



ICDAR
VIENNA 2026

**INTERNATIONAL CONFERENCE
ON DOCUMENT ANALYSIS
& RECOGNITION**

ICDAR 2026

Vienna, Austria

August 30 –
September 4, 2026

Program Booklet



Table of Contents

Welcome Message	5
Committees	10
Program Overview	14
Social Events	20
Keynotes	22
Awards	28
Tutorials	34
Doctoral Consortium	36
Scientific Program	39
Monday, August 31	39
Tuesday, September 1	48
Wednesday, September 2	51
Workshops	59
Thursday, September 3	61
Friday, September 4	71
Sponsors	80
About Vienna	87
Practical Information	90

Message from the General Chairs

Dear Colleagues,

Welcome to the 20th IAPR International Conference on Document Analysis and Recognition (ICDAR 2026), taking place from August 30 to September 4, 2026, at TU Wien in Vienna, Austria. Located in the heart of Vienna's city center, the venue places participants within walking distance of many of the city's world-famous historical and cultural attractions.

ICDAR 2026 continues the long-standing tradition of bringing together researchers, academics, and professionals from around the world in the field of Document Analysis and Recognition. Since the first ICDAR in Saint-Malo, France, in 1991, the conference has been held in Tsukuba, Japan (1993); Montreal, Canada (1995); Ulm, Germany (1997); Bangalore, India (1999); Seattle, USA (2001); Edinburgh, UK (2003); Seoul, Korea (2005); Curitiba, Brazil (2007); Barcelona, Spain (2009); Beijing, China (2011); Washington, DC, USA (2013); Nancy, France (2015); Kyoto, Japan (2017); Sydney, Australia (2019); Lausanne, Switzerland (2021); San Jose, USA (2023); Athens, Greece (2024); and Wuhan, China (2025). ICDAR is a leading international forum for researchers and practitioners, covering optical character recognition, handwriting and text recognition, layout analysis, document understanding, graphics recognition, historical documents, digital libraries, document forensics, and scene-text analysis.

TU Wien reflects ICDAR's academic character and close connection to university-based research. With a history of more than 200 years, it provides an excellent environment for scholarly exchange. As a university of technology, it covers a broad spectrum of scientific inquiry, from fundamental and applied research to technological development and industry collaboration. For more than two centuries, TU Wien has served research, teaching, and learning in the service of progress.

The scientific program will feature three keynote talks. C. V. Jawahar, Professor at IIIT Hyderabad, India, and Co-Founder of the Center for Visual Information Technology (CVIT), will discuss the changing landscape of Document Understanding. Christoph Rass, Professor at Osnabrück University, Germany, will address what historians need from Document Understanding, considering the topic from archives to algorithms and back. Tong Sun, Senior Director at Adobe, will share her perspective on the next frontier: from Document Understanding to Document Agency. During the conference, Silvia Cascianelli and Jorge Calvo-Zaragoza will receive the IAPR/ICDAR Young Investigator Award, which recognizes researchers under 40 who have made significant contributions to document analysis and recognition. In addition, Elisa Barney Smith, Laurence Likforman, and Jean-Marc Ogier will receive the 2026 IAPR/ICDAR Service Award. This newly established distinction honors individuals whose sustained contributions have significantly advanced the field through research, the mentorship of early-career researchers, and dedicated service to the community. Congratula-

tions to all award recipients!

In keeping with previous conferences, ICDAR 2026 is a rigorously peer-reviewed, single-track conference on document analysis and recognition. All 295 submissions underwent a double-blind review process with at least three reviews assessing originality, research quality, or analysis of experience, relevance to document analysis systems, and clarity of presentation. The three-day technical program will include 40 oral and 100 poster presentations. Following the tradition established by previous ICDAR editions, ICDAR 2026 once again features a Journal Track, combining the rapid turnaround and broad visibility of a conference with the depth, scope, and rigorous peer review of a journal. This year, the Journal Track had an acceptance rate of 24.5%, with 12 papers accepted for oral presentation at the conference and publication in a special issue of the *International Journal on Document Analysis and Recognition* (IJ-DAR). Eight workshops and four tutorials will offer an in-depth study of focused and emerging topics, while competitions will present new algorithms and results that continue to stimulate research in the community.

A conference such as ICDAR depends on volunteers who contribute tirelessly to its program, infrastructure, and organization. The General Chairs thank the Program Co-Chairs, Gernot A. Fink, TU Dortmund University, Germany; Alicia Fornés, Computer Vision Center, Spain; Koichi Kise, Osaka Metropolitan University, Japan; and Dan Lopresti, Lehigh University, USA, for assembling a first-class technical program. Special thanks are due to the Publication Chairs, Vincent Christlein, FAU Erlangen-Nürnberg, Germany, and Foteini Liwicki, Luleå University of Technology, Sweden; the Workshop Chairs, Jihad El-Sana, Ben-Gurion University of the Negev, Israel, Elisa Barney Smith, Luleå University of Technology, Sweden, and George Retsinas, NVIDIA, USA; and the Competition Chairs, Dimosthenis Karatzas, Computer Vision Center, Spain, and Sri-rangaraj Setlur, University at Buffalo, USA.

Gratitude is also extended to the Industry Chairs, Josep Lladós, Computer Vision Center, Spain, and Stefan Fiel; the Tutorial Chairs, Anna Scius-Bertrand, University of Applied Sciences and Arts Western Switzerland, Switzerland, Marcus Liwicki, Luleå University of Technology, Sweden, and Anna Zhu, Wuhan University of Technology, China; the Publicity Chairs, Jean-Marc Ogier, University of La Rochelle, France, and Umapada Pal, Indian Statistical Institute, India; the Doctoral Consortium Chairs, Anna Scius-Bertrand, University of Applied Sciences and Arts Western Switzerland, Switzerland, and Momina Moetesum, National University of Sciences and Technology (NUST), Pakistan; the Posters and Demo Chairs, Eyad Elyan, Robert Gordon University Aberdeen, UK, and Liangrui Peng, Tsinghua University, Beijing, China; the Awards Chair, Richard Zanibbi, Rochester Institute of Technology, USA; and the Sponsorship Chairs, Basilis Gatos, NCSR Demokritos, Greece, and Apostolos Antonacopoulos, University of Salford, Manchester, UK, for their valuable contributions.

The conference gratefully acknowledges its sponsors, Snowflake (Platinum sponsor),

Wacom and Micrografija (Silver sponsors), whose financial support helps reduce costs and fund various awards, including student travel awards.

The General Chairs also thank the local organizing team, especially Marco Peer (University of Fribourg, Switzerland), Marvin Burges, Katharina Pois, and Food and Beverage Chair Rafael Sterzinger, all from TU Wien.

Finally, the conference would not be possible without the excellent papers contributed by the authors. The Chairs thank all authors for their contributions to ICDAR 2026 and all external reviewers for their diligent work. This program is expected to stimulate further research and provide practitioners with improved techniques, algorithms, and deployment tools. It is an honor and a privilege to present the latest developments in Document Analysis and Recognition through this exciting program.

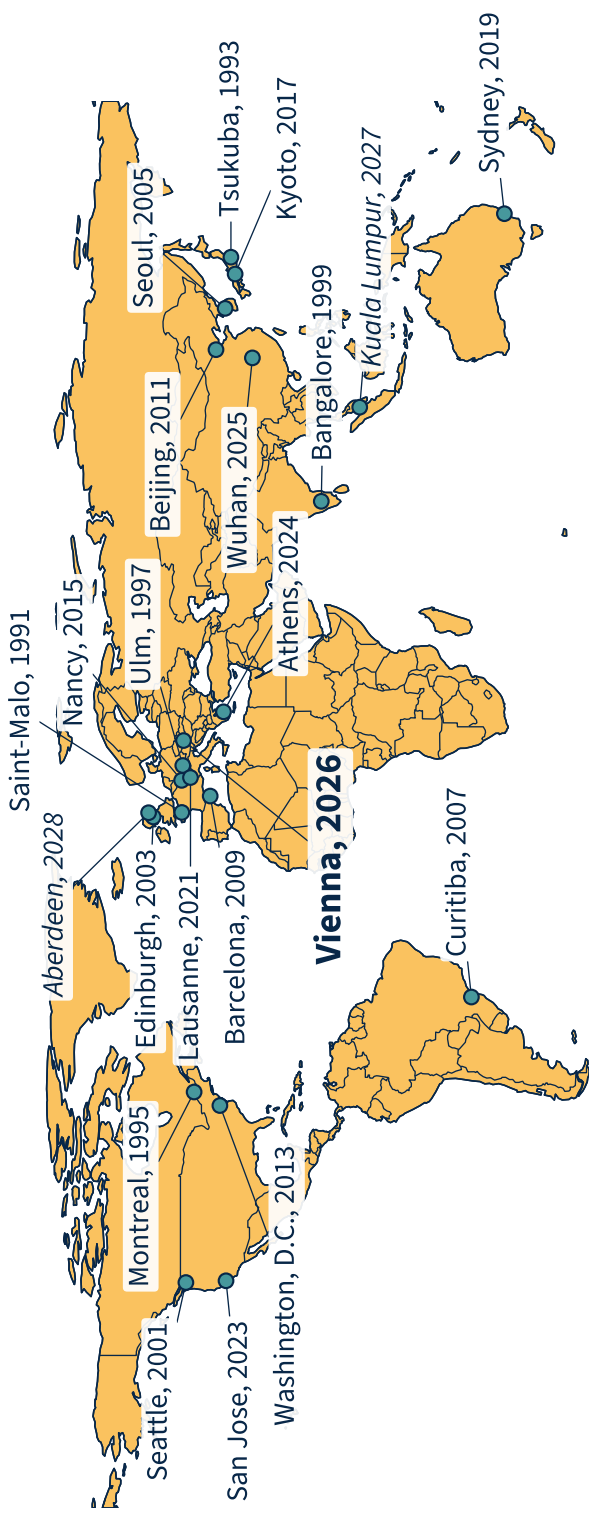
The Chairs hope that all participants enjoy the conference, the city of Vienna, its food and culture, and many stimulating exchanges with colleagues.

Robert Sablatnig and Florian Kleber

ICDAR 2026 General Chairs

Gernot A. Fink, Alicia Fornés, Koichi Kise, and Dan Lopresti

ICDAR 2026 Program Chairs



History of ICDAR

The Venue



The conference, including workshops and tutorials, will take place at **TU Wien, Campus Gußhaus – Faculty of Electrical Engineering and Information Technology**, Gußhausstraße 27–29, 1040 Vienna. The scientific program of ICDAR 2026 is held in the lecture halls E17, E18, E19 and E10 - marked in the plan above. The campus also includes a garden where participants can spend their breaks.

How to get there

Located in central Vienna, the venue is easily reached by public transport:

- **U1:** Taubstummengasse
- **Trams 1 and 62, and Badner Bahn:** Paulanergasse
- **U1/U4:** approximately five minutes on foot from Karlsplatz

Vienna International Airport can be reached from the city by the City Airport Train, the S7 railway or Vienna Airport Lines buses. Vienna's dense public transport network makes it easy to travel between the conference venue, hotels and social events; renting a car is not necessary.

Organizing Committee

General Chairs

Robert Sablatnig, TU Wien, Austria

Florian Kleber, TU Wien, Austria

Program Chairs

Gernot A. Fink, TU Dortmund, Germany

Alicia Fornés, Computer Vision Center, Spain

Koichi Kise, Osaka Metropolitan University, Japan

Dan Lopresti, Lehigh University, USA

Publication Chairs

Vincent Christlein, FAU Erlangen-Nürnberg, Germany

Foteini Liwicki, Luleå University of Technology, Sweden

Industry Chairs

Josep Lladós, Computer Vision Center, Spain

Stefan Fiel

Tutorial Chairs

Anna Scius-Bertrand, University of Applied Sciences and Arts Western Switzerland, Switzerland

Marcus Liwicki, Luleå University of Technology, Sweden

Anna Zhu, Wuhan University of Technology, China

Workshop Chairs

Jihad El-Sana, Ben-Gurion University of the Negev, Israel

Elisa Barney Smith, Luleå University of Technology, Sweden

George Retsinas, NVIDIA, USA

Publicity Chairs

Jean-Marc Ogier, University of La Rochelle, France

Umapada Pal, Indian Statistical Institute, India

Doctoral Consortium Chairs

Anna Scius-Bertrand, University of Applied Sciences and Arts Western Switzerland, Switzerland

Momina Moetesum, National University of Sciences & Technology (NUST), Pakistan

Posters & Demo Chairs

Eyad Elyan, Robert Gordon University Aberdeen, UK

Liangrui Peng, Tsinghua University, Beijing, China

Competition Chairs

Dimosthenis Karatzas, Computer Vision Center, Spain

Srirangaraj Setlur, University at Buffalo, USA

Awards Chair

Richard Zanibbi, Rochester Institute of Technology, USA

Sponsorship Chairs

Basilis Gatos, NCSR Demokritos, Greece

Apostolos Antonacopoulos, University of Salford, Manchester, UK

Local Organization Chairs

Marvin Burges, TU Wien, Austria

Marco Peer, University of Fribourg, Switzerland

Katharina Pois, TU Wien, Austria

Food & Beverage Chair

Rafael Sterzinger, TU Wien, Austria

Program Committees

Senior Program Committee

Apostolos Antonacopoulos	Liangcai Gao	Harold Mouchère
Xiang Bai	Daichi Haraguchi	Wataru Ohyama
Michael Blumenstein	Rajiv Jain	Umapada Pal
Łukasz Borchmann	C.V. Jawahar	Thierry Paquet
Jean-Christophe Burie	Lianwen Jin	Antonio Parziale
Silvia Cascianelli	Dimosthenis Karatzas	Liangrui Peng
Joseph Chazalon	Christopher Kermorvant	Faisal Shafait
Vincent Christlein	Bart Lamiroy	Elisa Barney Smith
Bertrand Couasnon	Cheng-Lin Liu	Jun Sun
Mickael Coustaty	Lu Liu	Seiichi Uchida
Kenny Davila	Josep Lladós	Jerod J Weinman
David Doermann	Muhammad Imran Malik	Richard Zanibbi
Véronique Eglin	Simone Marinai	Anna Zhu
Andreas Fischer	Ronaldo Messina	

