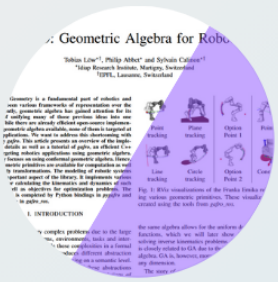
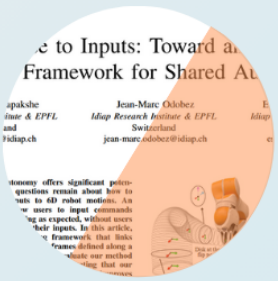
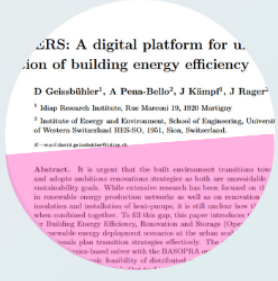
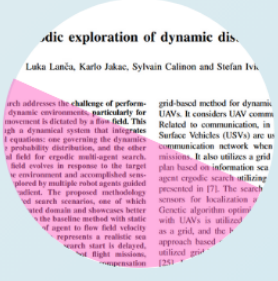
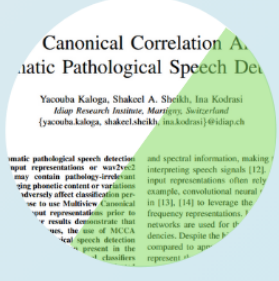
IDIAP RESEARCH INSTITUTE

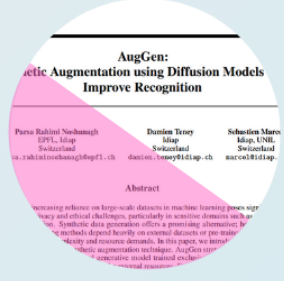


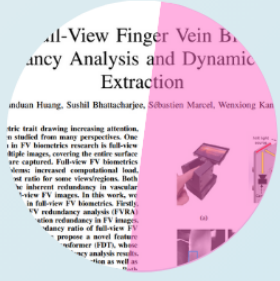


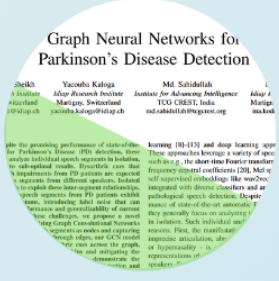


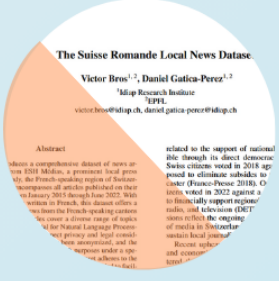


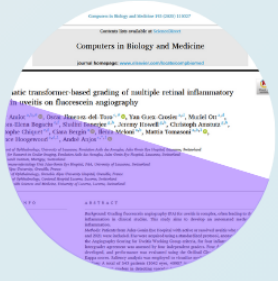


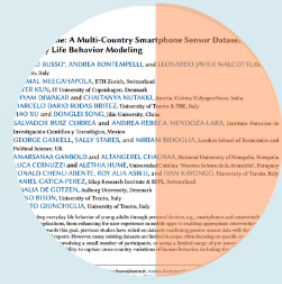


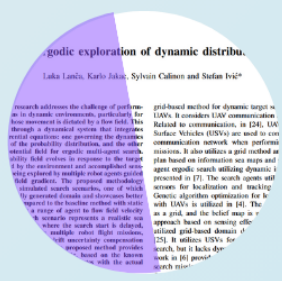


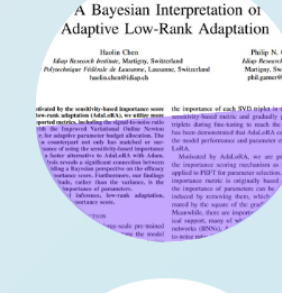


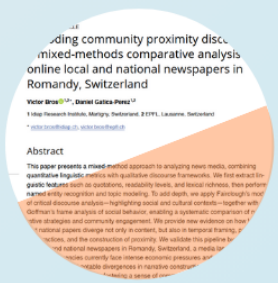


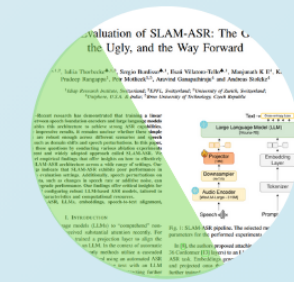












Scientific Report 2025

Contents

Introduction

Founded in 1991, the Idiap Research Institute is an independent, non-profit research foundation based in Martigny, Switzerland. Idiap is a Research Institute of National Importance, established by the State of Valais, the City of Martigny, the École Polytechnique Fédérale de Lausanne (EPFL), the University of Geneva, and Swisscom. Over more than three decades, Idiap has built extensive expertise in Artificial Intelligence (AI), producing a rich legacy of high-impact publications, widely used software, and curated datasets. This accumulated experience positions Idiap as a leading institution in AI research, with unique capabilities and a diverse technology portfolio that distinguishes it from more recent actors.

Idiap plays a central role in supporting Switzerland's societal and industrial needs through AI. Its research addresses pressing challenges, ranging from sustainable development and healthcare to secure digital infrastructures, while fostering innovation that strengthens Switzerland's competitiveness. By combining fundamental research with applied solutions, Idiap ensures that advances in AI are both scientifically robust and socially beneficial.

To achieve these objectives, Idiap organizes its research into four complementary Research Programs:

1. Human-AI Teaming: Designing collaborative AI systems that enhance human capabilities rather than replace them, enabling effective cooperation between humans and AI.
2. Sustainable and Resilient Societies: Developing AI technologies that promote environmental sustainability, social resilience, and security, such as models for climate adaptation, or technologies for secure societies and for combating disinformation.
3. AI for Life: Applying AI to life sciences and healthcare to improve diagnostics, treatment, and understanding of biological systems.
4. AI for Everyone: Ensuring AI is equitable, inclusive, accessible, transparent, and beneficial for all, supporting broad participation in and access to AI-driven technologies.

A transversal AI Fundamentals pillar supports these programs by advancing core AI technologies, such as machine learning, deep learning and foundation models, which are applied across all Research Programs to maximize scientific impact and societal relevance.

The selection of studies presented in this report provides an overview of Idiap's research activities carried out during the year and highlights the outcomes of collective research efforts.

These achievements are complemented by details of externally funded projects, publicly available code and data shared with the broader research community, as well as a list of published research articles.

In 2025, Idiap ran 84 externally funded projects. In addition, the Institute is active in technology transfer, collaborates with a wide range of companies, and creates and supports start-ups through its incubator, IdeArk S.A.

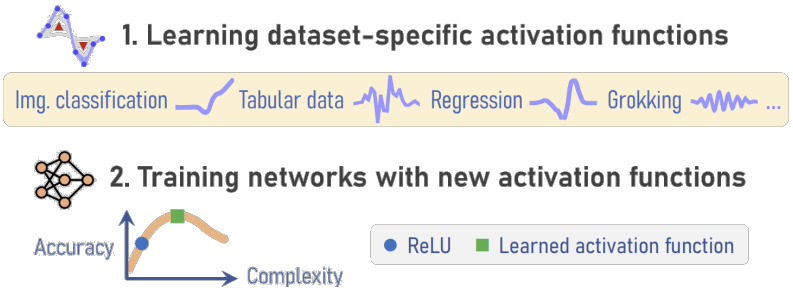
Idiap also pursues an active policy in support of reproducible science, with the release of 56 open-source software libraries in 2025. The Institute publishes data through FAIR (Findable, Accessible, Interoperable, Reusable) repositories. In 2025, the Idiap website offered 10 new publicly available datasets, bringing the total to 118.

AI Fundamentals

NEURAL NETWORK ARCHITECTURES AND THEIR INDUCTIVE BIASES

This work explores fundamental questions in machine learning and why particular neural architectures are effective on particular tasks. The findings of this study challenge the widely-held assumption that neural networks possess a *simplicity bias* (a tendency to represent simple functions) that is universally beneficial, as a form of Occam's Razor. We show instead that standard activation functions such as ReLUs are a design choice that is critical to obtain this simplicity bias, which also means that it can be modulated by designing different activation functions. Moreover, while the simplicity bias is beneficial for many classical tasks, e.g., image classification, many domains that are less studied can benefit from very different design choices.

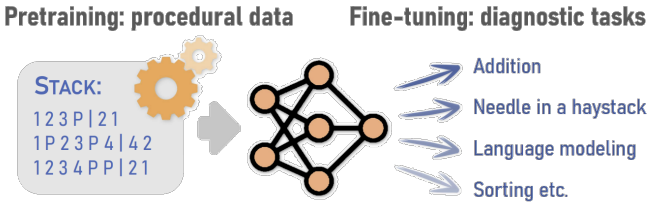
The paper introduces a novel meta-learning strategy to discover new activation functions that optimally modulate a model's inductive biases for a given target task. This method yields a series of new designs that outperform standard architectures on specific tasks, by favoring the models that represent higher-complexity solutions. The effects are clear improvements in generalization in domains such as tabular data and regression tasks, for example. This work highlights that there is no *one-size-fits-all* in neural network design. It opens the door to principled approaches for building architectures with inductive biases tailored to the data and problem at hand.



D. Teney, L. Jiang, F. Gogianu, and E. Abbasnejad, Do We Always Need the Simplicity Bias? Looking for Optimal Inductive Biases in the Wild, *The IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025

EFFICIENT PRETRAINING OF TRANSFORMERS ON ABSTRACT DATA

This work investigates how transformer and language models can acquire reasoning capabilities when trained on synthetic abstract data. We specifically study *procedural data*, which is made of sequences of abstract symbols with precise structure, generated with simple algorithms such as cellular automata and formal languages rather than natural language or code. The paper shows that pretraining on different forms of procedural data can instill distinct computational structures in a transformer model, which improve its reasoning skills, e.g., for memory recall, arithmetic, or sequence manipulation like sorting. Remarkably, these structures are modular: the attention and MLP layers of pretrained models can be recombined across models pretrained on different tasks to produce a single model with several improved capabilities. This demonstrates that reasoning skills can be learned independently of semantic information, opening a new perspective on the design of data-efficient pretraining of language models.

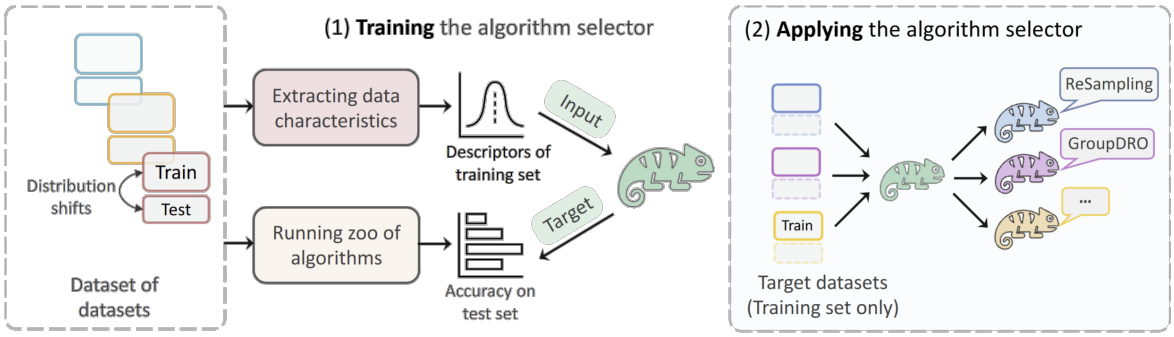


Pretraining task	Example sequence
k-DYCK	([{ }])
k-DYCK SHUFFLE	([{]) }
STACK	1 2 3 P | **2 1**
IDENTITY	1 2 3 | **1 2 3**
SET	1 2 2 | **1 2**

Z. Shinnick, L. Jiang, H. Saratchandran, A. van den Hengel, and D. Teney, Transformers Pretrained on Procedural Data Contain Modular Structures for Algorithmic Reasoning, *ICML Workshop on Methods and Opportunities at Small Scale*, 2025

LEARNING TO HANDLE DISTRIBUTION SHIFTS AND OUT-OF-DISTRIBUTION GENERALIZATION

Machine learning systems often struggle when the data they encounter at test time differs from the data seen during training, a challenge known as distribution shift. This can severely degrade performance and reliability in real-world applications. In this work, we investigate out-of-distribution (OOD) generalization from a new perspective. Instead of developing new algorithms designed to handle particular types of shifts, we ask whether it is possible to learn a better strategy to select among the many existing ones for handling a new given task or dataset. The main contribution of this paper is a method called *OOD-Chameleon*, which is an automatic selector trained on a "dataset-of-datasets" such that it can identify the most suitable method among a variety of existing options (e.g., ERM, GroupDRO, reweighting and debiasing strategies). This approach is a completely new take on the handling of distribution shifts, and it shows great promise on a range of use cases with synthetic, vision, and language data. The automatic selector consistently selects algorithms that outperform any single fixed or heuristics-based selection strategy. This shows that algorithm selection for OOD robustness is itself learnable. This provides a new avenue for deploying machine learning reliably in environments where data distributions are uncertain or evolving.

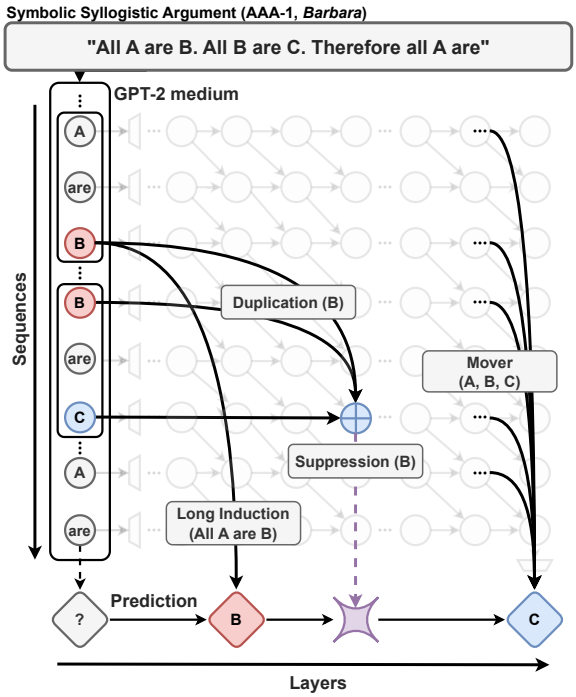


L. Jiang and D. Teney, OOD-Chameleon: Is Algorithm Selection for OOD Generalization Learnable? *Forty-Second International Conference on Machine Learning*, 2025

REASONING CIRCUITS IN LANGUAGE MODELS

This work tackles a central question in current LLM research: when a language model solves a logical problem, is it really implementing a systematic reasoning procedure, or is it just reusing surface patterns from pre-training? We focus on classical categorical syllogisms (e.g., “All A are B. All B are C. Therefore, all A are C”) because they are both ubiquitous in natural language and formally well-understood. This makes them an ideal testbed to separate genuine, content-independent reasoning from mere exploitation of world knowledge or lexical regularities.

The analysis reveals a three-stage reasoning circuit for syllogistic inference: an early phase that naively recites the first premise and initially biases the model toward the wrong conclusion, a suppression phase where duplicated middle-term information is actively damped, and a mediation phase where specialized *mover* attention heads propagate the remaining subject–predicate structure to the output, producing the correct conclusion. This circuit is both sufficient and necessary for all unconditionally valid syllogistic schemes on which models achieve high accuracy, and it generalizes across families of models. At the same time, we observe partial contamination from heads encoding commonsense and contextual knowledge, with this entanglement increasing in larger models. Overall, the findings indicate that language models do develop transferable, content-independent reasoning mechanisms, but these mechanisms are not isolated logical primitives; they are intertwined with the same world knowledge that supports the models’ broader linguistic competence.

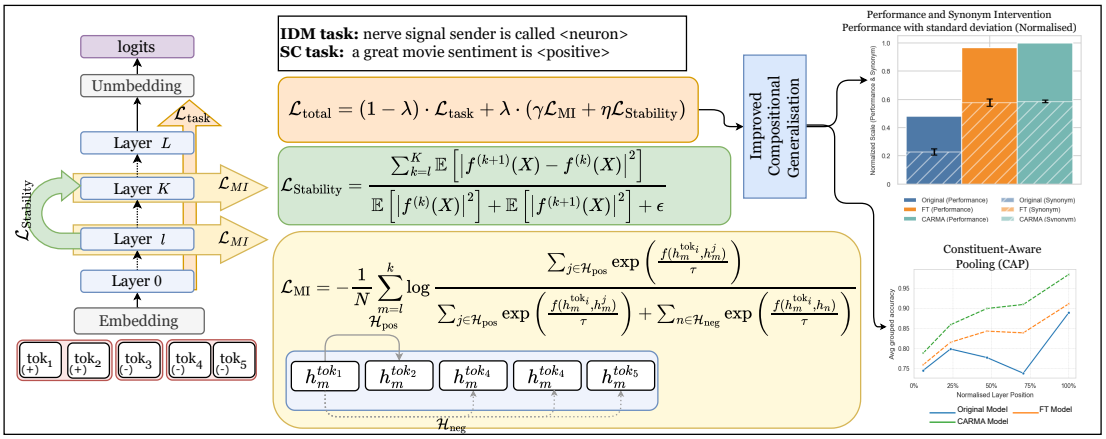


G. Kim, M. Valentino, and A. Freitas, A Mechanistic Interpretation of Syllogistic Reasoning in Auto-Regressive Language Models, *Findings of the ACL*, 2025

ENHANCED COMPOSITIONALITY IN LLMS

In this work, we address a persistent weakness of large language models: their limited *compositional generalization*, i.e., the ability to systematically recombine known pieces of meaning to handle novel inputs. Standard next-token training and fine-tuning tend to prioritize surface correlations and task accuracy, which fragments information across tokens and layers and makes models brittle under small input changes (e.g., synonym substitutions, rephrasings). Our goal is to design a scalable, architecture-agnostic intervention that strengthens the stability of compositional representations *inside* existing language models, without sacrificing downstream performance.

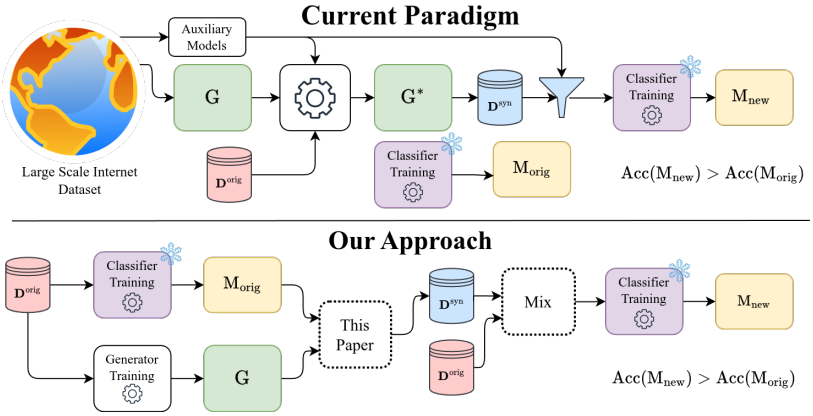
We introduce **CARMA** (Compositionality in LLMs via Advanced Regularization and Mutual Information Alignment), an auxiliary loss that we plug into standard fine-tuning. Our results show that CARMA consistently reduces representational fragmentation, stabilizes token and constituent representations, and improves compositional robustness under perturbations, especially in early and middle layers. Overall, the work demonstrates that we can improve compositional generalization in LLMs through non-intrusive, task-agnostic regularization that reinforces already-learned structures, rather than by redesigning architectures or relying on synthetic data augmentation.



N. Aljaafari, D. Carvalho, and A. Freitas, CARMA: Enhanced Compositionality in LLMs via Advanced Regularisation and Mutual Information Alignment, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025

Synthetic Augmentation using Diffusion Models Can Improve Recognition

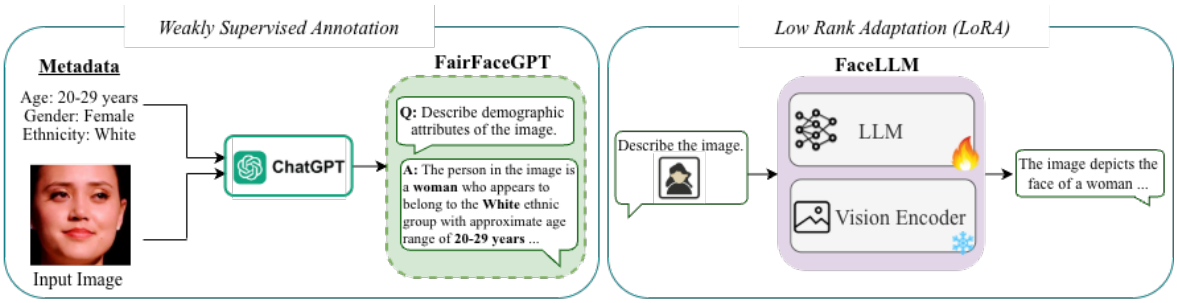
We study how to improve face recognition performance in data limited or privacy sensitive settings and we propose a method, which uses a class-conditional diffusion model trained exclusively on a target face recognition dataset, without relying on any external data or pre-trained generators. The method selectively generates "mixed" synthetic classes by interpolating between identity conditions in the latent space, then augments the original real-data training set with these novel synthetic samples to boost discrimination between identities and improve intra-class compactness. Empirical results across eight standard face recognition benchmarks demonstrate consistent gains over using real data alone and show that augmenting with synthetic classes can match the effect of adding up to 1.7 $\times$ more real images, thus offering a practical path toward enhanced recognition performance under constrained data regimes.



P. Rahimi, D. Teney, S. Marcel, AugGen: Synthetic Augmentation using Diffusion Models Can Improve Recognition, *Neural Information Processing Systems*, 2025.

A Multimodal Large Language Model for Face Understanding

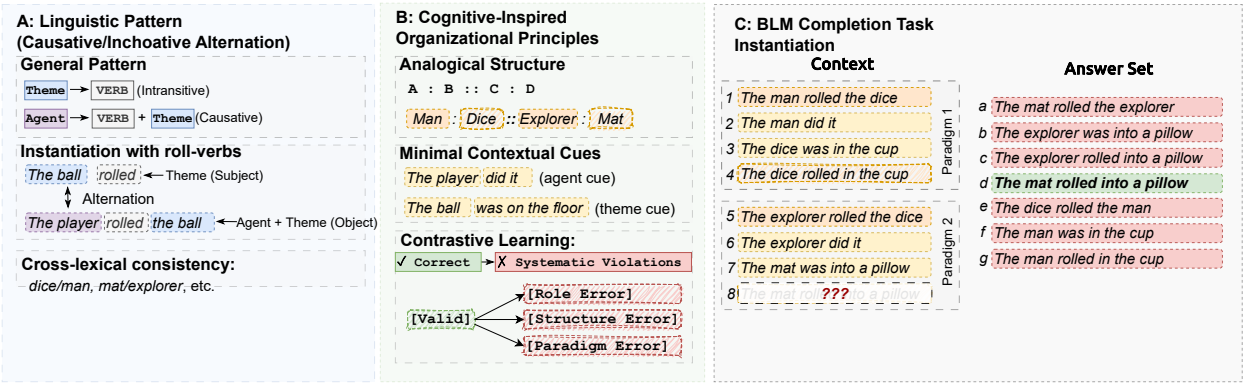
We study the limitation of current multimodal large-language models (MLLMs) in reasoning over the fine-grained structure, expressions, pose and demographic cues present in face images; we propose a domain-specialized solution by fine-tuning a general vision-language model via low-rank adaptation to produce FaceLLM. Our experiments demonstrate that FaceLLM achieves state-of-the-art performance on a wide variety of face-centric tasks (including high- and low-resolution face recognition, attribute prediction, head pose estimation, anti-spoofing and deep-fake detection).



H. Otroshi Shahreza and S. Marcel, FaceLLM: A Multimodal Large Language Model for Face Understanding, *IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, 2025

COGNITIVELY-INSPIRED SAMPLE-EFFICIENT LINGUISTIC RULE INDUCTION

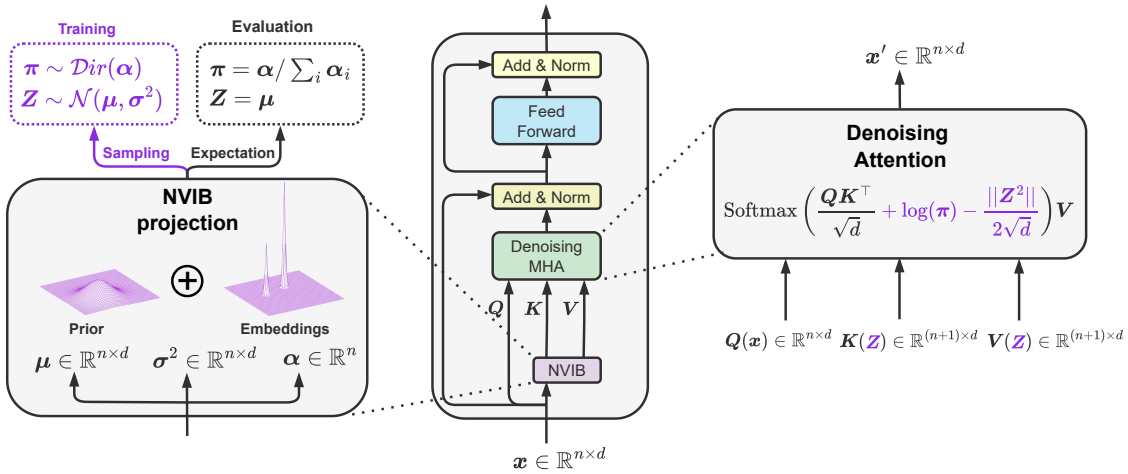
Large language models achieve strong performance through training on vast datasets. We ask: *Can analogical paradigm organization enable lightweight models to match this performance with minimal data?* We develop a computational approach implementing three cognitive-inspired principles: analogical structure, contrastive learning, and minimal contextual cues. We test this approach with structured completion tasks where models identify correct sentence completions from analogical patterns with contrastive alternatives. Ablation studies confirm that analogical organization and contrastive structure improve performance, consistently surpassing randomly shuffled baselines across architectures. Cross-phenomenon validation replicates these efficiency gains, confirming approach robustness. Our results show that analogical paradigm organization enables competitive linguistic rule learning with orders of magnitude less data than conventional approaches require.



