SCIENTIFIC REPORT 2023

idiap
RESEARCH INSTITUTE

Contents

1 Summary

↓ TODO for 2023. Copy from last year below.

Statutes and areas of activities: Founded in 1991 by the State of Valais, the City of Martigny, the Ecole Polytechnique Fédérale de Lausanne (EPFL), the University of Geneva, and Swisscom, the Idiap Research Institute[1] is an independent, non-profit, *Research Foundation* devoted to advanced research, training, developments and technology transfer in the areas of *Artificial and Cognitive Intelligence*, and is among the world-level leaders in the field.

Based in Martigny (Valais, Switzerland), the Institute is particularly active in advanced signal processing, machine learning, and knowledge management, applied to multiple domains, including voice and visual recognition, multimedia data mining, human-machine interaction, robotics, language analysis, and bio-imaging. Idiap is involved in numerous research projects at local, national and international levels. As part of its core missions, Idiap is also very active in academic training (Masters and PhD students), as well as in technology transfer, through collaboration with various industries, or through direct spin-off (and its incubator IdeArk S.A.[2]).

Affiliation and budget: Since its inception, Idiap has been a completely independent research institution, although academically affiliated with EPFL and the University of Geneva. Since July 2008, Idiap and EPFL are agreeing on a "Joint Development Plan", which is revisited and signed every 4 years. The last version was signed June 26, 2021, applicable until December 2024. Considered as part of the "ETH Strategic Domain" (Research Institution of National Importance[3]), Idiap is accredited and co-funded by the Federal Government, the State of Valais, and the City of Martigny, for a total averaging 45-50% of its annual budget, the remainder coming from competitive projects and industry. With an initial budget in 1991 of around 500 KCHF, the Idiap budget has been steadily growing over the last 30 years to be today well above 14.5 MCHF/year.

Projects: Idiap is active in numerous national and international projects (with an average of around 50 projects active simultaneously), while also being active in technology transfer and research contracts with industry. From 2001 to 2013, Idiap was also the Host Institution of one of the major National Centres of Competence in Research (NCCR) on "Interactive Multimodal Information Management" (IM2[4]). Project distributions and dynamics, including statistics of acceptance rates across all funding instruments, are presented in Section 4, page 33.

Staff and publications: With 16 research groups active in 2022 (Section 3.1, page 27) and a staff of approximately 150 people mainly composed of senior researchers, R&D engineers, post-docs, and PhD students, Idiap regularly generates a large amount of high-quality peer-reviewed international scientific publications per year (in addition to an equal amount of internal Research Reports), all available from a dedicated web site[5], automatically synchronised with EPFL Infoscience[6].

As part of our 2021-2024 Research Plan, "**Cross Research Groups**" (CRG) started to be created to foster collaborations between Idiap Research Groups. This new kind of CRG research structure was laid down a few years ago by Idiap's management with the objective to reinforce the institute's ability to foster collaborations, really critical in increasing technology transfer and impact on society. Their aim is to have long-term impacts on society thanks to an approach mixing both business oriented solutions

[1] www.idiap.ch
[2] www.theark.ch
[3] https://www.sbfi.admin.ch
[4] www.im2.ch
[5] publications.idiap.ch
[6] https://infoscience.epfl.ch/

and interdisciplinary scientific groundwork. In 2022, we initiated three CRGs (as part of the 16 Research Groups): (1) Neuro-Symbolic Learning & Reasoning (led by Dr. Andre Freitas), (2) AI for Trust (lead by Prof. Sébastien Marcel, and (3) Human-Centered Robotics & AI (led by Dr. Emmanuel Senft).

Open-source software libraries: Idiap pursues an active policy towards the open-source release of high-quality software libraries, after having been cleared through a strict quality control process and clear IPR and distribution rules. Idiap's web site[7] refers to 133 open-source libraries.

Public datasets: Idiap has always been driven towards open data and researchers are encouraged to have their data published through a FAIR repository [8] to promote reproductible science. All data distribution are centrally managed and comply with SNSF guidelines on open data as well as Swiss and European regulations on data protection (FADP and GDPR). Idiap web [9] site is currently referencing 80 publicly available datasets.

Reproducible research: All our work is always driven by fully reproducible frameworks[10] and Idiap regularly releases open source codes accompanying its publications[11].

Quality indicators: To calibrate and optimise its performance, Idiap regularly screens the progress of the Institute against multiple indicators related to its key missions, i.e., research, technology transfer, training, and competitive funding. Thus, in addition to the present Research Report, our 2022 Self-Assessment Report, provides a concise and factual picture of where we stand in terms of our organisation, group structure, human resources, quality of projects and research staff, academic and professional activities, publications, teaching and technology transfer activities.

Technology transfer is also one of Idiap's key missions. In our separate 2022 Self-Assessment report, we are reporting on the excellent performance and continuous increase of our technology transfer activities, patenting, collaboration with industries, also attracting startups, as well as initiating our own spin-off companies. In the present report, Section 5.5, page 58, we list the patents currently in our portfolio.

[7] https://www.idiap.ch/en/scientific-research/resources
[8] https://www.snf.ch/en/dMILj9t4LNk8NwyR/topic/open-research-data
[9] https://www.idiap.ch/en/scientific-research/resources
[10] See, e.g., https://gitlab.idiap.ch/bob/bob.ip.binseg
[11] See, e.g., https://www.idiap.ch/software/pbdlib/

2 Research Programs

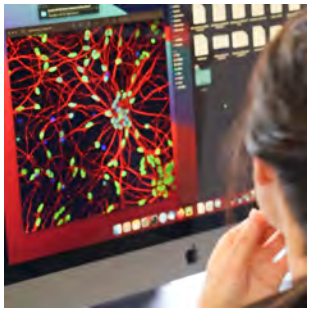
To remain competitive and sustainable in the fast-moving AI research domain, Idiap maintains a high density of technical expertise across complementary AI areas to be impactful at three levels:

1. internationally, indispensable to recruit global talent, to create close scientific links with top-tier research organizations, and to produce knowledge and research outcomes to disseminate through open-science scientific articles, code and data;
2. nationally, needed to lead and contribute to fundamental scientific projects within and across disciplines, to advance the state of art of AI for Society and to engage in a dialogue with both specialists and the general public around the opportunities and risks of AI development;
3. locally, to support and anticipate the specific needs by working with small and medium enterprises, cantonal and city government, and communities.

To achieve the objectives mentioned above at local, national and international scale, Idiap identified four broad Research Programs (RPs) whose objectives are driven by Switzerland's societal and industrial needs and innovation opportunities. The four research programs are Human-AI Teaming (RP1), AI for Sustainable and Resilient Societies (RP2), AI for Life (RP3), and AI for Everyone (RP4). These programs support targeted transfer and innovation and synergies across research groups.

Human-AI Teaming

Sustainable and Resilient Societies



AI for Life

AI for Everyone

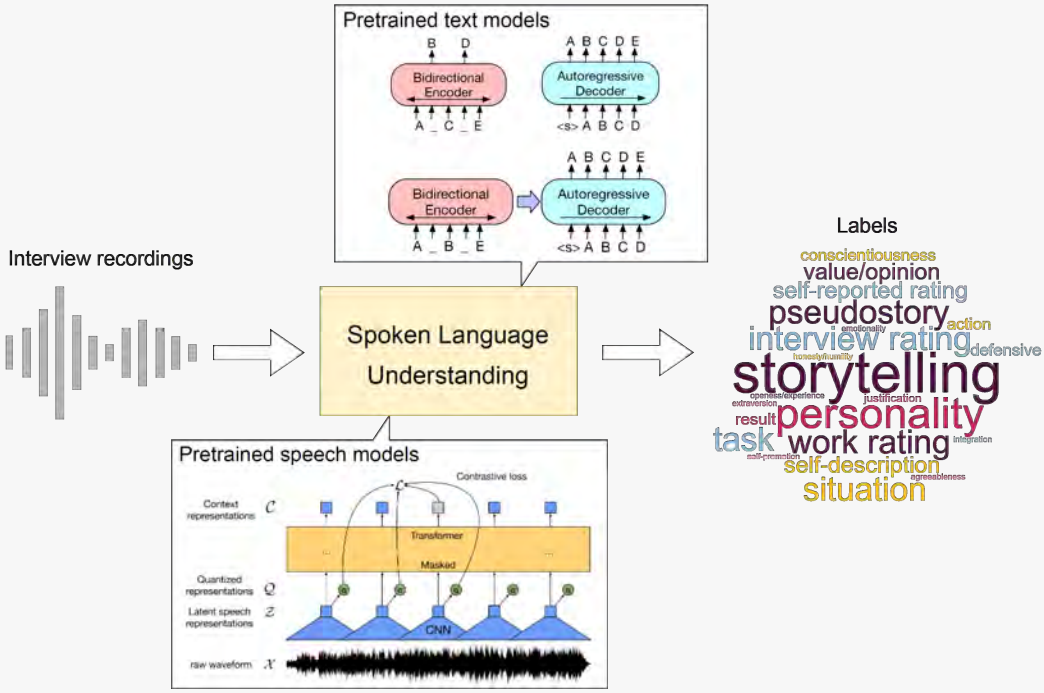
With its four Research Programs and its technology transfer activities, Idiap continues addressing fundamental challenges for the future of Switzerland, such as understanding how society should mobilize to interface with artificial intelligence systems and ensure positive human outcomes (RP1, RP4), acting upon the realization that climate change plays a pivotal role on setting up research priorities in artificial intelligence for a sustainable future (RP2), increasing trust in digital infrastructures and services with secure and privacy-preserving technologies (RP2), building upon national strengths in domains like health to address Switzerland's shifting demographic trends and strengthen the health industry ecosystem (RP1, RP3), and acknowledging the risk of dependency on foreign technology, and that national expertise and leadership is fundamental to mitigate such risk (knowledge and technology transfer activities). The aims and associated activities of the Research Programs align with the United Nations 2030 Agenda for sustainable development.

2.1 Human-AI Teaming

This research program capitalizes on the well-established expertise at Idiap on multimodal interaction. It leverages Idiap's unique ability to undertake in-depth multidisciplinary research across verbal and nonverbal communication, language processing, perceptual and cognitive systems, and human-robot interaction. The aim of this program is to expand human capabilities on several aspects (creativity, cognitive limitations, collaboration, knowledge). This research program aims to improve the sensing and understanding by machines of human activities, improve information access (e.g., through chatbots serving as on-demand domain experts), use human feedback for improving learning systems, and use robots to assist humans in everyday tasks at work and at home.

Spoken language understanding

In collaboration with psychologists at the universities of Lausanne and Neuchatel, we are attempting to predict the suitability of candidates for jobs via selection interviews. More generally, this is a spoken language understanding problem for which current solutions use foundation models. In 2023 we were able to show that translation can be used an an auxiliary task whilst fine-tuning such models [1], and that large language models such as ChatGPT are not far behind the bespoke state of the art [2].



[1] M. He and P. N. Garner, "The interpreter understands your meaning: End-to-end spoken language understanding aided by speech translation," in *Findings of EMNLP*, 2023

[2] M. He and P. N. Garner, "Can ChatGPT detect intent? evaluating large language models for spoken language understanding," in *Interspeech*, 2023

Inference of charisma from text

Previously with colleagues at the university of Lausanne, we have shown that charismatic traits of an individual can be inferred from text. More recently, that work has been used by colleagues in the US to show that charisma associated with speeches can be measured in COVID-mitigating behaviour.

U. Jensen, D. Rohner, O. Bornet, *et al.*, "Combating COVID-19 with charisma: Evidence on governor speeches in the United States," *The Leadership Quarterly*, May 2023

Blizzard challenge for speech synthesis

The Blizzard challenge is an annual competition to create state-of-the-art text-to-speech synthesis systems using data supplied by the organisers. The 2023 challenge was the first to be conducted for the French language, and the first in recent years to which Idiap has submitted a system. Our entry was evaluated quite favourably, behind much larger commercial labs but ahead of other francophone ones [1]. Some of the underlying techniques were published at the associated speech synthesis workshop [2].