Program Committee

Irfan Ahmad	Shanxiong Chen	Sheng He
Alireza Alaei	Zhineng Chen	Nina S. T. Hirata
Vlad Atanasiu	Christian Clausner	Kai Hu
Ahmad Montaser Awal	Mark Clement	Florent Imbert
Raphael Baena	Florence Cloppet	Donato Impedovo
Souhail Bakkali	Denis Coquenat	Brian K Iwana
Killian Barrere	Pengwen Dai	Maham Jahangir
Byron L D Bezerra	Axel De Nardin	Mohammed Javed
Ujjwal Bhattacharya	Claudio De Stefano	Ehsanollah Kabir
Sanket Biswas	Qi Dong	Lei Kang
Mélodie Boillet	Antoine Doucet	Aurélié Lemaitre
Carlos Boned Riera	Jonathan Fabrizio	Lzhenjiang Li
Alceu Britto	Francesco Fontanella	Xiao-Hui Li
Jorge Calvo-Zaragoza	Akio Fujiyoshi	Minghui Liao
Edwin Carlinet	Andrea Gemelli	Laurence Likforman
Sukalpa Chanda	Petra Gomez-Krämer	Yijun Lin
Clement Chatelain	Carlos-Emiliano González-	Chang Liu
Jin Chen	Gallardo	Yuliang Liu

Artemis Llabrés	Frédéric Rayar	Ernest Valveny
Muhammad Muzzamil Luqman	Christophe Rigaud	Emanuele Vivoli
Nam Tuan Ly	Verónica Romero	Da-Han Wang
Jiahao Lyu	Henry A Rowley	Qiufeng Wang
Carlos D. Martínez-Hinarejos	Joan Andreu Sánchez	Runmin Wang
Carlos Mello	Ravi Kiran Sarvadevabhatla	Zhihui Wang
Tomo Miyazaki	Martin Schall	Hongxi Wei
Momina Moetesum	Ayush Kumar Shah	Jin Wei
Hussein Mohammed	Anuj Sharma	Fei Wu
Adrià Molina Rodríguez	Baoguang Shi	Yejing Xie
Ajoy Mondal	Steven J Simske	Fei Xu
Carlos F Moreno-Garcia	Sukhdeep Dr. Singh	Pengyu Yan
Nibal Nayef	Mohamed Ali Souibgui	Chuang Yang
Clemens Neudecker	Yann Soullard	Chun Yang
Shinichiro Omachi	Daniel Stoekl	Fei Yin
Shivakumara Palaiahnakote	Suresh Sundaram	Gangyan Zeng
Mohammad T Parvez	Solène Tarride	Chongsheng Zhang
Marco Peer	Kengo Terasawa	Heng Zhang
Julien Perret	Iuliia Tkachenko	Yanming Zhang
Vincent Poulain d'Andecy	Pau Torras	Yaping Zhang
Ioannis Pratikakis	Pierrick Tranouez	Yuchen Zheng
Xugong Qin	Oliver Tüselmann	Guoqiang Zhong
Irina Rabaev	Kurban Ubul	Yuanping Zhu
Jean-Yves Ramel	Stephan M. Unter	
Oriol Ramos Terrades	Michal Uricar	
	Szilard Vajda	

Pre Conference • Tutorials

EI7		EI9
Sunday · August 30		
09:00	From Ink to AI #1	Neurosymbolic AI #1
10:30	Coffee Break	
11:00	From Ink to AI #2	Neurosymbolic AI #2
12:30	Lunch Break, Campus Gusshaus	
13:30	OCR-LLM #1	Agentic AI #1
15:00	Coffee Break	
15:30	OCR-LLM #2	Agentic AI #2
17:00		
19:00	Welcome Reception	

Main Conference • Day 1

E17		E19		Hallway	
Monday · August 31					
08:00		Registration			
09:00		Opening			
09:20		Keynote C.V. Jawahar			
10:20		Coffee Break			
10:40		Oral Session 1 Document Analysis Systems		Oral Session 2 Document Forensics	
12:00		Lunch Break, Campus Gusshaus			
13:20		Journal Track 1		Journal Track 2	
14:00		Oral Session 3 Document Image Processing		Oral Session 4 Benchmarks and Datasets	
15:20		Coffee Break			
15:40		Doctoral Consortium			
16:00				Posters 1	
16:40					
17:40					
18:00		TC10/11 Meeting			

Main Conference • Day 2

	EI7	EI9	Hallway
	Tuesday • September 1		
08:00	Registration		
09:00	Oral Session 5 Graphics Recognition	Oral Session 6 Handwriting Recognition	
10:20	Coffee Break		
10:40	Oral Session 7 Historical Document Analysis	Oral Session 8 Layout Analysis	
12:00	Lunch Break, Campus Gusshaus		
13:20	Keynote Christoph Rass		
14:20	Coffee Break		
14:40	Journal Track 3	Journal Track 4	
16:00	Keynote Young Investigator Award		
17:20			
19:00	Conference Banquet, Wiener Rathaus		

Main Conference • Day 3

E17		E19		Hallway
Wednesday • September 2				
08:00	Registration			
09:20	Keynote Tong Sun			
10:20	Coffee Break			
10:40	Oral Session 9 Recognition of Tables and Formulas		Oral Session 10 Text and Symbol Recognition	
12:00	Lunch Break, Campus Gusshaus			
13:20				Posters 2 Doctoral Consortium
15:00	Coffee Break			
15:20	Competitions 1		Competitions 2	
16:40	Closing			
17:00				

Post Conference • Workshops

	EI7	EI8	EI9	EI10
	Thursday • September 3			
08:00	Registration			
09:00	DAS	HIP	DALL	WML
10:30	Coffee Break			
11:00	DAS	HIP	DALL	WML
12:30	Lunch Break, Campus Gusshaus			
13:30	DAS	HIP	ADAPDA	WML
15:00	Coffee Break			
15:30	DAS	HIP	ADAPDA	WML
17:00				

	EI7	EI8	EI9	EI10
Friday · September 4				
08:00	Registration			
09:00	DAS	IWCP	VINALDO	WORMS
10:30	Coffee Break			
11:00	DAS	IWCP	VINALDO	WORMS
12:30	Lunch Break, Campus Gusshaus			
13:30	DAS	IWCP	–	WORMS
15:00	Coffee Break			
15:30	DAS	IWCP	–	WORMS
17:00				

Welcome Reception

Sunday, August 30, 2026 · 19:00

Ottakringer Brewery

Ottakringer Platz 1, 1160 Vienna

Conference Tram shuttle

Join us after the tutorials for a tram ride to the brewery.

Departure: 18:30

Meeting point: Karlsplatz tram stop

Travelling independently

U3	Ottakring
U6	Josefstädter Straße
Tram 2, 9 or 44	Johann-Nepomuk-Berger-Platz
Tram 46	Feßtgasse
Bus 10A	Familienplatz

All stops are a short walk from the brewery.

Photo: Ottakringer Brauerei, CC BY-SA 3.0, via Wikimedia Commons.

Conference Banquet

Tuesday, September 1, 2026 · 19:00

Festsaal, Vienna City Hall

Lichtenfelsgasse 2, 1010 Vienna

Join us for an evening of excellent cuisine, inspiring conversations and Viennese elegance in the historic Festsaal of Vienna City Hall.

How to get there

U2	Rathaus
U3	Volkstheater
Tram 1, 71 or D	Rathausplatz/Burgtheater
Tram 2	Rathaus

Approximately 25 minutes on foot from the conference venue through Vienna's city centre.

Photo: Gryffindor, via Wikimedia Commons. Public domain.



Keynote Talk

Monday • August 31



Prof. Dr. C. V. Jawahar

IIIT Hyderabad, India

The Changing Landscape of Document Understanding

Abstract

Document understanding has changed dramatically over the past two decades. Early systems were built as tightly connected pipelines combining optical character recognition, layout analysis, and rule based post processing. Over time, the field has moved toward richer representations that aim to capture not just text, but also structure, semantics, and intent. This shift is now accelerating, driven by large scale data, self-supervised learning, and foundation models that increasingly integrate vision, language, and reasoning.

This keynote reflects on how document understanding research has evolved and what has driven these changes. It revisits key shifts in perspective, from handcrafted features to learned representations, from narrowly defined tasks to more unified models, and from explicit structural modeling to implicit multimodal reasoning. Despite this progress, documents continue to pose distinctive challenges. Complex layouts, long range dependencies, multilingual content, and domain specific conventions still make document understanding resistant to simple end to end solutions.

The talk highlights emerging directions that are shaping both research and practice in the field. These include the growing influence of foundation models, new forms of supervision and data creation, and evolving evaluation practices that go beyond recognition accuracy alone. We also discuss how this impacts the ongoing research as well as researchers in the document community. These changes also impact the underrepresented languages and geographies.

The keynote also touches on open questions and opportunities for the community, emphasizing the importance of benchmarks that enable deeper and more meaningful document intelligence.

Bio

Dr. C. V. Jawahar is a Professor at IIIT Hyderabad and the co-founder of the Centre for Visual Information Technology (CVIT). His research spans document image analysis, computer vision, and multimodal learning, with a long-standing focus on building intelligent systems that go beyond recognition to enable higher-level understanding and reasoning over visual and textual data.

Prof. Jawahar has made significant contributions to document understanding, multilingual technologies, and real-world vision systems, particularly in the context of large-scale document processing for societal applications. He has been an active contributor to the ICDAR community through research, tutorials, and leadership roles.

His work has influenced both academic research and systems operating at a population scale. He is a recipient of several national and international recognitions, including the ACM's Outstanding Contribution to Computing Education (OCCE) and Outstanding Applied Innovation in Computing (OAIC). His current research interests include vision-language models, document intelligence, and next-generation AI systems that integrate perception with reasoning. He is passionate about applications in road safety, assistive technology, education, entertainment, and cultural heritage.

Keynote Talk

Tuesday • September 1



Prof. Dr. Christoph Rass

Osnabrück University, Germany

From Archives to Algorithms and Back: What Historians Need from Document Understanding

Abstract

Historians may be among the most demanding users of document analysis technologies, though this is often overlooked. The challenge differs by period. Modern history is overwhelmed by sources: reports, forms, and registries running into millions of pages, often legible yet impossible to process manually. Earlier periods present the opposite problem: fewer documents, but harder to read, written in historical hands and damaged by time. In both cases, understanding a document means more than recognizing characters. It requires reconstructing meaning, context, and intent – often recovering the categories through which contemporaries understood their world.

This keynote offers a historian's perspective on document understanding, drawing on nearly three decades of data-driven research. When I began building databases from Wehrmacht personnel files in 1996, “electronic data processing” meant relational databases and statistical analysis of manually transcribed records – laborious, yet transformative. Since then, the field has moved through several technological generations: geographic information systems, automated extraction from archival card files, and now Large Vision Language Models capable of processing document images directly. Each transition opened new possibilities, but also raised questions about what we gain and lose when machines mediate access to the past.

Three themes structure this reflection. First, scale: digital methods allow us to process mil-

lions of documents, changing which questions become tractable, while demanding new forms of judgment about what patterns mean and what they obscure. Second, representation: how do we model historical knowledge in graphs and ontologies without flattening ambiguity or reproducing the categories embedded in the documents themselves? How do we encode uncertainty in ways future researchers can still interpret? Third, validation: when extraction volumes exceed any possibility of comprehensive review, what role can domain expertise still play, and where does it reach its limits?

Bio

Christoph Rass is Professor of Modern History and Historical Migration Research at Osnabrück University, Germany. His path toward data-driven historical research began at Saarland University in the early 1990s, where he studied history alongside information science. His dissertation, a database-driven collective biography of Wehrmacht personnel records, established an approach he has pursued through several technological generations since: from relational databases and statistical social profile analysis, through geographical information systems for modeling migration and organized violence, to automated extraction from historical card files.

Projects on the Osnabrück Gestapo registry and the city's foreign resident registry developed pipelines for extracting structured data from administrative documents – typed, handwritten, and hybrid – at scale. Current work aims to extend these approaches to Large Vision Language Models. Funding has come from the German Research Foundation, the Volkswagen Foundation, and the Foundation Remembrance, Responsibility and Future. Collaborative work with the Pattern Recognition Group at TU Dortmund on few-shot information extraction was presented at ICDAR 2025.

Rass is Principal Investigator in the DFG Collaborative Research Center 1604 “Production of Migration,” where his subproject applies digital corpus analysis to trace the historical semantics of categories such as “Gastarbeiterkinder” in educational discourse. He contributes to the Lower Saxony research consortium FuturMig and serves on the board of the Institute for Migration Research and Intercultural Studies.

Keynote Talk

Wednesday • September 2



Dr. Tong Sun

Senior Director, Adobe

The Next Frontier: From Document Understanding to Document Agency

Abstract

For decades, the ICDAR community has advanced the state of the art in document analysis – transforming static pages into machine-readable and structured representations. Today, however, rapid progress in large language models, multimodal foundation models, and agentic AI systems is reshaping the landscape of knowledge work.

Understanding alone is no longer the end goal.

This keynote introduces the concept of document agency – a new paradigm in which documents evolve from passive information containers into dynamic, interactive, and executable knowledge environments. Rather than serving merely as inputs to AI systems, documents become structured state spaces that support reasoning, planning, personalization, and tool use.

I will explore the technical foundations required for this shift, including multimodal structural representations, long-context reasoning architectures, generative interface layers, and new evaluation frameworks that measure collaborative intelligence rather than extraction accuracy.

As AI transforms the knowledge workplace, ICDAR has an opportunity to lead the next frontier: building document systems that not only understand content, but actively participate in cognition and action.

Bio

Dr. Tong Sun is Senior Director of the Document Intelligence Lab at Adobe, where she leads research at the intersection of multimodal AI, document foundation models, and agentic systems. Her work focuses on transforming documents from static files into interactive, intelligent media that power the next generation of knowledge work.

Bridging cutting-edge research with large-scale product impact, she and her team at Adobe Research have advanced structured multimodal reasoning, efficient on-device models, and generative interfaces – shaping a future where document systems actively collaborate with humans in thinking, decision-making, and action.

Young Investigator Award

Tuesday • September 1

We have the honour to announce this year's recipients of the **IAPR/ICDAR Young Investigator Award**: **Dr. Silvia Cascianelli** and **Prof. Dr. Jorge Calvo-Zaragoza**.

Both recipients give keynote talks on their research in the ceremony session on Tuesday, September 1, at 16:00.

Congratulations!



Dr. Silvia Cascianelli

University of Modena and Reggio Emilia

The paraverbal language of text images: how your text looks matters as much as what it reads

Abstract

Every written artifact (a handwritten letter, a music score, a printed sign photographed in the wild) is a visual object, and how it looks carries meaning of its own beyond what it reads: the hand of a writer, the conventions of a notation system, an aesthetic, a historical trace. Learning to model this visual dimension is what allows us not only to read text, but to reproduce it, restyle it, and edit it, with direct consequences for document analysis, digital humanities, cultural heritage, and accessibility.