C. Jiang and P. Merlo, Analogical Structure, Minimal Contextual Cues and Contrastive Distractors: Input Design for Sample-Efficient Linguistic Rule Induction, *ArXiv cs.CL.2511.10441*

IMPROVING GENERALISATION IN DIVERSE ATTENTION-BASED MODELS WITH AN INFORMATION BOTTLENECK

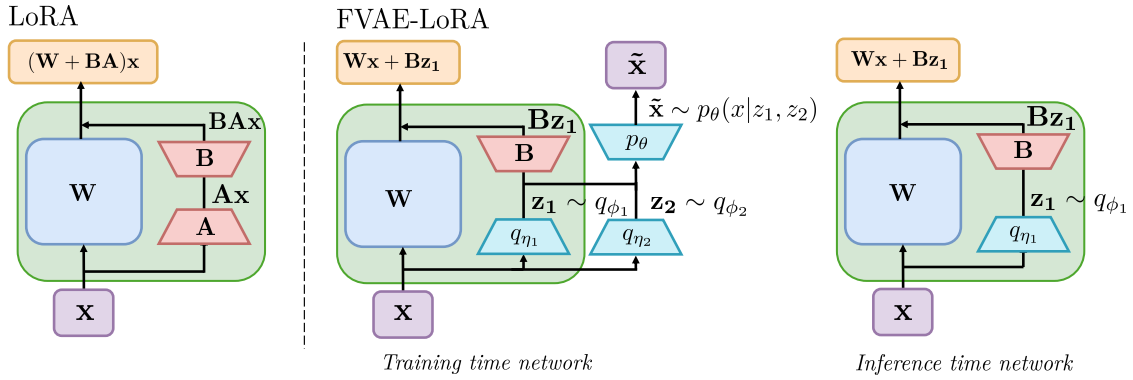
Pretrained and fine-tuned attention-based models are the state-of-the-art in many domains of AI, but they often struggle with generalization, leading to poor performance on tasks like out-of-domain transfer, distribution shifts, and few-shot learning. This limitation is prevalent across modalities such as speech, text, graphs, and vision. Nonparametric Variational Information Bottleneck (NVIB) is an attention-based information-theoretic regularizer applicable to pretrained models that has been shown to improve generalization. This work extends NVIB beyond our earlier work on the text modality without fine-tuning, leveraging NVIB's ability to avoid spurious correlations with noisy and superficial features. We extend NVIB for consistent performance across multiple diverse models and modalities, and find improved out-of-distribution generalization in speech quality assessment and language identification, text with induced attention sparsity, graph-based link prediction, and few-shot image classification.



F. Fehr, A. E. Baia, X. Chang, A. C. Coman, K. El Hajal, D. El Zein, S. Kumar, J. Zuluaga-Gomez, A. Cavallaro, D. Teney, and J. Henderson, Fine-Tuning Pretrained Models with NVIB for Improved Generalisation, *Workshop on Spurious Correlation and Shortcut Learning: Foundations and Solutions*, 2025

LATENT SPACE FACTORIZATION FOR FINE-TUNING LARGE MODELS

Automatic fine-tuning of large pretrained models using low-rank adaptation (LoRA) has become a cornerstone of efficient model specialization. However, existing LoRA techniques often fail to explicitly disentangle task-relevant information from residual variation in the learned low-rank subspace. Through cross-group collaboration at IDIAP, we introduce the Factorized Variational Autoencoder LoRA (FVAE-LoRA), a new parameter-efficient fine-tuning method. FVAE-LORA leverages a variational autoencoder to split the adaptation space into two distinct latent factors, where one factor captures task-salient features, while the other absorbs residual information. This design encourages a cleaner separation between useful and spurious signals during adaptation. We demonstrate the effectiveness of FVAE-LoRA through experiments across text, audio, and image tasks, which show that FVAE-LoRA consistently outperforms standard LoRA in both accuracy and robustness, even under distribution shifts.

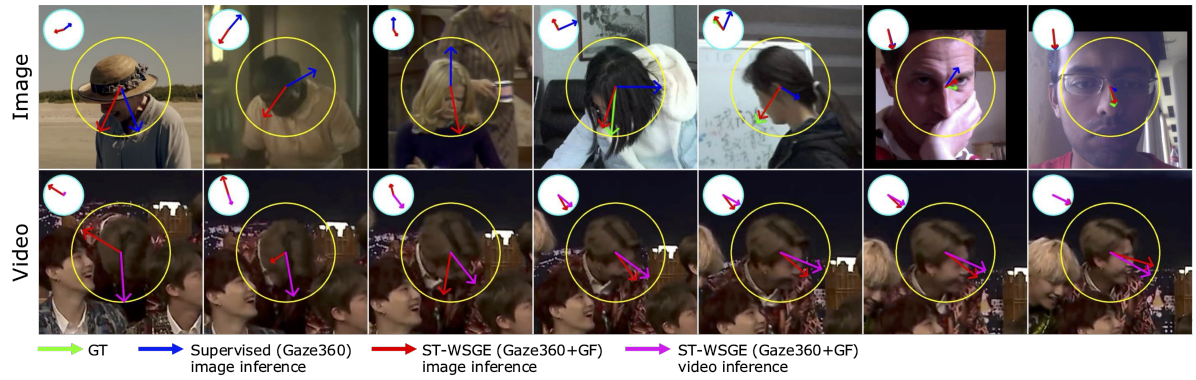


S. Kumar, Y. Kaloga, J. Mitros, P. Motlicek, and I. Kodrasi, Latent Space Factorization in LoRA, *Conference on Neural Information Processing Systems*, 2025

Human-AI Teaming

ADVANCING 3D GAZE ESTIMATION IN THE WILD

Gaze is central to human communication, revealing attention, intentions, emotions, and cognitive states. It aids in interpreting social cues, joint attention, and mental states, with applications in human-robot interaction, AR, education, and clinical research. However, estimating 3D gaze estimation "in-the-wild" is challenging because most real applications cannot assume near-frontal images (e.g., users looking straight towards a screen) due to camera placement constraints and people's natural behaviors. Existing 3D gaze datasets are typically small, rely on costly and intrusive labeling for reliable 3D gaze, and are biased toward controlled, near-frontal setups. They thus lack diversity in person appearance, lighting, head pose, occlusions, accessories (glasses, beards, hairstyles), and backgrounds. In addition, many datasets contain only single images, even though temporal information could significantly benefit video-based gaze estimation models.
The paper addresses these limitations with a modality-agnostic model and training strategy that can jointly exploit multiple heterogeneous datasets, from both 3D gaze data and simpler (weaker) 2D gaze annotations which can be obtained by indicating where someone is looking in a picture, a situation that allows to collect very large datasets. This enables learning a more robust, generalizable model that achieves state-of-the-art performance on many benchmarks and yields up to 30% error reduction in cross-dataset generalization evaluation, a substantial gain in a highly competitive field with numerous active research groups, including industrial players such as Nvidia.



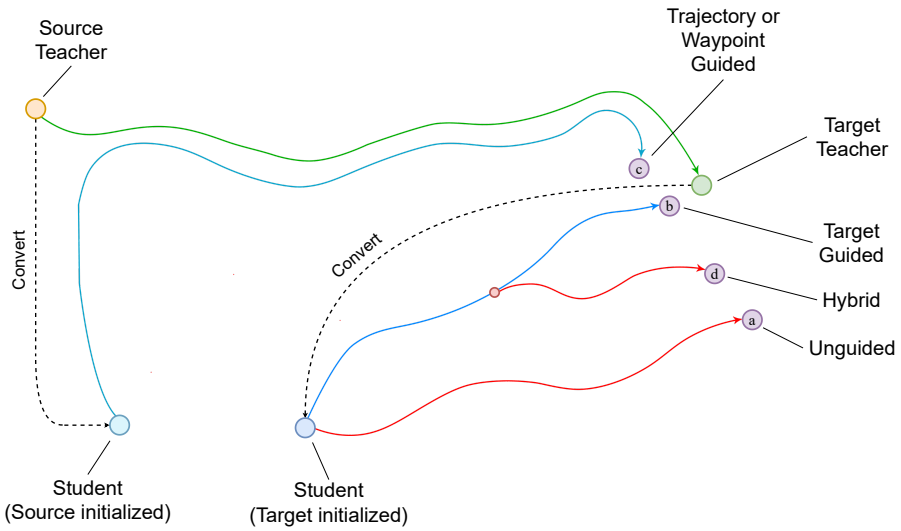
Significance of our ST-WSGE model. Our self-training based weakly-supervised framework for robust 3D gaze estimation in real-world conditions (e.g., varying appearance, extreme poses, resolution, and occlusion). All predictions used our image and video agnostic Gaze Transformer (GaT) model. Top row: importance of the training diversity using ST-WSGE and GazeFollow (GF) for generalization compared to standard supervised methods. Bottom row: influence of temporal context between image and video inference. Circles in images represent unit disks where 3D gaze vectors are projected onto the image plane (x, y in yellow) and a top-down view (x, z in blue).

P. Vuillecard and J.-M. Odobez, Enhancing 3D Gaze Estimation in the Wild using Weak Supervision with Gaze Following Labels, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025

MACHINE LEARNING FOR JOB SUITABILITY

With our collaborators, work psychologists at the Universities of Neuchâtel and Lausanne, we are interested in how to infer whether a candidate is suitable for a job or not. The hypothesis behind the SteADI project is that a candidate's ability to tell stories about past work experiences is indicative of their future ability in a similar job. Whilst our colleagues in Neuchâtel are able to look at data directly to establish ground-truths and investigate such hypotheses [1], here at Idiap we try to do it automatically.

We have found that the length of the interviews is a significant problem; conventional transformer architectures can rapidly exceed available memory on powerful devices. We have addressed the problem by developing linear transformer methods, and by converting large foundation models into linear models that can perform inference on long text or audio recordings [2].



[1] E. Germanier, M. He, A. M. Rufai, P. N. Garner, A. Bangerter, L. A. Renier, M. Schmid Mast, and K. Orji, Identifying storytelling in job interviews using deep learning, *Computers in Human Behavior Reports*, 2025

[2] M. He and P. N. Garner, Joint Fine-tuning and Conversion of Pretrained Speech and Language Models towards Linear Complexity, *13th International Conference on Learning Representations (ICLR)*, 2025

GEOMETRIC MODELS FOR ROBOTICS

Despite significant advances in AI, robots still struggle with tasks involving physical interaction. Robots can easily beat humans at board games such as chess or Go but are mostly incapable of skillfully moving the game pieces by themselves—the part of the task that humans subconsciously succeed in. Learning and transferring manipulation skills is hard because the movement behaviors that the robots need to acquire are tightly connected to our physical world and to embodied forms of intelligence.

Acquiring manipulation tasks encompasses a wide range of learning techniques, from foundational models with large-scale datasets to model-based optimization and frugal learning techniques that rely on a handful of demonstrations or exploration trials. Despite the difference in scale, these models can benefit from a clever use of the underlying model and algorithmic structures. Different terms are used in the literature to refer to such guidance, including prior knowledge, inductive biases, models, and representations. The research challenge is to design these representations without limiting the generalization, adaptation and processing speed capabilities of the system.

This work focuses on geometric priors, which can be viewed as a subset of physics-informed models to help robots acquire skills by imitation and self-refinement. We focus on three complementary geometric frameworks: signed distance fields, geometric algebra and Riemannian geometry, which provide representations facilitating learning, planning, control and optimization problems in robotics. The first consists of representing shapes in an implicit manner through the use of a distance function, where different approaches can be used to encode and learn this function. The second, geometric algebra, is linked to Clifford algebra and allows basic geometric primitives to be treated in a unified manner, including 6D poses, planes, lines, circles, and spheres, which can represent various forms of constraints in robot applications. The third leverages the use of Riemannian manifolds to extend models and algorithms originally developed for standard Euclidean data to curved spaces. These manifolds can represent a variety of geometric objects in robotics, not only for structured objects such as spheres, matrices and subspaces, but also for more generic smooth manifolds described by a Riemannian metric to measure distances.

[1] S. Calinon, Geometric Structures for Learning and Optimization in Robotics, *Annual Review of Control, Robotics, and Autonomous Systems*, 2025

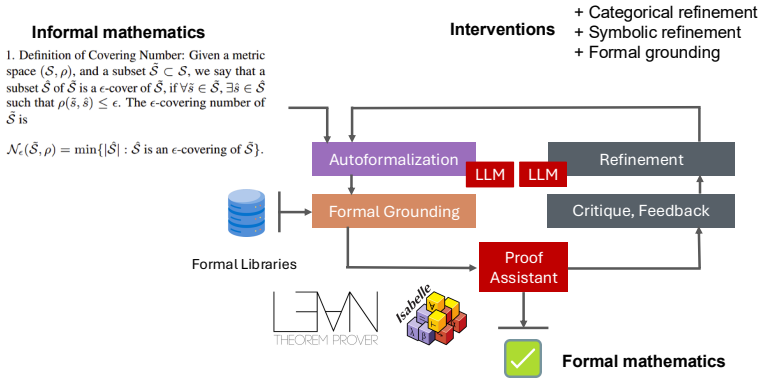
[2] T. Löw, P. Abbet, and S. Calinon, GAFRO: Geometric Algebra for Robotics [Tutorial], *IEEE Robotics and Automation Magazine*, 2025

[3] Y. Li and S. Calinon, From Movement Primitives to Distance Fields to Dynamical Systems, *IEEE Robotics and Automation Letters (RA-L)*, 2025

BUILDING AI MODELS FOR SUPPORTING MATHEMATICAL VERIFICATION AND DISCOVERY

We start from a simple question: can we get language models to read mathematics the way humans write it (e.g., in textbooks and research papers) and turn it into the precise "code" that proof assistants can check automatically? This process, called *autoformalization*, would bring the vision to facilitate mathematical discovery and verification. Yet most existing benchmarks in this area focus on short, exam-style problems. They tell us little about how well current models cope with the long, dense, and highly contextual arguments that lie at the heart of real mathematical practice.

To address this gap, we focus on autoformalizing *real-world mathematical definitions*. We construct curated resources based on how mathematics occur on real world proofs. We then assess the ability of several large language models (LLMs) to perform autoformalization and propose simple but effective interventions to improve the LLM-based autoformalization. The results show that real-world definitions are substantially more challenging than existing autoformalization benchmarks and that progress on autoformalization will not come from larger models alone, and that simple formal control mechanisms can deliver significant improvements.



L. Zhang, M. Valentino, and A. Freitas, Autoformalization in the wild: assessing LLMs on real-world mathematical definitions, *Conference on Empirical Methods in Natural Language Processing*, 2025

ERGODIC CONTROL IN ROBOTICS

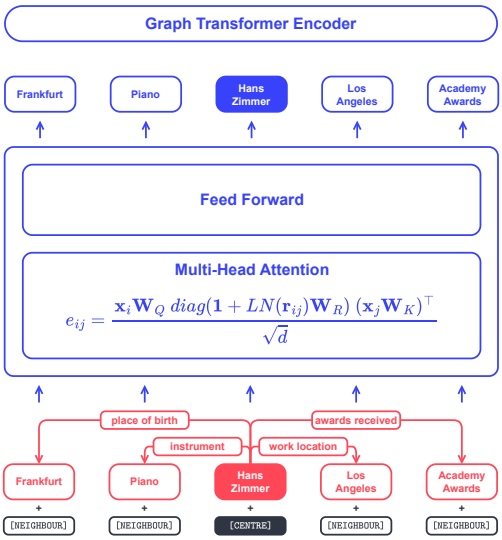
Ergodic control is a rapidly emerging technique for synthesizing search and coverage behaviors, initially developed from the field of control theory and rapidly spreading in robotics. It targets diverse robot applications in which some form of exploration or coverage is required, defined by a distribution, which arise naturally in applications such as surface cleaning, exploration, data collection for learning, localization, surveillance, or manipulation tasks.
Ergodic control generates control signals that drive a robot to spend time in areas proportional to the density of an arbitrary spatial distribution (e.g., distributed information). Such a controller produces natural exploration behavior that takes into account that there is a cost in moving from one sample to the next and the impact of physical dynamics of a system. This control approach is built on ergodic theory, which studies the connection between the time-averaged and space-averaged behavior of a dynamical system. By using it in a search and exploration context, it enables optimal exploration of an information distribution. Ergodic control can be extended in various ways, including a 3D coverage task from point cloud information, which we demonstrate in a dish washing application [1], or the handling of dynamically changing distributions in search & rescue applications with aerial robots [2].

[1] C. Bilaloglu, T. Löw, and S. Calinon, Tactile Ergodic Coverage on Curved Surfaces, *IEEE Transactions on Robotics (T-RO)*, 2025

[2] L. Lanča, K. Jakac, S. Calinon, and S. Ivić, Ergodic exploration of dynamic distribution, *IEEE Robotics and Automation Letters (RA-L)*, 2025

INTEGRATING INTERPRETABLE KNOWLEDGE GRAPHS IN TRANSFORMERS FOR TEXT

The transformer architecture underlies virtually all current approaches to text processing, including LLMs. Over a number of years, we have been developing methods which extend the attention layers in transformers to not only capture relationships between words, but also relationships between any kind of graph nodes. This work applies this approach to modeling the nodes and relations in knowledge graphs, which allows these models to use and predict human-editable representations of knowledge. In particular, our direct integration of graph relations into the same architecture used for processing text has allowed us to integrate the processing of knowledge graphs with the processing of the text which is attributed to the nodes and edges of these graphs. This is the first architecture which includes the integration of text labeling the edges of a knowledge graph, allowing the model to even process edge types it has never seen during training. We show that this architecture not only improves the model's ability to predict new edges in a knowledge graph, it also makes more effective use of the text processing abilities of transformers, thereby allowing us to use much smaller and more efficient pretrained LLM models. We also introduce new datasets for evaluating these abilities.

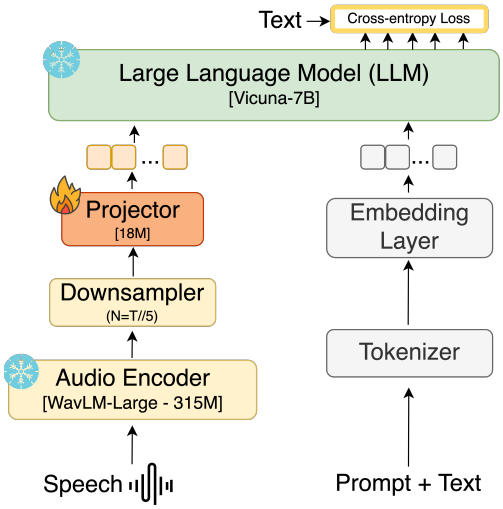


[1] A. C. Coman, C. Theodoropoulos, M.-F. Moens, and J. Henderson, Fast-and-Frugal Text-Graph Transformers are Effective Link Predictors, *Findings of the Association for Computational Linguistics*, 2025

[2] A. Coman, C. Theodoropoulos, M.-Fr. Moens, and J. Henderson. GADePo: Graph-assisted declarative pooling transformers for document-level relation extraction. *Workshop on Knowledge Augmented Methods for NLP*, ACL, 2024.

Performance Evaluation of SLAM-ASR: The Good, the Bad, the Ugly, and the Way Forward

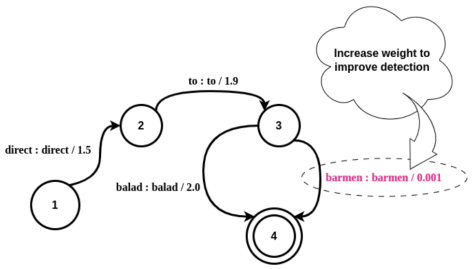
Recent research has demonstrated that training a linear connector between speech foundation encoders and large language models (LLMs) enables this architecture to achieve strong Automatic Speech Recognition (ASR) capabilities. Despite the impressive results, it remains unclear whether these simple approaches are robust enough across different scenarios and speech conditions, such as domain shifts and speech perturbations. Our work addresses these questions by conducting various experiments using a recent and widely adopted approach called SLAM-ASR. Concretely we present novel empirical findings that offer insights on how to effectively utilize the SLAM-ASR architecture across a wide range of settings. The main findings indicate that SLAM-ASR exhibits poor performance in cross-domain evaluation settings. Additionally, speech perturbations on in-domain data, such as changes in speech rate or additive noise, can significantly degrade performance. The work eventually offers critical insights for fine-tuning and configuring robust LLM-based ASR models, tailored to different data characteristics and computational resources.



S. Kumar, I. Thorbecke S. Burdisso, E. Villatoro-Tello, E. Manjunath, K. Hacioğlu, P. Rangappa, P. Motlicek, A. Ganapathiraju, and A. Stolcke. Performance Evaluation of SLAM-ASR: The Good, the Bad, the Ugly, and the Way Forward. *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2025.

MINIMUM EFFORT ADAPTATION OF AUTOMATIC SPEECH RECOGNITION SYSTEM IN AIR TRAFFIC MANAGEMENT

Advancements in Automatic Speech Recognition (ASR) technology are exemplified by ubiquitous voice assistants such as Siri and Alexa. Researchers have been exploring the application of ASR for Air Traffic Management (ATM) systems. Initial prototypes utilized ASR to pre-fill aircraft radar labels and achieved a technological readiness level before industrialization (TRL6). However, accurately recognizing infrequently used but highly informative domain-specific vocabulary is still an issue. This includes waypoint names specific to each airspace region and unique airline designators, e.g., “dexon” or “pobeda”. Traditionally, open-source ASR toolkits or large pre-trained models require substantial domain-specific transcribed speech data to adapt to specialized vocabularies. However, typically, a “universal” ASR engine capable of reliably recognizing a core dictionary of several hundreds of frequently used words suffices for ATM applications. The challenge lies in dynamically integrating the additional region-specific words used less frequently. These uncommon words are crucial for maintaining clear communication within the ATM environment. Our work proposes a novel approach that facilitates the dynamic integration of these new and specific word entities into the existing universal ASR system. This paves the way for “plug-and-play” customization with minimal expert intervention and eliminates the need for extensive fine-tuning of the universal ASR model. The proposed approach demonstrably improves the accuracy of these region-specific words by a factor of $\sim$7 (from 10% F1-score to 70%) for all rare words and $\sim$5 (from 13% F1-score to 64%) for waypoints.



[1] M. Bhattacharjee, P. Motlicek, S. Madikeri, H. Helmke, O. Ohneiser, M. Kleinert, and H. Ehr. Minimum effort adaptation of automatic speech recognition system in air traffic management. *European Journal of Transport and Infrastructure Research*, 2025.

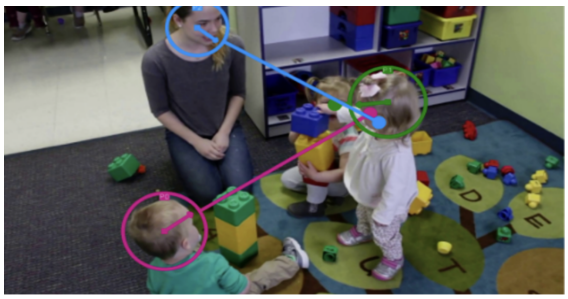
[2] P. Motlicek, S. Kumar, D. Khalil, and A. Prasad. Leveraging Untranscribed Data for End-to-End Speech and Callsign Recognition in Air-Traffic Communication. *15th edition, SESAR Innovation days conference*, 2025.

INVESTIGATING GAZE-FOLLOWING MODELS FOR ANALYSIS OF JOINT ATTENTION IN CHILD LANGUAGE DATA

This paper investigates how to automatically detect the visual gaze components of joint attention in naturalistic child language acquisition data. The core problem is that, although joint attention is crucial for early language learning, it is extremely time-consuming to code from video by hand; further, existing gaze-tracking tools are not tailored to messy, at-home recordings of young children.

To address this, the Zurich partner constructed a manually annotated dataset from a longitudinal Romansh-Tuatschin home video corpus and applied two state-of-the-art gaze-target estimation models from Idiap (Sharingan and MTGS), additionally fine-tuning MTGS on their data. They evaluated model outputs against human coders using reliability and distance metrics, and found that while manual coding is still superior, the models achieve acceptable reliability and their estimated gaze-point distances are already informative about whether joint attention is likely present in a clip.

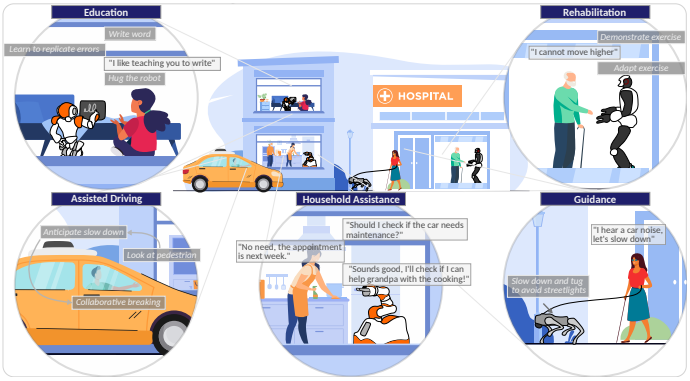
The work resulted from a collaboration between the ACQDIV laboratory of Prof. S. Stoll at University of Zurich and Idiap.



M. Dickerman, A. Gupta, S. Tafasca, X. Zhang, J.-M. Odobez, and S. Stoll, Automatic detection of the visual gaze components of joint attention in observational, naturalistic child language acquisition data, *Boston University Conference on Language Development*, 2025

HUMAN-INTERACTIVE ROBOT LEARNING: DEFINITION, CHALLENGES, AND RECOMMENDATIONS

In a world where robots are increasingly expected to learn and adapt to complex human environments, traditional robot learning paradigms that rely on large pre-collected datasets and isolated training fall short of capturing the richness of human expertise and interaction. In our paper “Human-Interactive Robot Learning: Definition, Challenges, and Recommendations”, we study Human-Interactive Robot Learning (HIRL), a research paradigm in which robots and humans engage during the learning process so that human teachers can directly shape behavior, share task knowledge, and help robots explore safely and efficiently, leading to faster adaptation and better alignment with human expectations. We provide a broad, consistent overview of HIRL research, clarify its core concepts at the intersection of artificial intelligence, robotics, and human–computer interaction, identify key open challenges across themes including human modeling, learning algorithms, interaction design and context integration, and offer concrete use cases and recommendations to guide future work and foster a cohesive research community.

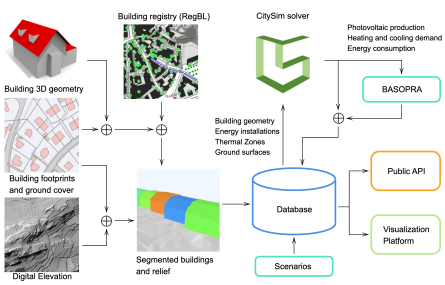


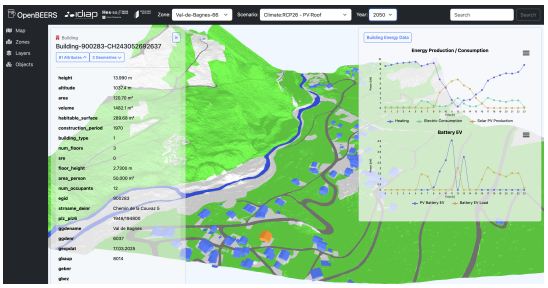
K. Baraka, I. Idrees, T. Kessler Faulkner, E. Biyik, S. Booth, M. Chetouani, D. H. Grollman, A. Saran, E. Senft, S. Tulli, A.-L. Vollmer, A. Andriella, H. Beierling, T. Horter, J. Kober, I. Sheidlower, M. E. Taylor, S. van Waveren, and X. Xiao, Human-Interactive Robot Learning: Definition, Challenges, and Recommendations, *Journal Human-Robot Interaction*, Dec. 2025

Sustainable & Resilient Societies

OpenBEERS: Standardizing Data for Urban Energy Transition

The OpenBEERS project tackles the critical challenge of urban building energy transition by developing a digital platform that aggregates scattered building and energy data in the Canton of Valais. It combines an open-access database, advanced simulation tools (CitySim and BASOPRA), and a 3D visualization interface to enable assessment and planning of building renovation and distributed energy resource (DER) deployment at the urban scale. The platform was validated through detailed scenario analyses in two case studies (Sierre and Val de Bagnes), using both open and synthetic data sources to quantify the effects of renovation, PV, heat pump, battery, and EV adoption. Results are delivered as comprehensive sets of key performance indicators (KPIs) on self-consumption, autarky, and peak demand, with actionable scenario comparisons for stakeholders. The project finds that data sharing remains a limiting factor, with annual aggregation impeding intraday temporal modeling. Electrification of heat and mobility will drive significant increases in peak load and annual electricity demand, with summer PV and short-term batteries unable to fully address seasonal gaps. Renovation is essential for system resilience, reducing winter peaks for heat pumps and aiding energy transition objectives. OpenBEERS provides a reproducible, open-source framework for urban-scale energy analysis and policy support, and establishes a structured basis for future data integration and methodological refinement within cantonal energy planning.