[1] H. Chen, M. He, L. Coppieters de Gibson, *et al.*, "The idiap speech synthesis system for the blizzard challenge 2023," in *Proceedings of the 18th Blizzard Challenge Workshop*, 2023
[2] H. Chen and P. N. Garner, "Diffusion transformer for adaptive text-to-speech," in *Proceedings of the 12th ISCA Speech Synthesis Workshop (SSW 12)*, 2023

Geometrically-consistent gaze following model

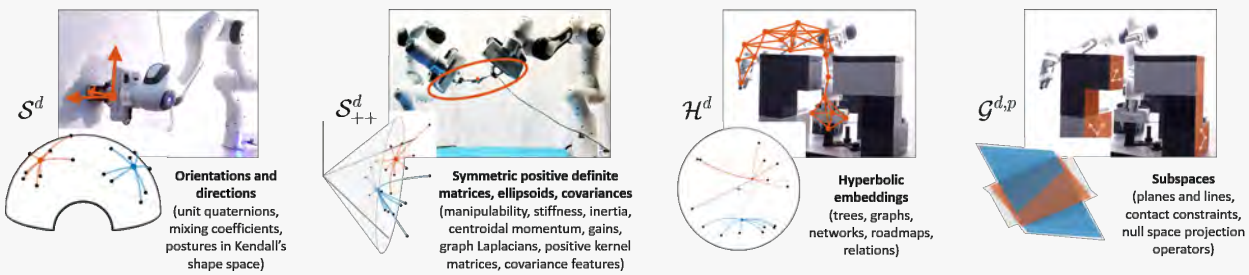
Instead of concentrating on gaze estimation in very specific settings, we focus on the gaze-following task—inferring people's gaze direction and 2D focus of attention—requiring a deep understanding of scene structure and context. Traditionally, depth is crucial for filtering out salient items not visible in 3D space along a person's 2D line of sight. However, prior works relied on depth estimated from pre-trained models, leading to stretched and distorted 3D scenes. To overcome this, we introduced a geometrically grounded approach using a novel depth estimation algorithm, producing geometry-preserving point clouds. Our method learned to predict a 3D gaze vector congruent with the point cloud, enhancing 3D Field of View (3DFoV) prediction and outperforming existing benchmarks.

Our 3D Field of View (3DFoV) highlights potential gaze targets, excluding objects where the depth does not match. Gaze target predictions in green and GT ones in red.

S. Tafasca, A. Gupta, and J.-M. Odobez, "Childplay: A new benchmark for understanding children's gaze behaviour," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023

Geometry-aware learning and control in robotics

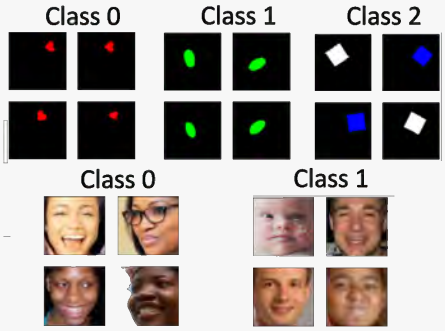
Data in robotics are essentially geometric. Our work leverages geometric representations for learning and optimization problems in robotics, contributing to new advances in the use of geometric algebra and Riemannian manifolds, demonstrated with diverse use cases from prosthetic devices to industrial robots. We showed that algorithms initially developed for standard Euclidean data could be extended to other manifolds while reducing the required amount of training data, effectively contributing to recent challenges in robotics requiring both uncertainty and geometry modeling.



[1] T. Löw and S. Calinon, "Geometric algebra for optimal control with applications in manipulation tasks," *IEEE Transactions on Robotics*, May 2023
[2] B. Ti, A. Razmjoo, Y. Gao, *et al.*, "A geometric optimal control approach for imitation and generalization of manipulation skills," *Robotics and Autonomous Systems*, 2023

Robustness of machine learning models

A key activity in the *machine learning* group at Idiap is to study robustness and reliability of models when they are applied on cases outside their training data. This is also known as their out-of-distribution (OOD) performance. In a paper presented at the NeurIPS conference [1], we established that OOD performance could not reliably predicted from the traditional in-distribution performance, contrary to common practice. This fundamental finding has important practical implications for assessing the reliability of models before their deployment on real applications.

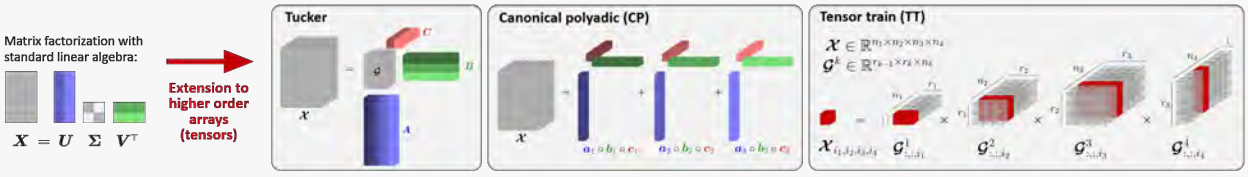


Examples of data used to diagnose biases in learned models. For example, a model trained to recognize geometric shapes (above) tends by default to rely on color instead of the shape's contours, since color in easier to identify. Similarly, a model trained to process face images may focus on age, gender, or ethinicity even when these attributes are not those we intend to capture. We developed methods to help detect and address such problematic situations [2].

[1] D. Teney, Y. Lin, S. J. Oh, *et al.*, "ID and OOD performance are sometimes inversely correlated on real-world datasets," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023
[2] L. Scimeca, A. Rubinstein, A. M. Nicolicioiu, *et al.*, "Leveraging diffusion disentangled representations to mitigate shortcuts in underspecified visual tasks," arXiv preprint, Under review, 2023

Tensor factorization for robotics

Our work considers the use of tensor train, more broadly known as tensor network, as a model to separate variables and functions, which are organized in arrays of more than 2 dimensions. This approach extends standard low-rank matrix factorization such as singular value decomposition to arrays of higher dimensions, which has the important merit of working with a low quantity of training data. In contrast to standard factorization techniques based on linear algebra (using data organized as matrices and vectors), we showed that tensor trains allows the robot to process data in a multilinear fashion. We leverage this principle in impactful problems in robotics, including timely challenges in robot manipulation learning and optimization.



S. Shetty, T. S. Lembono, T. Löw, *et al.*, "Tensor train for global optimization problems in robotics," *International Journal of Robotics Research*, p. 25, 2023

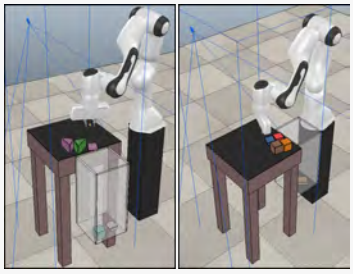
Multimodal behavior analysis and robotics

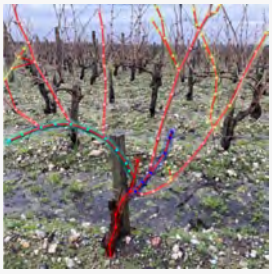
Gaze analytics. We have investigated methods for building user-specific 3D gaze direction prediction models from user samples, as well as designing a fully unsupervised gaze representation learning approach allowing to leverage internet-scale data, by jointly learning a gaze representation network and gaze redirection network only from pairs of eye images.

Audio analysis. We also recently studied different DNN architectures and unsupervised methods for sound processing: efficient multi-task approach for the joint localization and categorization (speech vs non-speech) of multiple sources from a microphone array (Fig a below).

Multitask and Kernel approach for Learning to Push Objects. Pushing, a vital motor skill for manipulation tasks, is a key focus in robotics research. Recent studies demonstrate that Deep Q-Networks (DQNs) can learn pushing policies, but often assume fixed settings and suffer from sparse-gradient backpropagation, making training challenging. To overcome these limitations, we propose a multi-head target-parameterized DQN for robotic manipulation, showing the benefits of predicting different reward and task aspects, exploring alternatives for policy generalization to new positions (Fig b below), and introducing a kernelized loss function for improved training performance [1].

From body landmarks to Vine Pruning. We extended our know-how on landmarks prediction to the challenging case of predicting the structure of vine trees, demonstrating high performance in real situations (Fig. c) and exploring performance-computation trade-offs. The output is exploited by our 3D2Cut partner within augmented reality glasses to assist operators in grapevine pruning [2,3].





[1] M. Ewerton, M. Villamizar, J. Jankowski, *et al.*, “A multitask and kernel approach for learning to push objects with a task-parameterized deep q-network,” in *IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2023

[2] T. Gentilhomme, M. Villamizar, J. Corre, *et al.*, “Efficient grapevine structure estimation in vineyards conditions,” in *International Conference on Computer Vision (ICCV)*, 2023, pp. 712–720

[3] T. Gentilhomme, M. Villamizar, J. Corre, *et al.*, “Towards smart pruning: Vinet, a deep-learning approach for grapevine structure estimation,” *Computers and Electronics in Agriculture*, vol. 207, p. 107 736, Apr. 2023

Modelling human creative thinking

Analogical reasoning has value regarding problem solving, human creative endeavors such as metaphor generation, scientific discovery, as well as in the field of education. In this paper, we test several ways to learn basic analogical reasoning, specifically focusing on analogies that are more typical of what is used to evaluate analogical reasoning in humans than those in commonly used NLP benchmarks. Our experiments find that models are able to learn analogical reasoning, even with a small amount of data. We additionally compare our models to a dataset with a human baseline, and find that after training, models approach human performance.

M. R. Petersen and L. van der Plas, “Can language models learn analogical reasoning? investigating training objectives and comparisons to human performance,” in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 2023

Human-robot teaming for manufacturing

Our ongoing collaboration with the University of Wisconsin–Madison has led to the development of new robotics systems for human-robot collaboration. In our first application, we evaluated the advantages of a robot-based shared control approach compared to a manual approach for sanding in aerospace manufacturing [1]. We also explored the scaling upof this approach to controlling multiple robots [2]. In our second application, we explored how a remote robot could provide adaptive view points in a dynamic collaboration interaction between two humans [3].

[1] A. Konstant, N. Orr, M. Hagenow, *et al.*, "Human-robot collaboration in a sanding task," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2023
[2] M. Hagenow, E. Senft, N. Orr, *et al.*, "Coordinated multi-robot shared autonomy based on scheduling and demonstrations," *IEEE Robotics and Automation Letters*, vol. 8, no. 12, pp. 8335 –8342, 2023
[3] P. Praveena, Y. Wang, E. Senft, *et al.*, "Periscope: A robotic camera system to support remote physical collaboration," *Proceedings of the ACM on Human Computer Interaction*, 2023

Putting sign language generation and sign language assessment in the loop

One of the challenges in sign language generation is the evaluation of the generated video content. One possible solution is to subjectively evaluate using human raters. This is time-consuming and costly. The other possible solution is objective evaluation. In the literature, video quality metrics such as PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index) and skeleton-based measures such as mean squared error (MSE) between sequences of 3D skeleton have been proposed. A limitation of these approaches is that they do not provide information about the generated video content. In the framework of SNSF Sinergia project SMILE-II, we have developed a novel phonology-based approach that evaluates the generated video along different channels, namely, hand movement and handshape, which convey the linguistic information in sign language. Our experimental studies demonstrate that the proposed objective scoring method yields a better correlation to subjective human ratings than PSNR, SSIM, and MSE-based metrics. This work was done in collaboration with International Institute of Information Technology Bangalore, India.

N. Tarigopula, P. Garg, S. Muralidhar, *et al.*, "Content-based objective evaluation of artificially generated sign language videos," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024

Multimodal dialogue modeling for applications such as contact centers and safety-critical domains

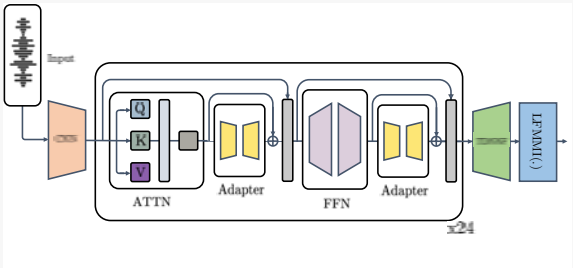
Idiap has been active in analysing unstructured dialogues for already several years. Both research and development, as well as technology transfer, have been pursued through many collaborations (i.e., European projects and direct contracts with industrial partners) in the domain of voice assistant-powered dialogue engines for commercial and governmental technological pipelines. The high-level goals are to translate dialogues into explainable, safe, knowledge-grounded, trustworthy and bias-controlled output. The innovative and mature technology is assumed to learn by its own, being adaptable from limited corpora across languages, use-cases, or business logic, as well as being sustainable w.r.t. existing computational architectures , including the support of safety-critical applications involving humans-in-the-loop.