This talk traces a research line that treats text as visual content to be modeled, generated, and edited. It will begin with handwritten text generation, from transformer-based GANs that render arbitrary content to more recent autoregressive formulations that improve controllability and faithfulness. I will then turn to controllable diffusion models to perform style transfer across images to support automatic recognition. Fi-

nally, I will present a pipeline for editing printed text directly within images, built on flow-matching models and extended to the challenging multi-instance setting.

The aim of the talk is to convey the idea that faithfully capturing how text looks is at once a demanding generative problem and a powerful instrument that complements recognition and opens new ways for the document community to preserve, understand, and interact with visually rich text.

Bio

Silvia Cascianelli is an Assistant Professor in the Department of Engineering "Enzo Ferrari" at the University of Modena and Reggio Emilia, Italy. She is a member of the Equality, Diversity and Inclusion Committee of the International Association for Pattern Recognition (IAPR) and of the European Laboratory for Learning and Intelligent Systems (ELLIS). She received her European PhD in Industrial and Information Engineering from the University of Perugia in 2019. Her research lies at the intersection of Computer Vision, Generative AI, Document AI, Multimodal AI, and Digital Humanities, with a particular focus on generating, editing, and understanding visually rich documents. She has published extensively in top-tier venues, including CVPR, ICML, ECCV, and IEEE T-PAMI, and fosters active research collaborations with international universities and industrial partners, including Google. She contributes to the community as an Associate Editor for IEEE Robotics and Automation Letters and as an Area Chair for conferences such as CVPR, ECCV, AAAI, and ICDAR. She is passionate about applying generative models to cultural heritage, digital humanities, and accessibility.



Prof. Dr. Jorge Calvo-Zaragoza

University of Alicante

Advances and Open Challenges in Optical Music Recognition

Abstract

Optical Music Recognition (OMR) addresses the automatic reading of music score images and their conversion into machine-readable representations. While often compared to OCR, OMR poses a distinctive document-analysis problem: music notation is not primarily a linear sequence of characters, but a structured visual language in which meaning depends on spatial organization, graphical relationships, simultaneity, notation conventions, and musical context. As a result, OMR requires not only recognizing visual elements on the page, but also reconstructing the musical structure that makes those elements interpretable and usable.

This keynote will introduce OMR from the perspective of document image analysis, outlining what makes music scores different from other document-reading problems and why they remain a challenging test case for visual recognition and structured interpretation. It will provide an overview of the evolution of the field, highlighting representative advances in approaches, methods, representations, datasets, evaluation, and practical tools. The talk will use selected examples to show how OMR has evolved from early recognition pipelines to current learning-based end-to-end approaches.

The talk will also discuss open challenges that continue to define the field, including the recovery of musical structure beyond symbol recognition, evaluation, data limitations, robustness across domains, and the role of human correction in real-world workflows. More broadly, it will present OMR as a relevant case study for document analysis problems in which reading means moving from visual objects to structured, domain-specific meaning.

Bio

Jorge Calvo-Zaragoza is Full Professor at the University of Alicante, Spain, where he leads the Pattern Recognition and Artificial Intelligence Group (PRAIG), devoted to machine learning, computer vision, document analysis, and music information retrieval. He received his PhD from the University of Alicante in 2016 and subsequently held postdoctoral positions at McGill University, Canada, and the Universitat Politècnica de València, Spain.

His research focuses especially on Optical Music Recognition (OMR), addressing how to transform digitised music scores into structured, machine-readable musical documents that support intelligent access to musical heritage. He has contributed to the consolidation of OMR through deep learning methods for printed and handwritten music recognition, structured representations, benchmarks, evaluation frameworks, software tools, and community-building initiatives such as the Workshop on Reading Music Systems (WoRMS). In this area, he has led several national and European research projects, supervised doctoral theses, and developed collaborations with cultural heritage institutions that have helped bring OMR closer to real-world applications.

IAPR/ICDAR Service Award

The newly established IAPR/ICDAR Service Award honors individuals whose sustained contributions have significantly advanced the field of document analysis through research, mentorship of early-career researchers, and service to the community.

We are proud to announce this year's recipients:



Elisa Barney Smith

Luleå University of Technology, Sweden



Laurence Likforman-Sulem

Telecom ParisTech, France



Jean-Marc Ogier

University of La Rochelle, France

The award ceremony takes place during the conference banquet.

Congratulations!

Outstanding Reviewer

We are pleased to recognize the following Outstanding Reviewers, whose high-quality reviews, as assessed by their Meta-Reviewers and the Program Committee Chairs, placed them among the top reviewers for ICDAR 2026. Congratulations!

Adrià Molina Rodríguez
Ajoy Mondal
Andrea Gemelli
Ayush Kumar Shah
Baole Wei
Byron Bezerra
Carlos Moreno-Garcia
Carlos-Emiliano González-Gallardo
Chang Liu
Claudio De Stefano
Denis Coquenat
Donato Impedovo
Fei Xu
Francesco Fontanella
Gonzalo Romero-García
Henry Rowley
Hongxi Wei
Irfan Ahmad
Jonathan Fabrizio
Jorge Calvo-Zaragoza
Kai Hu

Killian Barrere
Kurban Ubul
Lzhenjiang Li
Marco Peer
Maximiliano Hormazábal
Minghui Liao
Momina Moetesum
Nibal Nayef
Nina Hirata
Oliver Tüselmann
Pierrick Tranouez
Qi Dong
Qiufeng Wang
Ravi Kiran Sarvadevabhatla
Sanket Biswas
Stephan Unter
Xiao-Hui Li
Yann Soullard
Yuliang Liu
Zhineng Chen

Tutorials

Tutorials take place on Sunday, August 30, 2026.

From Ink to AI: Innovations and Integration in Forensic Handwriting Examination

E17 • Morning Session (09:00–12:30)

Organizers: Jonathan Heckerroth, Elisa van den Heuvel, Samiah Ibrahim, Erich Kupferschmid, and Angelo Marcelli.

This tutorial examines the role of handwriting in civil and criminal proceedings and discusses why automated handwriting and signature verification tools have only seen limited adoption in forensic casework. It connects forensic workflows with advances in pattern recognition and AI, highlighting challenges such as non-representative datasets, limited explainability, unsuitable output formats, and integration into operational practice. The session also presents practical pathways for collaboration, benchmarking, validation, and transparent tool development.

Document Understanding with Neuro-Symbolic AI: From Traditional Parsing to GraphRAG

E19 • Morning Session (09:00–12:30)

Organizer: Thomas M. Breuel, NVIDIA Research.

This tutorial traces the evolution of document understanding systems from traditional OCR and grammar-based parsing pipelines to neuro-symbolic architectures that combine large language models with structured knowledge representations. It discusses error propagation in classical pipelines, the limitations of purely statistical approaches, and emerging designs based on retrieval-augmented generation and GraphRAG. The focus is on practical patterns for building scalable, robust, and trustworthy document AI systems.

OCR in the Era of Large Language Models

E17 • Afternoon Session (13:30–17:00)

Organizers: Liangrui Peng and Yuliang Liu.

This tutorial focuses on how multimodal large language models are reshaping the traditional OCR pipeline, including layout analysis, text recognition, information extraction, and evaluation. It discusses end-to-end document recognition and understanding, memory mechanisms, text-centric multimodal models, long-document understanding, and benchmarking. The session offers a forward-looking perspective on methods, system design, and evaluation for OCR in the age of large language models.

Agentic AI for Document Analysis Systems

E19 • Afternoon Session (13:30–17:00)

Organizers: Ornella Vaccarelli, Olivier D’Ancona, Tim Hallyburton, Anna Scius-Bertrand, and Beat Wolf.

This hands-on tutorial introduces agent-based approaches for building adaptive document analysis workflows. Rather than relying on fixed processing pipelines, participants explore how agents can coordinate OCR modules, layout analysis models, data storage systems, and validation routines while adapting to intermediate outputs. The tutorial also discusses the Model Context Protocol as a structured interface for connecting models with external tools and data sources, supporting modularity, interoperability, and realistic document processing scenarios.

Doctoral Consortium

The Doctoral Consortium provides a forum for PhD students to present and discuss their research. It enables participants to get feedback from experienced researchers, exchange ideas with peers, and refine their research questions, methodology, and research plans. Each doctoral student is paired with one mentor who provides guidance and feedback throughout the consortium. Participants present their research in the pitch session and take part in the second poster session during ICDAR.

Sessions

- Monday, 15:40 – 16:40: Keynote Speaker + Pitch Session
- Wednesday, 13:20 – 15:00: Poster Session
- Wednesday, 16:40: Award Presentation during Closing session

Keynote Talk



Dr. Josep Lladós

Computer Vision Center, Barcelona, Spain

Beyond the Thesis: Developing a Research Vision and Navigating the PhD Journey

A PhD is more than a research project, it is a journey toward becoming an independent researcher. This keynote will explore how to develop a research vision, formulate research questions, and position one's work within a broader scientific landscape. Drawing on experience from research and doctoral supervision, the speaker will discuss approaches to defining a research problem, engaging with the state of the art, designing a research strategy, and communicating research results. The keynote will also reflect on the iterative nature of research and the role of critical thinking, feedback, and collaboration throughout the doctoral journey.

Participants

DC.1 Ankit Sinha

Some Aspects of Optical Music Recognition from Scanned Documents for Immersive Music Learning

Supervisor: Chiranjay Chattopadhyay and Rahul Kumar Ray

Mentor: Jorge Calvo-Zaragoza

DC.2 Catyana Heyne

Multimodal Information Extraction from Visually Rich Documents in Domain-Specific and Multilingual Settings

Supervisors: Jürgen Friel and Filippo Riccio

Mentor: Nicholas Howe

DC.3 Jan-Frederik Stock

Beyond Text and Pixels: Spatial Reasoning for Manuscript Exploration

Supervisors: Hussein Mohammed and Frank Steinicke

Mentor: Jean-Christophe Burie

DC.4 Mirjam Cuper

An Evaluation Framework for Automatic Text Recognition of Historical Documents

Supervisors: Cynthia Liem and Julián Urbano

Mentor: Mickael Coustaty

DC.5 Olivier D'Ancona

Agentic Systems for Complex Document Understanding

Supervisors: Andreas Fischer and Anna Scius-Bertrand

Mentor: Momina Moetesum

DC.6 Poonam Poonam

From Pixels to Data: Deep Learning for Understanding Information Visualizations

Supervisor: Timo Ropinski

Mentor: Andreas Fischer

DC.7 Seth Stewart

High-Level Retrieval from Historical Document Collections: From OCR to Evidence-Grounded Findings

Supervisor: David Wingate

Mentor: Vincent Christlein

DC.8 Syed Yaqub Shah

Deep Learning for Low-Resource Cursive Scripts: Advancing Urdu-Pashto Handwritten Text Recognition

Supervisor: Robert Sablatnig

Mentor: Faisal Shafait

DC.9 Tim Hallyburton

Combining Images and Graph Neural Networks

Supervisor: Gernot Fink

Mentor: Josep Lladós

DC.10 Welf Wustlich

Distributed Cognitive Architectures: From Biological Foundations to Formal Convergence Theory and Operational Validation

Supervisor: Thomas Martinetz

Mentor: Silvia Cascianelli

Monday • August 31

Oral Session 1

Document Analysis Systems

E17, Chair: Thierry Paquet, Josep Lladós

O.1.1 10:40 – 11:00

Spatially-Grounded Gaussian-Prior Attention for Handwritten Mathematical Expression Recognition

Ibtissem Haj Ali, Harold Mouchère

O.1.2 11:00 – 11:20

Bar-JEPA: Extracting Values from Bar Chart with Joint-Embedding Predictive Architecture

Poonam Poonam, Alexander Eppe, Timo Ropinski

O.1.3 11:20 – 11:40

GeoLogVQA: A Borehole Log Documents Dataset for Explicit and Implicit Spatial Reasoning

Stanislas Bagnol, Killian Barrere, Veronique Eglin, Elöd Egyed-Zsigmond, David Pitaval, Jean-Marie Côme

O.1.4 11:40 – 12:00

AdaNav: Query-Adaptive Multi-Granularity Navigation for Long Document Understanding

Yiming Xu, Eric López, Artemis Llabrés, Maximiliano Hormazábal, Ernest Valveny, Dimosthenis Karatzas

Oral Session 2

Document Forensics and Provenance

E19, Chair: Seiichi Uchida, Anna Scius-Bertrand

O.2.1 10:40 – 11:00

Doc-Protector: A Self-Healing Approach for Digital Documents

Sudev Padhi, Archana Tiwari, Umesh Kashyap, Sk. Subidh Ali

O.2.2 11:00 – 11:20

Temporal Modeling of Optically Variable Devices in Identity Documents

Glen Pouliquen, Joseph Chazalon, Guillaume Chiron, Oriol Ramos Terrades, Thierry Gerlaud, Ahmad Montaser Awal

O.2.3 11:20 – 11:40

DocQT: Improving Document Forgery Localization Robustness via Diverse JPEG Quantization Tables

Kylian Ronfleur Corail, Nicolas Sidere, Guillaume Bernard, Mickael Coustaty

O.2.4 11:40 – 12:00

Adversarial Attacks on Online Handwriting using Salience-based Temporal Editing

Yataro Tamura, Brian Kenji Iwana, Jiseok Lee

Journal Track Session 1

E17, Chair: Joseph Chazalon, Sanket Biswas

J.1.1 13:20 – 13:40

Bharat Scene Text: A Novel Comprehensive Dataset and Benchmark for Indian Language Scene Text Understanding

Anik De, Abhirama Subramanyam Penamakuri, Rajeev Yadav, Aditya Rathore, Harshiv Shah, Devesh Sharma, Sagar Agarwal, Pravin Kumar, Anand Mishra

J.1.2 13:40 – 14:00

Fidel: A Large-Scale Sentence Level Amharic OCR Dataset

Tunga Tessema Chamisso, Blessed Guda, Bereket Retta Adego, Carmel Prosper Sagbo, Gabriel Zencha Ashungafac, Assane Gueye

Journal Track Session 2

E19, Chair: Faisal Shafait, Jihad El-Sana

J.2.1 13:20 – 13:40

Reviving Medieval Byzantine Seals: A Synthetic-to-Real Approach to Character Recognition

