

LREC 20
Palma **26**



Tesla



Science Fund
of the Republic of Serbia

LREC 2026: Language Resources and Evaluation

Ranka Stanković¹, Anđelka Zečević², Jovana Rađenović¹

¹RGF & JERTEH, {ranka.stankovic@rgf.bg.ac.rs|jovana.radjenovic@rgf.bg.ac.rs}

²MISANU & JERTEH, andjelkaz@mi.sanu.ac.rs

Palau de Congressos de Palma

Palma, Mallorca (Spain)

11-16 May 2026

LREC 20 *Palma* 26

- LREC 2026 je jedna od najvažnijih međunarodnih konferencija iz oblasti obrade prirodnog jezika, računarske lingvistike i jezičkih resursa.
- LREC konferencije organizuje ELRA (Language Resources Association), a održavaju se na svake dve godine.
- LREC 2026 je održan u Palma de Majorci, u Španiji, od 11. do 16. maja 2026. godine, u kongresnom centru Palau de Congressos de Palma.
- Glavna konferencija je trajala od 13. do 15. maja, dok su radionice i tutorijali organizovani pre i posle glavnog programa.

<https://lrec2026.info/> <https://www.elra.info/>

O istorijatu: koliko i gde?

Od 1998. je
organizovano ukupno **15**
konferencija: Granada,
Atina, Las Palmas,
Lisabon, Genova,
Marakeš, Malta,
Istambul, Rejkjavik,
Portorož, Mijazaki,
Marsej (online), Marsej,
Torino, Palma de
Majorka



LREC-COLING 2024
Turin

LREC 2024 Proceedings



LREC 2022
Marseille
Palais du Pharo
June 20-25, 2022

LREC 2022 Marseille

LREC 2022 Proceedings



LREC 2020
Marseille
Palais du Pharo
May 11-16, 2020

LREC 2020 Marseille

LREC 2020 Proceedings



LREC 2018
MIYAZAKI
May 7-12, 2018
Phoenix Seagala
Resort
Miyazaki, Japan

LREC 2018 Miyazaki

LREC 2018 Proceedings



LREC 2016
Portoroz
May 23-28, 2016
Grand Hotel Bernardin
Portoroz (Slovenia)

LREC 2016 Portoroz

LREC 2016 Proceedings



LREC 2014
Reykjavik
Harpa Conference Centre
26 - 31 May, 2014

LREC 2014 Reykjavik

LREC 2014 Proceedings



lrec 2012
istanbul2
14th Middle East conference center
21-27 may 2012

LREC 2012 Istanbul

LREC 2012 Proceedings



lrec2010
malta
17-23 may 2010

LREC 2010 Malta

LREC 2010 Proceedings



lrec 2008
may 26-june 1
marrakech

LREC 2008 Marrakech

LREC 2008 Proceedings



LREC
2006
22 - 28 May
Genoa - ITALY

LREC 2006 Genoa

LREC 2006 Proceedings



LREC
2004
24-30 May
Lisbon - PORTUGAL

LREC 2004 Lisbon

LREC 2004 Proceedings



LREC
LAS PALMAS - MAY 2002
CANARY ISLANDS, SPAIN
2002

LREC 2002 Las
Palmas

LREC 2002 Proceedings



LREC 2000 Athens

LREC 2000 Proceedings



LREC 1998 Granada

Koje teme pokriva konferencija?

- Jezičke resurse (korpusi, leksikoni, WordNet, FrameNet, ontologije),
- Evaluaciju sistema za obradu prirodnog jezika,
- Manje zastupljene jezike i digitalnu humanistiku,
- Anotaciju i standarde,
- Prepoznavanje i povezivanje imenovanih entiteta,
- Grafove znanja i semantički veb,
- Multimodalne i govorene resurse,
- Velike jezičke modele i generativne modele za obradu jezika.

Tipični radovi na LREC-u često imaju:

- opis resursa i toka obrade (pipeline),
- anotaciju i evaluaciju,
- arhitekturu podataka,
- primere upotrebe,
- dostupnost resursa ili demonstraciju sistema.

Format

- usmeno izlaganje,
- panel sesije.