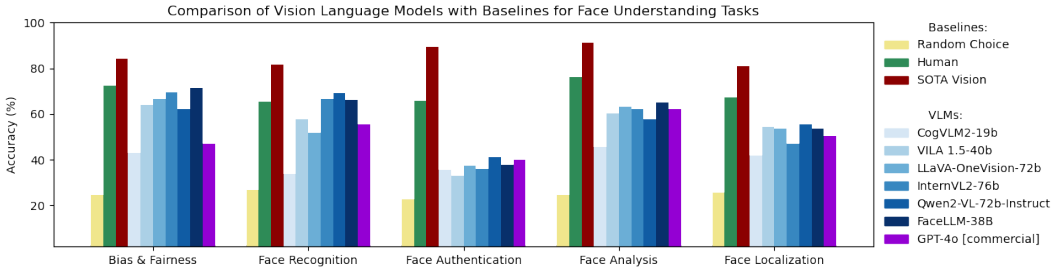




D. Geissbühler, A. Pena Bello, J. Kämpf, and J. Rager, OpenBEERS: A digital platform for urban scale simulation of building energy efficiency, *in Journal of Physics: Conference Series, IOP Publishing*, 2025.

Foundation Models and Biometrics

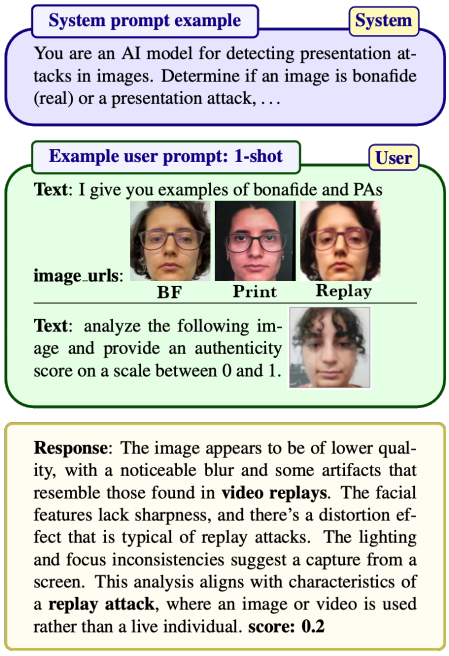
We study the emergence of large-scale "foundation models" as a paradigm in machine learning and their implications for biometric systems, and we systematically review how models originally developed for language, vision and multi-modal reasoning are being adapted to tasks such as recognition, anti-spoofing and synthetic-data generation in biometrics. We identify key successes where these approaches yield strong zero- or few-shot performance, analyze the technical and privacy challenges they introduce, and outline promising research directions for building biometric systems that fully leverage foundation models while preserving fairness, security and robustness.



H. Otroshi, and S. Marcel, Foundation Models and Biometrics: A Survey and Outlook, *in IEEE Transactions on Information Forensics and Security*, 2025.

LARGE VISION LANGUAGE MODELS FOR FACE PRESENTATION ATTACK DETECTION

We study how to enhance presentation attack detection (PAD) systems for face biometrics by leveraging large language models, and we propose a novel pipeline that uses ChatGPT to generate textual prompts describing attack modalities, which are then converted into synthetic images using a diffusion model and used to train PAD networks under extremely limited real-data conditions. Our experiments show that this combined language-to-image augmentation strategy significantly improves generalization to unseen attack types compared to conventional data-only or simulation-only baselines, pointing to a promising direction where language models act as meta-annotators to enrich biometric-security datasets at scale.



A. Komaty, H. Otroshi, and S. Marcel, Exploring ChatGPT for Face Presentation Attack Detection in Zero and Few-Shot in-Context Learning, *in Winter Conference on Applications of Computer Vision (WACV)*, 2025.

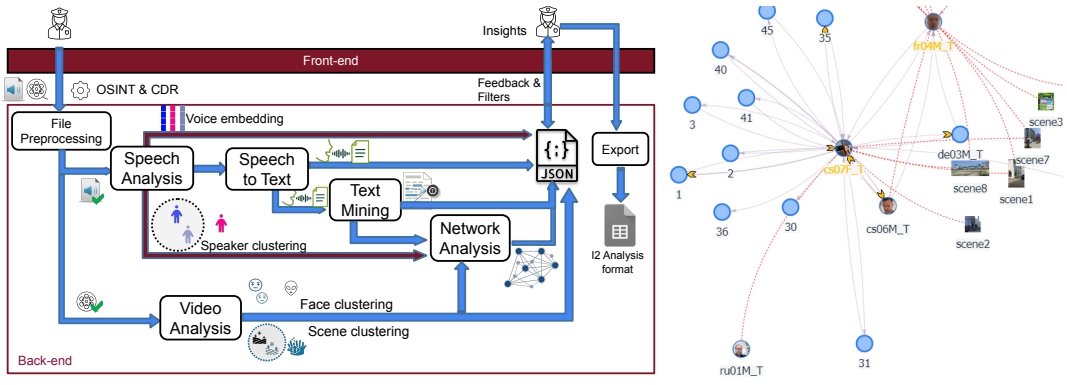
DETECTING DIGITAL MANIPULATIONS OF IDENTITY DOCUMENTS

We study how recent image generation technologies can be used to forge Identity Documents (IDs). We propose a novel publicly available dataset, FantasyID, which mimics real-world IDs but without tampering with legal documents. FantasyID contains ID cards with diverse design styles, languages, and faces of real people. We have emulated digital forgery/injection attacks that could be performed by a malicious actor to tamper with the IDs using existing generative tools. Current state-of-the-art forgery detection algorithms are challenged by the FantasyID dataset. Evaluations demonstrate that FantasyID is complex enough to be used as an evaluation benchmark for detection algorithms.

P. Korshunov, A. Mohammadi, V. Vidit, C. Ecabert, and S. Marcel, FantasyID: A dataset for detecting digital manipulations of ID-documents, *in IEEE International Joint Conference on Biometrics*, 2025.

AUTOCRIME - OPEN MULTIMODAL PLATFORM FOR COMBATING ORGANIZED CRIME

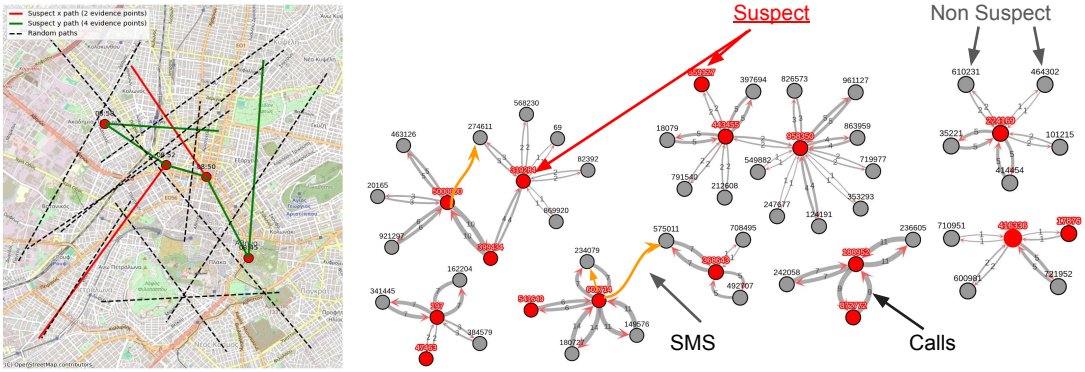
As part of the recently ended EC H2020 ROXANNE project, we have developed a multimodal AI-powered investigative platform, designed to accelerate and enhance the effectiveness of criminal network analysis. It is developed specifically for police practitioners to process large, heterogeneous data and automatically extract evidence to assist in making informed decisions. The platform integrates sophisticated technologies to analyze lawfully intercepted telephone conversations, extending this analysis with non-content data (metadata related to calls, time/spatial positions) and data from other modalities. The platform seamlessly combines state-of-the-art technologies to uncover hidden patterns and predictive insights in criminal activities and structures. The core method involves: (1) Multimodal Data Processing engine; (2) Information Extraction engine running speech, text, and video analysis technologies to extract crucial entities; and (3) Knowledge Graph Construction, which allows multiple knowledge graphs to be built from extracted information capturing interactions between actors.



S. Madikeri, P. Motlicek, D. Sanchez-Cortes, P. Rangappa, J. Hughes, J. Tkaczuk, A. Sanchez Lara, D. Khalil, J. Rohdin, D. Zhu, A. Krishnan, D. Klakow, Z. Ahmadi, M. Kováč, D. Boboš, C. Kalogiros, A. Alexopoulos,and D. Marraud: Autocrime - open multimodal platform for combating organized crime, *in Journal Forensic Science International: Digital Investigation*, Vol. 54, 2025.

TRACY: Accelerating Criminal Investigations

TRACY (A big-data analyTics from base-stations Registrations And CDR e-evidence sYstem) is a framework designed to analyze mobile signaling communication logs to accelerate criminal investigations. Effective criminal investigations increasingly depend on correlating content data (like audio or text) and non-content data, such as traffic/logs, subscriber details, and location information (metadata). Law enforcement agencies face a substantial challenge in processing the massive volumes of data lawfully obtained from communication service providers (CSPs). Even when data is restricted geographically and temporally (e.g., a 4-hour window within a 2x2 km area), mobile signaling records can exceed one million entries. TRACY focuses on overcoming the complexity and volume of this data to speed up the process of identifying relevant terminals near a crime scene. We address this challenge by scanning large-scale signaling communication logs to pinpoint terminals located near key evidence points during a crime. It relies solely on encrypted terminal identifiers and mobile network topology, operating without accessing call content. The framework utilizes a unified scoring system based on three core analytical methods, which integrate spatial-temporal heuristics with behavioral scoring in a unified system aimed at real-world criminal investigation scenarios.

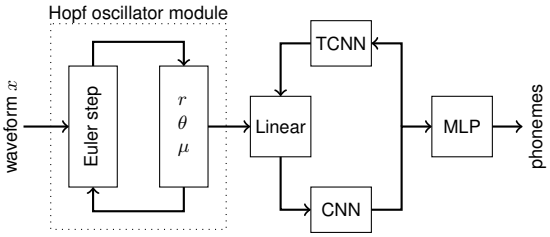


P. Rangappa, P. Motlicek, D. Sanchez-Cortes, A. Sanchez Lara, M. Antonopoulou, I. Fourfouris, N. Avgerinos and M. Tsangaris: Accelerating Criminal Investigations with TRACY, *in 16th EAI International Conference on Digital Forensics and Cyber Crime*, Miami, USA, 2025.

AI for Life

IMPROVED MODELS OF THE HUMAN AUDITORY SYSTEM

In speech processing it has been common practice for years to refer to the biology of the human production or perception mechanisms for clues about how to build artificial systems. The process has also been inverted: we can use artificial systems to create ideal perception mechanisms and find that they compare well to what we know of biology.
Recent understanding of the human auditory process is that it is active — it is controlled by the cognitive process. We know this because ears (counterintuitively) emit sounds. The mechanisms, however, are not well understood.
In an effort to both improve the artificial models and understand the biological ones, we have created more sophisticated hearing models and trained them in the context of modern speech to text applications. The experiments show that such models can work and that they do benefit from feedback coming from further up the auditory pathway.



[1] L. Coppieters de Gibson and P. N. Garner, Exploring auditory feedback mechanisms in speech recognition, *Interspeech 2025*, 2025
[2] L. Coppieters de Gibson and P. N. Garner, Training a filter-based model of the cochlea in the context of pre-trained acoustic models, Acoustics, vol. 6, 2024.

TEACHING AI TO UNDERSTAND THERAPIST–CHILD INTERACTIONS IN AUTISM DIAGNOSIS SESSIONS

This research is done in collaboration with the STARS team led by Dr Bremond at INRIA, France, and Prof. Suzanne Thümmler from the Nice Children's Hospital and Cote d'Azur University.

We have addressed the analysis of loose social interactions in real-world autism therapy sessions. In these settings, two people perform different actions that jointly define a global activity, often without contact or tight synchrony. Automatically analyzing such behavior can support social-skills assessment, therapy monitoring, and mental health diagnosis. The main difficulty is that loose dyadic interactions are long, complex, and asymmetric: a leader (clinician) and follower (child) act differently and asynchronously. Most action-recognition models target short, tightly synchronized or physically engaged interactions, so they miss this global joint behavior. Real clinical videos further suffer from varying viewpoints, occlusions, class imbalance, and heterogeneous child behavior.

To address these issues, we proposed a dual-path architecture that processes each participant's video with a 3D CNN to extract global features, then fuses them via a Global-Layer-Attention cross-attention module. Evaluations on a new Loose-Interaction dataset of ADOS-2 therapy sessions and on a public Autism dataset led to state-of-the-art results. A time-aware variant achieved the top performance on the NTU-RGB+D tight interactions, showing that different interaction types need different temporal modeling strategies.

A. Ali, R. Dai, A. Marisetty, G. Astruc, M. Thonnat, J.-M. Odobez, S. Thümmler, and F. Bremond, Loose Social-Interaction Recognition in Real-world Therapy Scenarios, *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025

AI TOOLS TO SUPPORT THE ECOLOGICAL ANALYSIS OF CHIMPANZEE BEHAVIORS

Thanks to a collaboration with the University of Neuchâtel, and as part of the NCCR Evolving Language, new AI-based tools were developed to automatically recognize chimpanzee behavior from videos.

Studying chimpanzee behavior in a non-intrusive way by analyzing video recordings is challenging, since – compared to humans – their appearance makes their pose more difficult to identify, AI models need large amounts of high-quality annotated data, and existing models often fail when they are used on videos from new environments (for example, models trained on zoo footage may perform poorly on forest footage). To address this, a public dataset (ChimpBehave) was created, comprising 1,362 annotated video segments of chimpanzees filmed at Basel Zoo, all labeled by an expert primatologist.

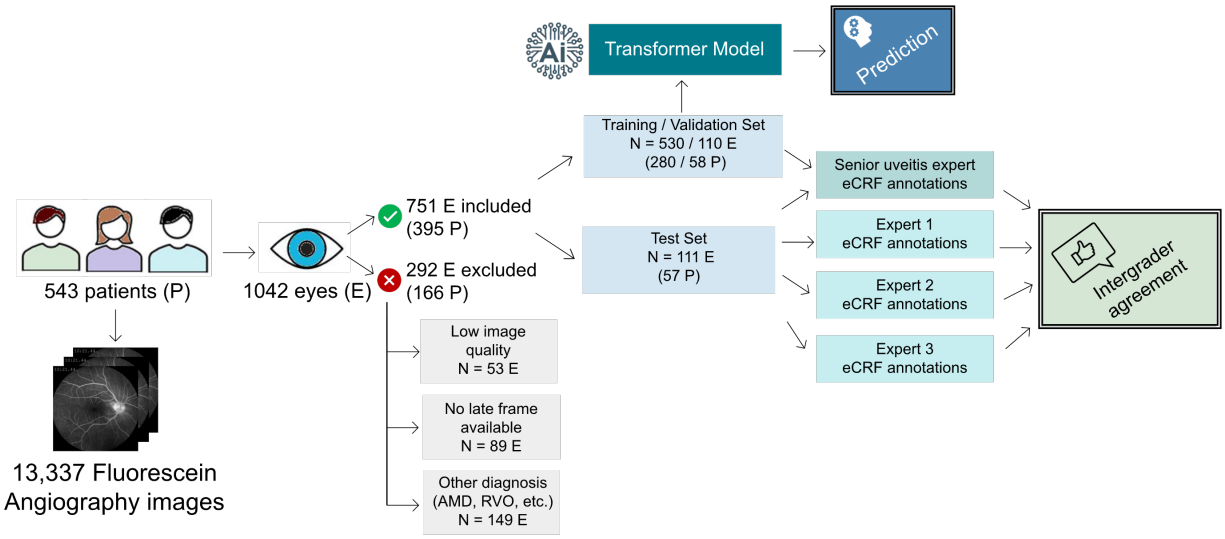
We investigated the design and training of five AI models: two using raw video and three using skeletal representations of movement. Results showed that video-based models were the most accurate, even in cross-dataset settings and in spite of the environment discrepancies, while lighter skeletal-based models were better suited for low-resource settings such as remote field stations. Overall, our approach demonstrates how AI can support ecological, large-scale, and non-invasive monitoring of endangered species and provide practical tools for wildlife conservation.

Sample video clips. Running and climbing down.

M. Fuchs, E. Genty, A. Bangerter, K. Zuberbühler, J.-M. Odobez, and P. Cotofrei, From forest to zoo: great ape behavior recognition with ChimpBehave, *International Journal of Computer Vision*, 2025

AUTOMATIC GRADING OF RETINAL INFLAMMATION USING TRANSFORMER MODELS

Accurate assessment of retinal inflammation is essential for guiding treatment in uveitis but remains a time-consuming and subjective task, even for experts. In collaboration with ophthalmology teams across Switzerland and France, we developed a transformer-based artificial intelligence model capable of automatically grading inflammatory signs on fluorescein angiography, the gold-standard imaging technique for this disease. The model was trained on the largest uveitis dataset of its kind — over 40,000 images from 1,000 eyes—graded by four expert clinicians following international standards. The system matched expert-level performance for vascular and capillary leakage, macular edema, and optic-disc hyperfluorescence, while highlighting the same retinal regions used by specialists for diagnosis. By automating this complex process, the approach enables faster, more consistent disease evaluation and paves the way for discovering imaging biomarkers and improving access to expert-level uveitis care worldwide.

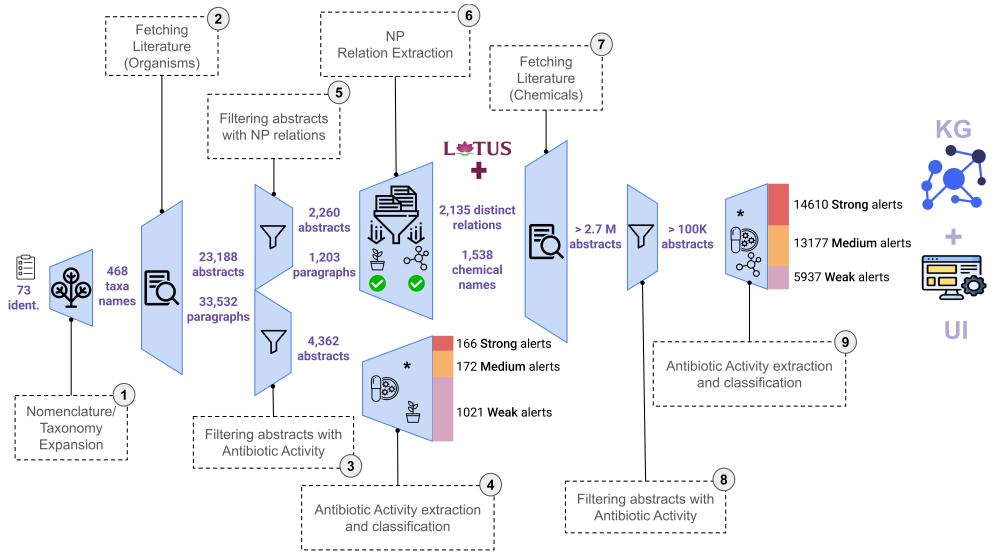


V. Amiot, O. Jimenez-del Toro, Y. Guex-Croisier, M. Ott, T.-E. Bogaciu, S. Banerjee, J. Howell, C. Amstutz, C. Chiquet, C. Bergin, I. Meloni, M. Tomasoni, F. Hoogewoud, and A. Anjos, Automatic transformer-based grading of multiple retinal inflammatory signs in uveitis on fluorescein angiography, *Computers in Biology and Medicine*, 2025

ACCELERATING ANTIBIOTIC DISCOVERY WITH LLMS AND KNOWLEDGE GRAPHS

We start from a practical bottleneck in early-stage antibiotic discovery. Pharmaceutical teams invest millions of dollars and years of work to explore poorly characterized organisms in search of new antibiotic-producing natural products. Yet checking for such prior evidence requires exhaustive, manual review across fragmented resources, complicated further by unstable taxonomy, synonymy, and the two-hop nature of the task (organism → compound → antibiotic activity). In this work we built an LLM-driven antibiotic discovery pipeline that systematically analyzes evidence of previously reported antibiotic activity, connecting organisms, compounds and antibiotic properties.

To do this, we design a pipeline that combines large language models with domain-specific knowledge graphs (KGs). The pipeline provides a systematic mechanism to disambiguate and link organisms, compounds and antibiotic evidence, allowing both a scalable and systematic approach for supporting evidence-based reasoning in antibiotic discovery. We evaluate the system on a real industrial use case: a private list of 73 candidate organisms. Running our pipeline over PubMed and LOTUS, we process tens of thousands of organism paragraphs and millions of chemical abstracts, yielding thousands of natural product relations and over 30,000 activity "alerts". Overall, the work shows that LLMs, when combined with targeted filtering, relation extraction, and a structured KG, can serve as a practical mechanism for supporting experts in large-scale literature review and helping to de-risk early-stage antibiotic discovery.

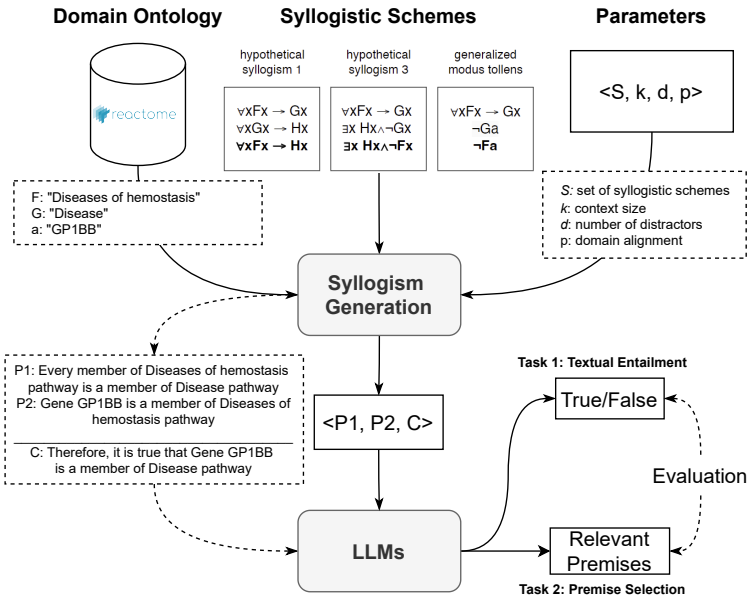


M. Delmas, M. Wysocka, D. Gusicuma, and A. Freitas, Accelerating Antibiotic Discovery with Large Language Models and Knowledge Graphs, *Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, 2025

LLMs AND COMPLEX BIOMEDICAL INFERENCES

We start from a safety-critical question: can current NLI models reliably perform *formal* syllogistic reasoning in specialized biomedical domains, or do they mostly lean on pattern-matching and world knowledge? In biomedicine, many inferences depend on drawing logically valid conclusions from pathway memberships, disease hierarchies, and gene-pathway relations. If a model cannot separate valid from invalid syllogisms in this setting, it becomes risky to use it for biomedical inference.

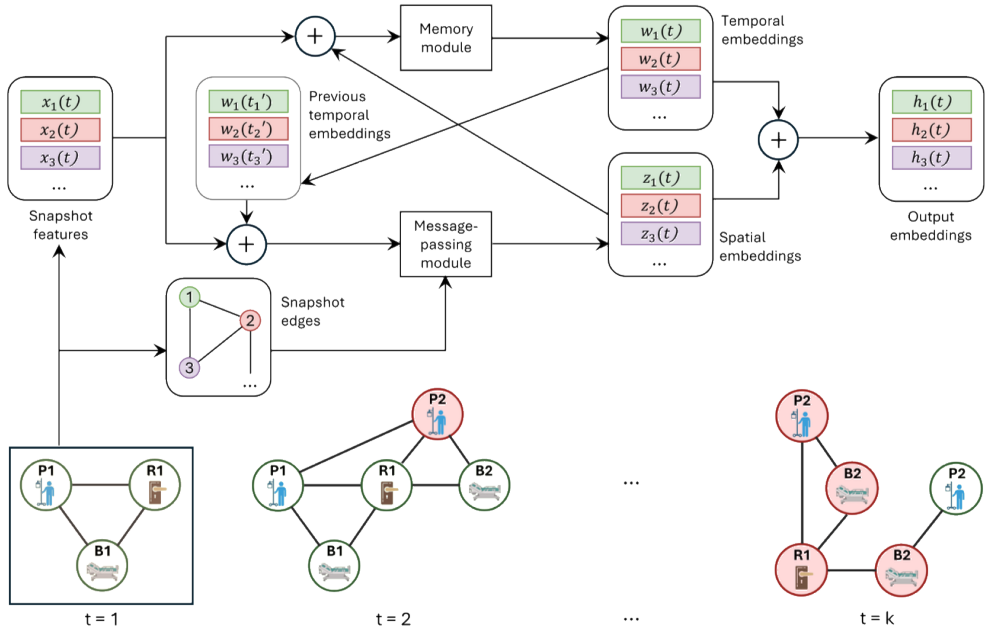
To this end, we introduce **SylloBio-NLI**, a benchmark and evaluation framework that combines formal syllogistic schemes with biomedical ontologies to automatically generate controlled natural language inference (NLI) datasets. The results show that LLMs models achieve surprisingly low accuracy on complex biomedical inferences, proving to be highly sensitive to surface and lexical variation. Overall, SylloBio-NLI shows that LLMs still confound formal and material inference patterns in biomedical NLI: even when examples help, models remain far from the robustness and consistency needed for safe deployment. At the same time, the framework itself offers a general recipe for generating logically controlled, ontology-grounded reasoning benchmarks in other specialized domains, with minimal manual annotation.



M. Wysocka, D. Carvalho, O. Wysocki, M. Valentino, and A. Freitas, SylloBio-NLI: Evaluating Large Language Models on Biomedical Syllogistic Reasoning, *Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics*, 2025

GRAPH NEURAL NETWORKS WITH MEMORY FOR PREDICTING HOSPITAL INFECTION RISKS

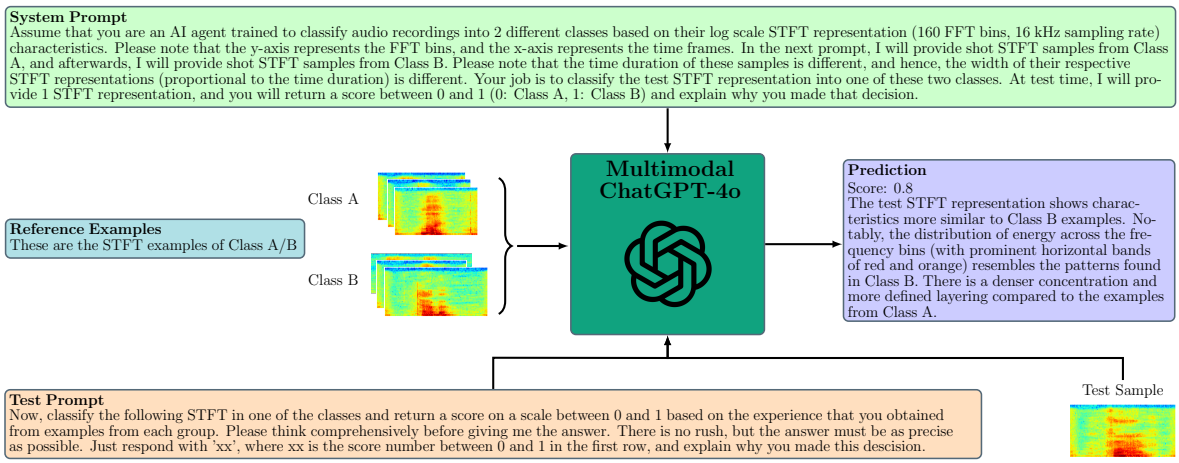
Preventing hospital-acquired infections caused by multidrug-resistant bacteria remains a major challenge for healthcare systems. In collaboration with the University of Geneva and the Polytechnic University of Leiria (Portugal), we developed the Space–Time–and–Memory Graph Neural Network (STM-GNN), a new model that predicts infection risks by analyzing patient interactions and environmental contamination over time. The model introduces a recurrent "memory" mechanism that tracks how bacteria spread across patients, rooms, and surfaces within healthcare units. Using a real infection-prevention dataset collected from a long-term care facility, STM-GNN achieved the highest predictive accuracy among several classical and deep-learning baselines, with an area under the ROC curve of 0.84. This study demonstrates how graph-based AI models can capture complex transmission patterns and support infection control by identifying high-risk patients and environments early.