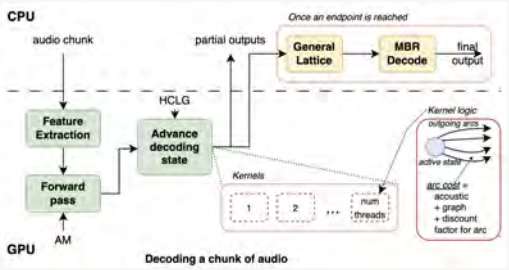
E. VILLATORO-TELLO, S. Madikeri, Z.-G. Juan, *et al.*, "Effectiveness of text, acoustic, and lattice-based representations in spoken language understanding tasks," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2023

E. VILLATORO-TELLO, S. Madikeri, B. Sharma, *et al.*, "Probability-aware word-confusion-network-to-text alignment approach for intent classification," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024

On-demand automatic speech recognition (ASR) for industrial applications

Although accurate ASR solutions are available from many technological providers worldwide, they do not always perform well for more specific use-cases. Idiap supports small and medium level enterprises in developing and integrating tailored (on-demand) solutions reaching low error-rates in specific tasks, deploying less resource demanding models, and enabling real-time response. Innovative concepts are pursued, including (i) contextual-biasing to dynamically integrate prior knowledge available from target domain to the ASR, (ii) integration of contextual biasing into GPU decoding to support multi-thread processing; (iii) combined ASR training with down-stream SLU tasks to obtain richer outputs, or (iv) efficient transfer learning of large pre-trained multilingual models.





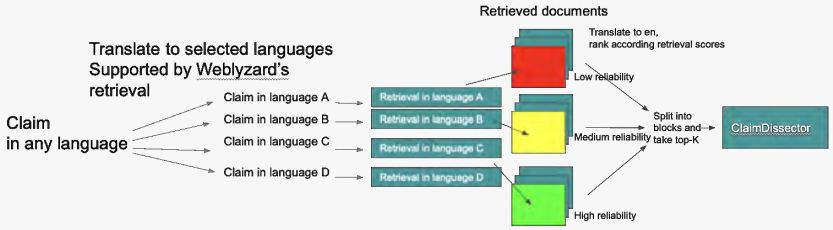
N. Iuliia, S. Madikeri, E. Villatoro-Tello, *et al.*, "Implementing contextual biasing in gpu decoder for online asr," in *Interspeech*, 2023
G. Vanderreydt, A. Prasad, D. Khalil, *et al.*, "Parameter-efficient tuning with adaptive bottlenecks for automatic speech recognition," in *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, 2023

Explainable fact checking

Idiap works on researching and developing multilingual text processing algorithms to perform explainable fact checking against news media services to help law enforcement and security analysts deal with the spread of misinformation within untrusted news sources.

A novel latent variable model for fact-checking and analysis has been developed, which, given a claim and a set of retrieved evidence, jointly learns to identify evidence relevant to the given claim, and the veracity of the claim.

As part of this, it is also important to distinguish between reliable and unreliable news sources, specially when the domains are not familiar to the analysts. This is solved by building a real-world graph showing how all news sources (used for the analysis) are connected to each other on the web. News sources are represented by nodes, and edges are weighted by citation/reference frequency. The reliability of each source in the graph in estimated based on its connections and interactions with other nodes.



M. Fajcik, P. Motlicek, and P. Smrz, "Claim-dissector: An interpretable fact-checking system with joint re-ranking and veracity prediction," *Association for Computational Linguistics*, vol. Findings of the Association for Computational Linguistics (ACL), 10184–10205, Jul. 2023

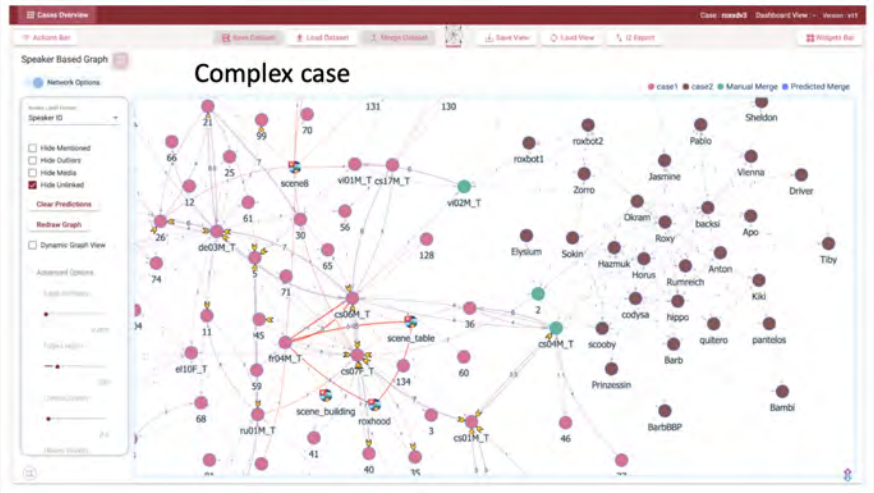
Voice forensics

Idiap is actively carrying research and development on continualy improving capabilities of voice technologies in person identification applicable for forensics cases. One of applications is a forensic voice comparison between two or many voice samples (e.g., obtained from intercepted telephone conversations, lawfully collected from mobile terminals, or from social media). The forensic voice analysis consists of comparing recordings that contain speech from a speaker with confirmed identity to those recordings where the identity of the speaker is questioned. Output of the analysis will be an objective report that provides a neutral expert opinion. The solution is built on top of power of deep learning combined with classical speech pre-processing and post-processing technologies. This semi-automatic approach allows to analyse relatively large number of samples while the human expert is involved to verify each of the processing stages. The system can be tuned through model adaptation and score calibration to reach target false-positive rate (e.g., below 0.01%).

S. Madikeri, S. Dey, and P. Motlicek, "A bayesian approach to inter-task fusion for speaker recognition," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2019, pp. 5786–5790

Multimodal data & network analytics to support law enforcement combatting organised crime

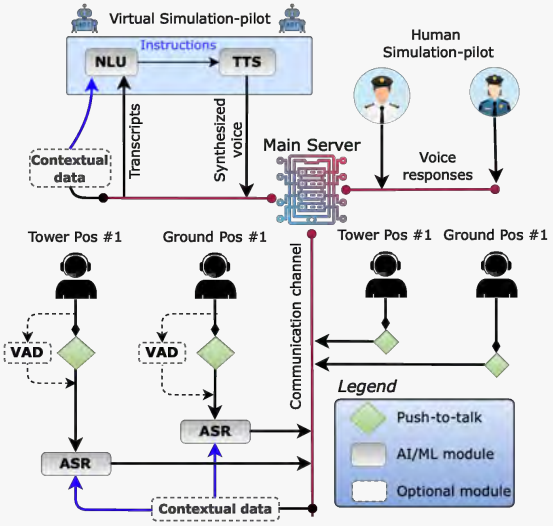
Criminal investigations require manual intervention of several investigators and police practitioners. However, the amount and the diversity of the data collected raises many challenges, and cross-border investigations against organized crime can quickly im-possible to handle. Idiap has been leading the activities on researching and developing an all-in-one investigative platform which processes intercepted phone calls and other data collected through the process of investigation, is translated into a knowledge graph of the extracted information. This open research platform regularly maintained, combining speech, text, and video processing algorithms with criminal network analysis for combating organized crime, is available for security practitioners and researchers.



Y. Solewicz, N. Cohen, J. Rohdin, et al., “Speaker recognition on mono-channel telephony recordings,” in The Speaker and Language Recognition Workshop, 2022
M. Fabien, S. Parida, D. Zhu, et al., “Roxanne research platform: Automate criminal investigations,” in Interspeech Show and Tell, 2021

Analysis of air traffic communication

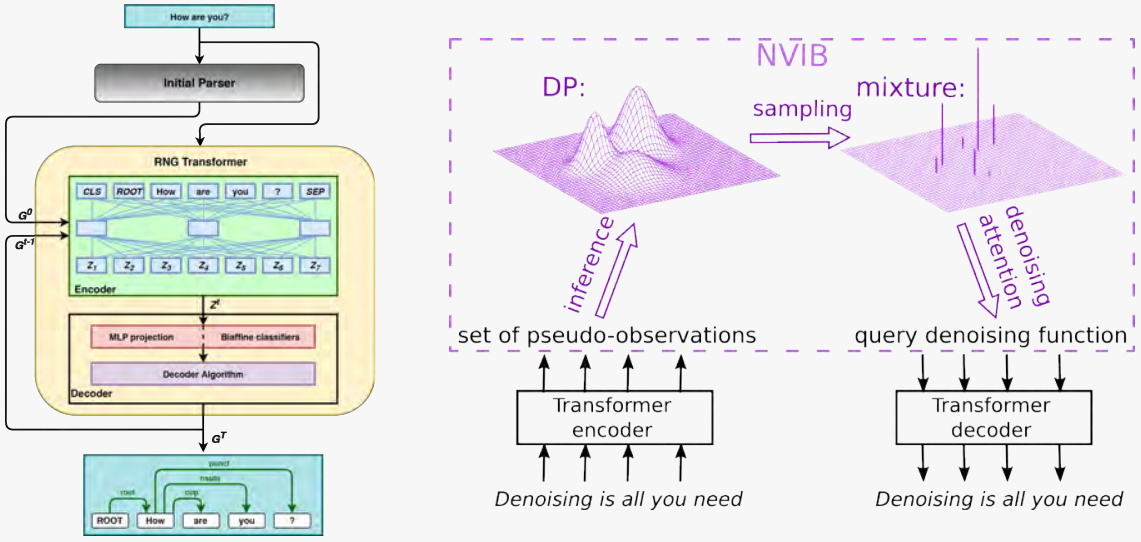
Idiap is among key organisations continuously contributing with long-term research in the field of air-traffic management. Both exploratory and industrial research is performed with overall goals to deploy artificial intelligence as one of the main enablers to overcome the current limitations in the current air-traffic management systems. Idiap specifically focuses on analysing communication and supporting air-traffic controllers with modern technologies to decrease their workload while increasing the safety.



Z.-G. Juan, A. Prasad, N. Iuliia, *et al.*, “A virtual simulation-pilot agent for training of air traffic controllers,” *Aerospace*, vol. 10, no. 5, 2023
H. Helmke, M. Kleinert, N. Ahrenhold, *et al.*, “Automatic speech recognition and understanding for radar label maintenance support increases safety and reduces air traffic controllers’ workload,” in *Fifteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2023)*, Eurocontrol (Europe), FAA (U.S.), 2023

Interpretable representation learning

Representation learning is the foundation of the latest AI advances, but these deep learning representations seem to have little to do with the way humans traditionally represent and communicate information, often in the form of graphs. In the *natural language understanding* group, we address this issue by developing novel deep learning architectures which integrate structured representations. The Graph2Graph Transformer architecture [1,2] uses the self-attention mechanism of transformers to input and output human-defined graph relations, and integrate them with automatically-induced graph relations. Both these graphs have human-defined nodes, so we are also investigating automatically inducing the nodes of the graph, called entity induction [3,4,5]. In particular, our nonparametric variational Bayesian approach to this problem provides both fundamental insights into transformers' latent representations (as nonparametric mixture distributions), and a practical, sparsity-inducing regulariser (NVIB) for transformers [3]. Recent results show that these discovered entities can be interpreted similarly to human-defined ones, such as words [4].



[1] J. Henderson, A. Mohammadshahi, A. C. Coman, *et al.*, "Transformers as graph-to-graph models," in *Big Picture Workshop at EMNLP*, 2023

[2] A. Mohammadshahi and J. Henderson, "Syntax-aware graph-to-graph transformer for semantic role labelling," in *Procceedings of 8th Workshop on Representation Learning for NLP*, 2023

[3] J. Henderson and F. Fehr, "A vae for transformers with nonparametric variational information bottleneck," in *The Eleventh International Conference on Learning Representations*, 2023

[4] M. Behjati, F. Fehr, and J. Henderson, "Learning to abstract with nonparametric variational information bottleneck," in *Conference on Empirical Methods in Natural Language Processing*, 2023

[5] M. Behjati and J. Henderson, "Inducing meaningful units from character sequences with dynamic capacity slot attention," *Transactions on Machine Learning Research*, 2023

Joining efforts from humanities and computer science for studying historical manuscripts

The Voynich Manuscript is an enigmatic medieval mystery. It is widely believed to be a 15th century tome which is written in a unknown script, in an unknown language whose contents, despite the efforts of many experts over the last century, has yet to be deciphered. Within its pages lie fantastical figures and illustrations of plants (many unidentified as of yet) and people, in addition to drawings of an astronomical nature, that baffle interpretation. Due to the many unknowns related to the manuscript, its study requires collaboration across disciplines. In collaboration with researchers from the University of Malta, Yale University and the Medieval Academy of America, researchers from Idiap organised a con-

ference that gathered scholars from different disciplines to help put some of the pieces of this puzzle together.