Organizacija

- veliki broj paralelnih sesija
- pre i posle konferencije
brojne radionice

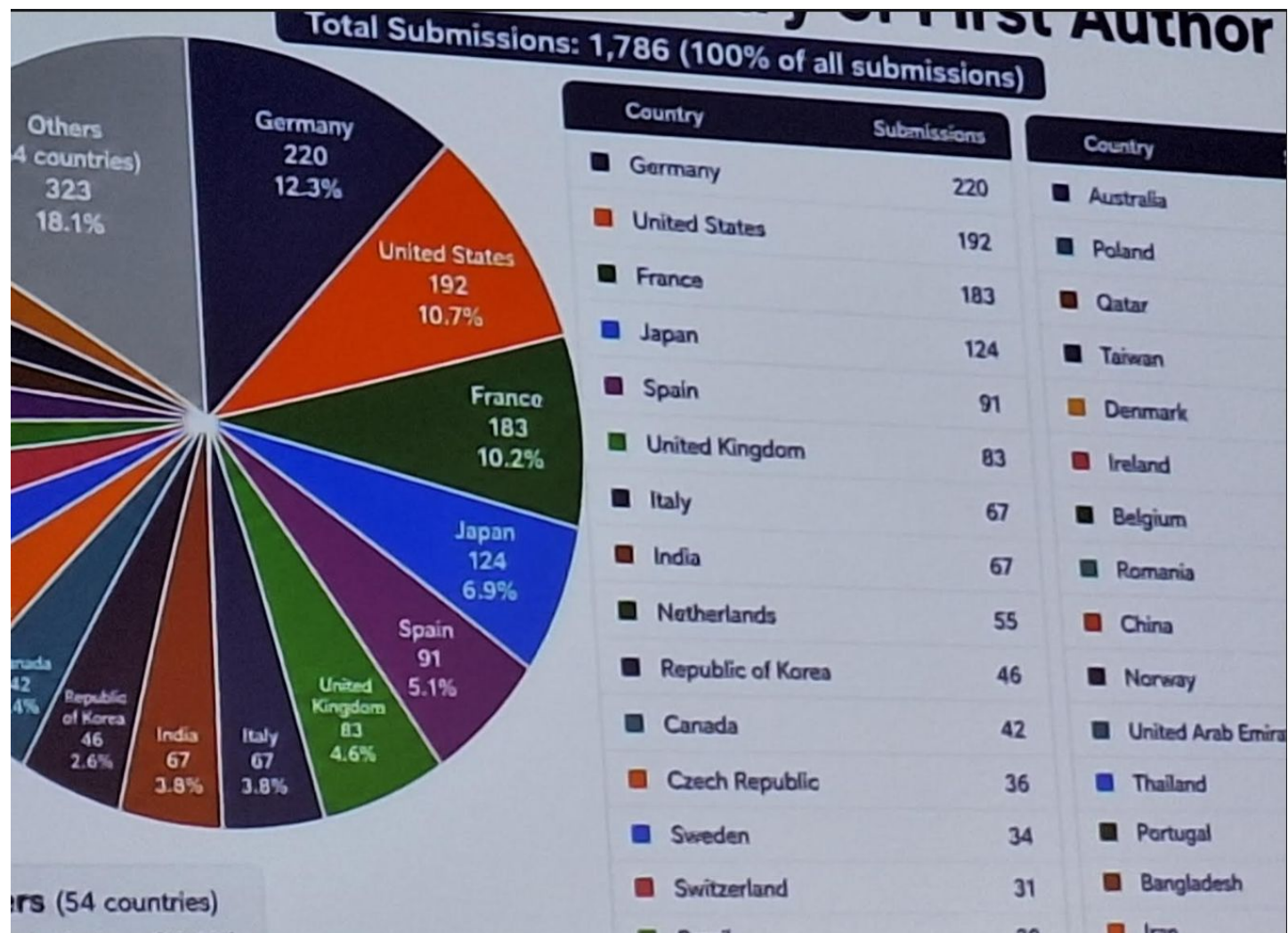
Ko su učesnici?

