

CONFERENCE PROGRAMME

eLex 2025

**Electronic lexicography in the 21st century:
Intelligent Lexicography**

v 1.0

Bled, Slovenia

18-20 November 2025

Tuesday, 18 November 2025 - morning

8:00-9:15	Registration		
9:10-9:30	Welcome address		
9:30-10:30	KEYNOTE 1 Large language models for lexicography (Marko Robnik-Šikonja)		
10:30-11:00	coffee break		
11:00-13:00	Arnold hall	Sonce hall	Zrak hall
11:00-11:30	Retention of English words from interaction with dictionaries and GenAI Chatbots (Robert Lew and Bartosz Ptasznik)	Compiling bilingual dictionaries: AI-assisted translation of Italian multiword expressions into English and French (Annalisa Greco, Matteo Delsanto, Andrea Di Fabio, Calogero Jerik Scozzaro, Lorenzo Mori, Cristina Onesti and Daniele Paolo Radicioni)	Parsing of explanatory dictionary (Iryna Ostapova, Yevhen Kupriianov and Mykyta Yablochkov)
11:30-12:00	Automating Adjectival Microstructures in Monolingual Dictionaries: A New Method Combining Embeddings and LLMs (Enikő Héja, László Simon and Veronika Lipp)	Neology in Practice: Lexicographic and Terminological Approaches to Lexical Innovation (Jelena Kallas, Kristina Koppel, Kris Heylen, Ilan Kernerman, Ana Ostroški Anić, Federica Vezzani, Hiwa Asadpour, Judit Freixa, Petar Božović and Špela Arhar Holdt)	Vision-Enabled Language Models in Lexicographical Digitisation: A Case Study of Anton Thor Helle's 1732 Dictionary (Madis Jürviste and Tiina Paet)
12:00-12:30	Automatic Non-recorded Sense Detection for Swedish through Word Sense Induction with fine-tuned Word-in-Context models (Dominik Schlechtweg, Emma Sköldberg, Shafqat Virk, James White and Simon Hengchen)	Using Large Language Models to Generate Distractors for Language Games (Iztok Kosem and Špela Arhar Holdt)	GramatiKat: A Corpus-Based Tool for Detecting Morphological Anomalies and Paradigm Variation (Dominika Kovarikova)

12:30-13:00	Automatic Detection of Word Sense Shift from Corpus Data (Ondřej Herman)	Enhancing Lexicographic Access for Deaf and Hard-of-Hearing Learners: A Digital Greek Sign Language Dictionary with AI-Powered Language Support (Isidora Despotidou and Zoe Gavriilidou)	Learner's reactions to false polysemy (Tomasz Michta and Ana Frankenberg-Garcia)
13:00-14:30	<i>lunch</i>		

Tuesday, 18 November 2025 - afternoon

14:30-16:00	Arnold hall	Sonce hall	Zrak hall
14:30-15:00	The Dictionary of Contemporary Serbian Language (RSSJ): Advanced Automation and Other Challenges (Ranka Stanković, Rada Stijović, Mihailo Škorić and Cvetana Krstev)	Documenting the final days of monolingual English learners' dictionaries using the archived web (Geraint Paul Rees)	Exploring the constructicographic potential of lexicographic data and language models: The case of the Estonian Nominal Quantifier Construction (Heete Sahkai, Geda Paulsen, Ene Vainik, Jelena Kallas, Ahto Kiil, Katrin Tsepelina, Kertu Saul and Arvi Tavast)
15:00-15:30	Lexical-Semantic Resources as a Culture-Aware Basis for LLM-Benchmarking (Nathalie Hau Norman, Bolette Sandford Pedersen, Sanni Nimb, Sussi Olsen and Nina Skovgaard Schneidermann)	Learner Corpora as a Lexicographic Resource: Mapping Slovene Learner Vocabulary to CEFR Scales (Mojca Stritar Kučuk)	The Mangalam dictionary of Buddhist Sanskrit: automating lexicography for a low-resource historical language with generative LLMs (Ligeia Lugli)

15:30-16:00	Do dictionary users prefer definitions by lexicographers or by LLM-s? (Maria Tuulik, Ene Vainik, Margit Langemets, Eleri Aedmaa, Lydia Risberg, Esta Prangel, Kristina Koppel and Sirli Zupping)	AI- and Corpus-Based Strategies for Identifying Phraseme Constructions: A Pilot Study on Croatian Repetitive Constructions (Ivana Filipović Petrović and Slobodan Beliga)	Why a dedicated dictionary device is more appropriate than an app for primary school learners (Lorna Morris)
16:00-16:30	<i>coffee break</i>		
16:30-17:00	Group discussions A - part 1	Group discussions B - part 1	Group discussions C - part 1
17:00-17:30	Group discussions A - part 2	Group discussions B - part 2	Group discussions C - part 2
17:30-18:00	Group discussions A - part 3	Group discussions B - part 3	Group discussions C - part 3
18:30-	Welcome reception		