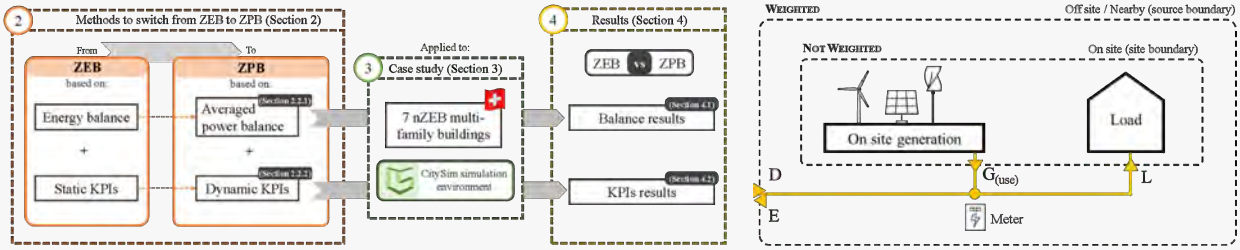
C. Layfield, R. Zandbergen, L. F. Davis, *et al.*, “International conference on the Voynich manuscript 2022,” in *Proceedings of the International Conference on Historical Cryptology*, 2023

2.2 Sustainable and Resilient Societies

This research program aims to anticipate harmful future changes such as wars, pandemic and climate events with cascading impacts on socio-economic systems: disrupting the global supply of commodities, sharply increasing food and energy prices, geopolitical tensions, and decreased security. Through inter-disciplinary approaches to AI, research in this program will model technological, natural, geopolitical and criminal threats based on evidence from heterogeneous sources. It will provide solutions based on the optimization of time-varying indicators, while striving for the sustainability and efficiency of AI technologies themselves.

On the road to Zero Power Buildings (ZPBs)

A collaboration with the Politecnico di Tornio has led to a new method for building performance assessment. The limits of the Zero Energy District are highlighted in a case-study with new buildings equipped with solar collectors and connected to a low-temperature district heating network [1]. A framework for the application and evaluation of Zero Power Buildings is introduced with a series of dedicated KPIs at fine temporal granularity [2]. The dynamic evaluation of the averaged power balance was realised using our Urban Energy Simulator CitySim.

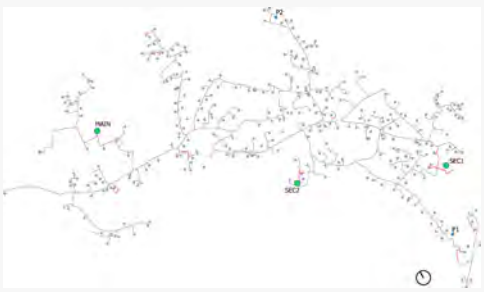


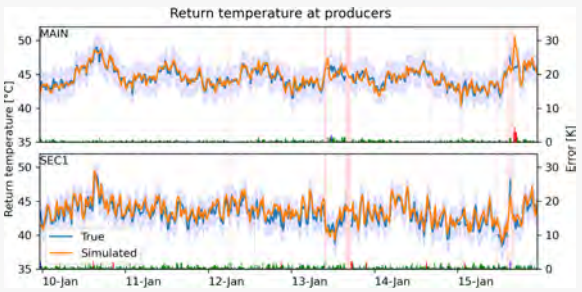
[1] M. Bilardo, J. Kämpf, and E. Fabrizio, "Energy assessment of a district by integrating solar thermal in district heating network: A dynamic analysis approach," in *Journal of Physics: Conference Series*, vol. 2600, IOP Publishing Ltd, 2023

[2] M. Bilardo, J. H. Kämpf, and E. Fabrizio, "From zero energy to zero power buildings: A new paradigm for a sustainable transition of the building stock," *Sustainable Cities and Society*, vol. 101, p. 105 136, 2024

Design, simulation, and optimisation of District Heating Networks (DHNs)

A Python library has been developed at Idiap for the simulation of district heating networks. The library named PyDHN was firstly verified against other simulation tools through the DESTEST [1]. Secondly, thanks to our industrial partner's monitoring data, we proposed an improved benchmark for the steady-state simulation of meshed district heating networks [2]. Finally, we improved the library to dynamically simulate the thermal behaviour in the pipes and verified it against 15 minute monitoring on a complex meshed network [3]. On the demand-side, we have prepared machine learning predictor models based on the CityGML standard [4].

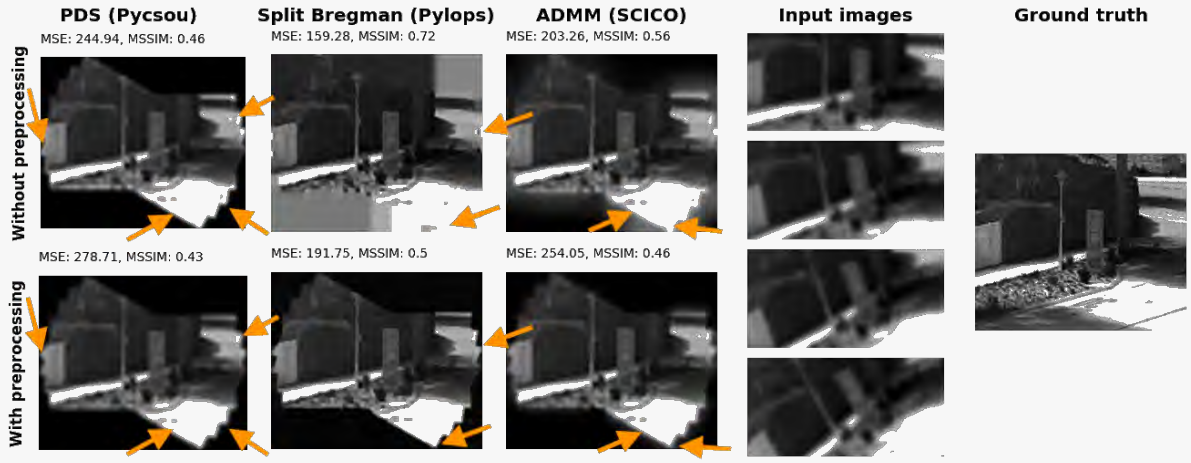




[1] R. Boghetti, G. Peronato, and J. Kämpf, “Verification of pydhn - a python library for the thermo-hydraulic simulation of district heating networks - through the destest,” in *Proceedings of Building Simulation 2023: 18th Conference of IBPSA*, ser. Building Simulation, IBPSA, vol. 18, IBPSA, 2023
[2] R. Boghetti and J. Kämpf, “A benchmark for the simulation of meshed district heating networks based on anonymised monitoring data,” in *Journal of Physics: Conference Series*, vol. 2600, IOP Publishing Ltd, 2023
[3] R. Boghetti and J. H. Kämpf, “Verification of an open-source python library for the simulation of district heating networks with complex topologies,” *Energy*, vol. 290, p. 130 169, 2024
[4] M. Tognoli, G. Peronato, and J. Kämpf, “A machine learning model for the prediction of building hourly heating demand from citygml files: Training workflow and deployment as an api,” in *Proceedings of Building Simulation 2023: 18th Conference of IBPSA*, 2023, pp. 2932 –2939

Multi-image deconvolution of thermal images

Handheld thermal imaging devices have a wide range of applications, from detecting heat signatures of objects or living beings in low visibility environments to identifying thermal leaks in buildings for energy efficiency. These devices can capture images in quick succession, each image with a slightly different orientiation, resulting in image series that can be combined to produce an improved image through multi-image deconvolution. Implementing deconvolution algorithms that take advantage of the entire information contained in the image series to produce an image with a field of view that is as large as possible given the coverage of all collected images is a challenging problem given that the image series covers a possibly non-square area. We have developed a multi-image deconvolution method that addresses this boundary condition problem. We solve a regularized minimization problem with two terms, (i) the residue between the forward transformation applied to the reconstructed candidate and the measured images and (ii) a regularization term to take into account image priors. In order to accommodate for a non-square coverage of the combined images, which results in boundary artifacts when the forward model is used during iterative minimization, we recast the problem into one where the original scene is masked, thereby mitigating the effects of unknown image values beyond image boundaries.

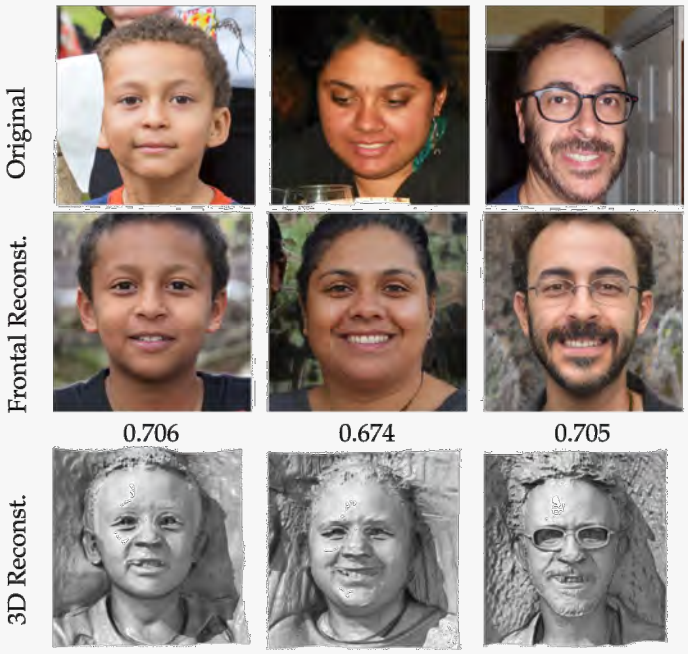


F. Piras, F. Marelli, E. De Moura Presa, *et al.*, “Multi-image deconvolution of thermal images with a boundary condition weighting scheme,” in *Target and Background Signatures IX*, International Society for Optics and Photonics, vol. 12736, SPIE, 2023, pp. 149–158

Vulnerability of Face Recognition to Template Inversion Attacks

We investigated the vulnerability of state-of-the-art face recognition systems to template inversion at-

tacks using 2D and 3D face reconstruction. We proposed several methods to reconstruct 2D and 3D faces from facial templates using a pretrained face generation network, and train a mapping from facial templates to the intermediate latent space of the face generator network. We train our mapping with a semi-supervised approach using real and synthetic face images. For real face images, we use a generative adversarial network (GAN)-based framework to learn the distribution of generator intermediate latent space. For synthetic face images, we directly learn the mapping from facial templates to the generator intermediate latent code. We proposed our methods in the whitebox and blackbox attacks against face recognition systems across various datasets. We also performed practical presentation attacks on face recognition systems using the digital screen replay and printed photographs, and evaluate the vulnerability of face recognition systems to different template inversion attacks.