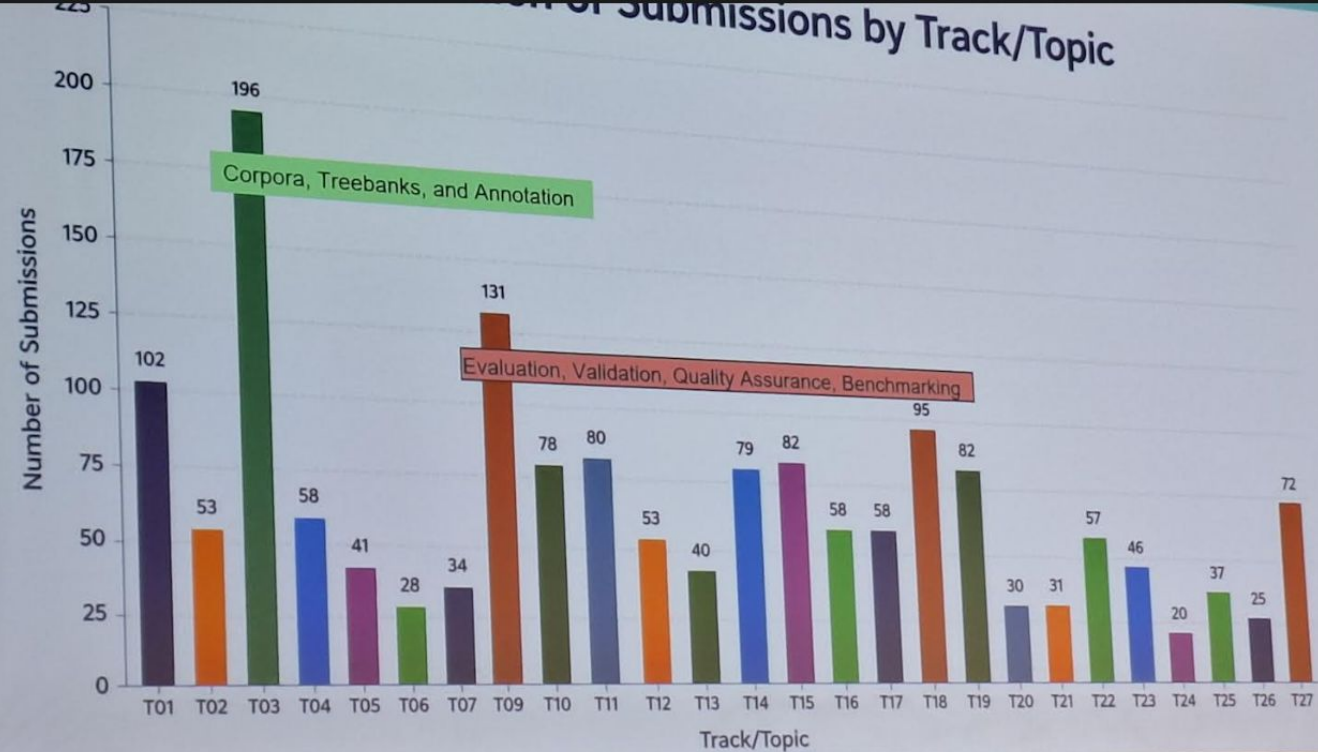
- Učesnici su iz svih krajeva sveta
- Naši predstavnici na konferenciji: trio Ranka, Anđelka i Jovana
- Ukupno 2388 registracija:
 - 1565 uživo
 - 823 online
- Ove godine su imali i najveći broj prispelih i prihvaćenih radova (942 od ~1786)
- Najveći broj predloga i prihvaćenih radionica (45 od 55)
- 12 tutorijala od 18 predloženih
- 19 industrijskih predstavljanja od 21 predloženi



54 zemlje

mi standardno
među manje
zastupljenim





Tipično popodne

Paralelno u 4 sale + posteri

Izabrati salu je poseban problem

Srećom, aplikacija sa unapred snimljenim izlaganjima je dostupna (i dalje)

Demo!

<https://conflux.events/download>

15:20 - 17:00	05 Inference, Reasoning, Question Answering I	06 Information Extraction and Text Mining I	07 Language Modeling and LRs I	08 Less-Resourced/Endangered/Less-studied Languages	Poster Area 2 • P2.1.1 Corpora and Treebanks I • P2.1.2 Corpora and Treebanks II • P2.1.3 Corpora and Treebanks III • P2.2 Discourse and Pragmatics I • P2.3.1 Interpretability, Explainability II • P2.3.2 Interpretability, Explainability III
17:00 - 17:20	Coffee break				
17:20 - 19:00	09 Corpora, Treebanks and Annotation; Tools, Systems and Platforms	010 Information Extraction and Text Mining II	011 Language Modeling and LRs II	012 Applications Involving LRs and Evaluation I	Poster Area 1 • P3.1.1 Dialogue, Conversational Systems I • P3.1.2 Dialogue, Conversational Systems II • P3.2.1 Less-Resourced/Studied Languages I • P3.2.2 Less-Resourced/Studied Languages II • P3.2.3 Less-Resourced/Studied Languages III

Speaker Author

ORAL 760

Integrating TEI, NER/NEL, Textometry, and Linked Data for a Semantically Enriched Interview Corpus

T15 Less-Resourced/Endangered/Less-studied Languages

📅 May 13, 4:00 PM - 4:20 PM (GMT+2)

[Irec2026-1.114.pdf](#)

Author

POSTER 1423

Development of Serbian QA Datasets through Prompt-Based Generation and Human Validation

T15 Less-Resourced/Endangered/Less-studied Languages

📅 May 13, 5:20 PM - 7:00 PM (GMT+2)

[Irec2026-1.267.pdf](#)

Author

ORAL 700

PARSEME 2.0 Multilingual Corpus of Multiword Expressions

T16 Lexicon and Semantics

📅 May 14, 11:00 AM - 11:20 AM (GMT+2)

[Irec2026-1.378.pdf](#)

Integrating TEI, NER/NEL, Textometry, and Linked Data for a Semantically Enriched Interview Corpus

Ranka Stanković, Tamara Vučenović, Biljana Rujević, Milica Ikonić Nešić, Mihailo Škorić



Integracija TEI-ja, NER/NEL-a, tekstometrije i povezanih podataka za semantički obogaćen korpus intervjua

- Korpus intervjua *Digitalne Ikone 20+* konvertovan u TEI XML format
- Strukturno obeležavanje: govornici, pasusi, teme, datumi i segmenti intervjua
- Primena NER-a za prepoznavanje osoba, lokacija, organizacija i događaja

Integracija TEI-ja, NER/NEL-a, tekstometrije i povezanih podataka za semantički obogaćen korpus intervjua