Gianluca Dalmaso, Patric Reineri, Mathieu Pscherer Noel, Ninon Achard, Beatrice Caseau, Laurence Likforman Sulem, Davide Cavagnino, Maurizio Lucenteforte, Attilio Fiandrott, I Victoria Eyharabide

J.2.2 13:40 – 14:00

DKDS: A Benchmark Dataset of Degraded Kuzushiji Documents with Seals for Detection and Binarization

Rui-Yang Ju, Kohei Yamashita, Hirotaka Kameko, Shinsuke Mori

Oral Session 3

Document Image Processing

E17, Chair: Vincent Poulain d'Andecy, Lukasz Borchmann

O.3.1 14:00 – 14:20

P-HTG: One-Shot Handwritten Text Generation via Prototype-Guided Adaptive Gated Fusion

Heng Wang, Yiming Wang, Hongxi Wei

O.3.2 14:20 – 14:40

Revisiting Structural Dependency in Autoregressive Multi-Task Table Recognition via Order-Independent Cell-Level Representations

Takaya Kawakatsu

O.3.3 14:40 – 15:00

Improving MLLM Historical Record Extraction with Test-Time Image Augmentation

Taylor Archibald, Tony Martinez

O.3.4 15:00 – 15:20

Handwriting Trajectory Recovery with Diffusion Models

Hiroki Nagamatsu, Shoji Toyota, Seiichi Uchida

Oral Session 4

Benchmarks and Datasets

E19, Chair: Anna Zhu, Veronique Eglin

O.4.1 14:00 – 14:20

JaWildText: A Benchmark for Vision-Language Models on Japanese Scene Text Understanding

Koki Maeda, Naoaki Okazaki

O.4.2 14:20 – 14:40

ADV-FORMS: A Dataset of Form-Based Historical Documents With Benchmarks for Layout Analysis, HTR and OCR

Bernhard Ortbauer, Tobias Doppler, Pauline Schmidt, Lukas Schilcher, Wolfgang Göderle, Malte Rehbein, Alexander Werth, Roman Kern

O.4.3 14:40 – 15:00

Structure-Aware Text Recognition for Ancient Greek Critical Editions

Nicolas Angleraud, Antonia Karamolegkou, Benoit Sagot, Thibault Clérice

O.4.4 15:00 – 15:20

A Text Recognition Dataset from Sahidic Coptic Ancient Manuscripts

Fabio Quattrini, Carmine Zaccagnino, Costanza Bianchi, Silvia Cascianelli, Rita Cucchiara

Poster Session 1 (16:00 – 17:40)

P.1.1

RoWeR: RoBERTa Word error Rate estimator for OCREd texts

Tomás Osório, Henrique Lopes Cardoso

P.1.2

A MambaVision-Based Cross-Modal Feature Enhancement Network for Scene Text Super-Resolution

Ruichang Zhu, Hongxi Wei, Bo Sun, Heng Wang

P.1.3

DocCenter: Center and Corner Aware Representation for Robust Multi-Document Localization

Mengyuan Zhao, Kun Xu, Xin Cheng, Ting Li, Qiuman Tan, Xinyao Zhang

P.1.4

EMBLEM: Enhancing Multi-script Table Detection through Masking

Dhruv Kudale, Udhay Brahmi, Ganesh Ramakrishnan

P.1.5

Diffusion-Based Multi-View Reasoning for Scene Text Detection

Debayan Das Gupta, Shivakumara Palaiahnakote, Palash Ghosh, Umapada Pal, Cheng-Lin Liu

P.1.6

MultiFOLD: A Multimodal Framework to correct OCR Lapses in cluttered Documents

Rajat Verma, Vriti Sharma, Manikandan Ravikiran, Rohit Saluja

P.1.7

BinDiffuser: Learning Binary Style Priors to Guide Diffusion Models for Palm-Leaf Document Binarization

Salman K H, Chakravarthy Bhagvati

P.1.8

Synthetic Training Data Generation for 3D Cuneiform Sign Recognition

Jan Philipp Bullenkamp, Florian Linsel, Lisa Wilhelmi, Hubert Mara

P.1.9

GSMP: Geometry-Structured Masked Pre-training with Multi-Granularity Masking and Curriculum Learning for Geometric Problem Solving

Ziming Li, Jie Zhang, Xingxiang Zhou, Minzhi Zhang, Zhi Chen, Guanglai Gao, Xiangdong Su

P.1.10

An Analysis of Lightweight Models for Document Image Machine Translation

Abantika Bose, Thomas Gorges, Lukas Hüttnert, Linda-Sophie Schneider, Mathias Seuret, Fei Wu, Vincent Christlein

P.1.11

Adaptive Hybrid Machine Translation for E-commerce: A Reinforcement Learning Approach to Arabic Localization

Ayman Hanafy, Farhan Khawar

P.1.12

Token Selection Strategies for Automatic Summarization of Historical Documents

Merlin Streilein, Tobias Steiner, Andreas Fischer, Kaspar Riesen

P.1.13

Doc2Doc: Structure-Aware Generative Rendering for Bi-Directional Document Translation

Fahad Alotaibi, Daulet Toibazar, Renad Almusaad, Ranya Alkahtani, Haneen Alhomoud, Asma Ibrahim, Yazeed Alharbi, Murtadha Aljubran, Pedro Moreno

P.1.14

TKPE: Topic-based Evaluation for Keyphrase Prediction

Bingke Li, Jinghan Li, Jinhao Chen, Wu Zhuang, Yuxiang Zhang

P.1.15

Active Reference Acquisition in Few-Shot Font Generation

Shinnosuke Matsuo

P.1.16

Synth-JDoc: Synthesizing a Japanese Document Image Dataset for OCR with Diverse Layouts and Embedded Images

Keito Sasagawa, Shuhei Kurita, Daisuke Kawahara

P.1.17

Ambiguity-Controlled Handwritten Mathematical Expression Generation via Harmonized Dual-Conditional Guidance

Ryo Ishiyama, Takaya Kawakatsu

P.1.18

Multi-Modal OMR for Heterogeneous Notations: A Collaborative Framework for Real-Time Symbolic-to-Immersive Mapping

Ankit Sinha, Atanu Saha, Chiranjoy Chattopadhyay, Rahul Kumar Ray

P.1.19

Meaning Lies in Structure: Fine-Grained Table-Centric Document Semantic Parsing

Xuan Li, Mengfei Xiao, Jingtian Wei, Jialiang Dong, Raymond Wong

P.1.20

Evaluating Feedback by Iterative Repair of Multi-Step Solution Documents

Tobias Lengfeld, Jakob Seitz, Radu Timofte

P.1.21

Towards Scalable Knowledge Graph Extraction from Piping and Instrumentation Diagrams

Sanket Deshmukh, Apurva Gala, David Blom, Detlef Hohl

P.1.22

Hierarchical Co-Embedding of Font Shapes and Impression Tags

Yugo Kubota, Kaito Shiku, Seiichi Uchida

P.1.23

Identify, Locate, Link: End-to-End Key-Value Extraction from Document Images

Abdurrahman Said Gürbüz, Ahmed Nassar, Christoph Auer, Maksym Lysak, Lucas Morin, Matteo Omenetti, Tim Strohmeyer, Panagiotis Vagenas, Nikolaos Livathinos, Michele Dolfi, Peter Staar

P.1.24

GraphVLM: Combining VLMs with GraphM-LLM for Document Understanding

Shuai Li, Xiao-Hui Li, Haijie Yuan, Fei Yin, Lin-Lin Huang

P.1.25

LMS-Retrieval: Layout-Aware, Modality-Aware, Structure-Aware Document Retrieval

Man Qin, Tim French, Wei Liu

P.1.26

CzechTopic: A Benchmark for Zero-Shot Topic Localization in Historical Czech Documents

Martin Kostelník, Michal Hradiš, Martin Dočekal

P.1.27

Patram-Bench: A Comprehensive Multi-task, Multi-domain and Multilingual Benchmark for Indian Document Understanding

Anirudh Srinivasan, Pratyush Jena, Arya Topale, Venkat Kesav, Ravi Kiran Sarvadevabhatla

P.1.28

DDD - A Diagnostic Dataset for Character Recognition and Detection on Ancient Egyptian Hieratic Characters and Words

Stephan Unter, Elena Hertel

P.1.29

TimeAgent: From Matches to Memories — Timeline Summarization for Sports Analytics

Swagata Mukherjee, Samar Kumar Srivastava, Sriparna Saha

P.1.30

Towards Breaking the Visual Perception Bottleneck for Geometry Problem Solving

Tianjiao Cao, Jiahao Lyu, Dongbao Yang, Weimin Mu, Zhou Yu

P.1.31

Synthetic Data from Simulated Lecture Environments for Handwritten Content Extraction

Min Song, Kenny Davila

P.1.32

Stringalign: Moving beyond summary statistics with a transparent Unicode-aware tool for evaluating automatic transcription models

Yngve Mardal Moe, Marie Roald

P.1.33

DeChart: A Benchmark and Text-Enhanced Chart-to-Table Conversion Method with Multimodal LLMs

Chen-Yu Xie, Xiao-Hui Li, Fei Yin, Cheng-Lin Liu

P.1.34

BullingerDB: A Dataset for Handwritten Text Recognition and Writer Retrieval

Marco Peer, Anna Scius-Bertrand, Patricia Scheurer, Andreas Fischer

P.1.35

n-gram injection into transformers for dynamic language model adaptation in handwritten text recognition

Florent Meyer, Laurent Guichard, Yann Soullard, Denis Coquenat, Guillaume Gravier, Bertrand Couasnon

P.1.36

G2I: A Progressive Structure-to-Detail Curriculum Training Strategy for Handwritten Mathematical Expression Recognition

Minzhi Zhang, Xingxiang Zhou, Ziming Li, Jie Zhang, Zhi Chen, Xiangdong Su

P.1.37

Few-Shot Writer Adaptation via Multimodal In-Context Learning

Tom Simon, Pierrick Tranouez, Stephane Nicolas, Clement Chatelain, Thierry Paquet

P.1.38

Specialized HTR vs Vision-Language Models: Evaluating DANIEL and Fine-Tuned Qwen on Historical Documents

Gabriel Frossard, Franck Gechter

P.1.39

A Millennium of Arabic Manuscripts in Three Styles: A Line-Level OCR Benchmark for Naskh, Taliq, and Nastaliq

Maxim Novopol'tsev, Ruslan Murtazin, Andrey Sakhovskiy, Emilia Bojarskaja, Vladimir Kokh, Ivan Ulitin, Botirjon Abdullayev, Khamidulla Aminov, Masudkhon Ismoilov, Semen Budennyi

P.1.40

DiffusionRec: Recognition-Guided Diffusion for Content-Aware Urdu Handwriting Generation

Saima Kausar, Ayesha Amjad, Ahmad Samad Ali, Momina Moetesum, Adnan Ul Hasan, Faisal Shafait

P.1.41

Online Urdu Text-Line Recognition by Bridging Stroke Dynamics and Offline Representations

Ali Hussain, Rafay Ahmad, Momina Moetesum, Adnan Ul-Hasan, Faisal Shafait

P.1.42

Enhancing IMU-Based Online Handwriting Recognition via Contrastive Learning with Zero Inference Overhead

Jindong Li, Dario Zanca, Vincent Christlein, Tim Hamann, Jens Barth, Peter Kämpf, Björn Eskofier

P.1.43

Reference-Free Handwritten Japanese Character Generation via CLIP-Conditioned Diffusion Models

Koki Fujita, Hideaki Yajima, Chee Siang Leow, Hiromitsu Nishizaki

P.1.44

METATR: A Multilingual, Evolving Benchmark for Automatic Text Recognition

Mélodie Boillet, Solène Tarride, Christopher Kermorvant

P.1.45

Writer Retrieval at Scale

Tim Raven, Tim Hallyburton, Gernot A. Fink

P.1.46

Benchmarking Information Retrieval for Large Archives of Historical Documents

Tobias Steiner, Merlin Streilein, Andreas Fischer, Kaspar Riesen

P.1.47

RAGXDoc: Structured Knowledge-guided Retrieval and Explainable Re-ranking for Academic Documents

Dipendra Sharma Kafle, Esma Talhi, Mickael Coustaty, Antoine Doucet

P.1.48

Recent Advances in Information Extraction from Historical Archival Records

Arthur Matei, Tim Hallyburton, Lukas Henies, Christoph Rass, Gernot A. Fink

P.1.49

An Exploratory Study of Text-to-Image Generation for Query-by-Example Retrieval of Historical Document Images

Melissa Cote, Alexandra Branzan Albu

Tuesday • September 1

Oral Session 5

Graphics Recognition

E17, Chair: Mickael Coustaty, Jorge Calvo-Zaragoza

O.5.1 09:00 – 09:20

Optical Music Recognition for Real-World Manuscripts with Synthetic Data

Jiří Mayer, Martina Dvořáková, Vojtěch Dvořák, Markéta Herzánová Vlková, Filip Jebavý, Pavel Pecina, Samuel Šomorjai, Petr Žabička, Jr., Jan Hajič

O.5.2 09:20 – 09:40

EAGLE: Explicit Anchoring and Graph Reasoning with Diagram Structure Priors for Multimodal Geometry Problem Solving

Jie Zhang, Xiangren Wang, Ziming Li, Minzhi Zhang, Xingxiang Zhou, Zhi Chen, Guanglai Gao, Xiandong Su

O.5.3 09:40 – 10:00

Stroke-Level Connectivity Verification: Grounding Vision-Language Models Against Topology Hallucination in Diagram Understanding

Abdullah Ibne Hanif Areean, Niamul Hassan Samin, Md Arifur Rahman, Renu Akter Sweetey, Juena Ahmed Noshin, Md Ashikur Rahman

O.5.4 10:00 – 10:20

CPAgent: A Tool-Augmented Agentic Framework for Chart Parsing

Chen-Yu Xie, Xiao-Hui Li, Boran Wang, Fei Yin, Cheng-Lin Liu

Oral Session 6

Handwriting Recognition

E19, Chair: Silvia Cascianelli, Raphael Baena

O.6.1 09:00 – 09:20

InkTree: A Unified Representation of Structured Online Ink

Jakob Seitz, Tobias Lengfeld, Radu Timofte

O.6.2 09:20 – 09:40

Democratizing the medieval English legal tradition

Michael Zhang, Elise Wang, Charlotte Whately, Seth Strickland, Dylan Bannon

O.6.3 09:40 – 10:00

Limits of Style-Transferred Synthetic Data for Ad-hoc Personalization in Handwritten Text Recognition