[1] H. Otroshi Shahreza and S. Marcel, “Comprehensive vulnerability evaluation of face recognition systems to template inversion attacks via 3d face reconstruction,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023

[2] H. Otroshi Shahreza and S. Marcel, “Template inversion attack against face recognition systems using 3d face reconstruction,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023, pp. 19 662–19 672

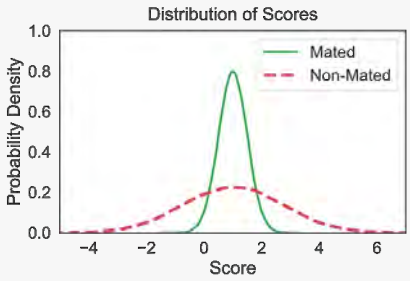
[3] H. Otroshi Shahreza and S. Marcel, “Blackbox face reconstruction from deep facial embeddings using a different face recognition model,” in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2023

[4] H. Otroshi Shahreza and S. Marcel, “Face reconstruction from facial templates by learning latent space of a generator network,” in *Thirty-seventh Conference on Neural Information Processing Systems*, 2023

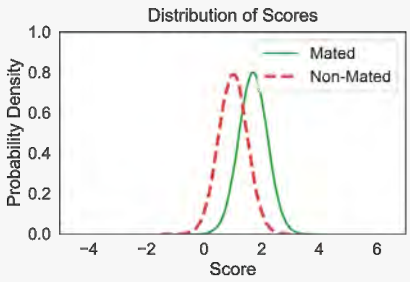
Measuring Linkability of Protected Biometric Templates Using Maximal Leakage

There is a growing need to secure the sensitive data used within biometric systems. Considering privacy challenges in such systems, different biometric template protection (BTP) schemes were proposed in the literature. We proposed a new method for measuring linkability of protected biometric templates based on maximal leakage, which is a well-studied measure in information-theoretic literature. We showed that the resulting linkability measure has a number of important theoretical properties and an operational interpretation in terms of statistical hypothesis testing. We used the proposed measure to

evaluate the linkability of biometric templates from different biometric characteristics (face, voice, and finger vein), which are protected with different BTP schemes.



mated: $\mathcal{N}(1, 0.5)$
non-mated: $\mathcal{N}(1, 1.75)$
$\mathrm{M}_{\leftrightarrow}^{sys}$ [ours]: 0.6177
$\mathrm{D}_{\leftrightarrow}^{sys}$ [17]: 0.3902

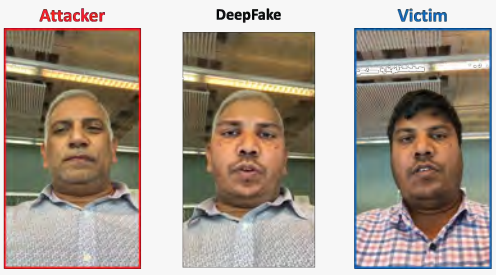


mated: $\mathcal{N}(1.7, 0.5)$
non-mated: $\mathcal{N}(1, 0.5)$
$\mathrm{M}_{\leftrightarrow}^{sys}$ [ours]: 0.5988
$\mathrm{D}_{\leftrightarrow}^{sys}$ [17]: 0.4289

H. Otroshi Shahreza, Y. Y. Shkel, and S. Marcel, "Measuring linkability of protected biometric templates using maximal leakage," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 2262 –2275, 2023

Vulnerability of Automatic Identity Recognition to Audio-Visual Deepfakes

We investigated audio-visual deepfakes, where lips and speech are well synchronized and video have high visual and audio qualities. We created audio-visual deepfakes using several models and blending techniques for face swapping and speech models for voice conversion. We demonstrated the vulnerability of a state of the art speaker recognition system to the synthetic voices. Similarly, we tested a face recognition system to several variants of our visual deepfakes. The vulnerability assessment show that by tuning the existing pretrained deepfake models to specific identities, one can successfully spoof the face and speaker recognition systems in more than 90% of the time and achieve a very realistic looking and sounding fake video of a given person.



P. Korshunov, H. Chen, P. N. Garner, *et al.*, "Vulnerability of automatic identity recognition to audio-visual deepfakes," in *IEEE International Joint Conference on Biometrics*, 2023

Characterising conspiracist texts

Conspiracy theories (CTs) are accounts of significant social and political events alleging secret plots by powerful actors. They compete for attention with mainstream news, but being freed from constraints of veridicality they can exploit cognitive biases for appealing information. Thus, false information is

more likely to be transmitted online. This transmission advantage might stem from ostensibly creative content features such as novelty, divergence from the norm, and sophistication that make narratives more attention-grabbing. We test whether CTs might be linguistically creative products, which would facilitate their transmission and thereby account for their widespread popularity. We analyzed nominal compounds (e.g., mind control, carbon dioxide; N = 1,713,568) from a large corpus of conspiracist and mainstream texts matched by topic. In conspiracist texts, compounds showed greater originality, divergence, and sophistication, but they were used with lower frequency and were more often repeated in different contexts. This pattern suggests a creative aspect in the generation of compounds, coupled with a rigidity in their use. Our findings not only contribute to the discourse on creativity in CTs but also provide insights into the communicative advantage of CTs.

A. Miani, L. van der Plas, and A. Bangerter, “Loose and tight: Creative formation but rigid use of nominal compounds in conspiracist texts,” *The Journal of Creative Behavior*, 2023

2.3 AI for Life

This research program aims to improve the quality of life and treatment options for individuals with complex disorders. This aim will be achieved with advanced learning and inference paradigms in biological, biomedical and behavioral sciences to integrate diverse, longitudinal and interventional data with prior scientific and expert knowledge. By applying deductive informative biology, research in this program will enhance understanding of complex biological systems and human diseases such as cancer, neuro-degenerative conditions, mental health disorders, and rare diseases. Idiap will develop AI-based assistive technologies and personalized medicine models that use community knowledge and continuously acquired data. These tools aim to promote patient engagement and provide comprehensive care. By empowering individuals with their own data, research in this program will address patient isolation and knowledge gaps in disease understanding and alternative therapeutic approaches. These technologies also aim to address communication barriers caused by disabilities, sub-optimal communication channels, language differences, and cultural diversity. The outcomes of this research will empower individuals to maintain and increase autonomy in their daily activities.

Computer-aided diagnosis for inflammatory eye diseases

In collaboration with the Fondation Asile des Aveugles (Jules-Gonin Hospital) and the Eye-clinic of the Hospital of Lucern, our study introduces a method to automate the grading of retinal inflammation - a condition that lead to blindness if untreated. This development, which involved analyzing over 3,000 images from various patients using new models and computer vision, aims to enhance efficiency in clinical settings. While our approach shows promising accuracy and is a step forward in applying such systems to real-world clinical data, it is an initial effort towards improving diagnostic processes in this field of ophthalmology.

V. Amiot, O. Jimenez-del Toro, P. Eyraud, *et al.*, “Fully automatic grading of retinal vasculitis on fluorescein angiography time-lapse from real-world data in clinical settings,” in *IEEE International Symposium on Computer-Based Medical Systems*, 2023, pp. 689–693

Automatic reading of chest X-rays for enhanced tuberculosis diagnosis and elimination

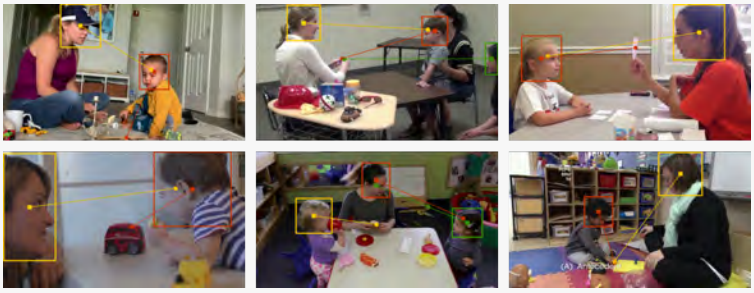
In collaboration with an international panel of Tuberculosis (TB) experts, we have explored the advanced capabilities of computer-aided detection (CAD) in diagnosing TB through chest radiographs (CXR). Our findings reveal CAD’s high sensitivity and accuracy, comparable to human analysis, in both symptomatic individuals and broader screenings. However, we also highlight significant implementation challenges, such as the need for tailored threshold scores and the complexity of integrating frequently updated CAD systems. Additionally, we emphasize the urgent need for regulatory measures to ensure equitable access to this technology, particularly in high-burden communities.

C. Geric, Z. Z. Qin, C. M. Denkinger, *et al.*, “The rise of artificial intelligence reading of chest x-rays for enhanced tb diagnosis and elimination,” *The International Journal of Tuberculosis and Lung Disease*, vol. 27, no. 5, pp. 367–372, May 2023

Childplay: a public dataset for children attention and behavior modeling

Analyzing clinical behavioral performance in autistic children necessitates training data. Unfortunately, privacy concerns limit access to such clinical data, hindering model training and comparison with the state of the art. Additionally, datasets containing typical children’s data are unavailable. To address this gap, we created the ChildPlay dataset, a publicly available resource featuring videos of children in

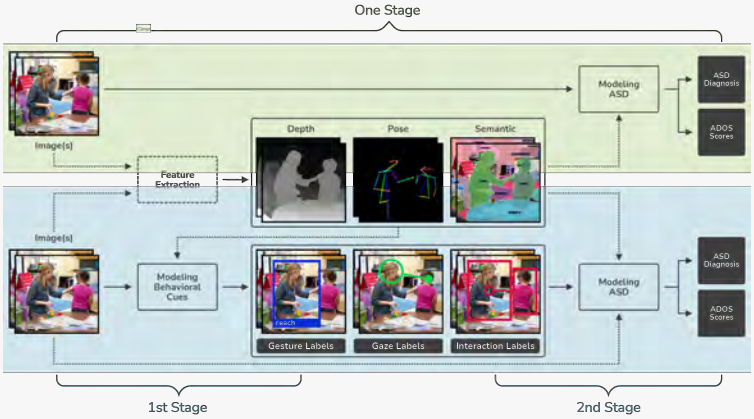
free-play environments. This dataset includes rich, unprompted social behaviors, communicative gestures, and interactions, with high-quality dense gaze annotations. Notably, it is the first gaze dataset designed to represent children, overcoming the scarcity of available resources for this specific purpose.The dataset is available on Zenodo



S. Tafasca, A. Gupta, and J.-M. Odobez, "Childplay: A new benchmark for understanding children's gaze behaviour," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023

The AI4Autism project: a multimodal and interdisciplinary approach to autism diagnosis and stratification

In the current landscape, 1 in 36 children receives an autism spectrum disorder (ASD) diagnosis, highlighting the urgent need for scalable screening tools. The AI4Autism project, funded by the Swiss National Science Foundation over four years, addresses this need by leveraging digital sensing. Using multimodal techniques such as computer vision and Internet of Things sensing, the project aims to provide automated measures of the extended autism phenotype. This includes stratifying autism subtypes for precision medicine applications. The paper provides an overview of key project outcomes, introducing datasets and annotations that will be made publicly available. The methods and algorithms for analyzing children's behaviors and facilitating ASD diagnosis are also presented.

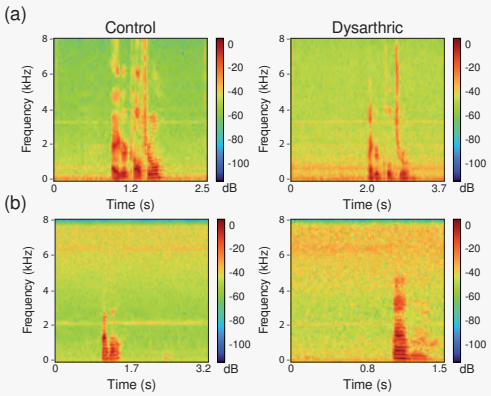


S. Tafasca, A. Gupta, N. Kojovic, *et al.*, "The ai4autism project: A multimodal and interdisciplinary approach to autism diagnosis and stratification," in *Companion Publication of the 25th International Conference on Multimodal Interaction*, ser. ICMI '23 Companion, Association for Computing Machinery, 2023, 414–425

Robust pathological speech detection in adverse environments

Despite the reported success of automatic pathological classification approaches, state-of-the-art literature typically relies on the underlying assumption that recordings from control and pathological speakers are obtained in the same noiseless environment using the same recording setup. If recordings

for one group of speakers are obtained in a consistently different environment than recordings for the other group of speakers, classifiers trained on such recordings would potentially learn characteristics of the recording environment instead of dysarthric speech characteristics. Unfortunately, such an assumption does not seem to be fulfilled for the commonly used UASpeech and TORGO databases. Although these databases are made available to the community to develop automatic speech recognition systems (ASR) (where different recording environments and setups can even be desirable in order to develop robust ASR systems), they have also been used to validate a considerable number of state-of-the-art automatic pathological speech detection approaches. To show that these databases are not suitable for evaluating pathological speech detection approaches, we have first estimated the utterance-level signal-to-noise ratio, confirming the large variability in recording conditions. Further, using voice activity detection, segments that contain only speech and segments that do not contain any speech are extracted from each utterance in these databases. State-of-the-art pathological speech classification approaches are then trained and validated using only the speech segments or using only the nonspeech segments. Remarkably, experimental results show that for both databases, the majority of the considered state-of-the-art approaches achieve the same or even a considerably better dysarthria classification performance when using only non-speech segments than when using only speech segments. These results confirm the hypothesis that dysarthria classification approaches validated on these databases may be learning characteristics of the recording environment rather than pathological characteristics.



G. Schu, P. Janbakhshi, and I. Kodrasi, "On using the UA-Speech and TORGO databases to validate automatic dysarthric speech classification approaches," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2023

Language Models to support treatment recommendations in Cancer Clinical Trials

There is a gap between cancer patients with complex clinical conditions and the set of available experimental treatments (clinical trials), as practicing oncologists have limited awareness on the full set of available clinical trials and most recent evidence. Bridge this gap can have significant impact on patients lives. In this project, we aim to extend the interpretation capabilities of Large Language Models (LLMs) to deliver complex, expert-level clinical trials recommendations, by interpreting a complex patient description and matching to the most relevant clinical trial. This involves extending the ability of LLMs to reliably reason over complex biomedical concepts and evidence, including the interpretation of complex eligibility criteria, patient state and previous treatment conditions. In this context, we developed new benchmarks and methods specialised for clinical trial matching, in close collaboration with oncology experts.