- Primena TESLA modela za NER/NEL obradu srpskog jezika
 - Integracija NER anotacija direktno u TEI dokumente
 - NEL povezivanje entiteta sa Wikidata identifikatorima
 - Dodavanje nedostajućih entiteta u Vikipodatke
- Transformacija korpusa u povezane podatke korišćenjem NIF okvira
- Tekstometrijska analiza u TXM okruženju uz JeRTeh resurse
- Retrieval-Augmented Generation (RAG)
 - za generisanje verodostojnih sažetaka intervjua
 - DSPy orkestracija LLM procesa sa izdvajanje dokaznih pasusa
 - Semantičke veze povećavaju interpretabilnost i mogućnosti pretrage
- Rešenje primenljivo na druge slične medijske i kulturne arhive
- Budući razvoj:
 - domen-specifični modeli, dopunjavanje grafova znanja i naprednije RAG arhitekture



↑ P3.2.2 Less-Resourced/Studied Languages II

LREC 20
Palma 26



Development of Serbian QA Datasets through Prompt-Based Generation and Human Validation

Jovana Rađenović¹, Olivera Kitanović¹, Ranka Stanković¹, Mihailo Škorić¹

¹University of Belgrade, Serbia, Faculty of Mining and Geology

Palau de Congressos de Palma
Palma, Mallorca (Spain)
11-16 May 2026

POSTER

Development of Serbian QA Datasets through Prompt-Based Generation and Human Validation

T15 Less-Resourced/Endangered/Less-studied Languages

May 13, 5:20 PM - 7:00 PM (GMT+2)

1423



Razvoj srpskih skupova za odgovaranje na pitanja pomoću upita i ručne validacije

Problem: Nedostatak velikih i pouzdanih QA skupova podataka za srpski jezik i Postojeći resursi sadrže prevodne i terminološke greške

SQuAD-sr-md

- Polazni resurs: **SQuAD-sr** (>87.000 QA parova)
- Ručno pregledano i korigovano: **7.000+ uzoraka**
- Ispravljene prevodne, terminološke i činjenične greške
- Rezultat: **SQuAD-sr-md** – kvalitetniji skup za obuku QA modela

SerbianQA-Gen

- Generisanje QA parova pomoću LLM modela
- Izvori: enciklopedijski članci, Vikipedija, apstrakti naučnih radova
- Veličina: **≈74.000 QA parova**

Ključni rezultati

- ✓ Više od **160.000 QA parova** za srpski jezik
- ✓ Kombinacija **LLM generisanja** i **ljudske validacije**
- ✓ Podrška razvoju QA sistema i velikih jezičkih modela za srpski jezik

Tok izgradnje skupova podataka

Tekstualni izvori → LLM generisanje QA parova → Ručna validacija → QA skupovi podataka → Obučavanje QA i LLM modela

sredinom 1960-ih godina.

1	U novoj komercijalnoj klimi, glam metal bendovi kao što su Jurop, Rat, Beli Lav i	Koji je Van H Hagar
1	U novoj komercijalnoj klimi, glam metal bendovi kao što su Jurop, Rat, Beli Lav i	Ko je I rouziz od str
1	U novoj komercijalnoj klimi, glam metal bendovi kao što su Jurop, Rat, Beli Lav i	Koji bi je krat Van H
1	U novoj komercijalnoj klimi, glam metal bendovi kao što su Jurop, Rat, Beli Lav i	Ko je I član k rouziz
1	U novoj komercijalnoj klimi, glam metal bendovi kao što su Jurop, Rat, Beli Lav i	U koj napus
1	Krajem 1960-ih godina, termin hevi metal se koristio naizmenično sa	Koji te za har 1960-
1	Krajem 1960-ih godina, termin hevi metal se koristio naizmenično sa	Koja h uvela, muzik
1	Krajem 1960-ih godina, termin hevi metal se koristio naizmenično sa	Koji ži ekstre godin.
1	Krajem 1960-ih godina, termin hevi metal se koristio naizmenično sa	Koji ži od he sadrži
1	Krajem 1960-ih godina, termin hevi metal se koristio naizmenično sa	Koji el metal rifovi?

#52764 zorica.gojak #44968 10 days ago

Hard_rok

Kontekst:

Krajem 1960-ih godina, termin hevi metal se koristio naizmenično sa hard rokom, ali postepeno je počeo da se koristi za opisivanje muzike koja se svira sa još većom jačinom i intenzitetom. Dok je hard rok zadržao bluz rokenrol identitet, uključujući izvesno swingovanje u pratećem ritmu i rifove sklone da ocrtavaju progresije akorda u svojim replikama, rifovi hevi metala su često funkcionisali kao samostalne melodije i nisu imali swinga u sebi. Hevi metal je poprimio „mračnije“ karakteristike nakon iznenađnog uspeha Blek Sabat-a početkom 1970-ih godina. Tokom 1980-ih godina razvio se niz podžanrova, često nazvanih ekstremni metal, od kojih su neki bili pod uticajem hardkor pank, što je dodatno razlikovalo ova dva stila. Uprkos ovoj razlici, hard rok i hevi metal su postojali jedan pored drugog, sa bendovima koji su često stajali na granici ili se ukrštali između žanrova.