D. Geissbuhler, A. Bornet, C. Marques, A. Anjos, S. Pereira, and D. Teodoro, STM-GNN: Space-Time-and-Memory Graph Neural Networks for Predicting Multi-Drug Resistance Risks in Dynamic Patient Networks, *International Conference on Artificial Intelligence in Medicine*, 2025

MULTIMODAL LARGE LANGUAGE MODELS FOR INTERPRETABLE AUTOMATIC PATHOLOGICAL SPEECH DETECTION

Automatic detection of pathological speech is essential for developing accessible and cost effective tools to support clinical assessment and monitoring of speech disorders. While recent deep learning approaches achieve strong performance, they often operate as black boxes, which limits their interpretability and acceptance in clinical contexts. We propose the use of large multimodal language models such as *ChatGPT-4o* to classify speech as pathological or typical in a few shot in context learning setting, while also generating natural language explanations for its decisions. Experiments on a benchmark dysarthric speech dataset show that the proposed approach reaches performance comparable to a state of the art convolutional neural network, while offering improved transparency through explanatory outputs. These results illustrate the potential of multimodal language models to combine competitive accuracy with interpretable reasoning for pathological speech analysis.



M. Amiri, H. Otroshi Shahreza, and I. Kodrasi, Exploring In-Context Learning Capabilities of ChatGPT for Pathological Speech Detection, *ITG Conference on Speech Communication*, IEEE, 2025

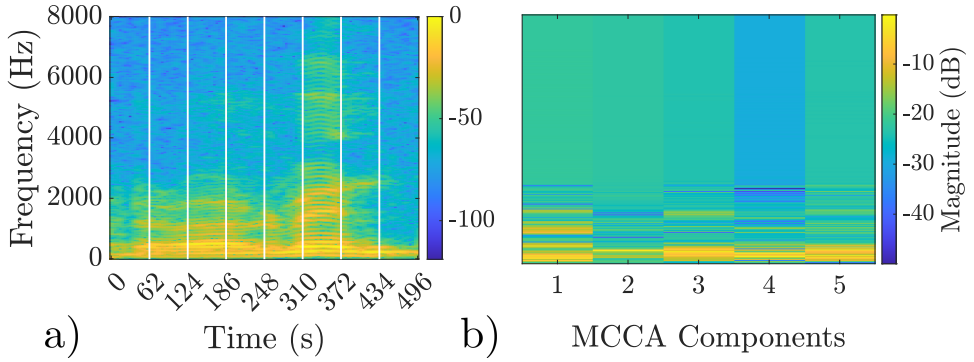
GRAPH NEURAL NETWORKS FOR PARKINSON'S DISEASE DETECTION

Automatic detection of Parkinson's disease from speech is a challenging task, as speech recordings are often heterogeneous, segmented, and affected by label noise, while disease-related cues may only emerge when information is aggregated across multiple utterances. In collaboration with TCG CREST, India, we propose a graph neural network framework in which speech segments are represented as nodes connected according to learned similarities, enabling the model to exploit relationships between segments rather than treating them independently. By propagating information across the graph, the approach captures consistent dysarthric patterns while reducing the impact of noisy or uninformative segments. Experiments on benchmark datasets demonstrate that this graph based modeling strategy improves robustness and accuracy compared to conventional segment level methods, highlighting the benefits of relational learning for speech-based Parkinson's disease detection.

S. Shakeel, Y. Kaloga, M. Sahidullah, and I. Kodrasi, Graph Neural Networks for Parkinson's Disease Detection, *International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2025

ENHANCING PATHOLOGICAL SPEECH DETECTION WITH MULTIVIEW REPRESENTATION LEARNING

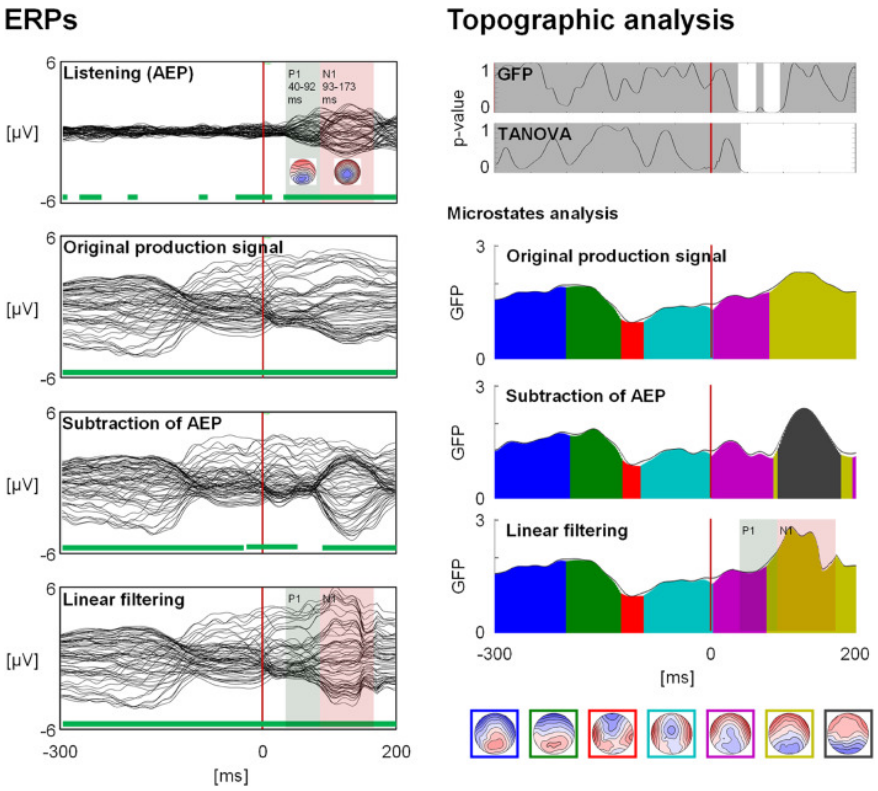
Automatic pathological speech detection systems often rely on spectrograms or wav2vec2 embeddings. These representations can carry uninformative variation unrelated to pathology, such as phonetic content or speaking style, which can degrade classification performance. In this work we introduce the use of multiview canonical correlation analysis (MCCA) as a preprocessing step to remove uncorrelated information from multiple input representations before classification, enabling models to focus on pathology-relevant features. We show that applying MCCA prior to traditional classifiers yields comparable or superior detection performance to more complex architectures, while preserving the structure of the learned representations and improving interpretability. These results suggest that multiview dimensionality reduction is a promising direction for building more robust pathological speech detection systems.



Y. Kaloga, S. Shakeel, and I. Kodrasi, Multiview Canonical Correlation Analysis for Automatic Pathological Speech Detection, *International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2025

Separating production and auditory feedback signals in speech-related EEGs

Speech production involves tightly coupled neural processes for planning, articulation, and auditory feedback, but disentangling these overlapping brain signals has long challenged researchers aiming to understand the neural bases of speech and its disorders. With the University of Geneva, we investigate how event-related potentials (ERPs) recorded during speech articulation contain contributions from both the production process and the auditory feedback that follows vocal onset. Further, we evaluate the use of a multichannel Wiener filter to separate these components. Participants produced and then listened to their own pseudoword productions, allowing us to estimate and filter the auditory response from the production ERP. Our analyses showed that the filtered ERP closely resembled the original production signal and that subtracting listening-related activity can alter signal topography, suggesting that removing auditory feedback may be unnecessary when signals are time-locked to production onset. These findings advance EEG-based analysis of speech motor planning and have implications for interpreting neurophysiological markers of speech in both typical and atypical speakers.

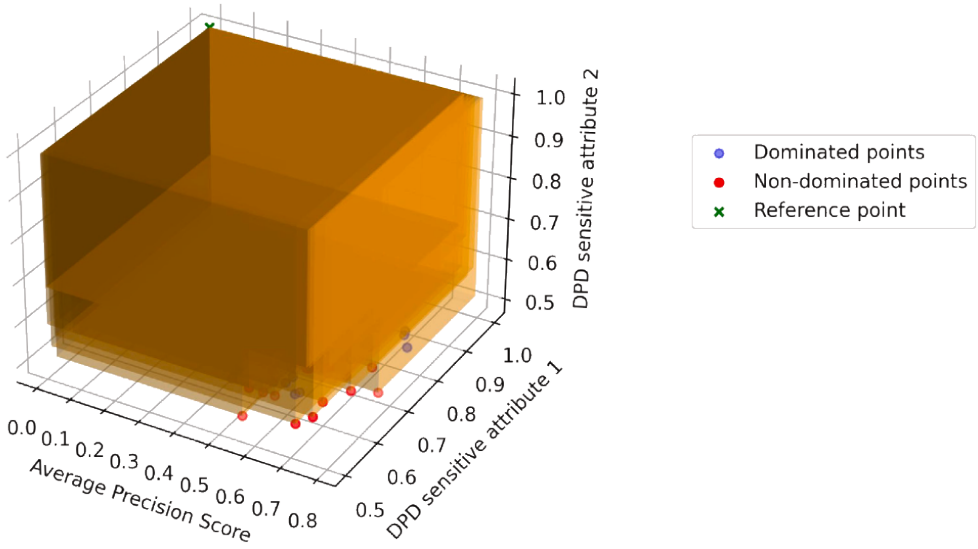


M. D. Pretto, M. Laganaro, and I. Kodrasi, ERP Signals During Speech Articulation: Does Auditory Feedback Mask Other Ongoing Cognitive-motor Processes? *Brain Topography*, 2025

AI for Everyone

OPTIMIZING FAIRNESS AND UTILITY IN MACHINE LEARNING FOR HEALTHCARE APPLICATIONS

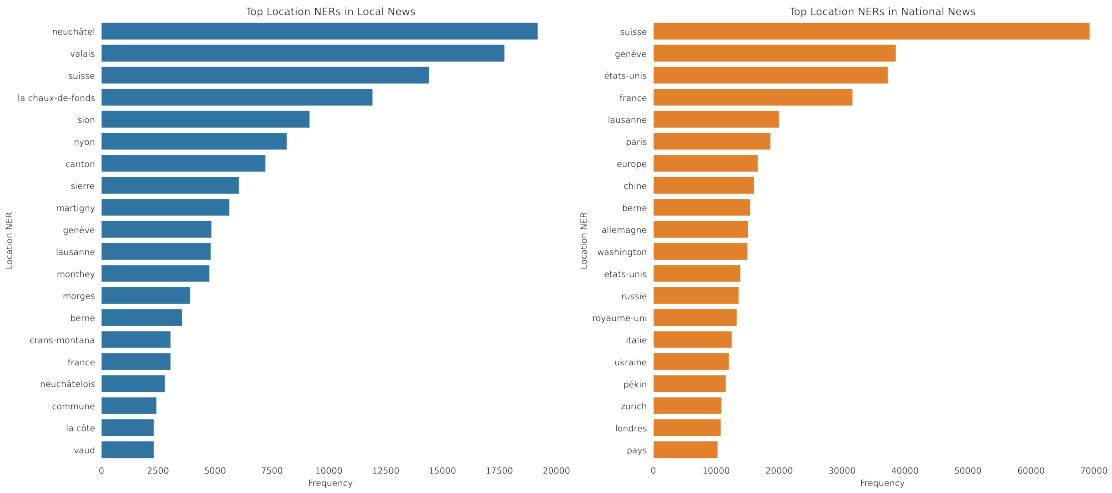
Ensuring that artificial intelligence benefits everyone requires models that perform well while treating individuals equitably across gender, age, and ethnicity. In collaboration between Idiap, and the Federal University of São Paulo, Brazil, we developed a mathematical optimization framework that explicitly enforces fairness constraints in machine learning models without sacrificing accuracy. The method extends classical logistic regression by incorporating multiple fairness conditions, such as demographic parity and equal opportunity, so that predictive differences between groups remain below specified thresholds. Applied to both simulated data and real-world healthcare datasets (MIMIC-III and Arrhythmia), the approach consistently reduced disparities in positive prediction rates among demographic groups while maintaining predictive performance. The framework supports intersectional fairness by addressing several sensitive attributes simultaneously, promoting transparent and accountable AI systems that serve diverse populations more equitably.



M. Fatoreto, G. Özbulak, L. Berton, and A. Anjos, Optimizing Fairness and Utility in Healthcare Machine Learning Models, *Applied Soft Computing*, 2025

Decoding Community Proximity Discourse: Local and National News Are Not Created Equal

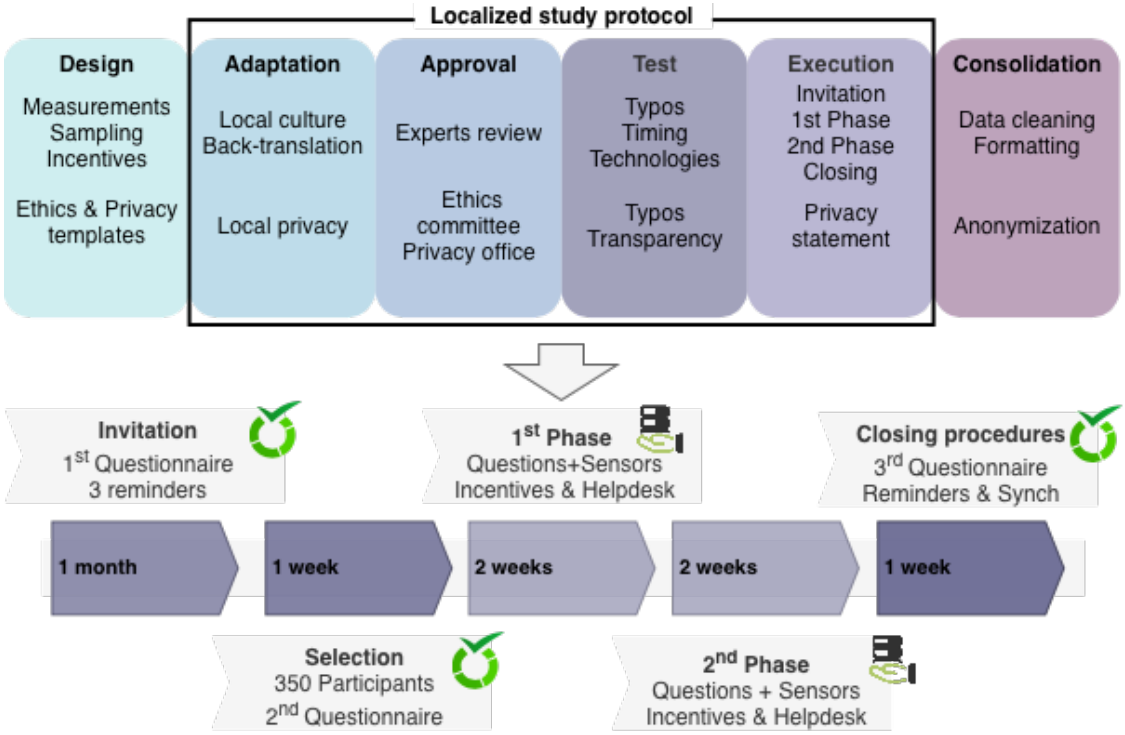
Understanding the specific characteristics of local news, and how these differ from other news sources, is important to develop human-centered approaches that support local news organizations and local readers. Through the combination of research methods, including the automatic extraction of linguistic indicators (readability level, entities, topics, etc.) and the use of qualitative methods from discourse and frame analysis, we conducted a comparison of narrative strategies and community engagement using articles published by Swiss local newspapers in the Romandy region and in Swiss national newspapers. Our analysis, which involved the generation of a new dataset of local news articles, showed important differences in narrative construction and audience engagement. While local news articles foster a sense of community and direct connection with readers, national news articles employ wider-ranging and more elaborate storytelling to appeal to broader audiences. Our work shows that different editorial choices contribute to shape identity and participation in media and highlights the importance of a multi-level perspective on media strategies and audience dynamics, at a time of rapid editorial and economic transformations.



V. Bros and D. Gatica-Perez, “Decoding Community Proximity Discourse: A Mixed-Methods Comparative Analysis of Online Local and National Newspapers in Romandy, Switzerland,” *PLOS ONE*, Vol. 20, No. 8, 2025.

DiversityOne: A Multi-Country Smartphone Sensor Dataset for Everyday Life Behavior Modeling

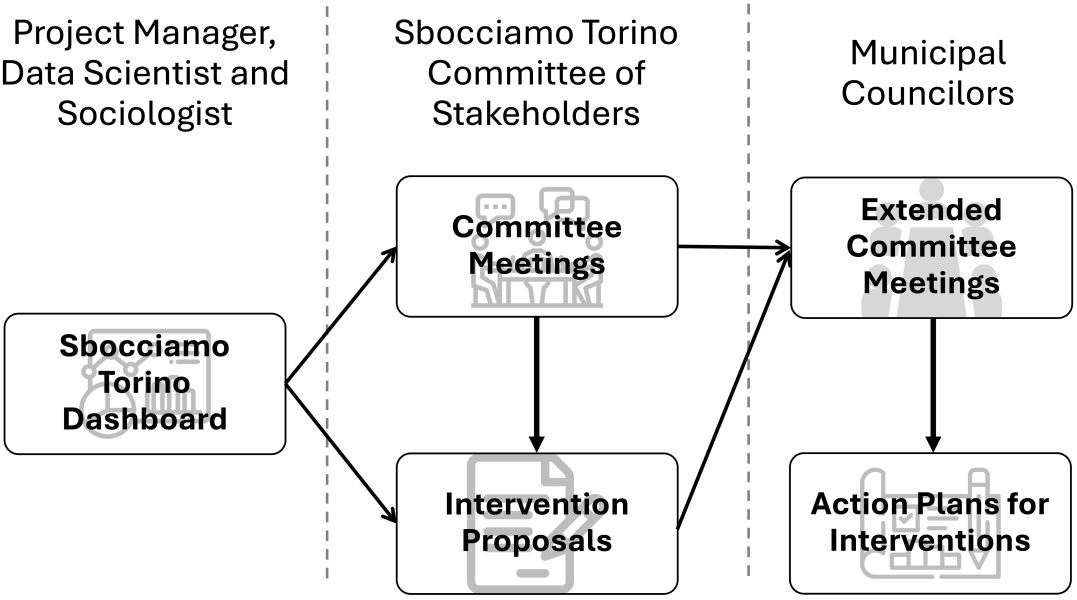
Understanding everyday life behavior of young adults through personal devices, e.g., smartphones and smartwatches, is key for various applications, from enhancing user experience to enabling interventions in digital health apps. Towards this goal, previous studies have relied on datasets combining passive sensor data with self-reports. However, such datasets have focused on specific countries primarily in the Global North, involved few participants, or used a limited range of sensors. These limitations restrict the ability to capture cross-country variations of human behavior, including the possibility of studying model generalization, and robustness. To address this gap, we introduced DiversityOne, a dataset which spans eight countries (China, Denmark, India, Italy, Mexico, Mongolia, Paraguay, and the United Kingdom) and includes data from 782 college students over a period of four weeks. DiversityOne contains data from 26 smartphone sensor modalities and 350,000 self-reports. The dataset is one of the largest and most diverse publicly available academic datasets, opening the possibility of studying research problems in ubiquitous computing, particularly in domain adaptation and generalization across countries.



M. Busso, A. Bontempelli, L.J. Malcotti, L. Meegahapola, P. Kun, S. Diwakar, C. Nutakki, M. D. Rodas Britez, H. Xu, D. Song, S. Ruiz-Correa, A. R. Mendoza-Lara, G. Gaskell, S. Stares, M. Bidoglia, A. Ganbold, A. Chagnaa, L. Cernuzzi, A. Hume, R. Chenu-Abente, R. A. Asiku, I. Kayongo, D. Gatica-Perez, A. de Gotzen, I. Bison, and F. Giunchiglia, DiversityOne: A Multi-Country Smartphone Sensor Dataset for Everyday Life Behavior Modeling. *PACM on Interactive, Mobile, Wearable, and Ubiquitous Technologies (IMWUT)*, Vol. 9, No. 1, Article 1, 2025.

Co-Designing a Community-Centered Process to Understand Youth Deviance with Multiple Stakeholders and Datasets

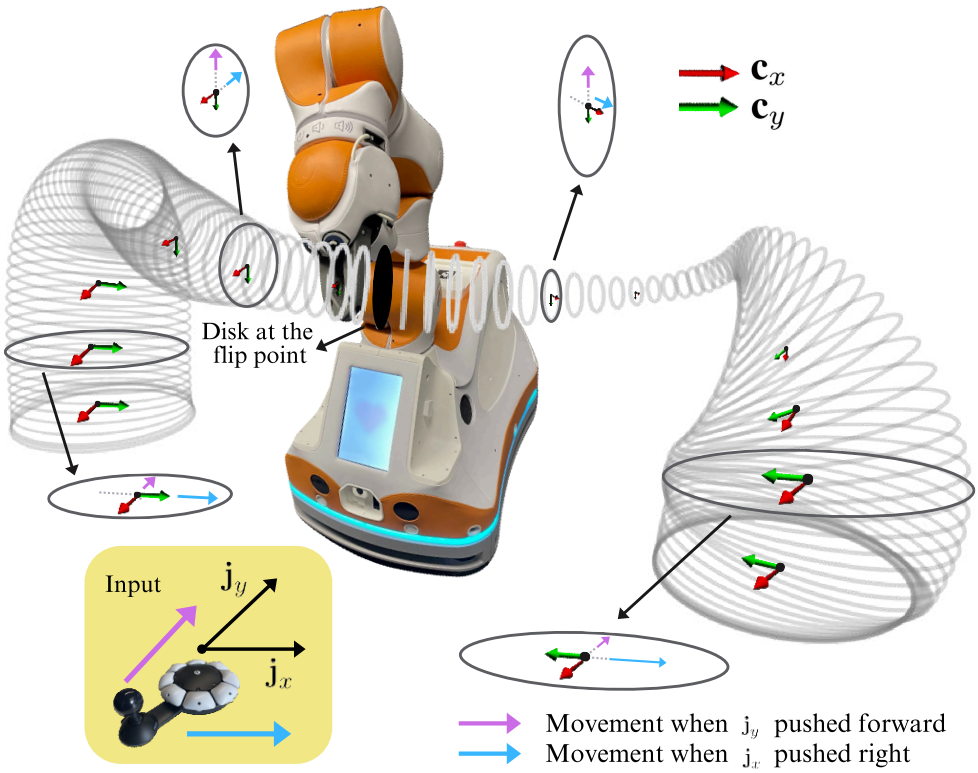
We created *Sbocciamo Torino*, a local civic tool that helps understand and act upon issues related to youth deviance in the city of Turin, Italy. The tool was developed in partnership with municipal authorities, law enforcement, NGOs, and social services, and reflects their institutional priorities while centering community knowledge. Rooted in research through design and participatory design methodologies, the tool integrates a data dashboard, a stakeholder committee, and structured co-design sessions to facilitate collaborative data analysis and intervention planning. The tool's impact on stakeholder trust, collaboration, and decision-making was assessed. Our findings show that stakeholders valued the inclusive design approach and data-driven collaboration, while revealing barriers in communication, data literacy, and operational coordination. Furthermore, political and institutional support were identified as critical to the tool's success. Our work contributes to research on community technologies by demonstrating how civic tools can be collaboratively developed to navigate wicked social problems through participatory design.



R. Annapureddy, A. Fornaroli, M. Fattori, V. Lacovara, E. Fiori, S. Vollmer, M. Konradi, B. E. Hecking, G. Todesco, and D. Gatica-Perez, "Co-Designing with Multiple Stakeholders and Datasets: A Community-Centered Process to Understand Youth Deviance in the Italian City of Turin," in *Proc. Int. Conf. on Communities and Technologies (C&T)*, Siegen, 2025.

Designing Shared Autonomy Framework for Assistive robots

We study how to make robot control more accessible for people with motor impairments, focusing on mapping simple 2D joystick inputs to complex 6D robot motions. We propose a dynamic input mapping framework that adapts joystick commands to robot trajectories through control frames defined along canal surfaces, making the interaction smoother and more intuitive. In a user study with 20 participants we show that our method lowers workload and improves performance compared to a baseline mapping. To ensure accessibility, we adapted our interfaces and validated our system with wheelchair users in realistic daily-living and creative tasks. This work contributes to the "AI for Everyone" program by promoting inclusive technology design, and to "Human-AI Teaming" by fostering intuitive and cooperative interaction between humans and robotic systems.

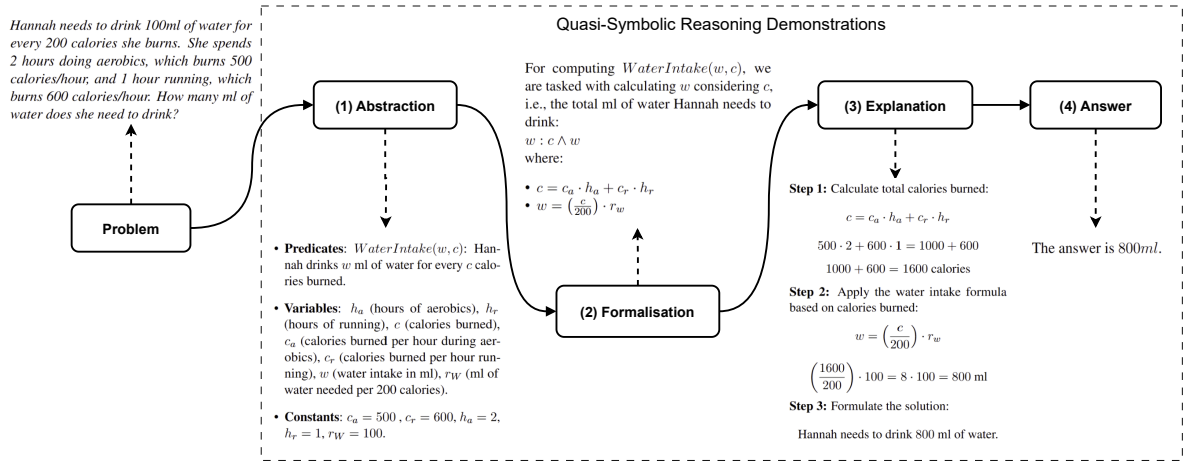


S. Rajapakshe, J.M. Odobez, and E. Senft. "Giving sense to inputs: Toward an accessible control framework for shared autonomy". *In proceedings of the 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 213-221), IEEE, 2025.

IMPROVING CHAIN-OF-THOUGHT FOR ABSTRACT REASONING

We start from a limitation of standard Chain-of-Thought (CoT) prompting: even when a language model produces step-by-step explanations, those steps are still written in free-form natural language and remain vulnerable to content effects, spurious associations, and shallow pattern-matching. Fully symbolic pipelines with external solvers can mitigate these issues, but they require complete translation from natural language into a formal language, which is costly, brittle, and hard to generalize across tasks. Our goal is to find a middle ground: a way to *partially* formalize the reasoning process so that we disentangle logical structure from surface content, while keeping the flexibility and coverage of natural language.