M. Jullien, M. Valentino, H. Frost, *et al.*, "Nli4ct: Multi-evidence natural language inference for clinical trial reports," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)*, 2023

Detecting severe adverse events events in cancer treatments

Experimental treatments in oncology can have a significant impact on disease management and patient survival. These treatments, however, come with an associated risk of severe adverse events (SAEs), which in extreme cases can lead to patient death. Recent therapies such as chimeric antigen receptor T-cell (CAR-T) involves reprogramming the patient's immune system, which is used as a mechanism to target the tumour, bring the associated risk of severe immune response events such as Cytokine Release Syndrome (CRS). Modelling the personalised response of a patient to such a treatment is a major challenge and requires the use of new analytical methods. In this project, we developed a Metareview-informed Machine Learning paradigm, which integrates previous domain knowledge from the specialised literature to facilitate the prediction of CRS events.

A. Bogatu, M. Wysocka, O. Wysocki, *et al.*, "Meta-analysis informed machine learning: Supporting cytokine storm detection during car-t cell therapy," *Journal of Biomedical Informatics*, vol. 142, Jun. 2023

Situated participatory design for interactions between older adults and robots

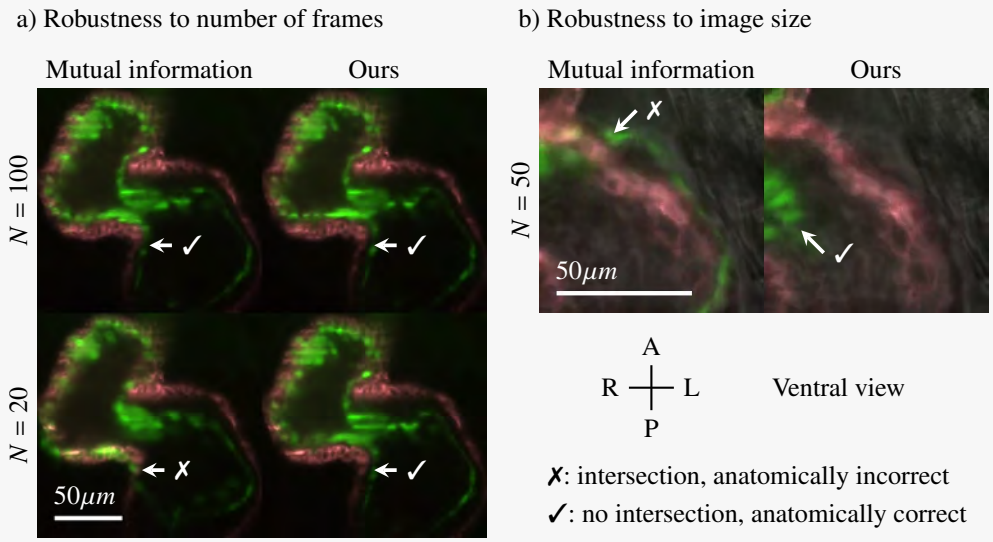
As populations age, the demand for robotic assistance in care homes is increasing. However, the intricacies of capturing the specific needs and preferences of older adults remain unclear. This study developed a participatory design approach that involved co-designing robotic behaviors with older adults through design sessions and simulated deployments. The results indicated that the method effectively captured evolving preferences, but further research is necessary to comprehensively understand the dynamic nature of older adults' desires.

L. Stegner, E. Senft, and B. Mutlu, "Situated participatory design: A method for in situ design of robotic interaction with older adults," in *CHI '23: Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2023

Light modulation microscopy for fast and compressed imaging

In vivo fluorescence microscopy is a powerful tool to image the beating heart in its early development stages. A high acquisition frame rate is necessary to study its fast contractions, but the limited fluorescence intensity requires sensitive cameras that are often too slow. Moreover, the problem is even more complex when imaging distinct tissues in the same sample using different fluorophores. We have developed the Paired Alternating AcQuisitions (PAAQ) a method to image cyclic processes in multiple channels, which requires only a single (possibly slow) camera. We generate variable temporal illumination patterns in each frame, alternating between channel-specific illuminations (fluorescence) in odd frames and a motion-encoding brightfield pattern as a common reference in even frames. After computationally estimating the reference frame positions, we assemble multichannel videos whose frame rate is virtually increased. On experimental cardiac images collected in zebrafish larvae, we obtained both quantitative and visual improvements in the reconstructed videos over existing non-gated sorting-based alternatives; Using a 15 Hz camera we produce video containing 2 fluorescence channels at an equivalent of 100 fps [1,2].

We further developped a reconstruction approach for OptoMechanical Modulation Tomography (OMMT), which exploits compressed sensing to reconstruct high resolution microscopy volumes from fewer measurement images compared to exhaustive section sampling in conventional light sheet microscopy. It uses a 1+2D Total Variation regularization that does not generate such artefacts and is amenable to efficient implementation using parallel computing [2,3].



[1] F. Marelli, A. Ernst, N. Mercader, *et al.*, "Paaq: Paired alternating acquisitions for virtual high frame rate multi-channel cardiac fluorescence microscopy," *Biological Imaging*, vol. 3, e20, Nov. 2023

[2] F. Marelli, "Practical computational imaging by use of spatiotemporal light modulation: From simulations to applications in biological microscopy," PhD thesis, EPFL, Oct. 2023

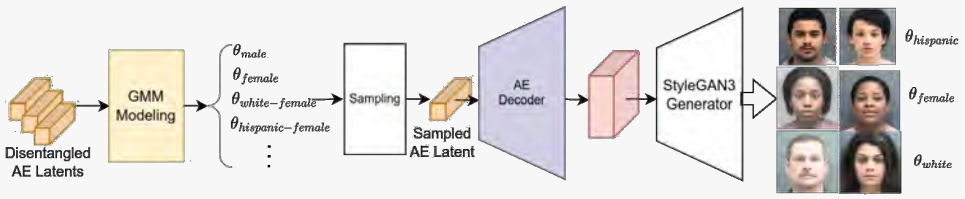
[3] F. Marelli and M. Liebling, "Efficient compressed sensing reconstruction for 3d fluorescence microscopy using optomechanical modulation tomography (ommt) with a 1+2d regularization," *Optics Express*, vol. 31, no. 20, pp. 31 718–31 733, Nov. 2023

2.4 AI for Everyone

AI has the potential to greatly benefit society by pooling of opinions, expertise and experiences, and adapting to individual needs. However, AI technology today largely remains one-size-fits-all: this is not in the best interest of everyone, and often ends up creating harm, even if unintended. This research program aims to ensure inclusiveness in this intelligent pooling for who is represented, who benefits, and who contributes. This research program implements a wide vision on who AI technology is not serving well today, and how this could be objectively done better in the future. This involves people who want to participate in the design and active use of AI for multiple purposes other than the current, often commercial ones, including the democratic process, people in world regions other than the Global North, and people everywhere in minorities, including special needs. This research program will lead to methodological contributions in designing AI with people, interactive methods and platforms for participatory AI design and deployment, novel data sampling and machine-learning methods that can leverage data both from minorities and majorities, opinion summarization and collective decision making technology, and the creation of sustainable communities of users (local but spread worldwide). One ambition is to systematize the study of multiple research territories (participatory approaches, designs for minorities and other world regions) that mainstream AI is not covering (and might not cover, given the main interests of the AI industry). This program is based on a participatory approach for the co-design and use of AI algorithms, within the different stages of the AI pipeline and that collaborate with specific groups of society towards satisfying real needs, moving from thinking about people as users towards people as agents.

Synthetic image generation for bias mitigation in face recognition

Recently, it has been exposed that some modern face recognition systems could discriminate specific demographic groups and may lead to unfair attention with respect to various facial attributes such as gender and origin. The main reason are the biases inside datasets, unbalanced demographics, used to train theses models. Unfortunately, collecting a large-scale balanced dataset with respect to various demographics is impracticable. We investigated as an alternative the generation of a balanced and possibly bias-free synthetic dataset that could be used to train, to regularize or to evaluate deep learning-based facial recognition models. We propose to use a simple method for modeling and sampling a disentangled projection of a StyleGAN latent space to generate any combination of demographic groups. Our experiments show that we can synthesis any combination of demographic groups effectively and the identities are different from the original training dataset.



P. Rahimi, C. Ecabert, and S. Marcel, "Toward responsible face datasets: Modeling the distribution of a disentangled latent space for sampling face images from demographic groups," in *IEEE International Joint Conference on Biometrics*, 2023

Speech translation for low-resource languages

This paper reports on the shared tasks organized by the 20th IWSLT Conference. The shared tasks address 9 scientific challenges in spoken language translation: simultaneous and offline translation, automatic subtitling and dubbing, speech-to-speech translation, multilingual, dialect and low-resource speech translation, and formality control. The shared tasks attracted a total of 38 submissions by 31

teams. The growing interest towards spoken language translation is also witnessed by the constantly increasing number of shared task organizers and contributors to the overview paper, almost evenly distributed across industry and academia.

M. Agarwal, S. Agarwal, A. Anastasopoulos, *et al.*, "Findings of the IWSLT 2023 evaluation campaign," in *International Conference on Spoken Language Translation (IWSLT)*, 2023

Children reading skill assessment

Reading is an acquired skill that is essential for integrating and participating in today's society. Yet, becoming literate can be particularly laborious for some children. Identifying reading difficulties early enough is the first, necessary step toward remediation. As part of the NCCR Evolang project, we investigated in collaboration with our collaborators from University of Geneva the opportunities and limitations of integrating commercial, off-the-shelf automatic speech recognition (ASR) services from IBM Watson to ease the administration and evaluation of children's reading assessment tests in French and Italian.

T. Piton, E. Hermann, A. Pasqualotto, *et al.*, "Using commercial asr solutions to assess reading skills in children: A case report," in *Interspeech*, 2023, pp. 4573–4577

Feature representations for distinguishing animal callers

Self-supervised learning (SSL) models tend to model the intrinsic structure of a given signal, independent of its acoustic domain, to extract essential information from the input to an embedding space. This implies that the utility of such representations is not limited to modeling human speech alone. Building on this understanding, in the framework of NCCR Evolang, we explored the cross-transferability of SSL neural representations learned from human speech to analyze bio-acoustic signals. We conducted a caller discrimination analysis and a caller detection study on Marmoset vocalizations using eleven SSL models pre-trained with various pretext tasks. The results showed that the embedding spaces carry meaningful caller information and can successfully distinguish the individual identities of Marmoset callers without fine-tuning. This demonstrates that representations pre-trained on human speech can be effectively applied to the bio-acoustics domain, providing valuable insights for future investigations in this field.

