepitsimple	0.4703	0.1601
13	Baseline	0.5373	0.013(	NotMSA	0.4631	0.4363
	(mark all)		13	Deloitte	0.4600	0.2986
14	TU Munich	0.5372	0.0054 14	HalluSearch	0.4534	0.4232
15	tsotsalab	0.5349	0.013( 15	TrustAl	0.4304	0.2503
16	ccnu	0.5045	0.505{ 16	Cantharellus	0.4011	0.4063
17	UMUTEAM	0.4393	0.393( 17	UMUTEAM	0.3875	0.4916
18	LCTeam	0.4183	0.370( 18	AILSNTUA	0.3866	0.4564
19	FunghiFunghi	0.4156	-0.117 19	ccnu	0.3834	0.4042
20	keepitsimple	0.3967	0.217( 20	NCL-UoR	0.3606	0.3540
21	Swushroomsia	0.3549	0.226! 21	BlueToad	0.2783	0.2262

	Team	loU	Сог	Team	loU	Сог
22	uir-cis	0.3080	0.365! 22	TU Munich	0.2160	0.0769
23	TrustAl	0.2484	0.255 23	Go Bison	0.2152	0.1119
24	CIMAT_NLP	0.2388	0.0547 24	Swushroomsia	0.2054	0.0966
25	FENJI	0.1154	0.013( 25	uir-cis	0.1913	0.3047
26	Go Bison	0.1110	0.066! 26	S1mT5v-FMI	0.0619	-0.0209
27	HalluciSeekers	0.0575	0.085( 27	FENJI	0.0371	0.0000
28	Baseline (neural)	0.0308	0.096{ 28	Baseline (neural)	0.0236	0.0884
29	Baseline (mark none)	0.0204	0.013( 29	Baseline (mark none)	0.0200	0.0000
30	S1mT5v-FMI	0.0204	0.0136			