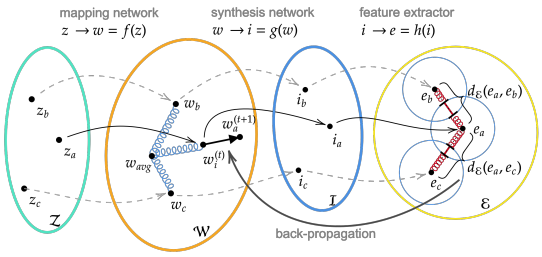
To achieve this, we introduce QuaSAR (Quasi-Symbolic Abstract Reasoning), a variation of CoT that asks models to operate through *quasi-symbolic* explanations. Our results show that quasi-symbolic abstractions consistently improve both accuracy and robustness on reasoning tasks, boosting performance on both mathematical and natural language tasks and improving resilience to adversarial variations for both Large and smaller language models. Overall, the work suggests that we can substantially strengthen CoT reasoning by injecting light-weight, quasi-symbolic structure into the explanation process, capturing many of the benefits of symbolic guidance without sacrificing the flexibility and scalability of natural language prompting.



L. Ranaldi, M. Valentino, A. Polonsky, and A. Freitas, Improving chain-of-thought reasoning via quasi-symbolic abstractions, *Annual Meeting of the Association for Computational Linguistics*, 2025

SYNTHETIC FACE DATASETS GENERATION VIA PHYSICS-INSPIRED MACHINE LEARNING

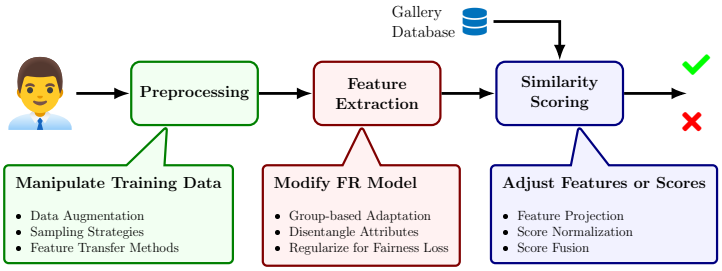
We study the ethical and technical challenges posed by using real face data in biometric systems, and we propose a novel physics-inspired method to generate large, diverse, and privacy-preserving synthetic face datasets. We treat synthetic identities as particles in a latent space, subjected to stochastic Brownian dynamics and repulsive interactions, and we develop algorithms to sample both inter-class diversity and intra-class variation. By training face recognition models on these synthetic datasets, we demonstrate that our method yields performance on par with state-of-the-art diffusion-based synthetic data, while substantially reducing the risk of memorizing real identities.



D. Geissbuhler, H. Otroshi, and S. Marcel, Synthetic Face Datasets Generation via Latent Space Exploration from Brownian Identity Diffusion, *in ICML*, 2025.

Demographic Fairness in Face Recognition

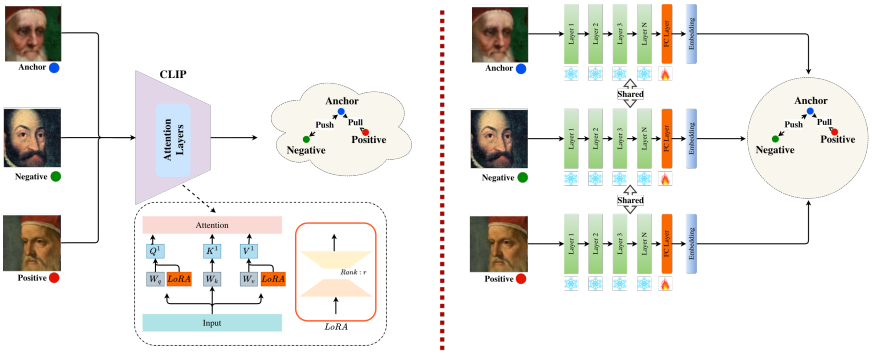
We study the challenge of demographic fairness in face recognition systems and propose a comprehensive review of the topic. We systematically examine how disparities in recognition performance across demographic groups arise, covering the root causes, available datasets, evaluation metrics, and mitigation strategies. Our work brings together a wide body of research into a unified framework that highlights current progress while identifying persistent gaps and emerging challenges in building equitable and trustworthy biometric systems.



K. Kotwal, and S. Marcel, Review of Demographic Fairness in Face Recognition, *in IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2025.

Historical Portrait Face Identification

We study the problem of identifying sitters in historical paintings, where conventional face recognition models struggle due to the large domain gap between photographic and artistic portraits. We propose ArtFace, a method that adapts foundation model embeddings to the domain of painted portraits and fuses them with classical face recognition networks, thereby improving recognition accuracy across stylistic variations. By carefully fine-tuning and combining representations, we show that our approach bridges the gap between photographic and artistic data and significantly advances the state of the art in sitter identification.

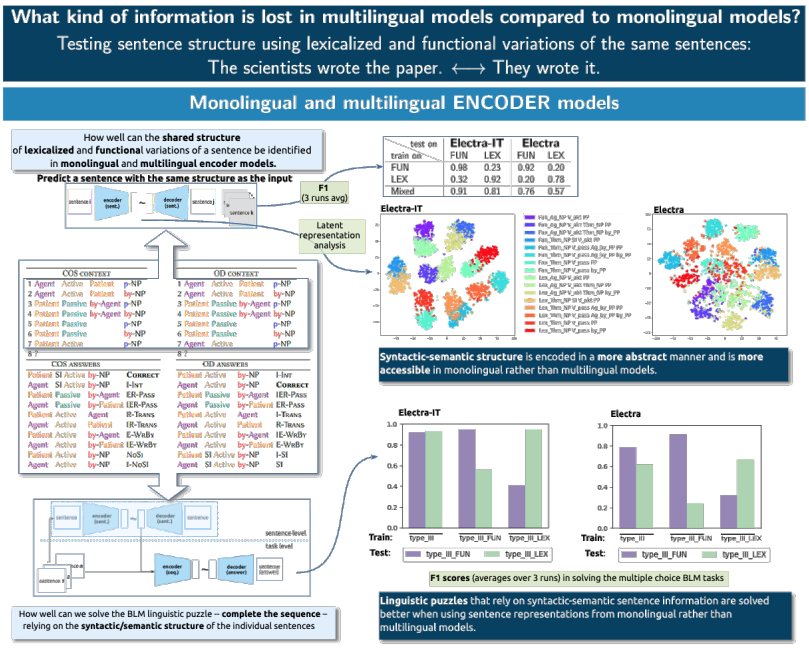


(a) CLIP-LoRA Adaptation (b) FR Model Adaptation

F. Poh, A. George, and S. Marcel, ArtFace: Towards Historical Portrait Face Identification via Model Adaptation, *in ICCC Workshop ArtMetrics*, 2025.

MULTILINGUAL VS. MONOLINGUAL MODELS: LINGUISTIC STRUCTURE AND LEXICAL ABSTRACTION

Multilingual language models are attractive, as they allow us to cross linguistic boundaries, and solve tasks in different languages in the same mathematical space. They come, however, at a cost: in the quest to find a shared space that satisfies (to a certain degree) all languages, the resulting representations lose, or fail to capture, properties specific to each language. We present an investigation into detecting linguistic structure through lexical abstraction. We study both a multilingual and a monolingual language model, and quantify the loss of information between them. Work investigating the impact of linguistic diversity contributes to RP4 as it sharpens our understanding of the variety of multilingual challenges for diverse and low-resource languages to widen access to technology.



Nastase, V., Samo, G., Jiang, C. and Merlo, P., Multilingual vs. monolingual transformer models in encoding linguistic structure and lexical abstraction, *CLiC-it 2025: Eleventh Italian Conference on Computational Linguistics*, Cagliari, Italy.

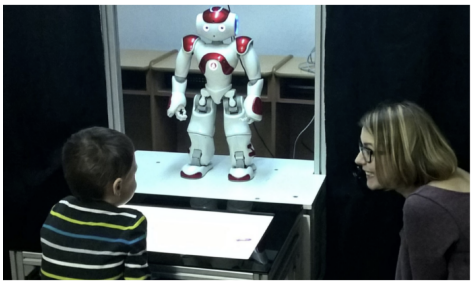
INCLUSIVE SPEECH ENHANCEMENT WITH FEW-SHOT MODEL ADAPTATION

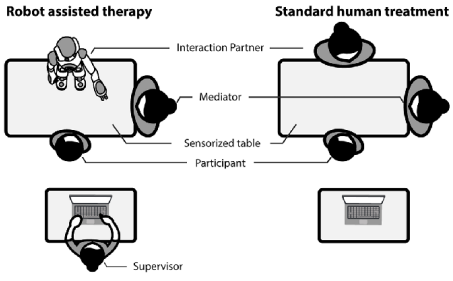
Speech enhancement systems are widely used to improve the clarity of speech in applications from hearing aids to telecommunication, but they often fail to generalize to voices with atypical acoustic patterns, such as those of speakers with neurological disorders. We examine the limitations of a hybrid variational autoencoder–non-negative matrix factorization (VAE–NMF) model for speech enhancement, showing that models trained on large datasets of typical speech degrade substantially when applied to pathological speech. To overcome this challenge, we investigate fine-tuning and personalized enhancement strategies in which a pretrained VAE is adapted using only a few seconds of clean speech from an individual speaker. Our results demonstrate that these personalized models significantly improve enhancement quality for both typical and pathological voices, yielding more consistent performance across speaker types and underscoring the potential of personalization to make speech enhancement more inclusive and effective for everyone.

M. Hou and I. Kodrasi, Variational Autoencoder for Personalized Pathological Speech Enhancement, *European Signal Processing Conference*, 2025

EFFICACY AND EFFECTIVENESS OF ROBOT-ASSISTED THERAPY FOR AUTISM SPECTRUM DISORDER: FROM LAB TO REALITY

Despite the multiple decades of research showing promising lab results in social robots in therapy for children with autism spectrum disorder, rigorous evidence of real-world impact remains limited. In our paper “Efficacy and effectiveness of robot-assisted therapy for autism spectrum disorder: From lab to reality,” we systematically evaluated robot-assisted interventions in both controlled clinical settings and more naturalistic environments such as schools and homes to close the gap between experimental designs and practical deployment. Through two randomized clinical trials involving young children, we show that robot-assisted therapy can achieve outcomes equivalent to standard treatment while significantly increasing engagement in the clinical context, and that simplified, portable robot setups can deliver comparable benefits outside the clinic, demonstrating both efficacy and effectiveness of these systems. These findings provide robust, empirical support for the potential of socially assistive robots to contribute meaningfully to developmental therapy and inform future translational research in human-robot interaction.





D. David, P. Baxter, T. Belpaeme, E. Billing, H. Cai, H.-L. Cao, A. Ciocan, C. Costescu, D. H. Garcia, P. G. Esteban, J. Kennedy, H. Liu, S. Matu, A. Mazel, M. Selescu, E. Senft, S. Thill, B. Vanderborght, D. Vernon, and T. Ziemke, Efficacy and effectiveness of robot-assisted therapy for autism spectrum disorder: From lab to reality, *Science Robotics*, vol. 10, no. 109, 2025

Research Staff

Group Leaders

Dr. André Anjos
Medical Artificial Intelligence Group

Olivier Bornet
R&D Team

Dr. Sylvain Calinon
Robot Learning & Interaction Group

Dr. André Freitas
Neuro-symbolic AI Group

Dr. Philip N. Garner
Audio Inference Group

Prof. Daniel Gatica-Perez
Social Computing Group

Dr. James Henderson
Natural Language Understanding Group

Dr. Jérôme Kämpf
Energy Informatics Group

Dr. Ina Kodrasi
Signal Processing for Communication Group

Dr. Mathew Magimai-Doss
Speech & Audio Processing Group

Prof. Sébastien Marcel
Biometrics Security & Privacy Group

Prof. Paola Merlo
Computational Learning and Computational Linguistics Group

Dr. Petr Motlicek
Speech & Audio Processing Group

Dr. Jean-Marc Odobez
Perception & Activity Understanding Group

Dr. Emmanuel Senft
Human-centered Robotics and AI Group

Dr. Damien Teney
Machine Learning Group

External Research Fellows

Prof. Raphaëlle Luisier
AI and Data Science for RNA Biology Lab
University of Bern

Prof. Lonneke van der Plas
Computation, Cognition & Language Group
Università della Svizzera italiana

R&D Engineers

Name	Position	Nationality	Supervisor
ABBET Philip	Senior R&D Engineer	Switzerland	BORNET Olivier
ALONSO DEL BARRIO David	R&D Engineer	Spain	BORNET Olivier
BEN MAHMOUD Imen	R&D Engineer	Tunisia	BORNET Olivier
CANÉVET Olivier	Senior R&D Engineer	Switzerland	BORNET Olivier
CARRON Daniel	R&D Engineer	Switzerland	BORNET Olivier
CLIVAZ Guillaume	Senior R&D Engineer	Switzerland	BORNET Olivier
DAYER Yannick	Senior R&D Engineer	Switzerland	BORNET Olivier
DROZ William	Senior R&D Engineer	Switzerland	BORNET Olivier
GAIST Samuel	Senior R&D Engineer	Switzerland	BORNET Olivier
GENTILHOMME Théophile	Senior R&D Engineer	France	BORNET Olivier
KAYAL Salim	Senior R&D Engineer	Switzerland	BORNET Olivier
KHALIL Driss	R&D Engineer	Morocco	BORNET Olivier
MACEIRAS Jérémy	Senior R&D Engineer	Switzerland	BORNET Olivier
MARCEL Christine	Senior R&D Engineer	Switzerland	BORNET Olivier
MAYORAZ André	R&D Engineer	Switzerland	BORNET Olivier
MICHEL Samuel	R&D Engineer	Belgium	BORNET Olivier
MOHAMMADI Amir	R&D Engineer	Iran	BORNET Olivier
NANCHEN Alexandre	Senior R&D Engineer	Switzerland	BORNET Olivier
VERZAT Colombine	Senior R&D Engineer	France	BORNET Olivier
ZANGGER Alicia	R&D Engineer	Switzerland	BORNET Olivier

PhD Students

Name	Position	Nationality	Supervisor
ABDENNADHER Yesmine	Doctoral Researcher	Tunisia	GARNER Philip
ALAVI Kimia	Doctoral Researcher	Iran	HENDERSON James
AMIRI Mahdi	Doctoral Researcher	Iran	KODRASI Ina
ANNAPUREDDY Ravinithesh Reddy	Doctoral Researcher	India	GATICA-PEREZ Daniel
BARANOUSKAYA Darya	Doctoral Researcher	Belarus	MARCEL Sébastien
BECKMANN Pierre	Doctoral Researcher	Switzerland	FREITAS André
BILALOGLU Cem	Doctoral Researcher	Turkey	CALINON Sylvain
BROS Victor	Doctoral Researcher	France	GATICA-PEREZ Daniel
CHEN Haolin	Doctoral Researcher	China	GARNER Philip
COMAN Andrei	Doctoral Researcher	Italy	HENDERSON James
COPPIETERS DE GIBSON Louise	Doctoral Researcher	Belgium	GARNER Philip
EL HAJAL Karl	Doctoral Researcher	Lebanon	MAGIMAI DOSS Mathew
EL ZEIN Dina	Doctoral Researcher	Lebanon	HENDERSON James
GANDARELA DE SOUZA João Pedro	Doctoral Researcher	Brazil	FREITAS André
ESPINOSA MENA Jose Rafael	Doctoral Researcher	Switzerland	HENDERSON James
FEHR Fabio	Doctoral Researcher	Switzerland	HENDERSON James
GUPTE Vivek	Doctoral Researcher	India	SENFT Emmanuel
HAIDAR Nawal	Doctoral Researcher	Lebanon	ODOBEZ Jean-Marc

HE Mutian	Doctoral Researcher	China	GARNER Philip
HOU Mingchi	Doctoral Researcher	China	KODRASI Ina
HRYNENKO Olena	Doctoral Researcher	Ukraine	MARCEL Sébastien
ISMAYILZADA Mahammad	Doctoral Researcher	Canada	VAN DER PLAS Lonneke
JIANG Chunyang	Doctoral Researcher	China	MERLO Paola
JIANG Liangze	Doctoral Researcher	China	TENEY Damien
JUNG Vincent	Doctoral Researcher	France	FREITAS André
LI Yiming	Doctoral Researcher	China	CALINON Sylvain
KUMAR Shashi	Doctoral Researcher	India	MOTLICEK Petr
LU Cen	Doctoral Researcher	China	MARCEL Sébastien
MARIĆ Ante	Doctoral Researcher	Croatia	CALINON Sylvain
MOHR Isabelle	Doctoral Researcher	Germany	FREITAS André
MUKHERJEE Anirban	Doctoral Researcher	India	ODOBEZ Jean-Marc
NAFEZ Mojtaba	Doctoral Researcher	Iran	HENDERSON James
ÖZBULAK Gokhan	Doctoral Researcher	Turkey	ANJOS André
PETERSEN Molly	Doctoral Researcher	United States	VAN DER PLAS Lonneke
RAHIMI NOSHANAGH Parsa	Doctoral Researcher	Iran	MARCEL Sébastien
RAJAPAKSHE Shalutha	Doctoral Researcher	Sri Lanka	SENFT Emmanuel
RAZMJOO FARD Amirreza	Doctoral Researcher	Iran	CALINON Sylvain
SARKAR Eklavya	Doctoral Researcher	Switzerland	MAGIMAI DOSS Mathew
SCHONGER Martin	Doctoral Researcher	Austria	CALINON Sylvain
SHIRAKAMI Haruki	Doctoral Researcher	Japan	HENDERSON James
TAFASCA Samy	Doctoral Researcher	Morocco	ODOBEZ Jean-Marc
TANG Yung-Chen	Doctoral Researcher	Taiwan	MARCEL Sébastien
TARIGOPULA Neha	Doctoral Researcher	India	MAGIMAI DOSS Mathew
THORBECKE (NIGMATULINA) Iuliia	Doctoral Researcher	Russian Federation	MOTLICEK Petr
VIDA Elija	Doctoral Researcher	Austria	GARNER Philip
VUILLECARD Pierre	Doctoral Researcher	Switzerland	ODOBEZ Jean-Marc
WATAWANA Hasindri	Doctoral Researcher	Sri Lanka	VILLATORO TELLO Esaú
XUE Teng	Doctoral Researcher	China	CALINON Sylvain
XU Lei	Doctoral Researcher	China	FREITAS André
ZHANG Yan	Doctoral Researcher	China	CALINON Sylvain

Post-doctoral Researchers

Name	Position	Nationality	Supervisor
AKSTINAITE Vita	Postdoctoral Researcher	Lithuania	GARNER Philip
BAÑERAS-ROUX Thibault	Postdoctoral Researcher	France	MOTLICEK Petr
CAROFILIS VASCO Roberto Andrés	Postdoctoral Researcher	Spain	MOTLICEK Petr
COLBOIS Laurent	Postdoctoral Researcher	Switzerland	MARCEL Sébastien
DELMAS Maxime	Postdoctoral Researcher	France	FREITAS André
FUCHS Michael	Postdoctoral Researcher	Switzerland	ODOBEZ Jean-Marc
GUPTA Anshul	Postdoctoral Researcher	India	ODOBEZ Jean-Marc
HERMANN Enno	Postdoctoral Researcher	Germany	MAGIMAI DOSS Mathew
HERMUS James	Postdoctoral Researcher	United States	CALINON Sylvain
KALOGA Yacouba	Postdoctoral Researcher	France	KODRASI Ina
KHAN Asif Hussain	Postdoctoral Researcher	Pakistan	MARCEL Sébastien
LÖW Tobias	Postdoctoral Researcher	Switzerland	CALINON Sylvain
KULKARNI Ajinkya	Postdoctoral Researcher	India	MAGIMAI DOSS Mathew
LUÉVANO GARCÍA Luis Santiago	Postdoctoral Researcher	Mexico	MARCEL Sébastien
MITRO Ioanni	Postdoctoral Researcher	Greece	TENEY Damien
OTROSHI SHAHREZA Hatef	Postdoctoral Researcher	Iran	MARCEL Sébastien
ÖZTÜRK Ünsal	Postdoctoral Researcher	Turkey	MARCEL Sébastien
PALLANCA Olivier	Postdoctoral Researcher	France	ANJOS André
PRASAD Amrutha	Postdoctoral Researcher	India	MOTLICEK Petr
RANALDI Leonardo	Postdoctoral Researcher	Italy	HENDERSON James
RANGAPPA Pradeep	Postdoctoral Researcher	India	MOTLICEK Petr
SANCHEZ-CORTES Dairazalia	Postdoctoral Researcher	Mexico	MOTLICEK Petr
TORNAY Sandrine	Postdoctoral Researcher	Switzerland	MAGIMAI DOSS Mathew
VALENTINO Marco	Postdoctoral Researcher	Italy	FREITAS André
VAN DER MEER Michiel	Postdoctoral Researcher	Netherlands	VAN DER PLAS Lonneke
VÁSQUEZ RODRÍGUEZ Laura	Postdoctoral Researcher	Costa Rica	VAN DER PLAS Lonneke
VIDIT Vidit	Postdoctoral Researcher	India	MARCEL Sébastien
WANG Hengfei	Postdoctoral Researcher	China	ODOBEZ Jean-Marc
WYSOCKI Oskar	Postdoctoral Researcher	Poland	FREITAS André
XU Zhi Ming	Postdoctoral Researcher	Canada	LUISIER Raphaëlle

Research Associates

Name	Position	Nationality	Supervisor
BHATTACHARJEE Sushil	Research Associate	Switzerland	MARCEL Sébastien
BURDISSO Sergio	Research Associate	Argentina	MOTLICEK Petr
ECABERT Christophe	Research Associate	Switzerland	MARCEL Sébastien
GEORGE Anjith	Research Associate	India	MARCEL Sébastien
GEISSBUHLER David	Research Associate	Switzerland	SENFT Emmanuel
HOVSEPYAN Sevada	Research Associate	Armenia	MAGIMAI DOSS Mathew
JIMÉNEZ DEL TORO Oscar	Research Associate	Mexico	ANJOS André
KOMATY Alain	Research Associate	France	MARCEL Sébastien

KORSHUNOV Pavel	Research Associate	Estonia	MARCEL Sébastien
KRIVOKUĆA HAHN Vedrana	Research Associate	Croatia	MARCEL Sébastien
MURALIDHAR Skanda	Research Associate	India	MAGIMAI DOSS Mathew
NASTASE Viviana-Antonela	Research Associate	Romania	MERLO Paola
SAMO Giuseppe	Research Associate	Italy	MERLO Paola
TIMONINA-FARKAS Anna	Research Associate	Russian Federation	SENFT Emmanuel
VILLAMIZAR Michael	Research Associate	Spain	ODOBEZ Jean-Marc
VILLATORO TELLO Esaú	Research Associate	Mexico	MOTLICEK Petr

Research Assistants

Name	Position	Nationality	Supervisor
AL AMINE Zeina	Research Assistant	France	MARCEL Sébastien
KIM Haeeun	Research Assistant	South Korea	GATICA-PEREZ Daniel
NEUGBER Samuel	R&D / Research Assistant	Germany	MARCEL Sébastien
OUTUMURO BUENO Alberto	Research Assistant	Spain	MARCEL Sébastien
POLAC Magdalena	Research Assistant	Poland	MARCEL Sébastien
PULVIRENTI Roberto	Research Assistant	Italy	ANJOS André
ULUCAN Ibrahim	Research Assistant	Turkey	MARCEL Sébastien

Research Interns and Academic Visitors

Name	Position	Nationality	Supervisor
ABDALAZIM Nouran Ayman Roushdy	Research Intern	Egypt	GATICA-PEREZ Daniel
ACHAKRI Maximilian	Research Intern	Switzerland	ANJOS André
AGRAWAL Sanat	Research Intern	India	MAGIMAI DOSS Mathew
ALMOMANI Ahmad Qasim Mohammad	Research Intern	Jordan	KÄMPF Jérôme
ARMANI RENZO Matheus	Research Intern	Brazil	ANJOS André
BAIA Alina-Elena	Research Intern	Romania	MARCEL Sébastien
BAKKER Saray	Research Intern	Netherlands	CALINON Sylvain
BARDEL Lola	Research Intern	France	LUISIER Raphaëlle
BAROUDI Séverin	Research Intern	France	MOTLICEK Petr
BEBER Luca	Research Intern	Italy	CALINON Sylvain
BISOGNO BERNARDINI Leonardo	Research Intern	Italy	KÄMPF Jérôme
BOGHETTI Roberto	Research Intern	Italy	KÄMPF Jérôme
BOLLENGIER Alexis	Research Intern	Belgium	ANJOS André
BRADICIC Manuel	Research Intern	Croatia	LUISIER Raphaëlle
CATIC Hana	Research Intern	Bosnia and Herzegovina	CALINON Sylvain
CESPEDES SARRIAS Berta	Research Intern	Spain	ROSSA Christophe
CHAPARINIYA Masoumeh	Research Intern	Iran	ODOBEZ Jean-Marc
CHEREMETIEV Vassiliy	Research Intern	Switzerland	ROSSA Christophe
COLLADO CAPELL Carlos	Research Intern	Spain	ROSSA Christophe
COSTA BARBOSA Lucas	Research Intern	Brazil	ANJOS André
DARWICHE Nael	Research Intern	Switzerland	CALINON Sylvain
PUROHIT Tilak	Research Intern	India	MAGIMAI DOSS Mathew

DE WAHA-BAILLONVILLE Gilles	Research Intern	Belgium	ODOBEZ Jean-Marc
DUPONT-ROC Maud	Research Intern	France	LUISIER Raphaëlle
EL-HOIYDI Felix	Research Intern	Switzerland	LUISIER Raphaëlle
ESMAEILY Abolghasem	Research Intern	Sweden	SENFT Emmanuel
EUTAMENE Camil	Research Intern	France	ODOBEZ Jean-Marc
FARKHONDEH Arya	Research Intern	Iran	ODOBEZ Jean-Marc
FERNANDEZ Solène	Research Intern	Switzerland	MARCEL Sébastien
FOURNIER Lisa	Research Intern	Switzerland	LUISIER Raphaëlle
FRATTI Riccardo	Research Intern	Italy	LUISIER Raphaëlle
GANSER Léo	Research Intern	Switzerland	LUISIER Raphaëlle
GATTI Martina	Research Intern	Italy	GATICA-PEREZ Daniel
GIRAUD Raphaël	Research Intern	France	ODOBEZ Jean-Marc
HIDRI Yasmine	Research Intern	Tunisia	LUISIER Raphaëlle
IBELINGS Roemer	Research Intern	Netherlands	GATICA-PEREZ Daniel
JAKAC Karlo	Research Intern	Croatia	CALINON Sylvain
JULLIEN Maël	Research Intern	France	FREITAS André
KAKAR Ahmad Bilal	Research Intern	Afghanistan	CAVALLARO Andrea
KOPP Damian	Research Intern	Switzerland	ANJOS André
KOTWAL Ketan	Research Intern	India	MARCEL Sébastien
KOVALENKO Sophia	Research Intern	Switzerland	GARNER Philip
LANCA Luka	Research Intern	Croatia	CALINON Sylvain
LANFRANCONI Michele	Research Intern	Italy	CALINON Sylvain
LIU Shiran	Research Intern	China	MOTLICEK Petr
LUO Yongkang	Academic Visitor	China	CALINON Sylvain
LUTZIGER Daniel	Research Intern	Switzerland	ANJOS André
MENENDEZ RUIZ DE AZUA Pablo	Research Intern	Spain	ANJOS André
MESSORI Elisa	Research Intern	Italy	LUISIER Raphaëlle
NGO Quang Long Ho	Research Intern	Switzerland	ROSSA Christophe
OLDENBURGER Laura	Research Intern	Netherlands	MARCEL Sébastien
PANNATIER Yvan	Research Intern	Switzerland	ANJOS André
PARK Gyeonglee	Research Intern	South Korea	LUISIER Raphaëlle
PIRO Matteo	Research Intern	Italy	KÄMPF Jérôme
POH François	Research Intern	United Kingdom	MARCEL Sébastien
RAISSI Tina	Research Intern	Italy	MAGIMAI DOSS Mathew
RAMIREZ CONTRERAS Hector Manuel	Research Intern	Mexico	GATICA-PEREZ Daniel
RES Jakub	Academic Visitor	Czech Republic	MOTLICEK Petr
RIMOLDI Emanuele	Research Intern	Italy	GARNER Philip
RODENAS RUIZ Pablo	Research Intern	Spain	ROSSA Christophe
RODRIGUES TORRAO Filipa	Research Intern	Portugal	LUISIER Raphaëlle
SAFTA Cristian	Research Intern	Romania	GATICA-PEREZ Daniel
SANCHEZ LARA Alejandra	Research Intern	Mexico	MOTLICEK Petr
SHAN Jennifer	Research Intern	Switzerland	GARNER Philip
SHAPOSHNIKOV Roman	Research Intern	Russian Federation	LUISIER Raphaëlle
SINGH VIRK Renuka	Research Intern	Switzerland	GATICA-PEREZ Daniel
SPAHR Guillaume	Research Intern	Switzerland	CALINON Sylvain
SYLA Valmir	Research Intern	Switzerland	KODRASI Ina
VERNHES Albin	Research Intern	France	LUISIER Raphaëlle

WANG Anlan	Research Intern	China	ODOBEZ Jean-Marc
WEN Jing	Research Intern	China	GATICA-PEREZ Daniel
WYSOCKA Magdalena	Research Intern	Poland	FREITAS André
XIE Qilin	Research Intern	China	LUISIER Raphaëlle
XI Lei	Research Intern	China	SENFT Emmanuel
XU Zhi Ming	Research Intern	Canada	LUISIER Raphaëlle
YAZBECK Rudolf	Research Intern	Lebanon	LUISIER Raphaëlle
YONG Zi Qian	Research Intern	Malaysia	MARCEL Sébastien
ZATEZALO Sara	Research Intern	Serbia	CAVALLARO Andrea
ZHANG Yingji	Research Intern	China	FREITAS André

Projects

Projects active in 2025 are grouped into three categories, namely Research Projects, Innovation Projects, and International Projects. We also list the projects accepted in 2025 and starting the following year.