Pitanje:

Koja hevi metal grupa je uvela „mračnije“ teme u muziku?

Kliknuti ispod na dugme Odgovori, a zatim selektovati najmanji deo teksta u okviru konteksta, koji predstavlja odgovor na dato pitanje.

Odgovori 1

Pomoć (odgovor u engleskoj bazi i srpskoj pre korekcija):

eng: [{"start":478,"end":491,"text":"Black Sabbath","labels":["Answer"]}]
sr-pre: [{"start":496,"end":508,"text":"probita Blek","labels":["Answer"]}]

Info History

Selection Details

Regions Relations

Manual By Time

1 probita Blek
2 Odgovori Blek Sabat

Update

CL4Health: Workshop on Patient-Oriented Language Processing

- **CL4Health: Workshop on Patient-Oriented Language Processing**
 - prva radionica - Torino 2024 @LREC-COLING
 - druga radionica - Novi Meksiko 2025 @ NAACL
- Organizatori: **Deepak Gupta** (US National Library of Medicine), **Paul Thompson** (National Centre for Text Mining and University of Manchester, UK), **Sophia Ananiadou** (National Centre for Text Mining and University of Manchester, UK & Archimedes/Athena Research Center, Greece), **Dina Demner-Fushman** (US National Library of Medicine)
- Motivacija:
 - Izražena upotreba velikih jezičkih modela za zdravstvene konsultacije, pojašnjenje dijagnostike i *patient-facing* upitnički sistemi.
 - Pojednostavljivanje i prilagođavanje kliničkih informacija kroz zadatke prezapisivanja, sumarizacije i personalizacije.
 - Prisutnost multimodalnih alata koji osnažuju puteve pacijenata (eng. patient journeys), tokove nege (eng. care pathways), i komunikaciju unutar zdravstvenih sistema.
 - Isticanje važnosti bezbednosti, temeljenja u činjenicama (eng. factual grounding) i standardizacije evaluacije sistema orijentisanih ka pacijentima.

CL4Health: Workshop on Patient-Oriented Language Processing

- Ciljevi radionice:

unapređivanje **jezičkih tehnologija** koje mogu da podrže **javno zdravstvo** uključujući **sumarizaciju** informacija iz **zdravstvenih kartona** za pacijente, **odgovaranje na pitanja** u domenu zdravstva korišćenjem **pouzdatih izvora**, prepoznavanje **netačnih informacija** ili sadržaja koji mogu **ugroziti** zdravlje, generisanje **multimodalnih odgovora** sa ciljem efektivnije **komunikacije** i boljeg **razumevanja**

- Naslovi:

- SimpliMed: Automatic Simplification of Cardiology Discharge Reports Using LLMs (Italija)
- Diagnostic Reasoning with LLMs for a Rare Disease: Case Study of Primary Ciliary Dyskinesia (USA)
- MedGore: An Approach and a Dataset for Identification of Sensitive Medical Images (NIH, USA)
- A Synthetic Conversational Dataset for Type 2 Diabetes Management (Holandija)
- Structured Radiology Intelligence: Extracting Structured Data from MRI Reports (Indija)
- ...

CL4Health: Workshop on Patient-Oriented Language Processing

Deljeni zadaci (eng. shared tasks):

- **Detecting Dosing Errors from Clinical Trials**
(University of Geneva, Switzerland)
- **Automatic Case Report Form Filling from Clinical Notes**
(Fondazione Bruno Kessler, Trento, Italy)
- **ArchEHR-QA 2026: Grounded Question Answering from Electronic Health Records**
(National Library of Medicine, USA)
- **FoodBench-QA 2026: Grounded Food & Nutrition Question Answering**
(Institut Jožef Stefan, Ljubljana, Slovenia)

Detecting Dosing Errors from Clinical Trials

Motivacija:

- preduprediti pojavu neželjenih događaja i unaprediti bezbednost pacijenata koji učestvuju u kliničkim studijama
- ubrzati razvoj novih lekova (svega oko 14% lekova se pojavi na tržištu)



**Clinical Trials Dosing
Error Benchmark 2026**
**Predicting Dosing Errors
in Interventional CTs**

Detecting Dosing Errors from Clinical Trials

Skup podataka:

- Izvor: [ClinicalTrials.gov](https://clinicaltrials.gov)
- U skupu se nalaze razne klinički relevantne informacije poput faze studije, opisa studije, intervencije, linkova do protokola studije (pdf dokumenti), ...
- Zadatak: predvideti mogućnost rizika greške doziranja (nebalansirana binarna klasifikacija)
- Ukupno **29,478** unosa u skupu za treniranje
- Ukupno **6,316** unosa u skupu za validaciju
- Skup je javno dostupan na [HuggingFace](https://huggingface.co/datasets/clinicaltrials)-u