Said Yasin, Torsten Zesch

O.6.4 10:00 – 10:20

Character Template Representation for Confidence Learning in Handwritten Text Recognition

Yangyang Liu, Heng Zhang, Fei Yin, Cheng-Lin Liu

Oral Session 7

Historical Document Analysis

E17, Chair: Elisa Barney Smith, Vincent Christlein

O.7.1 10:40 – 11:00

Blind Image Decomposition for Recovering Overlapping Text Layers on Palimpsests

Baharan Pourahmadi, Panagiotis Leonardidis, Paolo Scattolin, Mads Toudal Frandsen

O.7.2 11:00 – 11:20

Quality Prediction for Large Scale HTR - Confidence Is All You Need

Erik Lenas, Viktoria Lofgren, Olof Karsvall

O.7.3 11:20 – 11:40

Beyond Labels: Visual Invariance in Self-Supervised Learning for Aramaic Incantation Bowls

Nour Atamni, Boraq Madi, Islam Amar, Raid Saabni, Jihad El-Sana

O.7.4 11:40 – 12:00

Decipherment of Oracle Bone Inscription via Component Deconstruction and Alignment

Yuanhui Lin, Hetao Wu, Qingju Jiao, Yongge Liu, Da-Han Wang

Oral Session 8

Layout Analysis

E19, Chair: Momina Moetesum, Mathieu Aubry

O.8.1 10:40 – 11:00

REPLICA: An Agentic Framework for Visually Faithful Document Reconstruction

Raghuv eer R, Anirudh Srinivasan, Venkat Kesav Venna, Shanmukha Sreevatsa Tallapragada, Aryan Jain, Sahithi Kukkala, Ravi Kiran Sarvadev-abhatla

O.8.2 11:00 – 11:20

HKGC: A Hierarchical Knowledge Graph Construction Framework for Structure-Aware RAG

Yingxin Guan, Jian Xing, Zhaohua Zheng, Zhaofu Zeng, Bai Lei, Fanchen Meng, Haitao Guo

O.8.3 11:20 – 11:40

Complex Layout Classification in the Wild: A Low-Resource Approach with Layout-Preserving Augmentations

Sharva Gogawale, Iddo Hakim, Gal Grudka, Mohammad Suliman, Omer Ventura, Daria Shapira, Berat Barakat, Nachum Dershowitz

O.8.4 11:40 – 12:00

GRaF-Net: a Multi-Branch Gated Residual Architecture for Floor Plan Semantic Segmentation

Axel De Nardin, Silvia Zottin, Claudio Picciarelli, Gian Luca Foresti

Journal Track Session 3

E17, Chair: Umapada Pal, Andreas Fischer

J.3.1 14:40 – 15:00

TableSeq: Unified Generation of Structure, Content, and Layout

Laziz Hamdi, Amine Tamasna, Pascal Boisson, Thierry Paquet

J.3.2 15:00 – 15:20

PILOT: A Promptable Interleaved Layout-aware OCR Transformer

Laziz Hamdi, Amine Tamasna, Pascal Boisson, Thierry Paquet

J.3.3 15:20 – 15:40

Automatic Uncertainty-Aware Synthetic Data Bootstrapping for Historical Map Segmentation

Lukas Arzoumanidis, Julius Knechtel, Jan-Henrik Haunert, Youness Dehbi

J.3.4 15:40 – 16:00

AIKON: A Modular Computer Vision Platform for Historical Corpora

Ségolène Albouy, Somkeo Norindr, Paul Kervegan, Fouad Aouinti, Rémy Delanaux, Clara Grometto, Robin Champenois, Stavros Lazaris, Alexandre Guilbaud, Matthieu Husson, Mathieu Aubry

Journal Track Session 4

E19, Chair: Apostolos Antonacopoulos, Giuseppe de Gregorio

J.4.1 14:40 – 15:00

A Novel Domain Adaptation Based Pipeline for Character Classification and Handwritten Recognition

Florent Imbert, Simon Corbillé, Hui Han, Elisa Barney Smith

J.4.2 15:00 – 15:20

Unsupervised Document and Template Clustering using Multimodal Embeddings

Phillippe R. Sampaio, Helene Maxcici

J.4.3 15:20 – 15:40

HQ-Font: Few-shot Font Generation via Transferring Hierarchical Quantization Styles

Anna Zhu, Wei Pan, Guan Li, Hongyi Cai, Kenji Brian

J.4.4 15:40 – 16:00

Predicting Text Recognition Word Error Rate of Image Documents Without Ground Truth Transcripts

Enrique Vidal, Alejandro H. Toselli

Wednesday • September 2

Oral Session 9

Recognition of Tables and Formulas

E17, Chair: Bertrand Couasnon, Oriol Ramos Terrades

O.9.1 10:40 – 11:00

Beyond Rows to Reasoning: Agentic Retrieval for Multimodal Spreadsheet Understanding

Anmol Gulati, Sahil Sen, Waqar Sarguroh, Kevin Paul

O.9.2 11:00 – 11:20

Beyond the Page Break: An LLM-based Solution for Cross-Page Table Reconstruction

Yu Tang, Hongwei Li, Yixuan Cao, Ping Luo

O.9.3 11:20 – 11:40

PatentME: A Dataset and Reference-Free Post-OCR Verification Task for Printed Mathematical Expression Recognition

François Wicckowiak, Véronique Eglin, Tony Bonnet, Stéphane Bres, Laëtitia Rousseau

O.9.4 11:40 – 12:00

Multilingual Table Recognition: A Benchmark Dataset and A Local-Global Hybrid Model

Nam Tuan Ly, Atsuhiko Takasu, Masaki Nakagawa

Oral Session 10

Text and Symbol Recognition

E19, Chair: Jean-Christophe Burie, Enrique Vidal

O.10.1 10:40 – 11:00

What Can Low Resource Languages Teach Each Other?

Achyuth P, Kahaan Shah, Chetan Arora

O.10.2 11:00 – 11:20

A Benchmark of State-Space Models vs. Transformers and BiLSTM-based Models for Historical Newspaper OCR

Merveilles Agbeti-Messan, Thierry Paquet, Pierrick Tranouez, Clement Chatelain, Stephane Nicolas

O.10.3 11:20 – 11:40

Evolution-Guided Diffusion for Oracle Bone Script Decipherment

Haotian Chen, Hetao Wu, Qingju Jiao, Yongge Liu, Da-Han Wang

O.10.4 11:40 – 12:00

Towards Universal Khmer Text Recognition

Marry Kong, Rina Buoy, Sovisal Chenda, Nguonly Taing, Masakazu Iwamura, Koichi Kise

Competitions 1

E17, Chair: Dimosthenis Karatzas, Daichi Haraguchi

C.1.1 15:20 – 15:40

FalsID: Competition on Falsification and Imitation Detection

Elena Pierazzo, Franco Tomasi, Emilio Russo, Dominique Stutzmann, Riham Mokrani

C.1.2 15:40 – 16:00

Anyscript: Competition on Long-Term Handwriting Author Identification

Adrià Molina Rodríguez, Carles Boned Riera, Pau Torras Coloma, Oriol Ramos Terrades, Josep Lladós

C.1.3 16:00 – 16:20

CMMHWR26: Competition on Multilingual Medieval Handwriting Recognition

Agnes Boutreux, Bram Cars, Thibault Clérice, Matthias Gille Levenson, Mike Kestemont, Benjamin Kiessling, Ariane Pinche, Caroline Vandyck, Malamatenia Vlachou

C.1.4 16:20 – 16:40

CircleID: Competition on Writer Identification and Pen Classification from Hand-Drawn Circles

Thomas Gorges, Janne van der Loop, Fei Wu, Mathias Seuret, Vincent Christlein

Competitions 2

E19, Chair: Srirangaraj Setlur, Joan Andreu Sánchez

C.2.1 15:20 – 15:40

Sci-ImageMiner: Competition on Information Extraction from Atomic Layer Deposition/Etching (ALD/E) Scientific Figures

Fahad Ahmed, Jennifer D'Souza, Soren Auer, Ralph Ewerth, Alex Watkins, Bora Karasulu, Eleni Poupaki, Erwin Kessels

C.2.2 15:40 – 16:00

DocVQA2026: Competition on Multimodal Reasoning over Documents in Multiple Domains

Artemis Llabrés, Marc Serra Ortega, Tomás Ockier, Amritpal Singh, Christos Georgakilas, Andrey Barsky, Ernest Valveny, Dimosthenis Karatzas

C.2.3 16:00 – 16:20

HIPE-OCRepair: Competition on LLM-Assisted OCR Post-Correction for Historical Documents

Emanuela Boros, Juri Opitz, Andrianos Michail, Simon Clematide, Maud Ehrmann, Florian Wagner

C.2.4 16:20 – 16:40

Competition in Text Recognition on Greek Squeezes

Nicholas Howe, Aaron Hershkowitz

Poster Session 2 (13:20 – 15:00)

P.2.1

UniLipi: A Unified Multi-Script OCR for Historical Indic Manuscripts

Tathagata Ghosh, Sai Madhusudan Gunda, Simran Singh Sandral, Ravi Kiran Sarvadevabhatla

P.2.2

Learning Diachronic Representations of Ancient Greek Letterforms

John Pavlopoulos, Spyros Barbakos, Lavinia Ferretti, Dionysis Voulgarakis, Asimina Paparrigopoulou, Maria Konstantinidou, Giuseppe De Gregorio, Isabelle Marthot-Santaniello, Paraskevi Platanou, Holger Essler

P.2.3

Automatic Layout Detection in Historical Civil Records Using Deep Object Detection

Wissam Alkendi, Franck Gechter, Laurent Heyberger, Christophe Guyeux

P.2.4

Leveraging Morphology for Historical Script Metrological Analysis

Malamatenia Vlachou Efstathiou, Raphaël Baena, Dominique Stutzmann, Mathieu Aubry

P.2.5

One Model, Many Guidelines: Instruction Fine-Tuning for Historical Named Entity Recognition

Nam Nguyen, Emanuela Boros, Adam Jandt, Ahmed Hamdi, Mickael Coustaty, Antoine Doucet

P.2.6

Robust Interpretation of Historical Documents in Knowledge Graphs Through Query Inference and Execution

Sebastià Nicolau Orell, Adrià Molina Rodríguez, Oriol Ramos Terrades, Josep Lladós

P.2.7

Conversational Retrieval and On-the-Fly Knowledge Modeling from Historical Documents

Paula Font Solà, Adrià Molina Rodríguez, Josep Lladós

P.2.8

Revisiting how we access to historical archives: Auditing Gender Stereotypes and the Division of Labour in the Analysis of Historical Photography Collections

Francesc Net Barnes, Adrià Molina Rodríguez, Sofia Llacer-Caro, Lluís Gomez Bigorda

P.2.9

HME-Leibniz: A Multi-level Mathematical Expression Dataset from Leibniz's Manuscripts

Yejing Xie, Ze Qian, Yunfan Li, David Rabouin, Harold Mouchère

P.2.11

Parameter-Efficient and Adaptive Fine-Tuning for Long-Tailed Ancient Characters Recognition

Hao Wang, Aouaidjia Kamel, Konstantinos Kotropoulos, Chongsheng Zhang

P.2.13

Text region detection in historical astronomical diagrams

Zeynep Sonat Baltaci, Raphael Baena, Fei Meng, Somkeo Norindr, Florence Somer, Matthieu Husson, Mathieu Aubry

P.2.15

Angkorian-KSI: A Multi-Task Benchmark for Khmer Stone Inscription Analysis

Nimol Thuon, Jun Du, Ranyakol Thuon, Panhapin Theang

P.2.17

HIDRA: Hierarchical Ink-aware Dual-granularity Retrieval Architecture for Historical Fragments

Jihad Al Akl, Chady Abou Jaoude, Zahi Al Chami, Marianne Abi Kanaan, Abdallah Makhoul

P.2.10

EpiSAM: Character Segmentation in Challenging Stone Inscriptions

Arnav Sharma, Pratyush Jena, Amal Joseph, Ravi Kiran Sarvadevabhatla

P.2.12

Evaluating Vision-Language Models on Historical Postcards

Matthieu Pelingre, Salvatore Tabbone

P.2.14

Bridging the Gaps: Learning to Estimate Missing Text in Fragmentary Greek Inscriptions

Silvia Zottin, Axel De Nardin, Valentina Mignosa, Maddalena Zunino, Gian Luca Foresti

P.2.16

Vision-Language Model based Transfer Learning for Historical Document Recognition

Yifan Huang, Liangrui Peng, Tianqi Zhao, Di Wu, Kemeng Zhao, Shuo Li, Zhiyu Li, Yuyang Li

P.2.18

Theatre Chapbooks At Scale: A Statistical Comparative Analysis of Typography

Diego Belzarena, Seginus Mowlavi, Paula Casariego Castiñeira, Alejandra Ulla Lorenzo, Gregory Randall, Jean-Michel Morel

P.2.19

Generalized Open-set Single-shot Character Recognition on Ancient Egyptian Hieratic Characters

Stephan Unter, Chang Liu, Elisa Barney Smith

P.2.20

Automated Character-Level Annotation for Historical Nom Documents via an Iterative Self-Updating Radical-Aware Recognizer

Cuong Nguyen, Khoa Nguyen Tran, Ngoc Tuan Nguyen, Hung Tuan Nguyen, Nam Tuan Ly, Masaki Nakagawa

P.2.21

MaPE-Former: A Mask-Aware Position Encoding Network for Chinese Character Image Restoration

Wei Wei, Xinrui Liu, Jianxin Zhang, Xiaodong Duan

P.2.22

From Pixels to Structure: Lightweight Vision-Language Models for Document OCR and Structured JSON Extraction

Uddipan Basu Bir, Vincent Christlein, Andreas Maier, Mathias Zinnen

P.2.23

Towards Non-Latin Character and Layout Personalization for Enhanced Readability

Rina Buoy, Dylan Berkamp Fouepe Dongmo, Vesal Khean, Simone Marinai, Koichi Kise

P.2.24

Efficient Table QA via TableGrid Navigation and Progressive Inference Prompting

Amritansh Maurya, Navjot Singh, Mohammed Javed, Omar Moured

P.2.25

Can VLMs Understand Handwritten Mathematical Documents?