Research Projects

1. Name **AI2PUB** (AI Insights to the Public through bi-directional media channels)
   Funding SNSF Agora
   Coordinator HESSO - Geneve
   Duration 2025-04-01 - 2028-03-31
   Partner(s) RGB Project; Universita della Svizzera italiana; Compáz; Idiap Research Institute

2. Name **AI4AUTISM2** (Digital Phenotyping of Autism Spectrum Disorders in Children)
   Funding SNSF Sinergia
   Coordinator University of Geneva
   Duration 2021-11-01 - 2026-04-30
   Partner(s) University of Applied Sciences and Arts of Southern Switzerland (SUPSI); Idiap Research Institute

3. Name **AI-LITERACY** (AI Literacy and Literacy with AI: From Effectiveness to Efficiency)
   Funding Etat du Valais
   Coordinator Idiap Research Institute
   Duration 2025-04-15 - 2026-05-31
   Partner(s) Haute Ecole Pédagogique du Valais

4. Name **AIML-VISIT** (Research Visit at the University of Adelaide: Advancing Our Understanding of Modern AI and Machine Learning Models)
   Funding SNSF Scientific Exchanges
   Coordinator Idiap Research Institute
   Duration 2024-12-01 - 2025-03-31
   Partner(s) Unviersity of Adelaide

5. Name **BOVINE** (Nonparametric Variational Representation Learning for Natural Language Understanding)
   Funding SNSF Project Funding
   Coordinator Idiap Research Institute
   Duration 2024-10-01 - 2028-09-30
   Partner(s) –

6. Name **CHASPEEPRO** (Characterisation of motor speech disorders and processes)
   Funding SNSF Sinergia
   Coordinator University of Geneva
   Duration 2021-12-01 - 2026-11-30
   Partner(s) Geneva University Hospitals; New Sorbonne University Paris 3; Idiap Research Institute

7. Name **C-LING** (Towards Creative systems with LINGuistic modelling)
   Funding SNSF Division II: Mathematics, Natural sciences and Engineering
   Coordinator Idiap Research Institute
   Duration 2022-09-01 - 2025-02-28
   Partner(s) –

8. Name **COLLABCLOUD** (A Collaborative Research Cloud Infrastructure powering Discovery and Exchange)
   Funding SNSF R'Equip
   Coordinator Idiap Research Institute
   Duration 2023-06-01 - 2025-12-31
   Partner(s) –

9. Name **DISLING** (Disentangling linguistic intelligence: automatic generalisation of structure and meaning across languages)
   Funding SNSF Advanced Grant
   Coordinator Idiap Research Institute
   Duration 2022-08-01 - 2026-07-31
   Partner(s) –

10. Name **EMIL** (Emotion in the loop – a step towards a comprehensive closed-loop deep brain stimulation in Parkinson's disease)
    Funding SNSF Bridge Discovery
    Coordinator University of Bern
    Duration 2021-05-01 - 2025-10-31
    Partner(s) Centre suisse d'électronique et de microtechnique; Idiap Research Institute

11. Name **EVOLANG-2** (Evolving Language Phase 2)
    Funding SNSF NCCR 5th series
    Coordinator University of Geneva; University of Neuchâtel; University of Zurich
    Duration 2024-06-01 - 2028-05-31
    Partner(s) Eidgenoessische Technische Hochschule Zuerich; Universita della Svizzera italiana; University of Fribourg; University of Lausanne; Idiap Research Institute

12. Name **FACTCHECK** (FactCheck - Collaborative Fact Checking of Digital Media)
    Funding Hasler Foundation
    Coordinator Idiap Research Institute
    Duration 2024-06-01 - 2025-08-31
    Partner(s) –

13. Name **FAIRMI** (FairMI - Machine Learning Fairness with Application to Medical Images)
    Funding SNSF Division II: Mathematics, Natural sciences and Engineering
    Coordinator Federal University of Sao Paulo
    Duration 2024-03-01 - 2028-02-29
    Partner(s) Idiap Research Institute

14. Name **GAME** (Gaze and attention modeling in Evolang)
    Funding University of Zurich
    Coordinator University of Zurich
    Duration 2025-01-01 - 2025-12-31
    Partner(s) Idiap Research Institute

15. Name **HORACE** (Human-robot collaborative manipulation by considering geometry and uncertainty)
    Funding SNSF Weave
    Coordinator Idiap Research Institute
    Duration 2025-07-01 - 2029-06-30
    Partner(s) Jozef Stefan Institute; University of Zagreb - Faculty of Electrical Egineering and Computing

16. Name **INTERART** (Identification des portraits historiques par technologies de reconnaissance faciale hétérogène)
    Funding Loterie Romande
    Coordinator Idiap Research Institute
    Duration 2025-06-01 - 2026-11-30
    Partner(s) University of Lausanne; University of Oxford; Musée d'Art et d'Histoire de Genève, MAH

17. Name **KREATIV** (Installation at Phänomena: The roles of AI and robotics for creative and artistic applications in both digital and physical worlds)
    Funding SNSF Agora
    Coordinator Idiap Research Institute
    Duration 2024-07-01 - 2026-12-31
    Partner(s) Phänomena

18. Name **MELAS** (Alternative Splicing and Polyadenylation from single-cell RNA sequencing towards tumor subpopulation identification in melanoma)
    Funding Novartis Foundation
    Coordinator Idiap Research Institute
    Duration 2023-01-01 - 2025-05-31
    Partner(s) –

19. Name **MORPHYN** (MORPhyN: Modelling, Optimisation and Routing using Physiological Networks)
    Funding SNSF SPIRIT
    Coordinator Idiap Research Institute
    Duration 2025-07-01 - 2029-06-30
    Partner(s) Mahidol University

20. Name **M-RATIONAL** (Reconceiving and Improving Multi-Perspectivality and Rationality in Argumentative AI)
    Funding SNSF Project Funding
    Coordinator Idiap Research Institute; University of Zurich
    Duration 2025-10-01 - 2029-09-30
    Partner(s) University of St.Gallen

21. Name **NEUMATH** (NeuMath: Neural Discourse Inference over Mathematical Texts)
    Funding SNSF Division II: Mathematics, Natural and Engineering Sciences
    Coordinator Idiap Research Institute
    Duration 2022-04-01 - 2025-03-31
    Partner(s) –

22. Name **NEUROCIRT** (Investigating the role of cytoplasmic intronic sequences in ALS pathogenesis)
    Funding SNSF Division III: Biology and medicine
    Coordinator Idiap Research Institute
    Duration 2022-09-01 - 2025-05-31
    Partner(s) –

23. Name **NKBP** (Deep Learning Models for Continual Extraction of Knowledge from Text)
    Funding SNSF Division II Lead Agency
    Coordinator Idiap Research Institute
    Duration 2020-10-01 - 2025-03-31
    Partner(s) Katholieke Universiteit Leuven

24. Name **OKA** (Oka theory, CR-geometry, D-bar-problem and applications)
    Funding SNSF Vietnamese-Swiss Joint Research Projects
    Coordinator University of Bern
    Duration 2025-09-01 - 2029-08-31
    Partner(s) Phenikaa University; Idiap Research Institute

25. Name **OPENBEERS** (Open-data for Building Energy Efficiency, Renovation and Storage)
    Funding Etat du Valais
    Coordinator HES-SO Valais
    Duration 2024-08-01 - 2025-09-30
    Partner(s) Idiap Research Institute

26. Name **ORIENTER_2** (ORIENTER: tOwards undeRstanding and modelIng the language of mENTal health disordERs)
    Funding SNSF SPIRIT
    Coordinator Idiap Research Institute
    Duration 2025-07-01 - 2029-06-30
    Partner(s) Mathematics Research Center; National Institute of Psychiatry; Universita della Svizzera italiana

27. Name **PASS** (PaSS - Pathological Speech Synthesis)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2024-03-01 - 2028-02-29
    Partner(s) –

28. Name **PAUSE** (PAuSE: Pathological Speech Enhancement)
    Funding SNSF Division II: Mathematics, Natural sciences and Engineering
    Coordinator Idiap Research Institute
    Duration 2024-04-01 - 2028-03-31
    Partner(s) –

29. Name **RATIONAL** (Robust Augmented Retrieval for Natural Language Inference over Transformer-based Models)
    Funding SNSF Lead agency
    Coordinator State University of São Paulo
    Duration 2025-02-01 - 2029-01-31
    Partner(s) Idiap Research Institute

30. Name **REHABOT** (Specification and correction of physical human-robot behaviors for flexible rehabilitation robots)
    Funding SNSF Project Funding
    Coordinator Idiap Research Institute
    Duration 2025-10-01 - 2029-09-30
    Partner(s) Centre Neu'Rhône

31. Name **SAFER** (reSponsible fAir FacE Recognition)
    Funding Hasler Foundation
    Coordinator Idiap Research Institute
    Duration 2022-03-01 - 2026-02-28
    Partner(s) SICPA; University of Zurich

32. Name **SMILE-II** (SMILE-II Scalable Multimodal sign language technology for sIgn language Learning and assessmEnt Phase-II)
    Funding SNSF Sinergia
    Coordinator Idiap Research Institute
    Duration 2021-01-01 - 2025-06-30
    Partner(s) University of Applied Sciences of Special Needs Education; University of Surrey; University of Zurich

33. Name **SPIDER** (Spiking for Denoising and Regression)
    Funding SNSF Project Funding
    Coordinator Idiap Research Institute
    Duration 2025-11-01 - 2029-10-31
    Partner(s) –

34. Name **STEADI** (Storytelling and first impressions in face-to-face and algorithm-powered digital interviews)
    Funding SNSF Division I: Humanities and social sciences
    Coordinator University of Neuchâtel
    Duration 2021-02-01 - 2026-05-31
    Partner(s) University of Lausanne; Idiap Research Institute

35. Name **TUNASBE** (Towards a Unified understanding of Attention and Social Behaviors in the Wild)
    Funding SNSF Project Funding
    Coordinator Idiap Research Institute
    Duration 2025-09-01 - 2029-08-31
    Partner(s) –

36. Name **VALIDATE-H** (VALaIs aDaptation of Artificial inTelligence modEls in Healthcare)
    Funding Etat du Valais
    Coordinator Idiap Research Institute
    Duration 2025-06-30 - 2026-06-30
    Partner(s) HES-SO Valais; Observatoire Valaisan de la sante

37. Name **VALVISER_(B-POC)** (Transforming Waste Collection Fleet Management with AI-Driven Data Interpretation)
    Funding SNSF Bridge Proof of Concept
    Coordinator Idiap Research Institute
    Duration 2025-04-01 - 2026-03-31
    Partner(s) –

Innovation Projects

1. Name **BIPED** (Biped - Self-driving technology to guide blind pedestrians)
    Funding Innosuisse
    Coordinator Fusion Lab Technologies Sàrl
    Duration 2023-05-01 - 2025-04-30
    Partner(s) Idiap Research Institute

2. Name **EUROCONTROL** (Integrate the Automatic Speech Recognition system with eDEP, ESCAPE and audiolan)
    Funding Industrial
    Coordinator SkySoft-ATM
    Duration 2023-02-01 - 2025-12-31
    Partner(s) Eurocontrol; Idiap Research Institute

3. Name **IICT** (Inclusive Information and Communication Technologies)
    Funding Innosuisse
    Coordinator University of Zurich
    Duration 2022-03-14 - 2026-03-13
    Partner(s) CFS GmbH Capito; Federal Office for Civil Protection (FOCP); Federal Office for the Equality of Persons with Disabilities (OEPD); Institut Icare; Swiss Federation of the Deaf (SGB-FSS); Swiss Txt; University of Applied Sciences of Special Needs Education; Zürich Versicherungs-Gesellschaft AG; Idiap Research Institute

4. Name **MAVERICK** (MAVERICK: Maximum evidence platform for explainable predictions of risk)
   Funding Innosuisse
   Coordinator Basinghall Analytics Sàrl
   Duration 2024-03-01 - 2026-04-30
   Partner(s) Idiap Research Institute

5. Name **MECHANIC** (Automated End-To-End Mechanical Design Reasoning Supported by Large Language Models)
   Funding Industrial
   Coordinator Idiap Research Institute
   Duration 2024-12-01 - 2027-11-30
   Partner(s) Honda Research Institute

6. Name **MELAS-KPI** (Towards the development of predictive biomarkers for patient stratification and immunotherapy response in cancer)
   Funding Industrial
   Coordinator Idiap Research Institute
   Duration 2023-04-01 - 2025-05-31
   Partner(s) Centre Hospitalier Universitaire Vaudois; Hôpital du Valais; Novartis Institutes for BioMedical Research

7. Name **PHYGIBOT** (A phygital robotics education platform for industrial upskilling)
   Funding Innovation Booster in Robotics
   Coordinator Swiss Cobotics Competence Center (S3C)
   Duration 2025-09-01 - 2026-02-28
   Partner(s) Idiap Research Institute

8. Name **PMPM** (PmPm - Predictive Manufacturing Predictive Maintenance)
   Funding Innosuisse
   Coordinator HES-SO Valais
   Duration 2024-03-01 - 2026-08-31
   Partner(s) Constellium Valais SA; Eversys SA; LYSR Sàrl; Idiap Research Institute

9. Name **PRIMEAID** (PRiMEAiD: Privacy-pReserving bioMetric idEntification for humAnitarian aid Distribution)
   Funding Innosuisse
   Coordinator Idiap Research Institute
   Duration 2024-09-02 - 2026-09-01
   Partner(s) –

10. Name **ROSALIND** (Robust anti-fraud algorithms against malicious AI-generated face images and travel documents)
    Funding Innosuisse
    Coordinator PXL Vision AG
    Duration 2024-02-01 - 2025-11-30
    Partner(s) Idiap Research Institute

11. Name **SEM24** (SEM24)
    Funding Innosuisse
    Coordinator Arca24.com SA
    Duration 2023-07-01 - 2025-06-30
    Partner(s) EHL; Idiap Research Institute

12. Name **SENTINEL-UK** (Securing Age Estimation and Digital IDs against Presentation and Injection Threats)
    Funding Innosuisse
    Coordinator Privately SA
    Duration 2023-12-01 - 2025-09-30
    Partner(s) Age Check Certification Scheme; Age Verification Providers Association; Idiap Research Institute

13. Name **SINFONIA** (SINFONIA: Social IntelligeNce FOr strategic aNticIpAtion)
    Funding Innosuisse
    Coordinator Bloom Suisse Sàrl
    Duration 2023-04-17 - 2025-06-30
    Partner(s) Idiap Research Institute

14. Name **SMART-DISPENSING** (Smart Dispensing: Programming of a complex deposit task through learning from demonstration)
    Funding Innosuisse
    Coordinator Ciposa SA
    Duration 2023-03-01 - 2026-02-28
    Partner(s) Berner Fachhochschule; Idiap Research Institute

15. Name **WOLFCAM** (WolfCam)
    Funding Industrial
    Coordinator OPPAL Association
    Duration 2025-06-01 - 2026-02-28
    Partner(s) Idiap Research Institute

International Projects

1. Name **ALIGNAI** (value-ALIGNed socio-technical systems using large-language models (LLMs))
   Funding Horizon Europe
   Coordinator Munich University of Technology
   Duration 2024-09-01 - 2028-08-31
   Partner(s) Datacation V.V; DPG Media B.V.; Ecole Polytechnique Federale de Lausanne; ESH Médias; Fujitsu Technology Solutions GmbH; Region Hovedstaden; Technical University of Denmark; Technical University of Eindhoven; Technion Israel Institute of Technology; Idiap Research Institute

2. Name **BALM** (Interpretable Beliefs and Programmable Knowledge with Bayesian Attention in Large Language Models)
   Funding Horizon Europe
   Coordinator Idiap Research Institute
   Duration 2025-11-01 - 2030-10-31
   Partner(s) –

3. Name **CARMEN** (CarMen)
   Funding Horizon Europe
   Coordinator YNCREA MEDITERRANEE
   Duration 2024-09-01 - 2027-08-31
   Partner(s) ARCLAN'system; B.A.I Bretagne Angleterre Irlande SA; Cabinet Louis Reynaud SAS; Hellenic Police; Hochschule Darmstadt (University of Applied Sciences H-DA); Home Office; IDEMIA Identity & Security France; Kentro Meleton Asfaleias; Ministère de l'Interieur français; Notitia Ltd; Thales Dis France SAS; University of Reading; YNCREA Ouest; Idiap Research Institute

4. Name **CERTAIN** (Certification for Ethical and Regulatory Transparency in Artificial Intelligence)
   Funding Horizon Europe
   Coordinator IDEMIA Identity & Security France
   Duration 2025-01-01 - 2027-12-31
   Partner(s) EIT Digital; Maribor University Medical Centre; National Bank of Greece SA; Netcompany - Intrasoft; Tartu Ulikool; University of Luxembourg; University of Maribor; Charokopeio Panepistimio; Bürgerenergiegemeinschaft Empower; D.Tsakalidis-G.Domalis; Global Trust Foundation; Incom Simvuli Epihiriseon; N Vision systems and technologies SL; Red Alert LABS; Digital for Planet - d4p; Fachhochschule St. Pölten GmbH; Dexai - Etica Artificiale; Idiap Research Institute

5. Name **DREAM** (Detection and Research for Learning Sets on Electroencephalographic Patterns)
   Funding Fondation AP-HP
   Coordinator Idiap Research Institute
   Duration 2025-03-01 - 2027-02-28
   Partner(s) L'Assistance Publique - Hôpitaux de Paris

6. Name **ELIAS** (European Lighthouse of AI for Sustainability)
   Funding Horizon Europe
   Coordinator Rettorato Università degli Studi di Trento
   Duration 2023-09-01 - 2027-08-31
   Partner(s) Aalto University; Bitdefender SRL; Centre for Research and Technology Hellas; Copenhagen University; Czech Technical University Prague; Eberhard Karls University of Tübingen; Eidgenoessische Technische Hochschule Zuerich; Engineering - Ingegneria Informatica SPA; Fondazione Bruno Kessler; Fondazione Istituto Italiano di Tecnologia; Hasso Plattner Institute for Digital Engineering gGmbH; IBM Ireland Limited; Ideas NCBR ; Institut National de Recherche en Informatique et en Automatique; Institut Polytechnique de Paris; Jozef Stefan Institute; Max Planck Society for the Advancement of Sciences; Politecnico di Milano; Technical University of Bucharest; Robert Bosch GmbH; Umea University; Universita degli Studi di Milano; Université de Toulouse; University of Amsterdam; University of Genoa; University of Manchester; University of Modena and Reggio Emilia; University of Valencia; ELLIS Alicante; Bosch Hungary; Idiap Research Institute

7. Name **ELOQUENCE** (ELOQUENCE: Multilingual and Cross-cultural interactions for context-aware, and bias-controlled dialogue systems for safety-critical applications)
   Funding Horizon Europe
   Coordinator Telefonica Investigacion Y Desarrollo SA
   Duration 2024-01-01 - 2026-12-31
   Partner(s) Barcelona Supercomputing Center; Brno University of Technology; Brunel University London; Consiglio nazionale delle ricerche - Istituto di Scienza e Tecnologie dell'Informazione; Fondazione Bruno Kessler; Grantxpert Consulting Ltd.; Omilia Ltd; Privanova SAS; Transformation Lightouse, Poslovno Svetovanje, D.O.O; University Of Essex; University of Novi Sad; InoSens; Synelixis; European University Institute; Idiap Research Institute

8. Name **EVOICENET** (European Network to Advance the Development and Implementation of Vocal Biomarkers)
   Funding COST
   Coordinator Luxembourg Institute of Health
   Duration 2025-10-27 - 2029-10-26
   Partner(s) Armenian Association of Digital Health; Medical University of Vienna; KU Leuven; University of Banja Luka; Institute Of Information And Communication Technologies; Ecole Polytechnique de Ouagadougou; Simon Fraser University; Universidad De Antioquia; Catholic University of Croatia; Brno University of Technology; Scicake; Technical University of Denmark; Migrevention; Aalto University; Telecom SudParis; Eurecom; #ApresJ20 Association Covid Long France; INRIA; Deggendorf Institute of Technology; ki:elements GmbH; University Hospital Würzburg; ZANA Technologies GmbH; University of Erlangen-Nuremberg; Ruhr University Bochum; Otto-von-Guericke University Magdeburg; HYGEIA Hospital; Budapest University of Technology and Economics; Trinity College Dublin; University of Campania Luigi Vanvitelli; Lithuanian University Of Health Sciences; Ministry for Health - Government of Malta; University of Groningen; Delft University of Technology; Maastricht University; Ss. Cyril and Methodius University in Skopje; Medical University of Silesia; University of Krakow; University of Lisbon; University Of Nis; Research and Development Institute Life Activities Advancement Institute; University of Zilina; University of Maribor; Alpineon Research and Development; Technical University of Madrid; TecnoCampus Foundation; Autonomous University of Barcelona; Kapharos; Ecole Nationale D'ingenieurs De Tunis; dBsense; University College London; University Of Cambridge; University of Edinburgh; University of Southampton; King's College London; University Of Birmingham; University of South Florida; Massachusetts Eye and Ear; Idiap Research Institute

9. Name **GENAIDE** (Generative AI for Industrial Design Engineering)
   Funding Horizon Europe
   Coordinator Honda Research Institute
   Duration 2025-10-01 - 2029-09-30
   Partner(s) Altair Engineering GmbH; DATEV EG; Ecole Polytechnique Federale de Lausanne; Julius Maximilian University of Wuerzburg; University of Jyvaskyla; Kone Oyj; Leiden University; Massachusetts Institute of Technology; Nanyang Technological University; Rinf Outsourcing Solutions SRL; Samaya; Terra Quantum GmbH; The Mathworks Limited; University of Manchester; Idiap Research Institute

10. Name **HURO** (Leading in the brave new world: human - robot dynamics)
    Funding Horizon Europe
    Coordinator ISM University of Management and Economics
    Duration 2024-10-01 - 2026-09-30
    Partner(s) Idiap Research Institute

11. Name **INTELLIMAN** (AI-Powered Manipulation System for Advanced Robotic Service, Manufacturing and Prosthetics)
    Funding: Horizon Europe
    Coordinator: Alma Mater Studiorum - Università di Bologna
    Duration: 2022-09-01 - 2026-04-30
    Partner(s): Bavarian Research Alliance Gmbh; Deutsches Zentrum für Luft- und Raumfahrt e.V.; Elvez, Manufacture of cable assemblies and plastics processing D.o.o.; Eurecat Foundation; Friedrich-Alexander-Universität Erlangen-Nürnberg; National Institute for Industrial Accidents Insurance Inail; Ocado Innovation Limited; Universitat Politécnica de Cataluña; University of Campania Luigi Vanvitelli; University of Genoa; University of Zurich; Idiap Research Institute

12. Name **ORION** (Optimized Risk-based Intelligence-driven Operations for Next-generation Secure, Reliable and Privacypreserving Border Management)
    Funding: Horizon Europe
    Coordinator: Ubitech Limited
    Duration: 2025-09-01 - 2028-08-31
    Partner(s): Engineering - Ingegneria Informatica SPA; Eurostar International Limited; Finnish Interior Ministry; General Inspectorate of Romanian Border Police; Hellenic Mediterranean University; Institute for Corporative Security Studies; Interior Ministry of Slovenia; Iproov Netherlands B.V.; Main Directorate Of Bulgarian National Police; Netcompany - Intrasoft; Norwegian National Police Directorate; Quadible Greece IKE; Rapiscan Systems Limited; Risa Sicherheitsanalysen GmbH; Streamowl IKE; Telesto Information and Communication Technologies; Union Internationale des Chemins de Fer; University of Maribor; Free University of Brussels; Idiap Research Institute

13. Name **POPEYE** (robust Privacy-preserving biOmetric technologies for Passengers' identification and verification at EU external borders maximising the accuracY, reliability and throughput of the rEcognition)
    Funding: Horizon Europe
    Coordinator: Austrian Institute of Technology
    Duration: 2024-10-01 - 2027-09-30
    Partner(s): European Association for Biometrics; General Inspectorate of Romanian Border Police; Hogskolan i Halmstad; IDEMIA Identity & Security Germany AG; Katholieke Universiteit Leuven; Ministry of Interior of the Slovak Republic; Netcompany - Intrasoft; Norwegian University of Science and Technology; Quadible Greece IKE; Universiteit Twente; Free University of Brussels; Idiap Research Institute

14. Name **SESTOSENSO** (Physical Cognition for Intelligent Control and Safe Human-Robot Interaction)
    Funding: Horizon Europe
    Coordinator: University of Genoa
    Duration: 2022-10-01 - 2025-09-30
    Partner(s): Alma Mater Studiorum - Università di Bologna; Centre for Research and Technology Hellas; Centro Ricerche Fiat SCPA - Groupe Stellantis; Free University of Bolzano; Inertia Technology B.V.; Institut Franco-Allemand de Recherches de Saint-Louis; Ocado Innovation Limited; Rise Research Institutes of Sweden AB; Universidad de Zaragoza; University of Latvia, Institute Of Solid State Physics; University of Ljubljana; University of Oxford; Idiap Research Institute

15. Name **TRACY** (A big-data analyTics from base-stations Registrations And Cdrs e-evidence sYstem)
    Funding Horizon Europe
    Coordinator PERFORMANCE Technologies S.A.
    Duration 2023-06-01 - 2025-05-31
    Partner(s) Cosmote Mobile Telecommunications S.A; Galati County Police Inspectorate; General Police Inspectorate of the Republic of Moldova; Greek Center for Security Studies; Hellenic Police; Hochschule Fuer den Oeffentlischen Dienst in Bayern (HfoD); Time Lex CVBA; Idiap Research Institute

16. Name **UNDERSPEC-ROBUSTNESS** (Addressing Underspecification for Improved Fairness and Robustness in Conversational AI)
    Funding Amazon Research Awards
    Coordinator Idiap Research Institute
    Duration 2023-04-01 - 2025-12-31
    Partner(s) –

Projects awarded in 2025 and starting in 2026

1. Name **AIMANT** (Analysis of adult-child interactions for the objective analysis of neurodevelopmental disorders)
   Funding SNSF
   Coordinator Institut National de Recherche en Informatique et en Automatique
   Duration 2026-05-01 - 2029-10-31
   Partner(s) Côte d'Azur University; Université Lumière Lyon II; Idiap Research Institute

2. Name **DEMO-AI** (DEmocracy and Media Oversight in the era of AI and disinformation)
   Funding Hasler and Mercator Foundations
   Coordinator Idiap Research Institute
   Duration 2030-02-28 - 2030-02-28
   Partner(s) Nagravision Sàrl; RTS; SICPA; University of Lausanne; University of Neuchâtel

3. Name **GAMOE** (Gaze and Attention Modeling with Objects in Evolang)
   Funding SNSF
   Coordinator University of Zurich
   Duration 2027-01-31 - 2027-01-31
   Partner(s) Idiap Research Institute

4. Name **IMAGIN-AIR** (Des images aux connaissances : l'IA pour élucider les mécanismes et améliorer les stratégies thérapeutiques de l'inflammation rétinienne)
   Funding Carigest SA
   Coordinator Idiap Research Institute
   Duration 2030-02-28 - 2030-02-28
   Partner(s) –

5. Name **LIGHT** (Looking for Integrated Solutions for Healthy PEDs)
   Funding SNSF
   Coordinator Gdańsk University of Technology
   Duration 2028-12-31 - 2028-12-31
   Partner(s) Foton Oze SP ZOO; Gdansk Foundation for Social Innovation; Gdansk Medical University; Innsbrucker Immobilien Gmbh & Co Kg; Lund Commune; Lunds University; Municipality of Kartuzy; University of Gdansk; University of Innsbruck; Idiap Research Institute

6. Name **OSCAIL** (Open Science Communication through AI in EU Languages)
   Funding SNSF
   Coordinator Dublin City University
   Duration 2029-04-30 - 2029-04-30
   Partner(s) European Research Infrastructure for the development of open scholarly communication in the social sciences and humanities; Federation of Finnish Learned Societies; Laval University; PleIAs; Tartu Ulikool; University of Gent; Idiap Research Institute

7.	Name	**PREP-HERMUS-AMBIZIONE-1** (Confronting the Paradox of Human Motor Control: Leveraging Subtractive and Generative Models)
	Funding	SNSF
	Coordinator	Idiap Research Institute
	Duration	2026-09-01 - 2029-08-31
	Partner(s)	–

Code

The following software repositories were released in 2025 on Idiap's channels:
https://gitlab.idiap.ch and https://github.com/idiap.