E. Sarkar and M. Magimai-Doss, "Can self-supervised neural representations pre-trained on human speech distinguish animal callers?" In *Interspeech*, 2023

3 List of Research Staff

3.1 Group Leaders

Dr. André Anjos
Biosignal Processing Group

Dr. Sylvain Calinon
Robot Learning &
Interaction Group

Prof. Andrea Cavallaro
Idiap Director

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

Prof. Michael Liebling
Computational Bioimaging Group

Dr. Raphaëlle Luisier
Genomics & Health
Informatics Group

Dr. Mathew Magimai-Doss
Speech & Audio
Processing Group

Dr. Prof. Sébastien Marcel
Biometrics Security &
Privacy 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

Prof. Lonneke van der Plas
Computation, Cognition &
Language Group

3.2 Research Associates

Name		Nationality	Supervisor	Start	End
BHATTACHARJEE	Sushil	CH	Sébastien MARCEL	2016-01-01	
GEISSBUHLER	David	CH	Sébastien MARCEL	2018-06-01	
GEORGE	Anjith	IN	Sébastien MARCEL	2021-01-01	
HOVSEPYAN	Sevada	AM	Mathew MAGIMAI DOSS	2023-10-17	2024-10-16
KOMATY	Alain	FR	Sébastien MARCEL	2022-06-01	
KORSHUNOV	Pavel	EE	Sébastien MARCEL	2020-07-01	
KOTWAL	Ketan	IN	Sébastien MARCEL	2022-08-01	
KRIVOKUĆA HAHN	Vedrana	HR	Sébastien MARCEL	2021-01-16	
MADIKERI	Srikanth	IN	Petr MOTLICEK	2018-03-01	
MURALIDHAR	Skanda	IN	Mathew MAGIMAI DOSS	2023-04-01	2024-03-31
VILLAMIZAR	Michael	ES	Jean-Marc ODOBEZ	2020-11-01	
VILLATORO TELLO	Esaú	MX	Andrea CAVALLARO	2021-11-01	

3.3 Post-doctoral Researchers

Name		Nationality	Supervisor	Start	End
BAIA	Alina-Elena	RO	Andrea CAVALLARO	2023-09-16	2024-09-30
BHATTACHARJEE	Mrinmoy	IN	Petr MOTLICEK	2023-05-15	2024-05-15
DELMAS	Maxime	FR	André FREITAS	2022-10-01	2024-03-31
ECABERT	Christophe	CH	Sébastien MARCEL	2022-05-01	2024-04-30
HERMANN	Enno	DE	Mathew MAGIMAI DOSS	2023-02-01	2024-01-31
JIMÉNEZ DEL TORO	Oscar	MX	André ANJOS	2022-04-15	2023-04-14
KALOGA	Yacouba	FR	Ina KODRASI	2023-04-01	2024-03-31
MARELLI	François	BE	Michael LIEBLING	2023-07-01	2023-12-31
MARFURT	Andreas	CH	James HENDERSON	2023-06-01	2023-12-31
MITRO	Ioanni	GR	Damien TENEY	2023-09-01	2025-08-31
PERONATO	Giuseppe	IT	Jérôme KÄMPF	2021-12-01	2023-04-30
RANALDI	Leonardo	IT	André FREITAS	2023-05-15	2024-05-14
RANGAPPA	Pradeep	IN	Petr MOTLICEK	2023-09-01	2024-08-31
RAZEGHI	Behrooz	IR	Sébastien MARCEL	2022-12-01	2024-01-31
SANCHEZ-CORTES	Dairazalia	MX	Petr MOTLICEK	2023-03-01	2024-08-31
SHEIKH	Shakeel Ahmad	IN	Ina KODRASI	2023-11-12	2024-11-12
SIEGFRIED	Rémy	CH	Jean-Marc ODOBEZ	2021-10-01	2023-01-15
TKACZUK	Jakub	PL	Petr MOTLICEK	2022-01-15	2023-01-14
TORNAY	Sandrine	CH	Mathew MAGIMAI DOSS	2021-03-01	2024-08-31
VALENTINO	Marco	IT	André FREITAS	2022-05-15	2025-05-14
VÁSQUEZ RODRÍGUEZ	Laura	CR	Lonneke VAN DER PLAS	2023-08-01	2025-07-31
VLASENKO	Bogdan	DE	Mathew MAGIMAI DOSS	2021-02-15	2023-12-31
WYSOCKI	Oskar	PL	André FREITAS	2023-04-03	2025-04-03
XU	Zhi Ming	CA	Raphaëlle LUISIER	2023-09-15	2025-03-15

3.4 PhD Students

Name		Nationality	Supervisor	Start	End
AMIRI	Mahdi	IR	Ina KODRASI	2023-09-01	2027-08-31
ANNAPUREDDY	R. Reddy	IN	Daniel GATICA-PEREZ	2022-10-01	2026-09-30
BANDARU	Tagore Sanketh	IN	Raphaëlle LUISIER	2023-01-15	2023-05-31
BEHJATI	Melika	IR	James HENDERSON	2020-09-01	2023-08-31
BILALOGLU	Cem	TR	Sylvain CALINON	2022-09-01	2026-08-31
BITTAR	Alexandre	CH	Philip GARNER	2020-03-02	2024-03-01
BOGHETTI	Roberto	IT	Jérôme KÄMPF	2022-02-01	2025-05-31
BROS	Victor	FR	Daniel GATICA-PEREZ	2022-09-01	2026-08-31
CHEN	Haolin	CN	Philip GARNER	2021-08-30	2025-08-01
COLBOIS	Laurent	CH	Sébastien MARCEL	2021-01-01	2024-12-31
COMAN	Andrei	IT	James HENDERSON	2020-10-01	2024-09-30
COPPIETERS DE GIBSON	Louise	BE	Philip GARNER	2020-10-01	2024-09-30
COURDIER	Evann	FR	François FLEURET	2019-03-01	2023-02-28
DÉLITROZ	Maxime	CH	Raphaëlle LUISIER	2022-10-01	2023-12-31
EL HAJAL	Karl	LB	Mathew MAGIMAI DOSS	2023-03-01	2024-02-29
EL ZEIN	Dina	LB	James HENDERSON	2023-11-01	2027-09-30
FEHR	Fabio	CH	James HENDERSON	2021-02-01	2025-01-31
GARCIA SCHU PEIXOTO	Guilherme	BR	Ina KODRASI	2022-01-15	2023-01-31
GIRGIN	Hakan	TR	Sylvain CALINON	2018-09-01	2023-03-31
GUPTA	Anshul	IN	Jean-Marc ODOBEZ	2021-01-01	2024-12-31
HE	Mutian	CN	Philip GARNER	2022-04-01	2026-03-31
HRYNENKO	Olena	UA	Andrea CAVALLARO	2023-03-27	2027-03-26
ISMAYILZADA	Mahammad	CA	Lonneke VAN DER PLAS	2023-09-11	2027-09-10
JANKOWSKI	Julius	DE	Sylvain CALINON	2020-04-01	2024-03-31
JOHARI	Mohammad Mahdi	IR	François FLEURET	2020-01-01	2023-12-31
KARIMI MAHABADI	Rabeeh	IR	James HENDERSON	2018-10-01	2023-03-31
LI	Yiming	CN	Sylvain CALINON	2022-09-01	2026-08-31
LÖW	Tobias	CH	Sylvain CALINON	2021-01-04	2024-12-31
MAI	Florian	DE	James HENDERSON	2018-10-01	2023-05-31
MARIĆ	Ante	HR	Sylvain CALINON	2023-03-06	2027-03-05
MATOBA	Kyle	GB	François FLEURET	2020-08-01	2023-07-31
MEEGAHAPOLA	Iakmal	LK	Daniel GATICA-PEREZ	2019-06-24	2023-06-23
MESSORI	Elisa	IT	Raphaëlle LUISIER	2022-09-19	2026-09-18
MOHAMMADSHAHI	Alireza	IR	James HENDERSON	2019-09-01	2023-08-31
MOSTAANI	Zohreh	IR	Mathew MAGIMAI DOSS	2020-02-01	2024-01-31
NIGMATULINA	Iuliia	RU	Petr MOTLICEK	2021-01-18	2025-01-18
OTROSHI SHAHREZA	Hatef	IR	Sébastien MARCEL	2020-03-01	2024-02-29
PANNATIER	Arnaud	CH	François FLEURET	2020-03-01	2023-02-12
PETERSEN	Molly	US	Lonneke VAN DER PLAS	2022-09-01	2026-08-31
PRASAD	Amrutha	IN	Petr MOTLICEK	2020-08-01	2024-07-31
PUROHIT	Tilak	IN	Mathew MAGIMAI DOSS	2021-06-01	2025-05-31
RAHIMI NOSHANAGH	Parsa	IR	Sébastien MARCEL	2022-08-01	2026-07-31
RAJAPAKSHE	Shalutha	LK	Emmanuel SENFT	2023-09-15	2027-09-15

RAZMJOO FARD	Amirreza	IR	Sylvain	CALINON	2020-12-01	2024-11-30
SAJADMANESH	Sina	IR	Daniel	GATICA-PEREZ	2019-05-15	2023-05-15
SARKAR	Eklavya	CH	Mathew	MAGIMAI DOSS	2021-03-01	2025-02-28
SHETTY	Suhan	IN	Sylvain	CALINON	2019-06-03	2023-06-03
SIVAPRASAD	Prabhu	IN	François	FLEURET	2018-11-01	2023-03-31
TAFASCA	Samy	MA	Jean-Marc	ODOBEZ	2021-03-01	2025-02-28
TARIGOPULA	Neha	IN	Mathew	MAGIMAI DOSS	2021-03-08	2025-02-28
UNNERVIK	Alex	CH	Sébastien	MARCEL	2020-03-01	2024-02-29
VUILLECARD	Pierre	CH	Jean-Marc	ODOBEZ	2022-06-01	2026-05-31
XUE	Teng	CN	Sylvain	CALINON	2021-11-01	2025-10-31
ZHANG	Yan	CN	Sylvain	CALINON	2022-10-01	2026-09-30
ZULUAGA GOMEZ	Juan Pablo	CO	Petr	MOTLICEK	2020-01-01	2023-12-31

3.5 R&D Engineers

Name		Nationality	Supervisor		Start	End
BORNET	Olivier	CH	François	FOGLIA	2014-03-01	
CARRON	Daniel	CH	Olivier	BORNET	2020-08-01	
MAYORAZ	André	CH	Olivier	BORNET	2022-08-01	
ZANGGER	Alicia	CH	Olivier	BORNET	2023-08-01	
DAYER	Yannick	CH	Olivier	BORNET	2020-08-01	
VERZAT	Colombine	FR	Olivier	BORNET	2021-03-01	
MACEIRAS	Jérémy	CH	Olivier	BORNET	2020-08-01	
MARCEL	Christine	CH	Olivier	BORNET	2001-06-01	
GENTILHOMME	Théophile	FR	Olivier	BORNET	2018-03-01	
KAYAL	Salim	CH	Olivier	BORNET	2011-03-01	
TARSETTI	Flavio	CH	Olivier	BORNET	2008-01-01	2023-06-30
CLIVAZ	Guillaume	CH	Olivier	BORNET	2017-04-01	
GAIST	Samuel	CH	Olivier	BORNET	2013-11-01	
NANCHEN	Alexandre	CH	Olivier	BORNET	2008-02-01	
ABBET	Philip	CH	Olivier	BORNET	2006-07-17	
CANÉVET	Olivier	CH	Olivier	BORNET	2016-10-01	
DROZ	William	CH	Olivier	BORNET	2018-03-01	

3.6 Interns

Name		Nationality	Supervisor		Start	End
AGARLA	Mirko	IT	Andrea	CAVALLARO	2023-07-20	2024-01-20
AGNAOU	Zineb	MA	Damien	TENEY	2023-03-01	2023-07-15
AL-DABBOUSSI	Mustapha	CH	Ina	KODRASI	2023-02-01	2024-07-31
ALI	Abid	PK	Jean-Marc	ODOBEZ	2023-08-10	2023-11-09
ALONSO DEL BARRIO	David	ES	Daniel	GATICA-PEREZ	2023-01-01	2023-06-30
ARMANI RENZO	Matheus	BR	André	ANJOS	2022-09-01	2023-08-31
BARANOUSKAYA	Darya	BY	Andrea	CAVALLARO	2023-06-12	2023-10-11
BEN MAHMOUD	Imen	TN	Olivier	BORNET	2022-02-01	2023-07-31