Shree Mitra, Ajoy Mondal, C. V. Jawahar

P.2.26

GRACE: Gradient-Regulated Approach for Consistent Explanations

Sayantan Basu

P.2.27

GLiDRE: Generalist Lightweight Model for Document-level Relation Extraction

Robin Armingaud, Romaric Besançon

P.2.28

From Chunks to Graphs: Training-Free Multimodal Late Interaction for Document Understanding

Ayush Lodh, Souparni Mazumder, Sanket Biswas, Josep Lladós, Nisha Singh

P.2.29

Figures as Evidence: Multi-Image Scientific Generation

Jawad Ibn Ahad, Mritunjoy Chakraborty, Fuad Rahman, Sifat Momen, Shafin Rahman, Nabeel Mohammed

P.2.30

Prediction of Grade, Gender, and Academic Performance of Children and Teenagers from Handwriting Using the Sigma-Lognormal Model

Adrian Iste, Kazuki Nishizawa, Chisa Tanaka, Andrew Vargo, Anna Scius-Bertrand, Andreas Fischer, Koichi Kise

P.2.31

Hierarchical Stroke-Level Clustering and Step-Level Segmentation for Automatic Scoring of Geometric Construction Answers with an Electronic Drawing Compass

Thanh-Nghia Truong, Hung Tuan Nguyen, Nam Tuan Ly, Yoichi Tsuchida, Hiroshi Miyazawa, Tomo Asakura, Masamitsu Ito, Toshihiko Horie, Fumiko Yasuno, Masaki Nakagawa

P.2.32

Towards Khmer Scene Document Layout Detection

Marry Kong, Rina Buoy, Sovisal Chenda, Nguonly Taing, Masakazu Iwamura, Koichi Kise

P.2.33

Bounding Box Label Propagation for Re-Annotation of Document Layout Analysis Datasets

Nick Jochum, Tobias Alt-Veit, Christian Schön, Alexander Lück, René Schuster, Didier Stricker

P.2.34

Counterfeit Answers: Adversarial Forgery against OCR-Free Document Visual Question Answering

Marco Pintore, Maura Pintor, Battista Biggio, Dimosthenis Karatzas

P.2.35

Active Learning for Cascaded Object Detection: Balancing Coverage and Uncertainty in Table Extraction Pipelines

Eliott Thomas, Mickael Coustaty, Aurélie Joseph, Gaspar Deloin, Vincent Poulain D'andecy, Jean-Marc Ogier

P.2.36

FastTab: A Fast Table Recognizer with a Tiny Recursive Module and 1D Transformers

Laziz Hamdi, Amine Tamasna, Pascal Boisson, Thierry Paquet

P.2.37

ConRTF: Edge-Constrained Boundary Distribution Refinement for Realtime TransFormer Table Structure Recognition

Eliott Thomas, Tri-Cong Pham, Mickael Coustaty, Aurélie Joseph, Gaspar Deloin, Vincent Poulain D'andecy, Jean-Marc Ogier, Antoine Doucet

P.2.38

SCALES: Scalable Context-Aware Learning with Expert Specialization for Incremental Multilingual Text Recognition

Qing Lin, Xiaohui Li, Heng Zhang, Fei Yin, Chenglin Liu

P.2.39

Multi-Modal Deep Learning for Medieval Inscription Recognition: A Study of Saint Sophia Cathedral Graffiti

Aram Karimi, Jonathan Westine, Gunnar Almevik

P.2.40

AOSSig4000: A Real-World Chinese Handwritten Signature Dataset with Diverse Background Noise and Pixel-Level Annotations

Xunhui Qin, Desheng Wang, Kunpeng Gui, Fang Shi, Zhonghao Shen, Du Zhou, Ke Liu, Peirong Zhang, Yang Xue, Lianwen Jin

P.2.41

Online Signature Verification Using Augmented Path Signature and T-Mamba

Ruiling Li, Danyu Yang

P.2.42

Master Forgers, Fragile Detectors? A Forensic Study of Vision-Language Models for Signature Verification

Kumari Priya, Bibek Das, Chandranath Adak, Soumi Chattopadhyay

P.2.43

Preserving High-Fidelity Character Structure in Handwritten Text Generation via Multimodal Guidance

Heng Wang, Yiming Wang, Hongxi Wei

P.2.44

LLMSFL: LLM-Driven Smart Feedback Loop System for Target Document Generation

Ze Zhong Guo, Yongjian Zhang

P.2.45

Arbitrary Glyph and Multi-Resolution Font Generation with Mixed Content Representations

Xiaoge Chen, Shilin Li, Leilei Yao, Anna Zhu

P.2.46

MIDAS: Multi-LLM Iterative Data-Adaptive Summarization

Karen Lee, Dhanashree Balaram, Seojun Shon, Umair Rasheed

P.2.47

Agentic Document Reasoning for Evidence-Grounded Clinical Report Generation

Hira Masood, Momina Moetesum, Muhammad Imran Malik, Faisal Shafait, Hassan Aqeel Khan

P.2.48

Vision Language Models as OCR Correctors for Historical Texts

Radoslav Koynov, Triet Ho Anh Doan, Philipp Wieder

P.2.49

Error Patterns in Historical OCR: A Comparative Analysis of TrOCR and Qwen2.5-VL

Ari Vesalainen, Eetu Mäkelä, Laura Ruotsalainen, Mikko Tolonen

P.2.50

Structural Analysis of Character Identity at OCR Decision Boundaries in Visually Similar Pairs

Daichi Haraguchi

P.2.51

LiteDoc: Distilling Large Document Models into Efficient Task-Specific Encoders

Tayyab Raza, Syed Muhammad Taha Imam, Adrian Ulges, Ulrich Schwanecke, Momina Moetesum, Faisal Shafait

Workshops

DAS

**17th IAPR International Workshop
on Document Analysis Systems**

E17 • September 3 & 4 (Full Day)

<https://das2026.seecs.edu.pk>



HIP

**8th International Workshop on Historical Document
Imaging and Processing**

E18 • September 3 (Full Day)

<https://blog.sbb.berlin/hip2026/>



DALL

**2nd Workshop on Document Analysis
of Low-resource Languages**

E19 • September 3, 09:00 – 12:30

<https://icdar-dall.github.io/>



ADAPDA

**3rd Workshop on Automatically Domain-Adapted
and Personalized Document Analysis**

E19 • September 3, 13:30 – 17:00

<https://adapda.silviacascianelli.com/>



WML

**6th ICDAR International Workshop
on Machine Learning**

EI10 • September 3 (Full Day)

<https://icdarwml2026.midasoc.org/>



IWCP

**4th International Workshop
on Computational Paleography**

EI8 • September 4 (Full Day)

<https://www.csmc.uni-hamburg.de/iwcp2026.html>



VINALDO

**3rd International Workshop on Machine Vision
and NLP for Document Analysis**

EI9 • September 4, 09:00 – 12:30

[https://sites.google.com/view/
vinaldo3rdeditionofinternational/call-for-papers](https://sites.google.com/view/vinaldo3rdeditionofinternational/call-for-papers)



WoRMS

**7th International Workshop
on Reading Music Systems**

EI10 • September 4 (Full Day)

<https://sites.google.com/view/worms2026/>



Thursday • September 3

Morning Session 1 (09:00 – 10:30)

DAS

Location: EI7

09:00 – 09:15 **Opening**

09:15 – 10:00

Keynote: Parse, Reflect, Retrieve, Compile: An Agent Stack for Enterprise Document AI

Jordy van Landeghem

10:00 – 10:15

POTATR: A Lightweight Image-to-Graph Model for Page-Level Table Extraction

Brandon Smock, Libin Liang, Maxim Sokolov, Amrit Ramesh, Valerie Faucon-Morin, Tayyibah Khanam, Maury Courtland

10:15 – 10:30

Structured Handwritten Form Synthesis via Unsupervised Template Recovery and Field-Level Segmentation

Natalia Bottaioli, Yanhao Li, Thomas Gorges, Solène Tarride, Jean-Michel Morel, Gabriele Facciolo, Vincent Christlein

HIP

Location: EI8

09:00 – 09:10 **Welcome**

09:10 – 09:30

MEDUSA: A Curriculum-Based Vision–Language Framework for Multilingual Medieval Handwritten Text Recognition

Théo Moins, Brenna Hensley, Florian Cafiero, Jean-Baptiste Camps, Lilla Conte, Emilie Guidi, Katarzyna Kapitan, Carolina Macedo, Cecile Vermaas, Chahan Vidal-Gorène

09:30 – 09:50

Language Similarity and Cross-Lingual Transfer in Historical HTR: Evidence from Swedish, Norwegian, and Medieval Latin

Micaella Bruton, Crina Tudor, Wout Sinnaeve, Oreen Yousuf, Signe Rirdance, Raphaela Heil, Beáta Megyesi

09:50 – 10:10

Is a Generic Dataset and Foundation VLM for Arabic HTR Worth It? Lessons from AMIDDA

Chahan Vidal-Gorène, Noémie Lucas, Clément Salah, Aliénor Decours-Perez

10:10 – 10:30

Beyond Monolithic OCR: Mixture-of-Experts Specialisation for Ancient Devanagari Script Recognition

Vriti Sharma, Rajat Verma, Rohit Saluja

DALL

Location: E19

09:00 – 09:10 **Opening**

09:10 – 09:30

Keynote: Frontiers in Document Analysis for Low-Resource Languages: A Practical Perspective from Tibetan Ancient Document Digitization

Yong Tso

09:30 – 09:50

VedOCR: An Expert-Annotated Dataset and Transformer OCR Baseline for Vedic Sanskrit Manuscripts

Suyash Kumar Bhagat, N. Jayanthi, S. Indu

09:50 – 10:10

Age estimation from offline handwriting: a case study using regression-based modeling

Irina Rabaev, Lior Abergel, Maor Merling, Marina Litvak

10:10 – 10:30

Cross-Lingual Learning within Arabic Script for Low-Resource HTR

Sana Al-Azzawi, Marcus Liwicki, Elisa Barney Smith

09:15 – 09:30 **Opening**

09:30 – 09:50

Understanding Cross-Language Transfer Improvements in Low-Resource HTR: The Role of Sequence Modeling

Sana Al-Azzawi, Marcus Liwicki, Elisa Barney Smith, Chang Liu, Nudrat Habib

09:50 – 10:10

Scaling State-Space Models from Lines to Paragraphs: An Ablation of Mamba-based OCR

Merveilles Agbeti-Messan, Thierry Paquet, Pierrick Tranouez, Clément Chatelain, Stéphane Nicolas

10:10 – 10:30

Enhancing Hindi OCR Accuracy Using Color-Driven Sub-Character Segmentation

Sourava Behera, Krish Mittal, Ankit Maurya, Rajat Verma, Manikandan Ravikiran, Rohit Saluja

Morning Session 2 (11:00 – 12:30)

DAS

Location: EI7

11:00 – 11:45

Keynote: From Retrieval to Reasoning: The Evolution of RAG-based Systems for Document-Centric AI

Nibal Nayef

11:45 – 12:00

Modeling "Musical Words" for Efficient Sequence Search in Handwritten Music Manuscripts

Aitana Menárguez-Box, Enrique Vidal, Alejandro H. Toselli

12:00 – 12:15

Graph-based Aggregation for Writer Identification and Retrieval

Oliver Pilizar, Marco Peer, Florian Kleber

12:15 – 12:30

Reading Order Inference for Complex Document Layouts

Iddo Hakim, Sharva Gogawale, Omer Ventura, Gal Grudka, Daria Vasyutinsky-Shapira, Berat Kurar-Barakat, Nachum Dershowitz

HIP

Location: EI8

11:00 - 11:20

DIVA-UTP: A Pixel-Level Region Labeling Dataset for Complex Layout Analysis of Medieval Manuscripts

Najoua Rahal, Rolf Ingold

11:20 - 11:40

A Diachronic Multi-Type Dataset for Document Layout Analysis from the 17th Century to the Present

Juliette Janès, Sarah Bénéière, Benjamin Kiessling, Lucence Ing, Eric Astier, Simon Gabay, Benoît Sagot, Thibault Clérice

11:40 - 12:00

Generalization of Text Line Segmentation for HTR in Historical Documents

Gayana Pathirage, Stephan Unter, Simon Corbillé, Elisa Barney Smith

12:00 - 12:20

Consistent Line Extraction on Medieval Charters: Mask R-CNN and U-Net Architectures.

Nicolas Renet, Angelos Nikolaou, Georg Vogeler

DALL

Location: EI9

11:00 – 11:20

Neuro-Symbolic Framework for Verified Legal Inference of RAG-based AI from Law Document Analyzing

Dung Dang, Vuong Pham, Hung Nguyen, Nomi Baruah, Surajit Dutta, Debashish Gogoi, Hien Nguyen

11:20 – 11:40

KSE-Web: An Analysis of Hybrid Retrieval and LLM-Assisted Query Expansion for Low-Resource Khmer Semantic Search

Nimol Thuon, Panhapin Theang, Ratana Thuon

11:40 – 12:00

Semi-Supervised Learning using Ensembles based Pseudo-Labeling for Handwritten Text Recognition

Simon Corbillé, Thierry Paquet, Pierrick Tranouez

12:00 – 12:20

Synthetic Line Image Generation from Cropped Grapheme Images for Handwritten Marathi Text Recognition

Janhavi Vaishampayan, Kaushik Gopalan, Shagun Dwivedi

12:20 – 12:30 **Closing**

WML

Location: EI10

11:00 – 11:45 **Keynote Talk**

11:45 - 12:05

Tokenization vs. Augmentation: A Systematic Study of Writer Variance in IMU-Based Online Handwriting Recognition

Jindong Li, Dario Zanca, Vincent Christlein, Tim Hamann, Jens Barth, Peter Kämpf, Björn Eskofier

12:05 - 12:25

Comparing Vision Transformers and CNNs for Sex Classification from Handwriting

Andrea Morelli, Michela Baccini

Afternoon Session 1 (13:30 – 15:00)

DAS • Poster Session

Location: Hallway

P.1.1

Criticality-Aware Evaluation of Table Understanding Systems on Synthetic ESG Documents

Giorgia Bertacchini, Marcello Pietri, Nicola Bicocchi, Enrico Rossini, Marco Mamei

P.1.2

A Feasibility Study of Concept-Based XAI Classification Methods on Chart Images

Josh Knize, Kenny Davila, Min Song

P.1.3

Text Pixels Are Not Enough: A Controlled Study of Canvas Sensitivity in VLM-Based OCR

Daichi Haraguchi

P.1.4

BZK-LE: A Dataset for Zero-Shot and Few-Shot Archival Document Layout Classification