[C1] *A Multimodal Large Language Model for Face Understanding*,
https://github.com/idiap/facellm.

[C2] *A Web interface for the ABRoad demonstrator*,
https://github.com/idiap/abroad-demo-webapp.

[C3] *Adaptive neuro-symbolic reasoning framework that automatically identifies formal reasoning strategies from natural language problems and dynamically composes specialized logical solvers for multi-paradigm inference*,
https://github.com/idiap/adaptive_symbolic_reasoning.

[C4] *Alleviating Forgetfulness of Linear Attention by Hybrid Sparse Attention and Contextualized Learnable Token Eviction*,
https://github.com/idiap/hybrid-linear-sparse-attention.

[C5] *ArtFace: Towards Historical Portrait Face Identification via Model Adaptation*,
https://gitlab.idiap.ch/biometric/code.iccv2025artmetrics.artface.

[C6] *Building the ABRoad KG*,
https://github.com/idiap/abroad-kg-store.

[C7] *Buildpack for Python that consolidate all package manager installers under one roof*,
https://github.com/idiap/python-installers.

[C8] *Buildpack for Python that consolidate all package managers under one roof*,
https://github.com/idiap/python-packagers.

[C9] *ChildPlay: A New Benchmark for Understanding Children's Gaze Behaviour; code and checkpoints*,
https://github.com/idiap/geomgaze.

[C10] *ChildPlay-Hand: A Dataset of Hand Manipulations in the Wild*,
https://github.com/idiap/childplay_hand.

[C11] *Code for analyzing RNA/protein compartmentalization, 3'UTR isoform expression, and fluorescence microscopy data in healthy and ALS patient-derived cell lines*,
https://github.com/idiap/ALS-Compartmentalization.

[C12] *Comparing Self-Supervised Learning Models Pre-Trained on Human Speech and Animal Vocalizations for Bioacoustics Processing*,
https://github.com/idiap/ssl-human-animal.

[C13] *Demo application for Optical Projection Tomography*,
https://github.com/idiap/demo_opt.

[C14] *Diffusion PDE (heat equation) based methods on geometric manifolds including point clouds and meshes*,
https://github.com/idiap/diffused_fields.

[C15] *Digi2real: Bridging the realism gap in synthetic data face recognition via foundation models*,
https://gitlab.idiap.ch/biometric/code.datacv2025.

[C16] *EdgeDoc: a document-forgery detection and localization system designed to protect KYC and remote onboarding by combining a lightweight convolutional transformer with auxiliary noiseprint features to expose subtle manipulations*,
https://gitlab.idiap.ch/biometric/code.edgedoc_iccv2025.

[C17] *Enhancing 3D Gaze Estimation in the Wild using Weak Supervision with Gaze Following Labels*, https://github.com/idiap/gaze3d.

[C18] *Exploring ChatGPT for Face Presentation Attack Detection in Zero and Few-Shot in-Context Learning*, https://gitlab.idiap.ch/bob/bob.paper.wacv2025_chatgpt_face_pad.

[C19] *Extended pre-training of histopathology foundation models uncovers co-existing breast cancer archetypes characterized by RNA splicing or TGF-beta dysregulation*, https://github.com/idiap/TumorArchetype-FM.

[C20] *FaceLLM: a multimodal large language model trained specifically for facial image understanding*, https://gitlab.idiap.ch/biometric/code.iccv2025_facellm.

[C21] *Fairical: a Python library to assess adjustable demographically fair Machine Learning (ML) systems*, https://gitlab.idiap.ch/medai/software/fairical.

[C22] *From Movement Primitives to Distance Fields to Dynamical Systems*, https://github.com/idiap/mp-df-ds.

[C23] *Functions to deal with genome and annotation files*, https://github.com/idiap/genomemanager.

[C24] *HyperFace: Generating Synthetic Face Recognition Datasets by Exploring Face Embedding Hypersphere*, https://gitlab.idiap.ch/biometric/code.iclr2025_hyperface.

[C25] *Identity-Preserving Aging and De-Aging of Faces in the StyleGAN Latent Space, accepted at the IEEE Joint Conference on Biometrics (IJCB) 2025*, https://gitlab.idiap.ch/biometric/code.ijcb2025.agesynth.

[C26] *Inductive Learning of Logical Theories with LLMs: A Expressivity-graded Analysis*, https://github.com/idiap/ilp-llm.

[C27] *Interface to ROS2 to visualize geometric primitives and robots using the gafro library*, https://github.com/idiap/gafro_ros2.

[C28] *Investigation of accuracy and bias in face recognition trained with synthetic data*, https://gitlab.idiap.ch/bob/bob.paper.ijcb2025_synthbias.

[C29] *Joint Fine-tuning and Conversion of Pretrained Speech and Language Models towards Linear Complexity*, https://github.com/idiap/linearize-distill-pretrained-transformers.

[C30] *Listening to Hypoglycemia – Voice as a Biomarker for Detection of a Medical Emergency using Machine Learning*, https://github.com/idiap/SustainedVowelBoundaries.

[C31] *MM-HSD: Multi-Modal Hate Speech Detection in Videos*, https://github.com/idiap/mm-hsd.

[C32] *OTTC: Optimal Temporal Transport Classification*, https://github.com/idiap/OTTC.

[C33] *Package to help with bulk RNA sequencing data analysis*, https://github.com/idiap/bulkanalysis.

[C34] *Package to help with the analysis of kallisto outputs*, https://github.com/idiap/kallistomanager.

[C35] *Package to help with the analysis of single-cell RNA sequencing data*, https://github.com/idiap/scanalysis.

[C36] *Reducing Prompt Sensitivity in LLM-based Speech Recognition Through Learnable Projection*, https://github.com/idiap/llm-asr-prompt.

[C37] *Robotics package for object-centric robot manipulation applications*, https://github.com/idiap/diffused_fields_robotics.

[C38] *Sampling-Based Constrained Motion Planning with Products of Experts*, https://github.com/idiap/smpc_poe.

[C39] *Self-contained tutorial projects about TT-Cross & TT-SVD*, https://github.com/idiap/tt-sandbox.

[C40] *Snakemake pipeline to quantify relative expression of 3'UTR isoforms from RNA-seq data*, https://github.com/idiap/3UTR-Quant.

[C41] *Snakemake pipeline to reannotate 3'UTR isoforms based on read-coverage changes from RNA-Seq data*, https://github.com/idiap/3UTR-Annot.

[C42] *Soft Skills in the Wild: Challenges in Multilingual Classification*, https://github.com/idiap/multilingual_skill_extraction.

[C43] *Specializing General-purpose LLM Embeddings for Implicit Hate Speech Detection across Datasets*, https://github.com/idiap/implicit-hsd.

[C44] *Structured-augmented RAG on PubMed abstracts*, https://github.com/idiap/biotopg.

[C45] *Synthetic Dialog Generation and Analysis with LLMs*, https://github.com/idiap/sdialog.

[C46] *Synthetic face dataset generation and related tooling*, https://github.com/idiap/face_synthetics.

[C47] *Synthetics: Synthetic face image generation, inversion, editing and database generation*, https://gitlab.idiap.ch/biometric/code.icml2025_synthetics_disco.

[C48] *The Invisible Threat: Evaluating the Vulnerability of Cross-Spectral Face Recognition to Presentation Attacks*, https://gitlab.idiap.ch/bob/bob.paper.ijcb2025_vulncfr.

[C49] *Tools to analyze Alternative Polyadenylation (APA) from RNA sequencing data*, https://github.com/idiap/apatools.

[C50] *Tools to analyze CLIP-sequencing data*, https://github.com/idiap/cliptools.

[C51] *Toward Semantic Gaze Target Detection*, https://github.com/idiap/semgaze.

[C52] *Traversal Over Proposition Graph (TOPG): a hybrid RAG (Retrieval-Augmented Generation) framework that builds a graph from passages, entities, and propositions*, https://github.com/idiap/ToPG.

[C53] *Tutorial Biomedical Knowledge Graphs meet LLM (preliminary version)*, https://github.com/idiap/tutorial-biokg-meet-llm.

[C54] *Unsupervised Rhythm and Voice Conversion Framework*, https://github.com/idiap/RnV.

[C55] *When Specialization Helps (and Hurts): Cross-Modality Transfer in ophthalmic imaging with foundation models*, https://github.com/idiap/fm-overspecialization.

[C56] *xEdgeFace: Efficient Cross-Spectral Face Recognition for Edge Devices*, https://gitlab.idiap.ch/bob/bob.paper.ijcb2025_xedgeface.

Datasets

The following datasets were released in 2025 by Idiap on the Zenodo platform for open science: https://zenodo.org/communities/idiap/.

[D1] M. Delmas, M. Wysocka, D. Gusicuma, and A. Freitas, *ABROAD-KG-DATA: a demonstrator knowledge graph to support a new pipeline designed to assist reviewers in curating evidence of antibiotic activity associated with target organisms*, https://doi.org/10.34777/YR1S-ZH98.

[D2] A. Farkhondeh, S. Tafasca, and J.-M. Odobez, *ChildPlay-Hand: a video dataset for modeling hand-object interaction in the wild*, https://doi.org/10.34777/QG7C-AW54.

[D3] A. Gupta, *VSGaze: annotations across multiple video datasets for gaze following and social gaze*, https://doi.org/10.34777/JGGS-1128.

[D4] P. Korshunov, K. Kotwal, C. Ecabert, V. Vidit, A. Mohammadi, and S. Marcel, *FairFaceGen: a dataset of synthetic faces generated to study the impact of synthetic data generation on the performance and bias of face recognition models*, https://doi.org/10.34777/SD5B-2153.

[D5] P. Korshunov, A. Mohammadi, V. Vidit, C. Ecabert, and S. Marcel, *FantasyID: a dataset for detecting digital manipulations of ID-documents*, https://doi.org/10.34777/C966-NN94.

[D6] L. S. Luevano Garcia, P. Korshunov, and S. Marcel, *AgeSynth: a dataset of aged and de-aged synthetic samples of face images.* https://doi.org/10.34777/B2DN-BW71.

[D7] H. Otroshi Shahreza and S. Marcel, *FairFaceGPT: a dataset of question-answer pairs generated with ChatGPT, based on images from the FairFace dataset.* https://doi.org/10.34777/CQGR-3712.

[D8] H. Otroshi Shahreza and S. Marcel, *HyperFace: a collection of datasets of synthetic face images - up to 50,000 unique synthetic identities and 3.2 million images*, https://doi.org/10.34777/R9AG-E273.

[D9] L. Vásquez-Rodríguez, B. Audrin, S. Michel, S. Galli, J. B. Martinez Atencio, J. Negro Cusa, and L. van der Plas, *Skill Extraction Green: additional annotations related to skills and occupations for the Green dataset, designed to support entity recognition in job descriptions*, https://doi.org/10.34777/TR37-ER53.

[D10] L. Vásquez-Rodríguez, B. Audrin, S. Michel, J. B. Martinez Atencio, J. Negro Cusa, L. van der Plas, and S. Galli, *Skill Extraction: collection of hard skill entities extracted from a corpus of resumes*, https://doi.org/10.34777/PEAY-XE10.

Publications

PhD Theses

[T1] H. Chen, Efficient Adaptation for Speech Technology, PhD thesis, École polytechnique fédérale de Lausanne (EPFL), 2025.

[T2] L. Colbois, Face morphing attacks in the era of deepfakes: risks, detection & source attribution, PhD thesis, Université de Lausanne (UNIL), 2025.

[T3] A. C. Coman, Text-Graph Encoders and Retrieval-Augmented Generation, PhD thesis, École polytechnique fédérale de Lausanne (EPFL), 2025.

[T4] F. Fehr, Nonparametric Variational Information Bottleneck: Attention-based Architectures as Latent Variable Models, PhD thesis, École polytechnique fédérale de Lausanne (EPFL), 2025.

[T5] L. Coppieters de Gibson, Modelling cochlea and its interaction with the auditory path for speech processing, PhD thesis, École polytechnique fédérale de Lausanne (EPFL), 2025.

[T6] A. Gupta, Unified and Multimodal Learning for Gaze Prediction in Naturalistic Settings, PhD thesis, École polytechnique fédérale de Lausanne (EPFL), 2025.

[T7] T. Löw, Robot Manipulation with Geometric Algebra: A Unified Geometric Framework for Control and Optimization, PhD thesis, École polytechnique fédérale de Lausanne (EPFL), 2025.

[T8] Z. Mostaani, Towards Integrated Processing of Physiological Signals and Speech, PhD thesis, Ecole polytechnique fédérale de Lausanne (EPFL), Switzerland, 2025.

[T9] E. Sarkar, Transferability of Learnt Speech Representations for Decoding Non-Human Vocal Communication, PhD thesis, École polytechnique fédérale de Lausanne (EPFL), 2025.

[T10] N. Tarigopula, Advancing Phonology-Based Sign Language Assessment: From Learner to Machine-Generated Videos, PhD thesis, Ecole polytechnique fédérale de Lausanne (EPFL), 2025.

Journal Papers

[J1] V. Amiot, O. Jimenez-del Toro, Y. Guex-Croisier, M. Ott, T.-E. Bogaciu, S. Banerjee, J. Howell, C. Amstutz, C. Chiquet, C. Bergin, I. Meloni, M. Tomasoni, F. Hoogewoud, and A. Anjos, Automatic transformer-based grading of multiple retinal inflammatory signs in uveitis on fluorescein angiography, *Computers in Biology and Medicine*, 2025.

[J2] W. Bangamuarachchi, A. Chamantha, L. B. Meegahapola, H. Kim, S. Ruiz-Correa, I. Perera, and D. Gatica-Perez, Inferring Mood-While-Eating with Smartphone Sensing and Community-Based Model Personalization, *ACM Transactions on Computing for Healthcare*, 2025.

[J3] K. Baraka, I. Idrees, T. Kessler Faulkner, E. Biyik, S. Booth, M. Chetouani, D. H. Grollman, A. Saran, E. Senft, S. Tulli, A.-L. Vollmer, A. Andriella, H. Beierling, T. Horter, J. Kober, I. Sheidlower, M. E. Taylor, S. van Waveren, and X. Xiao, Human-Interactive Robot Learning: Definition, Challenges, and Recommendations, *Journal Human-Robot Interaction*, Dec. 2025.

[J4] D. Berio, M. Stroh, S. Calinon, F. F. Leymarie, O. Deussen, and A. Shamir, Neural Image Abstraction Using Long Smoothing B-splines, *ACM Transactions on Graphics (ToG)*, 2025.

[J5] M. Bhattacharjee, P. Motlicek, S. Madikeri, H. Helmke, O. Ohneiser, M. Kleinert, and h. Ehr, Minimum effort adaptation of automatic speech recognition system in air traffic management, *European Journal of Transport and Infrastructure Research*, 2025.

[J6] C. Bilaloglu, T. Löw, and S. Calinon, Tactile Ergodic Coverage on Curved Surfaces, *IEEE Transactions on Robotics (T-RO)*, 2025.

[J7] V. Bros and D. Gatica-Perez, Decoding community proximity discourse: A mixed-methods comparative analysis of online local and national newspapers in Romandy, Switzerland, *PLOS One*, 2025.

[J8] M. Busso, A. Bontempelli, L. J. Malcotti, L. B. Meegahapola, P. KUN, S. Diwakar, C. Nuttakki, M. R. Britez, H. Xu, D. Song, S. Ruiz-Correa, A. Mendoza-Lara, G. Gaskell, S. Stares, M. Bidoglia, A. Ganbold, A. Chagnaa, L. Cernuzzi, A. Hume, R. Chenu-Abente, R. A. Asiku, I. Kayongo, D. Gatica-Perez, A. de Götzen, I. Bison, and F. Giunchiglia, DiversityOne: A Multi-Country Smartphone Sensor Dataset for Everyday Life Behavior Modeling, *ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2025.

[J9] S. Calinon, Geometric Structures for Learning and Optimization in Robotics, *Annual Review of Control, Robotics, and Autonomous Systems*, 2025.

[J10] D. David, P. Baxter, T. Belpaeme, E. Billing, H. Cai, H.-L. Cao, A. Ciocan, C. Costescu, D. H. Garcia, P. G. Esteban, J. Kennedy, H. Liu, S. Matu, A. Mazel, M. Selescu, E. Senft, S. Thill, B. Vanderborght, D. Vernon, and T. Ziemke, Efficacy and effectiveness of robot-assisted therapy for autism spectrum disorder: From lab to reality, *Science Robotics*, vol. 10, no. 109, 2025.

[J11] M. Fatoreto, G. Özbulak, L. Berton, and A. Anjos, Optimizing Fairness and Utility in Healthcare Machine Learning Models, *Applied Soft Computing*, 2025.

[J12] M. Fuchs, E. Genty, A. Bangerter, K. Zuberbühler, J.-M. Odobez, and P. Cotofrei, From forest to zoo: great ape behavior recognition with ChimpBehave, *International Journal of Computer Vision*, 2025.

[J13] E. Germanier, A. Bangerter, K. Orji, L. A. Renier, M. Schmid Mast, M. He, and P. N. Garner, Responses to Past-Behavior Questions in Face-To-Face and Asynchronous Video Interviews: Storytelling, Interview Performance and Criterion Validity, *Human Performance*, 2025.

[J14] E. Germanier, M. He, A. M. Rufai, P. N. Garner, A. Bangerter, L. A. Renier, M. Schmid Mast, and K. Orji, Identifying storytelling in job interviews using deep learning, *Computers in Human Behavior Reports*, 2025.

[J15] O. Hrynenko and A. Cavallaro, Identifying Privacy Personas, *Proceedings on Privacy Enhancing Technologies*, 2025.

[J16] J. Huang, S. Bhattacharjee, S. Marcel, and W. Kang, Study of Full-View Finger Vein Biometrics on Redundancy Analysis and Dynamic Feature Extraction, *IEEE Transactions on Information Forensics and Security*, 2025.

[J17] J. Jankowski, L. Brudermuller, N. Hawes, and S. Calinon, Robust Pushing: Exploiting Quasi-static Belief Dynamics and Contact-informed Optimization, *International Journal of Robotics Research (IJRR)*, 2025.

[J18] K. Kotwal and S. Marcel, Review of Demographic Fairness in Face Recognition, *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2025.

[J19] L. Lanča, K. Jakac, S. Calinon, and S. Ivić, Ergodic exploration of dynamic distribution, *IEEE Robotics and Automation Letters (RA-L)*, 2025.

[J20] V. Lehmann, M. Hilpert, Z. Mostaani, S. Hovsepyan, E. Wallace, C. Verzat, S. Feuerriegel, M. Kraus, J. Rosenthal, G. Yilmaz, M. Magimai-Doss, and C. Stettler, Listening to Hypoglycemia: Voice as a Biomarker for Detection of a Medical Emergency Using Machine Learning, *Diabetes Care*, 2025.

[J21] Y. Li and S. Calinon, From Movement Primitives to Distance Fields to Dynamical Systems, *IEEE Robotics and Automation Letters (RA-L)*, 2025.

[J22] J. Liu, Z. Li, M. Yu, Z. Dong, S. Calinon, D. G. Caldwell, and F. Chen, Human-Humanoid Robots Cross-Embodiment Behavior-Skill Transfer Using Decomposed Adversarial Learning from Demonstration, *IEEE Robotics and Automation Magazine (RAM)*, 2025.

[J23] T. Löw, P. Abbet, and S. Calinon, GAFRO: Geometric Algebra for Robotics [Tutorial], *IEEE Robotics and Automation Magazine*, 2025.

[J24] W. Lu, H. Yu, J. Wang, D. Teney, H. Wang, Y. Chen, Q. Yang, X. Xie, and X. Ji, ZooPFL: Exploring Black-box Foundation Models for Personalized Federated Learning, *IEEE Transactions on Neural Networks and Learning Systems*, 2025.

[J25] S. Madikeri, P. Motlicek, D. Sanchez-Cortes, P. Rangappa, J. Hughes, J. Tkaczuk, A. Sanchez Lara, D. Khalil, J. Rohdin, D. Zhu, A. Krishnan, D. Klakow, Z. Ahmadi, M. Kovac, D. Boboš, C. Kalogiros, A. Alexopoulos, and D. Marraud, Autocrime - open multimodal platform for combating organized crime, *Forensic Science International: Digital Investigation*, 2025.

[J26] P. Melzi, H. Otroshi Shahreza, C. Rathgeb, R. Tolosana, R. Vera-Rodriguez, J. Fierrez, S. Marcel, and C. Busch, Cancelable Face Biometrics With Soft-Biometric Privacy Enhancement, *IEEE Access*, 2025.

[J27] P. O'Regan, F. Butt, L. Carter, D. M. Graham, A. Le Blanc, R. Hoskins, L. Stephenson, A. Patil, M. Shabbir, D. Eken, S. Singh, A. Villa, L. Agnelli, S. Damian, C. Grave, G. Pretelli, E. Garralda, H. Frost, F. de Braud, A. Freitas, C. Dive, and H. Unsworth, UpSMART: five years of digital innovation in cancer clinical research—achievements, challenges, and recommendations, *Frontiers in Digital Health*, 2025.

[J28] H. Otroshi Shahreza, A. George, R. Parsa, A. Unnervik, and S. Marcel, Second FRCSyn-onGoing: Winning Solutions and Post-Challenge Analysis to Improve Face Recognition with Synthetic Data, *Information Fusion*, 2025.

[J29] H. Otroshi Shahreza and S. Marcel, Foundation Models and Biometrics: A Survey and Outlook, *IEEE Transactions on Information Forensics and Security*, 2025.

[J30] M. D. Pretto, M. Laganaro, and I. Kodrasi, ERP Signals During Speech Articulation: Does Auditory Feedback Mask Other Ongoing Cognitive-motor Processes? *Brain Topography*, 2025.

[J31] A. Razmjoo, T. Xue, S. Shetty, and S. Calinon, Sampling-Based Constrained Motion Planning with Products of Experts, *International Journal of Robotics Research*, 2025.

[J32] E. Sarkar, K. Wierucka, A. B. Bosshard, J. Burkart, and M. Magimai-Doss, On feature representations for marmoset vocal communication analysis, *Bioacoustics: The International Journal of Animal Sound and its Recording*, 2025.

[J33] L. Schewski, M. Magimai-Doss, G. Beldi, and S. Keller, Measuring negative emotions and stress through acoustic correlates in speech: A systematic review, *PLoS One*, 2025.

[J34] S. Shakeel, M. Sahidullah, and I. Kodrasi, Overview of Automatic Speech Analysis and Technologies for Neurodegenerative Disorders: Diagnosis and Assistive Applications, *IEEE Journal of Selected Topics in Signal Processing*, 2025.

[J35] D. Šipka, B. Vlasenko, M. Stein, T. Dierks, M. Magimai-Doss, and Y. Morishima, Multidisciplinary characterization of embarrassment through behavioral and acoustic modeling, *Scientific reports*, 2025.

[J36] N. Tarigopula, S. Tornay, O. Mercanoglu Sincan, R. Bowden, and M. Magimai-Doss, Posterior-based analysis of spatio-temporal features for Sign Language Assessment, *IEEE Open Journal of Signal Processing*, 2025.

[J37] M. Valentino and A. Freitas, On the Nature of Explanation: An Epistemological-Linguistic Perspective for Explanation-Based Natural Language Inference, *Philosophy & Technology*, 2025.

[J38] T. Xue, A. Razmjoo Fard, S. Shetty, and S. Calinon, Robust Contact-rich Manipulation through Implicit Motor Adaptation, *International Journal of Robotics Research*, 2025.

[J39] Y. Zhang, T. Xue, A. R. Fard, and S. Calinon, Learning problem decomposition for efficient sequential multi-object manipulation planning, *IEEE Robotics and Automation Letters*, 2025.

Conference Papers

[1] A. Ali, R. Dai, A. Marisetty, G. Astruc, M. Thonnat, J.-M. Odobez, S. Thümmler, and F. Bremond, Loose Social-Interaction Recognition in Real-world Therapy Scenarios, *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025.