Wednesday, 19 November 2025 - morning

9:00-10:30	Arnold hall	Sonce hall	Zrak hall
9:00-10:00	KEYNOTE 2 LLMs and Lexicography at the Dutch Language Institute (Carole Tiberius & Jesse de Does)		
10:05-10:30	The lemma dilemma, Slovene version (Polona Gantar, Cyprian Laskowski and Simon Krek)	Lexicography and Generative Artificial Intelligence for contextualised meaning (Theo Bothma and Rufus Gouws)	An electronic Ukrainian dictionary as a derussification tool (Vasyl Starko and Andriy Rysin)
10:30-11:00	<i>coffee break</i>		
11:00-11:30	Making Sense of the Past: AI-Assisted Historical Word Sense Disambiguation and the OED (Elinor Hawkes, Phoebe Nicholson and Will Rogers)	Exploring the power of generative artificial intelligence for automatic term extraction from small samples (Lena De Pourcq, Marie Grégoire and Leonardo Zilio)	Navigating linguistic diversity: modelling diatopic and bibliographic information with TEI Lex-0 (Michaela Rausch-Supola, Veronika Engler, Stephan Procházka, Daniel Schopper and Karlheinz Mörth)
11:30-12:00	Bridging Human and AI Perspectives: Semantic Annotation of Generic Nouns in German (Ivan Arias-Arias and Elena Martín-Cancela)	Lexicom at 25: reflections on the changing world of lexicography and language technology (Michael Rundell, Miloš Jakubíček, Vojtech Kovar and Ondrej Matuska)	Artificial intelligence in English dictionary entries compiled in Slovar krajšav (Mojca Kompara Lukancic)

12:00-13:00	POSTER & DEMO SESSION
	POSTERS: A Bazaar Among Cathedrals – Leveraging Wikidata as an Open Marketplace for Lexicographic Data (Gregor Middell) Implementing Frames in the Phrase-based Active Dictionary: why Frames are needed but FrameNet can only be a partial solution (Laura Rebosio) Project of a Specialized Dictionary Website (Mykyta Yablochkov, Alona Dorozhynska, Iryna Ostapova, Iuliia Verbynenko) The Hare and the Tortoise: Pipeline for Latvian Information and Communication Technologies Secondary Term Formation (Dace Šostaka, Inguna Skadiņa) Exploring Derivational Families through Intelligent Lexicography (Krešimir Šojat, Kristina Kocijan) Comparative Analysis of Medical Adjectives in Croatian General Dictionaries (Martina Pavić, Daša Farkaš) Handling abstract constructions in a dictionary-based constructicon (Bálint Sass, Éva Dömötör, Balázs Indig, Mátyás Lagos Cortes, Veronika Lipp, Márton Makrai, Gergely Pethő) Qualitative Evaluation of LLM Translation of MWEs for Developing a Croatian Sense Repository (Ana Ostroški Anić, Jaka Čibej, Ivana Filipović Petrović, Martina Pavić, Siniša Runjaić, Robert Sviben) Up to No Good: Exploiting Word Embeddings for an Automatic Extraction of Candidates for a Lexicon of Slovene Taboo Language (Jaka Čibej) Introducing DigiMet: a Psycholinguistic Database for Croatian Multi-Word Expressions (Jasmina Jelčić Čolakovac) Accelerating the lexicographic process with automatic methods and AI (Nathalie Norman, Nicolai Hartvig Sørensen, Jonas Jensen, Kirsten Appel, Sanni Nimb) Presenting verbal aspect data in a learner’s dictionary: Devices and usage scenarios (Sarah Piepkorn) Digitalization of Romanian dictionaries (Mititelu Catalin) The Challenges of Syntactic Descriptions of Multiword Expressions in Electronic Lexicography (Verginica Barbu Mititelu, Voula Giouli, Gražina Korvel, Chaya Liebeskind, Irina Lobzhanidze, Rusudan Makhachashvili, Stella Markantonatou, Alexandra Markovic, Ivelina Stoyanova) The dictionary of pluricentric Portuguese project: theoretical aspects (Tanara Zingano Kuhn)
	DEMOS: CJVT Igre: New Word Games Based on the Digital Dictionary Database of Slovene (Špela Arhar Holdt, Iztok Kosem) lexicographR: R infrastructure to develop and deploy digital dictionaries from scratch (Ligeia Lugli)
13:00-14:30	<i>lunch</i>