Marzieh Malekzadeh Mahani, Genet Asefa Gesese, Mahsa Vafaie, Inger Banse-Strobel, Kevin Dubout, Harald Sack

P.1.5

Towards Zero-Shot Domain Generalization for ID Cards Presentation Attack Detection

Mario Nieto, Juan Manuel Espin, Juan Tapia

P.1.6

PeopleGator: A Multimodal Benchmark for Person Identification in Historical Documents

Marek Vako, Martin Kiss, Adam Herout, Michal Hradiš

P.1.7

Scale-Aware Feature-Level Encoding for Historical Document Stamp Detection

Mohamed Chelali, Florence Cloppet, Camille Kurtz, Isabelle Bloch, Daniel Foliard

P.1.8

TrOCR for Medieval HTR: A Systematic Ablation Study with Cross-Dataset Validation

Sachin Sharma, Michele Flammini, Federico Simonetta

P.1.9

DocScribe: Handwriting-Aware Document Templating with Compact OCR-Free Models

Marcel Lamott, Nauman Riaz, Sheraz Ahmed, Andreas Dengel, Ulrich Schwanecke, Adrian Ulges

P.1.10

Evaluating multimodal large language models for end-to-end diagram recognition

Moritz Burmester, Jonas Scharfenberger, Burkhardt Funk

P.1.11

Lect-Former: Localized Self-Attention Skip Connections for Whiteboard Content Extraction

Kenny Davila, Fei Xu, Sunyoung Ro, Min Song

HIP

Location: E18

13:30 – 13:50

Postcorrecting OCR with LLMs: a low resource approach

Valentina Vavassori, Harry Lloyd

13:50 – 14:10

Reconstruction Error Ratios for Prototype-Anchored Unsupervised Learning in Optical Character Recognition

Tim Hallyburton, Anna Scius-Bertrand, Arthur Neto, Andreas Fischer, Gernot Fink

14:10 – 14:30

Cross-View Retrieval of Byzantine Monograms with Contrastive Representation Learning

Saranga Mahanta, Victoria Eyharabide, Béatrice Caseau, Isabelle Bloch

14:30 – 14:50

Diagram Recognition for Byzantine Manuscripts Using a Pretrained nnU-Net

Lilly Osburg, Germaine Götzelmann, Felix Kraus, Danah Tonne

ADAPDA

Location: E19

13:30 **Welcome**

14:00 – 14:20

Page-Level Synthetic Russian Handwriting via Decomposed Diffusion

Maksim Golyadkin, Innokentiy Humonen, Ilya Makarov

14:20 – 14:40

Test-Time Adaptation of Frozen Scene-Text Recognizers to Documents and Handwriting

Kunal Purkayastha, Josep Lladós, Umapada Pal

14:40 – 15:00

Age-Adaptive Handwriting Reconstruction from an IMU-Based Digital Pen through Shared Representations and Domain-Specific Heads

Florent Imbert, Yann Soullard, Eric Anquetil, Hui Han

WML

Location: EI10

13:30 – 13:50

Resolving Visual Ambiguity in Dense Layouts via Magnitude Pruning and Semantic Rectification

Amisha Mishra, Mayurakshi Chaudhuri, Chiranjoy Chattopadhyay, Shrimoyee Guha Thakurta

13:50 – 14:10

Cross-Dataset Transfer in Open-Set Writer Identification: From Hand-Drawn Circles to Handwritten Pages

Md Raihan, Thomas Gorges, Lukas Hüttner, Vincent Christlein

14:10 – 14:30

A Spatio-Temporal Graph Refinement Framework for Robust Video Text Tracking

Anant Sinha, Shivakumara Palaiahnakote, Umapada Pal, Mo Saraee

14:30 – 14:50

An Efficient Weighted Semi-Supervised End-to-End Approach for Scene Text Spotting

Abhinav Shukla, Shivakumara Palaiahnakote, Sukalpa Chanda, Umapada Pal, Daniel P. Lopresti

14:50 – 15:10

Incremental Document Image Classification: A Comparative Study of Content and Layout Representations

S. M. Abdul Ahad, Thomas Gorges, Lukas Hüttner, Linda-Sophie Schneider, Mathias Seuret, Fei Wu, Vincent Christlein

15:10 – 15:30 **Closing Ceremony**

Afternoon Session 2 (15:30 – 17:00)

DAS • Demo Session

Location: Hallway + E17

D.1 CrimeNER Demo: Named-Entity Recognition in the Crime Domain

Miguel Lopez-Duran, Julian Fierrez, Marcos Grande, Aythami Morales, Daniel DeAlcala, Gonzalo Mancera, Javier Irigoyen, Ruben Tolosana, Oscar Delgado, Francisco Jurado, Alvaro Ortigosa

D.2 AlriskEval-edu Demo: Explainable Auditing of Pedagogical Risks in Educational Explanations

Javier Irigoyen, Roberto Daza, Francisco Jurado, Julian Fierrez, Ruben Tolosana, Alvaro Ortigosa, Miguel Lopez-Duran, Aythami Morales

D.3 EpiSAM Demo: An Interactive System for Character Segmentation and Text-Line Extraction in Challenging Stone Inscriptions

Arnav Sharma, Pratyush Jena, Amal Joseph, Ravi Kiran Sarvadevabhatla

D.4 Multi Script OCR System for Handwritten Indic Manuscripts

Tathagata Ghosh, Sai Madhusudan Gunda, Simran Singh Sandral, Ravi Kiran Sarvadevabhatla

D.5 Arkindex: An Open-Source Infrastructure for AI-Based Processing of Cultural Heritage Collections

Christopher Kermorvant, Erwan Rouchet, Bastien Abadie

16:15 – 17:00 **Discussion Group**

HIP

Location: E18

15:30 – 15:50

Reading Order Inference for Complex Document Layouts

Iddo Hakim, Sharva Gogawale, Omer Ventura, Gal Grudka, Daria Vasyutinsky-Shapira, Berat Kurar-Barakat, Nachum Dershowitz

15:50 – 16:10

Towards Hierarchical Structure Understanding of Newspaper Images

William Mocaër, Solène Tarride, Thomas Constum, Merveilles Agbeti-Messan, Tom Simon, Clément Chatelain, Stéphane Nicolas, Pierrick Tranouez, Sébastien Cretin, Thierry Paquet

16:10 – 16:30

Mapping Armenian Paris: Extracting and Geocoding Commercial Advertisements from the 20th-Century Diaspora Press

Chahan Vidal-Gorène, Seda Kirakosyan, Edita Matevosyan

16:30 – 16:50

A Prompt Optimization Framework for Parsing Historical Addresses

Amel Gader, Rafael Patronilo, Mahsa Vafaie, Genet-Asefa Gesese, Harald Sack

16:50 – 17:00 **Closing**

ADAPDA

Location: EI9

15:30 – 15:50

When Low CER is Not Enough: An Analysis of Hallucinations in Vision-Language OCR Systems on Historical Uruguayan Documents

Marina Gardella, Camilo Mariño, Diego Belzarena, Ignacio Ramírez, Gregory Randall, Jean-Michel Morel

15:50 – 16:10

Comparative Study of Domain-adapted VLMs for General Document Visual Question Answering

Miguel Lopez-Duran, Elena Marrero, Julian Fierrez, Marta Robledo-Moreno, Ruben Vera-Garcia, Daniel DeAlcala, Aythami Morales, Ruben Tolosana, Oscar Delgado, Alvaro Ortigosa, Javier Ortega-Garcia

16:10 – 16:50 **Discussion Session**

16:50 – 17:00 **Closing**

Friday • September 4

Morning Session 3 (09:00 – 10:30)

DAS

Location: EI7

09:00 – 09:45 **Keynote: Brandon Smock**

09:45 – 10:00

Nutri-TabNet: Detection of Non-planar Nutrition Information Panels in Real-World Food Packaging

Shiza Maham, Abdullah Tariq, Claire Pulker, Lauren Blekkenhorst, Martin Masek, Faisal Shafait, Syed Zulqarnain Gilani

10:00 – 10:15

CollabScore OMR: A Hybrid System for Music Score Recognition

Aur lie Lemaitre, Bertrand Cou asnon, Philippe Rigaux

10:15 – 10:30

IndicPage-OCR: Robust Low-Resource Adaptation for Multi-Script Indic Handwritten Page Recognition

Shaon Bhattacharyya, Ajoy Mondal, C. V. Jawahar

IWCP

Location: EI8

09:00 – 09:15 **Welcome**

09:15 – 09:40

Unsupervised Domain Adaptation for Symbol Spotting in Historical Encrypted Manuscripts

Giuseppe De Gregorio, Alicia Forn s, Lei Kang, Be ta Megyesi

09:40 – 10:05

Evaluating Hierarchy-Aware Deep Learning for the Recognition of Tironian Notes

Yule Kang, Thomas Gorges, Janne van der Loop, Franziska Marske, Nikolaus Weichselbaumer, Tino Licht, Vincent Christlein

10:05 – 10:30

ANACD: Synthetic Data for Neo-Assyrian Cuneiform Sign Detection via Physically Plausible Diffusion Conditioning

George Argirov, Innokentiy Humonen, Maksim Golyadkin, Ilya Makarov

VINALDO

Location: EI9

09:00 – 09:15 **Welcome**

09:15 – 09:45

Ontology-grounded Chunking for Parser-Independent Retrieval

Yeonjee Han, Jaemin Na, Rock Sakong

9:45 – 10:15

Accurate Text Recognition with Character Position Prediction via Dual Decoders

Ruifeng Li, Liangrui Peng, Changsong Liu

10:15 – 10:30 **Discussion**

WORMS

Location: EI10

09:00 – 09:05 **Welcome**

09:05 – 09:45

Keynote: Towards Computational Piano Rehearsal Analysis: The Rach3 Project

Carlos Eduardo Cancino-Chacón

09:45 – 09:55

Evaluating General-Purpose Vision-Language Models on Music Score Images

Noelia Luna-Barahona, Antonio Ríos-Vila, Andrea García Torres, José M. Iñesta, Jorge Calvo-Zaragoza, David Rizo

09:55 – 10:05

Impact of BPE-based Melodic Units in Optical Music Recognition

Aitana Menárguez-Box, Enrique Vidal, Alejandro H. Toselli

10:05 – 10:15

Towards a foundational model for recognising diastematic Gregorian notation

Daniel Kurek, Jan Hajič, Jr.

10:15 – 10:25

MuSViT: A Foundation Vision Model for Sheet Music Representation

Carlos Pérez-Sancho, Alejandro Galan-Cuenca, Juan P. Martínez-Esteso, Francisco J. Castellanos and Antonio Javier Gallego

Morning Session 4 (11:00 – 12:30)

DAS • Poster Session

Location: EI7

P.2.1

Rethinking Text Color Estimation in Screen Documents: A Reliability-Aware Approach

Liang Zhang, Zhenfeng Lan, Huajian Ni, Pei Guo, Qinjie Hu, Jie Liu, Zhongxin Wang

P.2.2

Decoupled Prototype Memory Mechanism for Complex Table Recognition

Di Wu, Liangrui Peng, Tianqi Zhao

P.2.3

A Recognition-Aware Multimodal Framework for Actual and Intended Text Transcription of Child Handwriting

Sumi Suresh M. S., Sahana Rangasrinivasan, Abbie Olszewski, Srirangaraj Setlur, Venu Govindaraju

P.2.4

Decoding Romanian Law: A Fine-Tuned LLM Approach

Marinela Adam, Alexandru Berindeie, Darian Onchis, Todor Ivascu, Corina Negrutiu

P.2.5

SciXplain: Structure-Aware Representation Learning for Scientific Method Diagrams

Zhicheng He, Zheyi Deng, Junhao Ren, Xiaoqing Lyu

P.2.6

Text-Based Classification of Historical Notarial Documents from Noisy HTR Transcripts

Daniela Perdomo, Veronica Romero, Miriam Gil

P.2.7

Handwritten Devanagari Character Recognition: Conjunct Consonants Dataset Creation and Benchmarking with Transfer Learning

Sudip Tiwari, Satyasa Khadka, Sandhya Baral, Sanjeeb Prasad Panday

P.2.8

Mind the Gaps: A Curated Benchmark for Form Field Detection

Iheb Brini, Omar Moured, Hamza Gbada, Elisa Barney Smith

P.2.9

From Pen Strokes to Sleep States: Detecting Low-Recovery Days Using Sigma-Lognormal Handwriting Features

Short Paper

Chisa Tanaka, Andrew Vargo, Anna Scius-Bertrand, Andreas Fischer, Koichi Kise

P.2.10

Reconstructing Historical Manuscripts through MSI: The Potential of Contrast in Assessing Image Quality and Legibility

Short Paper

Anna Breger

P.2.11

YOTO - You Only Train Once: Template-Specific Table Extraction from a Single Example

Short Paper

Gaspar Deloin, Aurélie Joseph, Vincent Poulain d'Andecy

P.2.12

Recall-First Stroke Grouping: A 53K-Parameter Symbol-Candidate Proposal Stage for Graph-Based Online Handwritten Mathematical Expression Recognition

Short Paper

Eunpyo Lee

P.2.13

French PDF-to-Markdown Benchmark

Short Paper

Bruno Rigal, Victor Dupriez, Alexis Mignon, Ronan Le Hy, Nicolas Mery

IWCP (Poster Session)

Location: E18 + Hallway

11:00 – 11:40 **Flash Presentation of the Posters**

11:40 – 12:30 **Poster Session**

P.2.14

Automatic Transcription of Brazilian Manuscripts: a preliminary study

Mauricio Maia Vinhas de Azevedo, Carlos Henrique Marcondes

P.2.15

A Jibo-Labeled Classical Hiragana Image Dataset for Computational Paleography: Bootstrapped Variant Annotation at Scale

Michael R. Zeng, Herman Chau, Paul S. Atkins

P.2.16

Automating annotation and classification of palaeographic types in ancient Sicilian Inscriptions

Simona Stoyanova, Miguel Arana-Catania

P.2.17

What transcription approach HTR should learn? HTR, Normalization, and Editorial Tradition in Medieval Manuscript Studies

Anna Michalcová, Jan Odstrčilík

P.2.18

Computational Palaeography and Early Modern Italian Scripts: Methods and Workflows for Letterform Analysis

Serena Carlamaria Crespi, Ambra Stefanello, Carlos-Emiliano González-Gallardo

P.2.19

Bluepond: A Local-First, Browser-Based OCR/HTR Workflow for Historical Manuscripts

So Miyagawa

P.2.20

Visual Explainability for Writer Identification through Connected Component Similarity Mapping

Konstantinos Palaiologos, Basilis Gatos

P.2.21

Transformer-based OCR and LLM Post-Correction for Greek Cursive Papyri

Elena Chepel, Anton Repushko

VINALDO

Location: E19

11:00 – 11:30

Letter Lemmatization: One-to-one and Banded RNNs for Reversing Character-Set Simplification and Abbreviation in Medieval Text

Angelos Nikolaou, Maria Pia Iseo, Tamás Kovács, Nicolas Rennet, Georg Vogeler

11:30 – 12:15 **Panel Discussion**

12:15 – 12:30 **Closing**

WORMS

Location: EI10

11:00 – 11:10

The Music Notation Graph (MuNG) ecosystem 2.0: ready for production

Jiří Mayer, Vojtěch Dvořák, Martina Dvořáková, Markéta Herzánová Vlková, Jan Hajič, Jr.