Field	Availability (%)	Mean chars
allocation	99.62%	8.46
armDescriptions	99.40%	388.24
briefSummary	100%	502.40
conditions	100%	39.21
conditionsKeywords	63.94%	81.17
detailedDescription	61.93%	1,610.83
interventionDescriptions	100%	286.18
interventionModel	99.67%	9.34
interventionNames	100%	82.72
locationDetails	94.38%	1,310.26
protocolPdfText	42.33%	223,247.14

Detecting Dosing Errors from Clinical Trials

Dizajn sistema:

- reprezentacija numeričkih atributa (broj učesnika u studiji, broj intervencija, broj lokacija, ...)
- reprezentacija kategoričkih atributa (faza kliničke studije, pol učesnika, ...)
- reprezentacija tekstualnih atributa (opis studije, opis intervencije, protokol studije)
korišćenjem jezičkih modela kao što su BioBERT i MedCPT
- agregiranje reprezentacija
- modeli klasifikacije: XGBoost

Task:				Test Task				
#	Participant	Date	ID	ROC-AUC	F1-Macro	Recall	Precision	Balanced. Accuracy
	msmadi	2026-02-08 16:17	530456	0.8466	0.5072	0.0235	0.2667	0.5099

Dodatni medicinski sadržaji

[tutorial] **Healthcare NLP**: Where we are and what's next?

[workshop] **8th Clinical NLP**

Radovi na glavnoj konferenciji:

- **Mental Health** Disorder Detection beyond Social Media: A Systematic Review of Available Datasets
- Building Effective Japanese Medical LLMs with an Open Recipe for **Domain Adaptation** through Continued Pre-training
- **Small LLMs** for Medical NLP: A Systematic Analysis of Few-Shot, Constraint Decoding, Fine-Tuning and Continual Pre-Training in Italian
- Injecting **Structured Biomedical Knowledge** into Language Models: Continual Pretraining vs. GraphRAG
- Is Biomedical Specialization Still Worth It? Insights from **Domain-Adaptive** Language Modelling with a New French Health Corpus
- MedInjection-FR: Exploring the Role of Native, Synthetic, and Translated Data in Biomedical **Instruction Tuning**
- Developing the German Medical Text Corpus (GeMTeX): **Legal Compliance** and **Semantic Enrichment**
- ...

Šta kaže Džurafski?

LLM-ovi su:

1. Previše samouvereni
2. Rasno pristrasni
3. Dodvoravajući
4. Epistemički egocentrični
(posmatraju znanje i istinu uglavnom iz sopstvene perspektive, bez stvarnog razumevanja tuđih ograničenja, iskustava ili konteksta)

Posledične štete:

- preterano oslanjanje, iskrivljeno rasuđivanje
(+ zablude, emocionalna zavisnost)

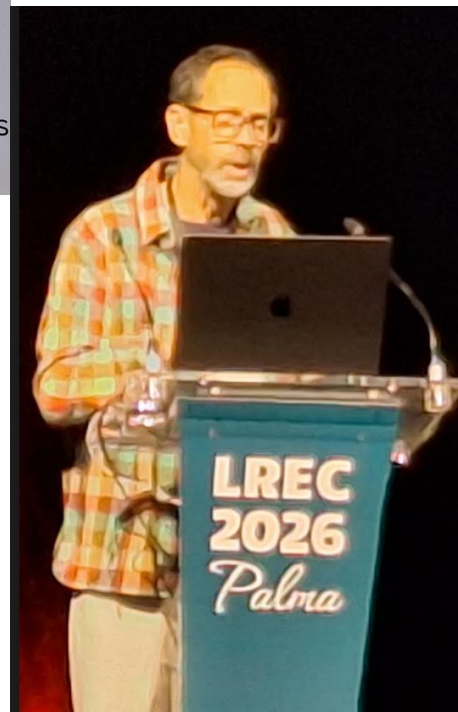
4 Vignettes about LLMs as conversational pa

LLMs are:

1. Overconfident
2. Racially biased
3. Sycophantic
4. Epistemically ego-centric

Resulting harms:

- overreliance, distorted judgment (+delusions, emotional dependence)



Za kraj izlaganja Džurafski:

Prema Herbertu H. Clarku i Michaelu F. Schoberu, 1992. (blago izmenjeno)

„Uobičajena zabluda je da se jezik prvenstveno odnosi na reči i na ono što one znače.

Nije tako.

Jezik se prvenstveno odnosi na ljude i na ono što oni znače.”

Thanks and concluding thought!



From Herbert H. Clark & Michael F. Schober, 1992 (slightly edited)

"The common misconception is that language has primarily to do with **words** and what they mean.

It doesn't.

It has primarily to do with **people** and what **they** mean.