Wednesday, 19 November 2025 - afternoon

14:30-16:00	Arnold hall	Sonce hall	Zrak hall (ONLINE)
14:30-15:00	Choosing Suitable Text Corpora for Identifying Collocations – A Case Study of a Large Reference Dictionary of Contemporary German (Luise Köhler, Gregor Middell and Alexander Geyken)	Automated Transcription of Mixed-Script Dialectal Materials (Markus Kunzmann)	Modeling and structuring of a bilingual French-Chinese phraseological dictionary: neural automatic approach for ontology and lexicography. (Lian Chen)
15:00-15:30	The role of subjectivity in lexicography: Towards a data-driven approach to dictionary labels (Lydia Risberg, Eleri Aedmaa, Maria Tuulik, Margit Langemets, Ene Vainik, Esta Prangel, Kristina Koppel and Hanna Pook)	Identifying the Most Representative Phraseological Units Using Language Corpora and Artificial Intelligence for Lexicography: The Case of Slovenian Comparative Phrasemes (Matej Meterc and Nataša Jakop)	Matching meaning: Evaluating ChatGPT's ability to assign corpus examples to dictionary senses of polysemous sound-related verbs (Sylwia Wojciechowska)
15:30-16:00	So Close but Still Far: Case Study on Application of LLMs in Idioms Identification, Definition, and Generation of illustrative examples (Aleksandra Marković and Ranka Stanković)	Information seeking behavior of the English learner in the AI era (Anna Dziemianko and Mojca M. Hocevar)	The DICI-A: A Learner Dictionary of Italian Collocations (Stefania Spina, Fabio Zanda, Irene Fioravanti, Luciana Forti, Damiano Perri and Osvaldo Gervasi)
16:00-16:30	coffee break		
16:30-17:00	How Effective is AI as a Language Consultant? (Urška Vranjek Ošlak)	Corpus-Based Methods and AI-Assisted Terminography for Contextonym Analysis (Antonio San Martín)	Compiling a candidate list of taboo constructions for an under-resourced language (Monique Rabe, Martin Puttkammer and Gerhard B. Van Huyssteen)

17:00-17:30	Corpus-Based Vocabulary Profiling for Ukrainian: From Lexical Analysis to the PULS Digital Learning Platform (Olena Synchak, Vasyl Starko, Mariana Burak and Mykhaylo Svystun)	User interaction with assistive technology for a thorough evaluation of a WCAG 2-compliant e-dictionary: assessing the accessibility of the Diccionario de la Lengua Española, v. 23.8 (Jesús Torres del Rey and María García Garmendia)	You get it through lexicography: extracting suppressed language from LLMs using lexicographic scenarios as jailbreaking tools (Esra Abdelzaher and Agoston Toth)
17:30-18:00	Meaning comprehension and learning with sound effects in dictionaries (Anna Dziemianko)	From Word of the Year to Word of the Week: Daily-updated Monitor Corpora for 25 Languages (Ondřej Herman, Vit Suchomel, Miloš Jakubíček and Jan Kraus)	Corpus-based Dictionary for the Endangered Megrelian Language (Irina Lobzhanidze and Rusudan Gersamia)
19:00-	Gala dinner		

Thursday, 20 November 2025 - morning

	Arnold hall	Sonce hall	Zrak hall
9:00-10:00	KEYNOTE 3 We need to talk about data structures in lexicography (Michal Měchura)		
10:05-10:30	Contrasting a new AI-powered dictionary designed for on-screen reading with electronic dictionaries that have evolved from print editions (Ana Frankenberg)	LLM-Assisted Dialect Lexicography: Challenges and Opportunities in Processing Historical Bavarian Dialects (Philipp Stöckle, Daniel Elsner, Wolfgang Koppensteiner and Katharina Korecky-Kröll)	Toward a corpus-based multilingual terminology database for Intercultural Communication (María Iglesias Vázquez, Marie Steffens and Charlotte Venema)
10:30-11:00	coffee break		

11:00-12:25	Arnold hall	Sonce hall	Zrak hall
11:00-11:30	DMLEX on Wikibase: Legacy dictionaries as collaboratively editable dataset (Simon Krek, Primož Ponikvar, Andraž Repar, Iztok Kosem and David Lindemann)	Inductive Categorization for Conceptual Analysis with LLMs: A Case Study from the Humanitarian Encyclopedia (Loryn Isaacs, Santiago Chambó and Pilar León-Araúz)	A Pipeline for Automated Dictionary Creation with Optional Human Intervention (Thomas Widmann)
11:30-12:30	Image-to-Sense Alignment Using AI Tools (Andrej Perdih, Dejan Gabrovšek and Janoš Ježovnik)	Passive Lexicon: A Statistical Approach (František Kovařík, Vojtěch Kovář, Marek Blahuš and Miloš Jakubíček)	Better something than nothing: Analysis of GPT-4 performance in identifying Croatian proverbs (Nikola Bakarić)
12:00-12:30	Woordpeiler: A New Tool for Visualizing and Analyzing Lexical Trends in Contemporary Dutch (Kris Heylen, Vincent Prins, Katrien Depuydt, Jesse de Does, Laura van Eerten and Thomas Haga)	Automatically Updated Corpora of EU National Parliaments with Terminology Extraction in Twenty Languages (Marek Blahuš and Ota Mikušek)	
12:30-12:45	Closing		
14:00-18:00	GWLN-7 WORKSHOP		