11:10 – 11:20

Completing Music Notation Datasets for Multimodal Learning: Conversion Pipelines and Fusion Baselines

Antonio Hidalgo-Centeno, Daniel Espinosa-Canovas, Josmar Suarez, Alejandro Galan-Cuenca, Antonio Javier Gallego

11:20 – 11:30

The Optical Music Recognition Bibliography 2.0 (non-archival track)

Alexander Pacha

11:40 – 12:30

Community Discussion I

Afternoon Session 3 (13:30 – 15:00)

DAS

Location: EI7

13:30 – 14:15 **Keynote: Sheraz Ahmed**

14:15 – 14:30

Edge-Aware Relation Prediction for Symbol-Level Structure Refinement in Online Handwritten Mathematical Expressions

Thanh-Nghia Truong, Hung Tuan Nguyen, Cuong Tuan Nguyen, Masaki Nakagawa

14:30 – 14:45

End-to-End Camera-Based Analysis of Egyptian Hieroglyphic Textual Artifacts

David Konieczny, Said Yasin, Torsten Zesch

14:45 – 15:00 **Closing and Awards**

13:30 – 13:55

A Hybrid Deep Learning Pipeline for Text Recognition on Greek Squeezes

Monika Radhakisan Chavan, Pratiksha Manik Patil, Vineetha Mulamoottil Mathews, Thomas Gorges, Janne van der Loop, Vincent Christlein

13:55 – 14:20

Reading the Spiral: Deep Feature Clustering for Exploratory Script Typology in Aramaic Incantation Bowls

Nour Atamni, Boraq Madi, Said Naamneh, Jihad El-Sana

14:20 – 14:45

Timeline Embedding for Papy-Date, a new Dataset of Greek Papyri dated to the Year

Dinh Dung Van, Rodrigo Cerqueira Gonzalez Pena, Lavinia Ferretti, Giuseppe De Gregorio, Olga Serbaeva Saraogi, Pedro Garcia-Baro, Andreas Fischer, Isabelle Marthot-Santaniello

14:45 – 15:05

Evidence-Aware Recognition of Paleo-Hebrew Seal Inscriptions: Structured Sign Hypotheses, Visual Hints, and Deployable Review

Alex Gorbulev, Innokentiy Humonen, Maksim Golyadkin, Ilya Makarov

13:30 – 13:40

Optical Music Generation: Leveraging Off-The-Shelf Diffusion Models to Synthesise Handwritten Scores

Pau Torras, Vittorio Pippi, Fabio Quattrini, Gerard Asbert, Silvia Cascianelli, Rita Cucchiara, and Alicia Fornés

13:40 – 13:50

Data-Driven Glyph Positioning for Handwritten Music Synthesis

Elona Shatri, Jiří Mayer, and Jan Hajič jr.

13:50 – 14:00

One-shot Handwritten Music Generation with Spiking Neural Networks

Aksel Serret-Llopis, Marc Pomar-Pallarès, Gerard Asbert, Pau Torras, Alicia Fornés, and Xavier Otazu

14:10 – 14:20

DEMUN: Fast and Accurate Discovery of Music Notation in Very Large Collections

Vojtěch Dvořák, Filip Bím, Jiří Mayer, Martina Dvořáková, Markéta Herzánová Vlková, Pavel Pecina, Petr Žabička, and Jan Hajič, Jr.

14:20 – 14:30

PixelSmart: An AI-Powered Pixel-Level Annotation Tool for Active Learning in Music Document Analysis

Carlos Pérez-Sancho, Alejandro Galan-Cuenca, Juan P. Martinez-Esteso, Francisco J. Castellanos, and Antonio Javier Gallego

14:30 – 14:40

A Low-Resource Template-Matching Baseline for Handwritten Pianoform OMR (non-archival track)

Igor Sok, Matija Marolt, Nejc Sukljan, and Matevž Pesek

14:40 – 14:50

Interpreting and Encoding the Melodic Variability: The Case of Latvian Traditional Recitatives (non-archival track)

Davis Engelis and Ieva Weaver

Afternoon Session 4 (15:30 – 17:00)

IWCP

Location: EI8

15:30 – 15:55

Within the Hands of Amitāyus: Line-Level Self-Supervised Encoders for Intra-Style Scribal Identification

Colin Brisson

15:55 – 16:20

TextileNet: Towards Zero-shot Text-style Segmentation of Manuscripts

Angelos Nikolaou, Antonella Ambrosio, Desirée Di Donato, Georg Vogeler

16:20 – 17:00 **General Discussion and Closing**

WORMS

Location: EI10

15:30 – 15:45

A Novel Dataset and Framework for Schenkerian Optical Music Recognition (non-archival track)

Stephen Hahn, Yixuan Huang, Nathan Lowe, Isis Cabrera, Simon Mak, Yue Jiang, and Cynthia Rudin

15:45 – 17:00

Community Discussion II

Sponsors

We gratefully acknowledge the generous support of our sponsors and partners.

Platinum Sponsor



Silver Sponsor



Other Sponsors and Partners



SPRINGER NATURE



 **Get FlowBuddy**



Platinum Sponsor



Snowflake is the platform for the AI era. More than 13,900 customers around the globe, including hundreds of the world's largest companies, use Snowflake's AI Data Cloud to build, use, and share data, applications, and AI. Our mission: empower every enterprise to achieve its full potential through data and AI. And that mission starts with you.

This is where cutting-edge AI meets enterprise scale. You'll work alongside world-class researchers and engineers to build the infrastructure, models, and intelligent products shaping the future of AI.

Learn more at <https://careers.snowflake.com/ai-research>.



MEET THE AI RESEARCH TEAM

WHERE CUTTING-EDGE RESEARCH MEETS GLOBAL IMPACT.

 **WARSAW**



➤ WHO WE'RE LOOKING FOR

- **PhD Candidates:** computer science, data science, or related AI fields
- **Postdoctoral Researchers:** passionate about advancing AI
- **Experienced AI Engineers:** with expertise in ML and distributed systems

➤ KEY FOCUS AREAS

- AI Products
- Foundation Models
- AI Infrastructure
- Distributed Systems

➤ STOP BY OUR BOOTH

- Meet the team and learn how we're transforming AI research into intelligent products that power the AI Data Cloud.

CONNECT WITH US



@_snowflake_inc



@snowflake



@snowflakedevelopers

WHERE DATA DOES MORE





LET'S BUILD THE NEXT BIG THING TOGETHER.

careers.snowflake.com

Joining Snowflake means joining our mission: to empower every enterprise to achieve its full potential through data and AI. And that mission starts with you.

Snowflake's AI Research team explores the frontier of artificial intelligence, conducting open, foundational, and applied research to turn big questions into bigger breakthroughs. This is where research across agents, post-training, and AI training and inference systems moves beyond theory and into real-world impact. Where ideas become easy, connected, and trusted AI features that power our platform and inspire innovation.

Join our team and expand what's possible in enterprise AI — at Snowflake and beyond.

STAY CONNECTED

Scan the QR code to unlock opportunities at Snowflake



GLOBAL ENGINEERING HUBS

- Menlo Park, California
- Berlin, Germany
- Bellevue, Washington
- Warsaw, Poland
- Toronto, Canada
- San José, Costa Rica

WHERE DATA DOES MORE



Notes

Notes

Notes

ICDAR Vienna Dictionary

Schlawina	pfiffiger, gerissener, lebhafter Mensch (a sharp, shrewd, lively person)
Hallodri	sorgloser Mensch (Carefree person)
Hawara	guter Freund, Kamerad, Kumpel (good friend, buddy, pal)
Lauser	liebvolle, scherzhafte Bezeichnung für einen frechen, vorlauten oder zu Streichen aufgelegten Menschen (an affectionate, playful term for a cheeky, sassy, or mischievous person)
Bazi	Schlitzohr, ein durchtriebener, aber meist sympathischer Lausbub (Sly Fox, a crafty but mostly likeable rascal)
Jo eh	Typisch österreichischer Ausdruck, der oft als Füllwort dient. Er drückt meist eine beiläufige Bestätigung aus, kann aber auch Ironie oder Desinteresse signalisieren (a typical Austrian expression that is often used as a filler. It usually conveys a casual acknowledgment, but can also signal irony or disinterest)
Moch ma scho	Das kriegen wir schon hin; Keine Sorge, das wird geregelt (we'll figure it out; don't worry, it'll be taken care of)
Es is wie's is	Bedingungslose Akzeptanz; es bedeutet, dass man eine gegenwärtige Situation oder Tatsache als unveränderlich anerkennt und aufhört, sich darüber zu ärgern (unconditional acceptance; it means acknowledging a current situation or fact as unchangeable and stopping getting upset about it)
Wennst manst	Ich stehe natürlich hinter dir, wenn du dich mit Anlauf ins Unglück stürzen willst (Of course, I'll stand by you if you want to run headlong into disaster)
Nur ned hudln	Nur nichts überstürzen; nur nicht zu eilig (Don't panic; just don't rush into anything)

Discover Vienna

Discover Vienna during ICDAR 2026, starting at Karlsplatz, next to the conference venue.

Karlsplatz and Karlskirche

Begin at Karlsplatz, an important cultural and transport hub surrounded by museums, historic architecture and green spaces. Walk through Resselpark and enjoy the view across the reflecting pool towards Karlskirche. With its monumental green dome and two decorated columns, this magnificent Baroque church is one of Vienna's most distinctive landmarks.

Vienna State Opera

Walk towards the Ringstrasse and the Vienna State Opera. Opened in the nineteenth century, it is one of the world's leading opera houses and an important symbol of Vienna's musical heritage. Guided tours are available, although schedules and availability should be checked in advance.

Hofburg and Heldenplatz

Continue through the Burggarten to the Hofburg, the former imperial residence of the Habsburg dynasty. Explore its grand courtyards and walk to Heldenplatz, one of Vienna's most impressive public spaces. The palace complex reflects several centuries of Austrian and European history.

Stephansplatz and St. Stephen's Cathedral

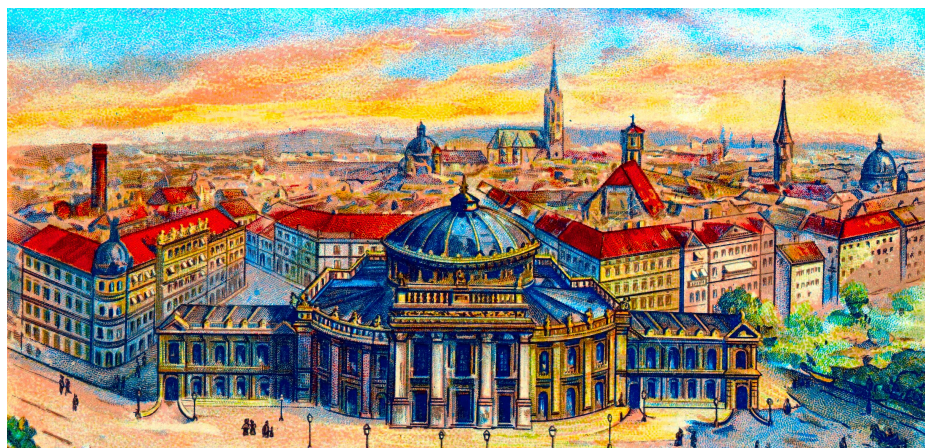
Walk through Vienna's historic centre to Stephansplatz, one of the liveliest squares in the city. At its centre stands St. Stephen's Cathedral, Vienna's most recognisable landmark, known for its Gothic architecture, colourful tiled roof and prominent south tower.

Prater and the Giant Ferris Wheel

Finish your visit at the Prater, Vienna's famous public park and amusement area. Its best-known attraction is the historic Giant Ferris Wheel, which offers panoramic views across the city. The Prater can be reached easily from the city centre by underground and is a pleasant destination for an evening walk after the conference.

Viennese Cuisine

No visit to Vienna is complete without trying some of its local food. Popular dishes include **Wiener Schnitzel**, **Tafelspitz** and **Schwammerlgulasch** (vegetarian). **Apfelstrudel**, **Kaiserschmarrn** and **Sachertorte** are perfect for dessert. Vienna is also famous for its traditional **coffee houses**, which are great places to relax, enjoy a coffee and pastry, and chat after a busy conference day.



Make the most of your stay during ICDAR 2026
and discover Vienna beyond the conference venue.

ivie City Guide



Download Vienna's free city-guide app for maps, themed walks, information about famous sights and tips for discovering hidden corners of the city.

wien.info/en/travel-info/ivie-app

Vienna's Neighborhoods



Explore Vienna's diverse neighborhoods, known locally as *Grätzl*, and discover markets, cafés, local shops and places away from the usual tourist routes.

wien.info/en/see-do/neighborhoods

Scan a QR code and start exploring!

Practical Information

Venue	TU Wien, Campus Gusshaus, Gusshausstrasse 27, 1040 Wien
Registration	Open daily from 08:00
Lunch	Served daily at Campus Gusshaus
Emergency	Contact the registration desk
Wi-Fi	tunet, tunetguest
...username	uname
...password	pwd

Credits



Booklet Design and Layout

Marco Peer

Logo and Booklet Review

Florian Kleber

Image Credits

Vienna motifs from Wien Museum (free admission, located 5min from the conference venue)

Welcome Reception: Ottakringer Brewery

Conference Banquet: Gryffindor via Wikimedia Commons

Acknowledgements

We thank all members of the committees, reviewers, volunteers, sponsors and contributors that made ICDAR 2026 possible.

Copyright

© 2026 ICDAR 2026 Organizing Committee. All rights reserved.



<https://icdar2026.org>