LLMs misuse **expressions of certainty**, incorrectly asserting confidence

Vruće teme

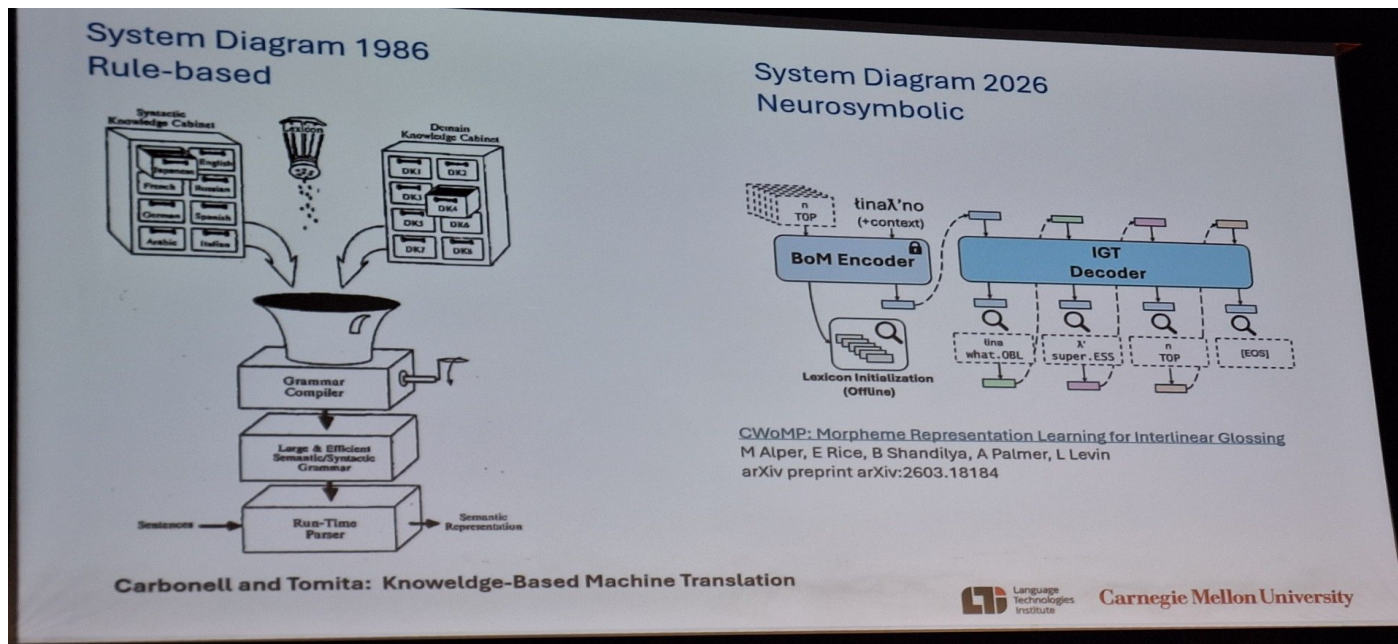
- Zajednički engrami u prihvaćenim radovima
 - large scale corpus
 - linguistics knowledge
 - percentage points
 - publicly available dataset

Rank	N-gram	Count
1	large language models	89
2	named entity recognition	16
3	evaluating large language	11
4	automatic speech recognition	9
5	using large language	7
6	hate speech detection	6
7	multimodal large language	6
8	large language model	6
9	natural language inference	5
10	lowresource machine translation	5

Dobitnica Zampolijeve nagrade ove godine

Lori Levin, Carnegie Mellon University, osnivač

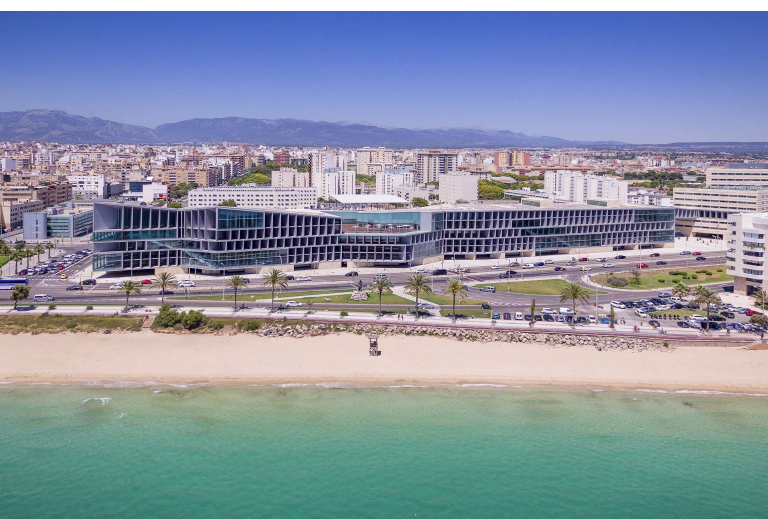
North American Computational Linguistics Open Competition