[2] N. Aljaafari, D. Carvalho, and A. Freitas, CARMA: Enhanced Compositionality in LLMs via Advanced Regularisation and Mutual Information Alignment, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025.

[3] N. Aljaafari, D. Carvalho, and A. Freitas, TRACE: Training and Inference-Time Interpretability Analysis for Language Models, *Demonstration at the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025.

[4] T. Almeida, G. Leticio, L. Pascotti, A. Freitas, and D. Pedronette, Effective Graph and Rank-based Contextual Embeddings for Textual and Multimedia Data, *International Joint Conference on Neural Networks*, 2025.

[5] M. Amiri, H. Otroshi Shahreza, and I. Kodrasi, Exploring In-Context Learning Capabilities of ChatGPT for Pathological Speech Detection, *ITG Conference on Speech Communication*, IEEE, 2025.

[6] R. Annapureddy, A. Fornaroli, M. Fattori, V. Lacovara, E. Fiori, S. Vollmer, M. Konradi, B. E. Hecking, G. Todesco, and D. Gatica-Perez, Co-Designing with Multiple Stakeholders and Datasets: A Community-Centered Process to Understand Youth Deviance in the Italian City of Turin, *International Conference on Communities & Technologies*, ser. C&T '25, New York, NY, USA: Association for Computing Machinery, 2025.

[7] O. Beker, N. Gürtler, J. Shi, A. René Geist, A. Razmjoo, G. Martius, and S. Calinon, A Smooth Analytical Formulation of Collision Detection and Rigid Body Dynamics with Contact, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025.

[8] D. Berio, S. Calinon, R. Plamondon, and F. F. Leymarie, Differentiable rasterization of minimum-time sigma-lognormal trajectories, *Conference of the International Graphonomics Society (IGS)*, 2025.

[9] D. Berio, G. Clivaz, M. Stroh, O. Deussen, R. Plamondon, S. Calinon, and F. F. Leymarie, Image-driven robot drawing with rapid lognormal movements, *IEEE International Symposium on Robot and Human Interactive Communication (Ro-Man)*, 2025.

[10] R. Boghetti and J. Kämpf, Optimizing Supply Temperature Control in District Heating Networks via Differentiable Dynamic Simulation and Gradient Descent, *International Conference on Construction, Energy, Environment and Sustainability*, 2025.

[11] V. Bros and D. Gatica-Perez, The Suisse Romande Local News Dataset, *International AAAI Conference on Web and Social Media*, 2025.

[12] S. Calinon, Movement Generation and Drawing in Robotics, *Conference of the International Graphonomics Society (IGS)*, 2025.

[13] A. Carofilis, P. Rangappa, S. Madikeri, S. Kumar, S. Burdisso, J. Prakash, E. Villatoro-Tello, P. Motlicek, B. Sharma, K. Hacioğlu, S. Venkatesan, S. Vyas, and A. Stolcke, Better Semi-supervised Learning for Multi-domain ASR Through Incremental Retraining and Data Filtering, *Interspeech 2025*, 2025.

[14] D. Carvalho, E. Manino, J. Rozanova, L. Cordeiro, and A. Freitas, Montague semantics and modifier consistency measurement in neural language models, *31st International Conference on Computational Linguistics*, 2025.

[15] D. Carvalho, Y. Zhang, H. Unsworth, and A. Freitas, LangVAE and LangSpace: Building and Probing for Language Model VAEs, *Demonstration at Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025.

[16] B. Céspedes-Sarrias, C. Collado-Capell, P. Rodenas-Ruiz, O. Hrynenko, and A. Cavallaro, MM-HSD: Multi-Modal Hate Speech Detection in Videos, *ACM International Conference on Multimedia*, 2025.

[17] H. Chen and P. N. Garner, A Bayesian Interpretation of Adaptive Low-Rank Adaptation, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2025.

[18] V. Cheremetiev, Q. L. H. Ngo, C. Y. Kot, A. E. Baia, and A. Cavallaro, Specializing General-purpose LLM Embeddings for Implicit Hate Speech Detection across Datasets, *International Workshop on Diffusion of Harmful Content on Online Web (DHOW)*, 2025.

[19] X. Chi, H. Girgin, T. Löw, Y. Xie, T. Xue, J. Huang, Z. Liu, and S. Calinon, Efficient and Real-Time Motion Planning for Robotics Using Projection-Based Optimization, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025.

[20] A. C. Coman, I.-T. Sorodoc, L. F. R. Ribeiro, B. Byrne, J. Henderson, and A. de Gispert, RAGferee: Building Contextual Reward Models for Retrieval-Augmented Generation, *Empirical Methods in Natural Language Processing*, 2025.

[21] A. C. Coman, C. Theodoropoulos, M.-F. Moens, and J. Henderson, Fast-and-Frugal Text-Graph Transformers are Effective Link Predictors, *Findings of the Association for Computational Linguistics*, 2025.

[22] M. Delmas, M. Wysocka, D. Gusicuma, and A. Freitas, Accelerating Antibiotic Discovery with Large Language Models and Knowledge Graphs, *Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, 2025.

[23] M. Dickerman, A. Gupta, S. Tafasca, X. Zhang, J.-M. Odobez, and S. Stoll, Automatic detection of the visual gaze components of joint attention in observational, naturalistic child language acquisition data, *Boston University Conference on Language Development*, 2025.

[24] B. G. Doan, A. Shamsi, X.-Y. Guo, A. Mohammadi, H. Alinejad-Rokny, D. Teney, D. Ranasinghe, and E. Abbasnejad, Bayesian low-rank learning (Bella): A practical approach to bayesian neural networks, *AAAI Conference on Artificial Intelligence*, 2025.

[25] K. El Hajal, E. Hermann, S. Hovsepyan, and M. Magimai-Doss, Unsupervised Rhythm and Voice Conversion to Improve ASR on Dysarthric Speech, *Interspeech*, ISCA, 2025.

[26] K. El Hajal, E. Hermann, A. Kulkarni, and M. Magimai-Doss, Unsupervised Rhythm and Voice Conversion of Dysarthric to Healthy Speech for ASR, *Proceedings of Workshop on Speech Pathology Analysis and DEtection (SPADE)*, IEEE, 2025.

[27] K. El Hajal, A. Kulkarni, E. Hermann, and M. Magimai-Doss, kNN Retrieval for Simple and Effective Zero-Shot Multi-speaker Text-to-Speech, *Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL)*, ACL, 2025.

[28] F. Fehr, CoRet: Improved Retriever for Code Editing, *Annual Meeting of the Association for Computational Linguistics*, 2025.

[29] F. Fehr, A. E. Baia, X. Chang, A. C. Coman, K. El Hajal, D. El Zein, S. Kumar, J. Zuluaga-Gomez, A. Cavallaro, D. Teney, and J. Henderson, Fine-Tuning Pretrained Models with NVIB for Improved Generalisation, *Workshop on Spurious Correlation and Shortcut Learning: Foundations and Solutions*, 2025.

[30] J. Gandarela, D. Carvalho, and A. Freitas, Inductive Learning of Logical Theories with LLMs: A Complexity-graded Analysis, *The Annual AAAI Conference on Artificial Intelligence*, 2025.

[31] D. Geissbuhler, A. Bornet, C. Marques, A. Anjos, S. Pereira, and D. Teodoro, STM-GNN: Space-Time-and-Memory Graph Neural Networks for Predicting Multi-Drug Resistance Risks in Dynamic Patient Networks, *International Conference on Artificial Intelligence in Medicine*, 2025.

[32] D. Geissbuhler, H. Otroshi Shahreza, and S. Marcel, Synthetic Face Datasets Generation via Latent Space Exploration from Brownian Identity Diffusion, *The Forty-second International Conference on Machine Learning (ICML)*, 2025.

[33] D. Geissbuhler, A. Pena-Bello, J. Kämpf, and J. Rager, OpenBEERS: A digital platform for urban scale simulation of building energy efficiency, *Journal of Physics: Conference Series*, IOP Publishing, 2025.

[34] A. George and S. Marcel, Digi2Real: Bridging the Realism Gap in Synthetic Data Face Recognition via Foundation Models, *IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, 2025.

[35] A. George and S. Marcel, EdgeDoc: Hybrid CNN-Transformer Model for Accurate Forgery Detection and Localization in ID Documents, *ICCV*, 2025.

[36] A. George and S. Marcel, Enhancing Domain Diversity in Synthetic Data Face Recognition with Dataset Fusion, *IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, 2025.

[37] A. George and S. Marcel, The Invisible Threat: Evaluating the Vulnerability of Cross-Spectral Face Recognition to Presentation Attacks, *International Joint Conference on Biometrics (IJCB 2025), IEEE.*, 2025.

[38] A. George and S. Marcel, xEdgeFace: Efficient Cross-Spectral Face Recognition for Edge Devices, *International Joint Conference on Biometrics (IJCB 2025), IEEE.*, 2025.

[39] L. Coppieters de Gibson and P. N. Garner, Exploring auditory feedback mechanisms in speech recognition, *Interspeech 2025*, 2025.

[40] M. He and P. N. Garner, Joint Fine-tuning and Conversion of Pretrained Speech and Language Models towards Linear Complexity, *13th International Conference on Learning Representations (ICLR)*, 2025.

[41] E. Hermann, K. El Hajal, A. Kulkarni, and M. Magimai-Doss, Idiap kNN-TTS System for the Blizzard Challenge 2025, *Blizzard Challenge Workshop*, 2025.

[42] M. Hou and I. Kodrasi, Variational Autoencoder for Personalized Pathological Speech Enhancement, *European Signal Processing Conference*, 2025.

[43] S. Hovsepyan and M. Magimai-Doss, Speech power spectra: a window into neural oscillations in Parkinson's disease, *Interspeech*, 2025.

[44] T. Inada, E. Villatoro-Tello, J. Park, J. Pulcrano, and B. F. Leleux, The Greatest Challenge For Startups: Computational Text Analysis on Swiss Ventures, *Academy of Management Proceedings 2025.*, 2025.

[45] J. Jankowski, A. Marić, P. Liu, D. Tateo, J. Peters, and S. Calinon, Distilling Contact Planning for Fast Trajectory Optimization in Robot Air Hockey, *Proceedings of Robotics: Science and Systems*, 2025.

[46] I. M. De la Jara, C. Rodriguez-Opazo, D. Teney, D. Ranasinghe, and E. Abbasnejad, Mysteries of the Deep: Role of Intermediate Representations in Out of Distribution Detection, *Advances in neural information processing systems*, 2025.

[47] L. Jiang and D. Teney, OOD-Chameleon: Is Algorithm Selection for OOD Generalization Learnable? *Forty-Second International Conference on Machine Learning*, 2025.

[48] Y. Kaloga and I. Kodrasi, Towards interpretable emotion recognition: Identifying key features with machine learning, *Forum Acusticum/EuroNoise*, 2025.

[49] Y. Kaloga, S. Shakeel, and I. Kodrasi, Multiview Canonical Correlation Analysis for Automatic Pathological Speech Detection, *International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2025.

[50] G. Kim, M. Valentino, and A. Freitas, A Mechanistic Interpretation of Syllogistic Reasoning in Auto-Regressive Language Models, *Findings of the ACL*, 2025.

[51] A. Komaty, H. Otroshi Shahreza, A. George, and S. Marcel, Exploring ChatGPT for Face Presentation Attack Detection in Zero and Few-Shot in-Context Learning, *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025.

[52] P. Korshunov, K. Kotwal, C. Ecabert, V. Vidit, A. Mohammadi, and S. Marcel, Investigation of accuracy and bias in face recognition trained with synthetic data, *Proceedings of IEEE International Joint Conference on Biometrics*, 2025.

[53] P. Korshunov, A. Mohammadi, V. Vidit, C. Ecabert, and S. Marcel, FantasyID: A dataset for detecting digital manipulations of ID-documents, *Proceedings of IEEE International Joint Conference on Biometrics*, 2025.

[54] P. Korshunov, V. Vidit, A. Mohammadi, C. Ecabert, N. Shamoska, S. Marcel, Z. Yu, Y. Tian, J. Ni, L. Lazarevic, R. Khizbullin, A. Evteeva, A. Tochin, A. Grishin, A. George, D. DeAlcala, T. Endrei, J. Munoz-Haro, R. Tolosana, R. Vera-Rodriguez, A. Morales, J. Fierrez, G. Cserey, H. Sharma, S. Chaudhary, A. Dudhane, P. Hambarde, A. Shukla, P. Shaily, J. Kumar, A. Hase, S. Maurya, M. Sharma, and P. Dwivedi, DeepID Challenge of Detecting Synthetic Manipulations in ID Documents, *International Conference on Computer Vision (ICCV)*, 2025.

[55] V. Krivokuca, G. Stragapede, S. Merrick, J. Sukaitis, and V. Graf Narbel, Securing Face and Fingerprint Templates in Humanitarian Biometric Systems, *International Joint Conference on Biometrics (IJCB)*, IEEE, 2025.

[56] A. Kulkarni, D. Sandipana, M. Magimai-Doss, and T. alumae, Unveiling Audio Deepfake Origins: A Deep Metric learning And Conformer Network Approach With Ensemble Fusion, *Interspeech*, 2025.

[57] A. Kulkarni, F. Teixeira, E. Hermann, T. Rolland, I. Trancoso, and M. Magimai-Doss, Children's Voice Privacy: First Steps and Emerging Challenges, *Interspeech*, 2025.

[58] S. Kumar, Y. Kaloga, J. Mitros, P. Motlicek, and I. Kodrasi, Latent Space Factorization in LoRA, *Conference on Neural Information Processing Systems*, 2025.

[59] S. Kumar, S. Madikeri, E. Villatoro-Tello, S. Burdisso, P. Rangappa, A. Carofilis, P. Motlicek, K. P. D. S, S. Venkatesan, K. Hacioğlu, and A. Stolcke, TokenVerse++: Towards Flexible Multitask Learning with Dynamic Task Activation, *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, IEEE, 2025.

[60] S. Kumar, S. Madikeri, J. Zuluaga-Gomez, E. Villatoro-Tello, I. Thorbecke, P. Motlicek, M. K. E, and A. Ganapathiraju, XLSR-Transducer: Streaming ASR for Self-Supervised Pretrained Models, *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, IEEE, 2025.

[61] S. Kumar, I. Thorbecke, S. Burdisso, E. Villatoro-Tello, M. K. E, K. Hacioğlu, P. Rangappa, P. Motlicek, A. Ganapathiraju, and A. Stolcke, Performance Evaluation of SLAM-ASR: The Good, the Bad, the Ugly, and the Way Forward, *SALMA Workshop, Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, IEEE, 2025.

[62] C. Li, J. Liu, T. Teng, S. Wang, S. Calinon, and F. Chen, Whole-Body Impedance Control of a Humanoid Robot Based on Human-Human Demonstration for Human-Robot Collaboration, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025.

[63] Z. Li, J. Liu, D. Li, T. Teng, M. Li, S. Calinon, D. G. Caldwell, and F. Chen, ManiDP: Manipulability-Aware Diffusion Policy for Posture-Dependent Bimanual Manipulation, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025.

[64] C. Liu, Y. Zhang, D. Zhang, W. Zhang, C. Gong, H. Li, Y. Lu, S. Zhou, Y. Lu, Z. Gan, Z. Wang, J. Liao, H. Wu, J. Liu, A. Freitas, Q. Wang, Z. Xu, R. Zhang, and Y. Dai, Nexus: An Omni-Perceptive And-Interactive Model for Language, Audio, And Vision, *ACM Multimedia*, 2025.

[65] G. Marshall, A. Freitas, and C. Jay, An evidence-based guidance framework for neural network system diagrams, *PLOS One*, 2025.

[66] J. Meadows, M. Valentino, and A. Freitas, Controlling Equational Reasoning in Large Language Models with Prompt Interventions, *The Annual AAAI Conference on Artificial Intelligence*, 2025.

[67] M. van der Meer, P. Korshunov, S. Marcel, and L. van der Plas, HintsOfTruth: A Multimodal Checkworthiness Detection Dataset with Real and Synthetic Claims, *Annual Meeting of the Association for Computational Linguistics*, 2025.

[68] P. Motlicek, S. Kumar, D. Khalil, A. Prasad, and S. Christof, Leveraging Untranscribed Data for End-to-End Speech and Callsign Recognition in Air-Traffic Communication, *SESAR Innovation Days 2025*, 2025.

[69] V. Nastase, G. Samo, C. Jiang, and P. Merlo, Multilingual vs. monolingual transformer models in encoding linguistic structure and lexical abstraction, *Italian Conference on Computational Linguistics (CLiC-it)*, 2025.

[70] H. Otroshi Shahreza, L. Colbois, and S. Marcel, 3D Face Morph Generation Using Geometry-Aware Template Inversion, *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2025.

[71] H. Otroshi Shahreza, A. George, and S. Marcel, Face Reconstruction from Face Embeddings using Adapter to a Face Foundation Model, *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2025.

[72] H. Otroshi Shahreza and S. Marcel, FaceLLM: A Multimodal Large Language Model for Face Understanding, *IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, 2025.

[73] H. Otroshi Shahreza and S. Marcel, Generating Synthetic Face Recognition Datasets Using Brownian Identity Diffusion and a Foundation Model, *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2025.

[74] H. Otroshi Shahreza and S. Marcel, HyperFace: Generating Synthetic Face Recognition Datasets by Exploring Face Embedding Hypersphere, *The Thirteenth International Conference on Learning Representations*, 2025.

[75] R. Parsa, D. Teney, and S. Marcel, AugGen: Synthetic Augmentation using Diffusion Models Can Improve Recognition, *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

[76] M. Piro, J. Kämpf, I. Ballarini, and V. Corrado, Assessing the reliability of archetype-based Urban Building Energy Simulations: A case study analysis in Turin (Italy), *Journal of Physics: Conference Series*, IOP Publishing, 2025.

[77] F. Poh, A. George, and S. Marcel, ArtFace: Towards Historical Portrait Face Identification via Model Adaptation, *(Non-Archival)*, 2025.

[78] T. Purohit and M. Magimai-Doss, Emotion information recovery potential of wav2vec2 network fine-tuned for speech recognition task, *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, IEEE, 2025.

[79] T. Purohit, B. Ruvolo, J. R. Orozco-Arroyave, and M. Magimai-Doss, Automatic Parkinson's disease detection from speech: Layer selection vs adaptation of foundation models, *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, IEEE, 2025.

[80] X. Quan, M. Valentino, D. Carvalho, D. Dalal, and A. Freitas, PEIRCE: Unifying Material and Formal Reasoning via LLM-Driven Neuro-Symbolic Refinement, *Demonstration at Annual Meeting of the Association for Computational Linguistics*, 2025.

[81] X. Quan, M. Valentino, L. Dennis, and A. Freitas, Faithful and Robust LLM-Driven Theorem Proving for NLI Explanations, *Annual Meeting of the Association for Computational Linguistics*, 2025.

[82] S. Rajapakshe, J.-M. Odobez, and E. Senft, Giving Sense to Inputs: Toward an Accessible Control Framework for Shared Autonomy, *ACM/IEEE International Conference on Human-Robot Interaction*, ACM, 2025.

[83] L. Ranaldi, M. Valentino, and A. Freitas, Eliciting Critical Reasoning in Retrieval-Augmented Language Models via Contrastive Explanations, *Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics*, 2025.

[84] L. Ranaldi, M. Valentino, A. Polonsky, and A. Freitas, Improving chain-of-thought reasoning via quasi-symbolic abstractions, *Annual Meeting of the Association for Computational Linguistics*, 2025.

[85] P. Rangappa, A. Carofilis, J. Prakash, S. Kumar, S. Burdisso, S. Madikeri, E. Villatoro-Tello, B. Sharma, P. Motlicek, K. Hacioğlu, S. Venkatesan, S. Vyas, and A. Stolcke, Efficient Data Selection for Domain Adaptation of ASR Using Pseudo-Labels and Multi-Stage Filtering, *Interspeech*, 2025.

[86] P. Rangappa, P. Motlicek, D. Sanchez-Cortes, A. S. Lara, M. Antonopoulou, I. Fourfouris, N. Avgerinos, and M. Tsangaris, Accelerating Criminal Investigations with TRACY, *16th EAI International Conference on Digital Forensics & Cyber Crime*, 2025.

[87] P. Rangappa, J. Zuluaga-Gomez, S. Madikeri, A. Carofilis, J. Prakash, S. Burdisso, S. Kumar, E. Villatoro-Tello, N. Iuliia, P. Motlicek, K. P. D. S, and A. Ganapathiraju, Speech Data Selection for Efficient ASR Fine-Tuning using Domain Classifier and Pseudo-Label Filtering, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2025)*, 2025.

[88] H. Rauf, A. Bogatu, N. Paton, and A. Freitas, Gem: Gaussian Mixture Model Embeddings for Numerical Feature Distributions, *8th International Conference on Extending Database Technology*, 2025.

[89] H. Rauf, A. Freitas, and N. Paton, TableDC: Deep Clustering for Tabular Data, *ACM SIGMOD International Conference on Management of Data*, 2025.

[90] A. Razmjoo, S. Calinon, M. Gienger, and F. Zhang, CCDP: Composition of Conditional Diffusion Policies with Guided Sampling, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025.

[91] A. Razmjoo, S. Calinon, M. Gienger, and F. Zhang, CCDP: Model-free Failure Recovery via Guided Diffusion Sampling, *Workshop on The Art of Robustness: Surviving Failures in Robotics, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025.

[92] D. Rodrigue, R. Boghetti, J. Kämpf, B. Pasdeloup, M. T. Mabrouk, P. Meyer, and B. Lacarrière, Validation of two distinct simulation models of district heating networks: application to efficient looping analysis, *Journal of Physics: Conference Series*, IOP Publishing, 2025.

[93] C. Rodriguez-Opazo, E. Abbasnejad, D. Teney, H. Damirchi, E. Marrese-Taylor, and A. van den Hengel, Synergy and diversity in CLIP: Enhancing performance through adaptive backbone ensembling, *International Conference on Learning Representations*, 2025.

[94] B. Ruvolo, T. Purohit, B. Vlasenko, J. R. Orozco-Arroyave, and M. Magimai-Doss, Exploring the Complexity of Parkinson's Patient Speech for Depression Detection task: A Qualitative Analysis, *Proceedings of Workshop on Speech Pathology Analysis and DEtection (SPADE)*, IEEE, 2025.

[95] L. S. Luevano, P. Korshunov, and S. Marcel, Identity-Preserving Aging and De-Aging of Faces in the StyleGAN Latent Space, *IEEE International Joint Conference on Biometrics (IJCB)*, IEEE, 2025.

[96] E. Sarkar and M. Magimai-Doss, Comparing Self-Supervised Learning Models Pre-Trained on Human Speech and Animal Vocalizations for Bioacoustics Processing, *International Conference on Acoustics, Speech and Signal Processing*, 2025.

[97] E. Sarkar and M. Magimai-Doss, Towards Leveraging Sequential Structure in Animal Vocalizations, *Neural Information Processing Systems workshop: AI for Non-Human Animal Communication*, 2025.

[98] S. Shakeel, Y. Kaloga, M. Sahidullah, and I. Kodrasi, Graph Neural Networks for Parkinson's Disease Detection, *International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2025.

[99] Z. Shinnick, L. Jiang, H. Saratchandran, A. van den Hengel, and D. Teney, Transformers Pretrained on Procedural Data Contain Modular Structures for Algorithmic Reasoning, *ICML Workshop on Methods and Opportunities at Small Scale*, 2025.

[100] K. Song, X. Wang, X. Tan, H. Jiang, C. Zhang, Y. Shen, C. Lu, Z. Li, Z. Song, C. Shan, Y. Wang, K. Ren, X. Zheng, T. Qin, Y. Yang, D. Li, and L. Qiu, Chain-of-Model Learning for Language Model, *Conference on Neural Information Processing Systems*, 2025.

[101] J. E. Tapia, N. Mario, J. Espin, A. Sanchez, N. Damer, C. Busch, M. Ivanovska, L. Todorov, R. Khizbullin, A. Grishin, L. Lazarevic, D. Schulz, S. Gonzalez, A. Mohammadi, K. Kotwal, S. Marcel, R. Mudgalgundurao, K. B. Raja, P. Schuch, P. Couto, J. Pinto, M. Xavier, A. Valenzuela, B. Batagelj, J. Barrachina, M. Peterlin, P. Peer, A. Muhammed, D. Nunes, N. Gonçalves, S. Patwardhan, and R. Ramachandra, Second Competition on Presentation Attack Detection on ID Card, *IEEE International Joint Conference on Biometrics (IJCB)*, 2025.

[102] D. Teney, L. Jiang, F. Gogianu, and E. Abbasnejad, Do We Always Need the Simplicity Bias? Looking for Optimal Inductive Biases in the Wild, *The IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025.

[103] S. Tornay and M. Magimai-Doss, Towards Dynamic Skeleton-based Handshape Subunits for Sign Language Assessment, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025.

[104] L. Vásquez-Rodríguez, B. Audrin, S. Michel, S. Galli, J. Rogenhofer, J. Negro Cusa, and L. van der Plas, Soft Skills in the Wild: Challenges in Multilingual Classification, *Swiss Text Analytics Conference*, 2025.

[105] V. Vidit, P. Korshunov, A. Mohammadi, C. Ecabert, K. Kotwal, and S. Marcel, Detecting Text Manipulation in Images using Vision Language Models, *36th British Machine Vision Conference 2025*, 2025.

[106] M. Vitek, D. Tomašević, A. Das, S. Nathan, G. Özbulak, G. A. Tataroğlu Özbulak, J.-P. Calbimonte, A. Anjos, H. Hemant Bhatt, D. Dhirendra Premani, J. Chaudhari, C. Wang, J. Jiang, C. Zhang, Q. Zhang, I. Iyappan Ganapathi, S. Sadaf Ali, D. Velayudan, M. Assefa, N. Werghi, Z. A Daniels, L. John, R. Vyas, J. Nourmohammadi Khiarak, T. Akbari Saeed, M. Nasehi, A. Kianfar, M. Pashazadeh Panahi, G. Sharma, P. Raj Panth, R. Ramachandra, A. Nigam, U. Pal, H. Pedrini, and V. Struc, Privacy-enhancing Sclera Segmentation Benchmarking Competition: SSBC 2025, *International Joint Conference on Biometrics*, IEEE, 2025.

[107] B. Vlasenko and M. Magimai-Doss, Multimodal Prosody Modeling: A Use Case for Multilingual Sentence Mode Prediction, *Interspeech*, 2025.

[108] P. Vuillecard and J.-M. Odobez, Enhancing 3D Gaze Estimation in the Wild using Weak Supervision with Gaze Following Labels, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.

[109] M. Wysocka, D. Carvalho, O. Wysocki, M. Valentino, and A. Freitas, SylloBio-NLI: Evaluating Large Language Models on Biomedical Syllogistic Reasoning, *Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics*, 2025.

[110] L. Zhang, M. Valentino, and A. Freitas, Autoformalization in the wild: assessing LLMs on real-world mathematical definitions, *Conference on Empirical Methods in Natural Language Processing*, 2025.

[111] L. Zhang, M. Valentino, and A. Freitas, Formalizing Complex Mathematical Statements with LLMs: A Study on Mathematical Definitions, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025.

[112] L. Zhang, M. Valentino, and A. Freitas, MASA: A Modular Framework for LLM-Driven Multi-Agent Systems for Autoformalization, *Demonstration at the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025.

[113] Y. Zhang, D. Carvalho, and A. Freitas, Quasi-symbolic Semantic Geometry over Transformer-based Variational AutoEncoders, *Conference on Computational Natural Language Learning*, 2025.

#AIforSociety

Idiap Research Institute | Rue Marconi 19 | 1920 Martigny | Switzerland