- T02 Tutorial Low-Resource, High-Impact: Building Corpora for Inclusive Language Technologies
- T03 Tutorial Under-Resourced Studies of Under-Resourced Languages: LLM-as-Annotator PipelinesW48 DELITE 2026 2nd Workshop on Language-driven Deliberation Technology
- T04 Tutorial From Shared Tasks to Benchmark Datasets: Lessons from NTCIR's 26-Year Journey
- T06 Tutorial Linguistic Creativity in the Age of LLMs: From Diversity to InnovationW31 RAIL 7th Workshop on Resources for African Indigenous Languages
- T07 Tutorial Distributed Compositionality: Formal Foundations for Distributional Semantics
- T11 Tutorial Meaning Representation in NLP: Foundations, Frameworks, and Evaluation
- T12 Tutorial Biases we Research by
- T13 HealthcareNLP HealthcareNLP: where we are and what's next?
- T18 Tutorial Multilingual & Multimodal LLMs in the Wild: Building for Low-Resource LanguagesW45 InDor The Information Disorder Workshop

- W01 FNP 2026 7th Financial Narrative Processing Workshop
- W02 CL4Health 3rd Workshop on Patient-oriented Language Processing
- W03 OSACT7 7th Workshop on Open-Source Arabic Corpora & Processing Tools
- W04 NLP4Ecology 2026 2nd Workshop on Ecology, Environment, and NLP
- W05 DMR 2026 7th Int'l Workshop on Designing Meaning Representations
- W06 CMCL 2026 15th Workshop on Cognitive Modeling and Computational Linguistics
- W07 Clinical NLP 8th Clinical NLP Workshop
- W10 FAIRTAG4ALL FAIR Terminology-Augmented Generation for AIW55 RESOURCEFUL 2026 Role of Resources in the Age of Large Language Models
- W12 LT4HALA 2026 4th Workshop on LT for Historical and Ancient Languages
- W13 Resources & Processing of Data from People with Cognitive/Psychiatric Impairments
- W14 DialRes Dialects in NLP: A Resource Perspective
- W14 Tutorial Metaphor Processing: Methods, Theories and ResourcesW08 NakbaNLP 2026 2nd Int'l Workshop on Nakba Narratives as Language Resources
- W15 Tutorial Hallucination in Large Foundation ModelsW15 NSLP 2026 3rd Int'l Workshop on Natural Scientific Language Processing
- W16 UDW 2026 9th Workshop on Universal Dependencies
- W17 DeTermIt! 2026 2nd Workshop on Evaluating Text Difficulty in a Multilingual Context
- W18 ISA 2026 22nd Joint ACL–ISO Workshop on Interoperable Semantic Annotation
- W20 Social Context The Social Context Workshop
- W21 CHiPSAL-2026 2nd Workshop on Challenges in Processing South Asian Languages
- W24 CMLC-12 12th Workshop on Challenges in Management of Large Corpora
- W25 CAS 2026 1st Workshop on Computational Affective Science
- W26 NLPerspectives 5th Workshop on Perspectivist Approaches to NLP
- W27 Neology & LLM Neology and Large Language Models

- W28 PressMint 1st Workshop on Creating Interoperable Corpora of Historical Newspapers
- W29 LDL 2026 10th Workshop on Linked Data in Linguistics
- W30 SLiDE Workshop on Structured Linguistic Data
- W32 BUCC 19th Workshop on Building and Using Comparable Corpora
- W33 PoliticalNLP 2026 3rd Workshop on NLP for Political Sciences
- W34 DTF Leveraging Derived Text Formats to Unlock Copyrighted Collections for Open Science
- W35 KG&LLM Knowledge Graphs and Large Language Models
- W36 SIGUL 2026 Towards Inclusivity: LR&T for Under-Resourced & Endangered Languages
- W37 Non Literal Learning Non-Literal Expressions with Small Data
- W38 ParlaCLARIN V Workshop on Interoperability, Multilinguality & Multimodality in Parliamentary Corpora
- W39 SignLang 12th Workshop on Representation & Processing of Sign Languages: Language in Motion
- W40 SPEAKABLE Speech Language Models in Low-Resource Settings
- W42 Gaze4NLP Eye-Tracking Resources & Evaluation for Human-Aligned NLP
- W43 LLMs4SSH Shaping multilingual, multimodal AI for social sciences & humanities
- W44 LEGAL2026 & CALD-pseudo Joint Workshop on Legal & Ethical Issues in HLT + Data Pseudonymization
- W46 Identity-Aware NLP Ethical & Technical Challenges for Identity-Aware NLP
- W47 WILDRE-8 8th Workshop on Indian Language Data: Resources & Evaluation
- W49 READiXTSAR READiXTSAR
- W51 CAWL 2026 3rd Workshop on Computation and Written Language
- W53 HTRes-2026 Holocaust Testimonies as Language Resources



<https://lrec2026.info/>

LREC 20
Palma **26**

Organized by ELRA Language Resources Association

Palau de Congressos de Palma

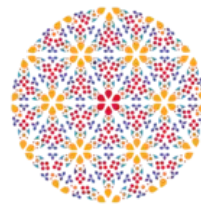
Palma, Mallorca (Spain)

11-16 May 2026

LREC2026 PROCEEDINGS ARE NOW AVAILABLE!

WORKSHOPS PROCEEDINGS ARE NOW AVAILABLE!





Catedral
de Mallorca





vreme za opuštanje i uživanje u Palmi