BI	Yuhan	CN	Daniel	GATICA-PEREZ	2023-10-09	2024-04-09
BILARDO	Matteo	IT	Jérôme	KÄMPF	2023-01-09	2023-01-27
BONTEMPELLI	Andrea	IT	Daniel	GATICA-PEREZ	2022-11-14	2023-01-27
BOUTON-BESSAC	Emma	FR	Daniel	GATICA-PEREZ	2022-10-01	2023-01-31
BRAGLIA	Giovanni	IT	Sylvain	CALINON	2023-06-26	2023-09-01
BRAUD	Lucas	FR	Sébastien	MARCEL	2022-09-20	2023-01-06
BRUNINI	Gabriele	IT	André	ANJOS	2022-09-01	2023-03-31
BURDISSO	Sergio	AR	Petr	MOTLICEK	2023-02-15	
CAROFILIS VASCO	Roberto Andrés	ES	Petr	MOTLICEK	2023-05-15	2023-08-17
CHAMOT	Anna	FR	Raphaëlle	LUISIER	2023-02-06	2023-08-06
CHANG	ChunTzu	CH	Emmanuel	SENFT	2023-09-21	2023-12-22
CHANG	Xiaoguang	CN	Damien	TENEY	2023-12-01	2024-11-30
CHAVANNE	Olivier	CH	Jérôme	KÄMPF	2023-03-06	2023-09-06
CHI	Xuemin	CN	Sylvain	CALINON	2022-11-01	2024-10-31
CHU	Dong	CN	Mathew	MAGIMAI DOSS	2023-02-20	2023-06-30
CHUTISILP	Naravich	TH	Jean-Marc	ODOBEZ	2023-04-12	2023-09-12
DALAL	Dhairya	US	André	FREITAS	2023-09-01	2023-12-31
DASTENAVAR	Atharva	FR	Emmanuel	SENFT	2023-09-11	2024-03-11
DE CAMPOS	Ruben	CH	Jean-Marc	ODOBEZ	2022-02-01	2023-07-31
DRAMÉ	Victor	FR	Sylvain	CALINON	2023-02-27	2023-07-07
EGLITIS	Teodors	LV	Sébastien	MARCEL	2023-05-01	2023-07-31
EL AYOUCH	Aliyasin	FR	Emmanuel	SENFT	2023-09-04	2024-03-01
ELBANNA	Gasser	EG	Mathew	MAGIMAI DOSS	2023-04-15	2023-08-15
ESPUÑA FONTCUBERTA	Aleix	ES	Petr	MOTLICEK	2023-01-16	2023-07-16
FARKHONDEH	Arya	IR	Jean-Marc	ODOBEZ	2022-10-01	2023-05-31
FORNAROLI	Alessandro	IT	Daniel	GATICA-PEREZ	2023-10-01	2024-03-31
FOURNIER	Lisa	CH	Raphaëlle	LUISIER	2022-09-19	2023-01-23
FRITSCH	Julian	DE	Mathew	MAGIMAI DOSS	2023-01-01	2023-02-28
FUERTES COIRAS	Daniel	ES	Andrea	CAVALLARO	2023-06-01	2023-09-01
GE	Chang	CN	Ina	KODRASI	2023-09-18	2023-12-31
GE	Wanying	CN	Sébastien	MARCEL	2023-04-10	2023-07-02
GRAFFAGNINO	Andrea	IT	Jérôme	KÄMPF	2023-05-01	2023-10-31
GÜLER	Özgür Acar	CH	André	ANJOS	2023-02-01	2023-09-30
HAEFLIGER	Garance	CH	Raphaëlle	LUISIER	2023-09-18	2024-03-18
HASSOUNE	Hamza	MA	Daniel	GATICA-PEREZ	2023-02-20	2023-07-20
HENKING	Caspar	CH	Michael	LIEBLING	2023-03-09	2023-06-30
HENRY	Oscar	FR	Raphaëlle	LUISIER	2023-11-27	2024-06-07
HUANG	Junduan	CN	Sébastien	MARCEL	2022-12-15	2024-12-14
JUNG	Vincent	FR	Lonneke	VAN DER PLAS	2023-04-01	2023-09-30
KALAAJI	Dana	FR	Ina	KODRASI	2023-09-18	2023-12-31
KHALIL	Driss	MA	Petr	MOTLICEK	2022-11-01	2023-10-31
KIM	Haeeun	KR	Daniel	GATICA-PEREZ	2023-03-01	2023-05-31
KOEHN	Edouard Erwan	CH	Raphaëlle	LUISIER	2022-08-01	2023-01-31
KUMAR	Shashi	IN	Petr	MOTLICEK	2023-04-15	2023-08-15
LESUR	Jean	FR	Sylvain	CALINON	2023-02-20	2023-06-30
LINGHU	Yu	CN	Sébastien	MARCEL	2023-03-01	2023-05-31
MARIÉTHOZ	Cédric	CH	Sylvain	CALINON	2022-02-01	2023-07-31

MICHEL	Tobias	CH	Sylvain	CALINON	2023-02-20	2023-06-30
MICHEL	Samuel	BE	Olivier	BORNET	2022-02-01	2023-07-31
MOZO	Rafael	ES	Daniel	GATICA-PEREZ	2023-02-20	2023-07-20
NADERI	Maryam	IR	Mathew	MAGIMAI DOSS	2023-02-01	2024-07-31
NICOLICIOIU	Armand-Mihai	RO	Damien	TENEY	2023-05-01	2023-09-15
NIKRAY	Tiyam	CH	Daniel	GATICA-PEREZ	2023-03-15	2023-09-15
NIU	Zhenwei	CN	Sylvain	CALINON	2023-01-02	2023-04-30
ÖZBULAK	Gökhan	TR	Sébastien	MARCEL	2023-06-19	2023-12-18
PANG	Yik Lung	GB	Andrea	CAVALLARO	2023-06-15	2023-12-15
PARK	Jung	KR	François	FOGLIA	2022-03-01	2023-02-28
PARRA GALLEGO	Luis Felipe	CO	Mathew	MAGIMAI DOSS	2023-10-01	2024-01-31
PIRAS	Florian	BE	Michael	LIEBLING	2022-12-01	2023-03-31
POCARD	Valentin	FR	Philip	GARNER	2022-11-01	2023-04-30
RUFAI	Amina	NG	Philip	GARNER	2023-02-01	2024-01-31
RUVOLO	Barbara	IT	Mathew	MAGIMAI DOSS	2023-02-20	2023-06-30
SANCHEZ LARA	Alejandra	MX	Petr	MOTLICEK	2023-10-01	2024-01-31
SCHNEUWLY	Christelle	CH	Raphaëlle	LUISIER	2022-08-01	2023-01-31
STEL	Lucas	CH	Jean-Marc	ODOBEZ	2022-09-15	2023-01-31
SYFRIG	Guillaume	CH	Sébastien	MARCEL	2023-02-13	2023-07-13
TAVERNIER	Clara	CH	Daniel	GATICA-PEREZ	2023-02-20	2023-07-20
VANDERREYDT	Geoffroy	BE	Petr	MOTLICEK	2023-03-13	2023-07-13
VAUCHER	Karine	CH	Sébastien	MARCEL	2022-08-02	2023-01-31
VAUTHEY	Adrien	CH	Daniel	GATICA-PEREZ	2023-09-15	2024-03-15
VIDAL	Maxime	FR	Petr	MOTLICEK	2023-06-01	2024-05-31
XU	Yixuan	CN	Lonneke	VAN DER PLAS	2023-10-11	2024-01-11
ZANELLA	Riccardo	IT	Sylvain	CALINON	2023-06-19	2023-09-19

3.7 Academic Visitors

Name		Nationality	Supervisor		Start	End
DE REZENDE ROCHA	Anderson	BR	Sébastien	MARCEL	2023-05-29	2023-08-31
GUSICUMA	Danilo	BR	André	FREITAS	2023-03-15	2023-04-30
SABER	Rhéa	LB	Emmanuel	SENFT	2023-03-09	2023-06-09
NORIEGA POLO	Carlos	ES	Michael	LIEBLING	2023-01-20	2023-02-17
MONTAZERI	Ahad	IR	Jérôme	KÄMPF	2023-08-16	2023-08-25
XOMPERO	Alessio	IT	Andrea	CAVALLARO	2023-04-29	2023-05-25

4 List of Projects

An overview of the projects that have been active during the year 2023 is presented below. The projects are grouped in four categories, namely National Research Projects, National Innovation Projects, European/International Research Projects, and Industry-oriented Projects. Section 4.2 lists the projects accepted in 2023 and starting the following year.

4.1 Projects in Progress during 2023

4.1.1 National Research Projects

1. Name **2000LAKES** (alpine research and citizen science toward the microbial conservation of high-mountain lakes in Switzerland)
   Funding EPFL-UNIL
   Coordinator EPFL
   Duration 2022-03-01 – 2023-12-31
   Partner(s) University of Lausanne; Idiap Research Institute; EPFL

2. Name **AI4AUTISM2** (Digital Phenotyping of Autism Spectrum Disorders in Children)
   Funding SNSF Sinergia
   Coordinator University of Geneva
   Duration 2021-11-01 – 2025-10-31
   Partner(s) SUPSI; Idiap Research Institute

3. Name **CAD4IED** (Computer assisted detection and grading of inflammatory eye diseases via fluorescein angiograms and novel biomarkers)
   Funding Idiap Research Institute
   Coordinator Idiap Research Institute
   Duration 2022-12-01 – 2024-03-31
   Partner(s) Fondation Asile des aveugles Hôpital Jules-Gonin; Lucerne Cantonal Hospital; Idiap Research Institute

4. Name **CHASPEEPRO** (Characterisation of motor speech disorders and processes)
   Funding SNSF Sinergia
   Coordinator University of Geneva
   Duration 2021-12-01 – 2025-11-30
   Partner(s) Idiap Research Institute; HUG; New Sorbonne University Paris 3

5. Name **C-LING** (Towards Creative systems with LINGuistic modelling)
   Funding SNSF Division II
   Coordinator Idiap Research Institute
   Duration 2022-09-01 – 2026-08-31
   Partner(s) -

6. Name **CODIMAN** (A future that works: Cobotics, digital skills and the re-humanization of the workplace)
   Funding SNSF NRP77
   Coordinator Bern University of Applied Sciences
   Duration 2020-05-01 – 2024-12-31
   Partner(s) Idiap Research Institute

7. Name **COLLABCLOUD** (CollabCloud)
   Funding Loterie Romande
   Coordinator Idiap Research Institute
   Duration 2023-06-01 – 2024-05-31
   Partner(s) -

8. 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 – 2025-04-30
   Partner(s) Idiap Research Institute; CSEM

9. Name **EPOC** (A personalized speech recognition framework for audio messaging on the edge)
   Funding SNSF Bridge Proof of Concept
   Coordinator Idiap Research Institute
   Duration 2023-06-01 – 2024-05-31
   Partner(s) -

10. Name **EVOLANG** (Evolving Language)
    Funding SNSF
    Coordinator University of Zurich
    Duration 2020-06-01 – 2024-05-31
    Partner(s) University of Neuchâtel; Zurich University of the Arts; University of Lausanne; ETHZ; University of Basel; University of Geneva; University of Fribourg; EPFL; Idiap Research Institute

11. Name **EYE-TRACKING** (Creating a Multilingual Eye-Tracking Corpus for Human and Machine-Based Language Processing)
    Funding University of Zurich
    Coordinator University of Zurich
    Duration 2022-09-01 – 2025-08-31
    Partner(s) University of Zurich; Idiap Research Institute

12. Name **FAIRMI** (Machine Learning Fairness with Application to Medical Images)
    Funding SNSF Division II
    Coordinator Federal University of Sao Paulo
    Duration 2023-11-01 – 2027-10-31
    Partner(s) Idiap Research Institute

13. Name **INTREPID** (Automated interpretation of political and economic policy documents: Machine learning using semantic and syntactic information)
    Funding SNSF Sinergia
    Coordinator Graduate Institute of International and Development Studies
    Duration 2019-01-01 – 2023-12-31
    Partner(s) Idiap Research Institute

14. Name **LAOS** (Learning Representations of Abstraction in Text)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2018-10-01 – 2023-03-31
    Partner(s) SNSF

15. Name **LEARN-REAL** (LEARNing physical manipulation skills with simulators using REAListic variations)
    Funding SNSF ERA-NET
    Coordinator Idiap Research Institute
    Duration 2019-04-01 – 2023-03-31
    Partner(s) Italian Institute of Technology; Lyon Central School

16. 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 – 2023-12-31
    Partner(s) NOVARTIS FOUNDATION

17. Name **MELAS-KPI** (Towards the development of predictive biomarkers for patient stratification and immunotherapy response in cancer)
    Funding Idiap Research Institute
    Coordinator Idiap Research Institute
    Duration 2023-04-01 – 2024-03-31
    Partner(s) Hôpital du Valais; CHUV; Idiap Research Institute

18. Name **NAST** (Neural Architectures for Speech Technology)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2020-02-01 – 2024-09-30
    Partner(s) -

19. Name **NATAI** (The Nature of Artificial Intelligence)
    Funding SNSF Agora
    Coordinator Idiap Research Institute
    Duration 2020-12-01 – 2023-11-30
    Partner(s) Musee de la Main UNIL/CHUV

20. Name **NEUMATH** (Neural Discourse Inference over Mathematical Texts)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2022-04-01 – 2025-03-31
    Partner(s) -

21. Name **NEUROCIRT** (Investigating the role of cytoplasmic intronic sequences in ALS pathogenesis)
    Funding SNSF Division III
    Coordinator Idiap Research Institute
    Duration 2022-09-01 – 2026-08-31
    Partner(s) -

22. Name **NEWSONAI** (Scientific voice to latest news of AI technology 2022)
    Funding SNSF Agora
    Coordinator Idiap Research Institute
    Duration 2023-04-01 – 2024-03-31
    Partner(s) -

23. Name **NKBP** (Deep Learning Models for Continual Extraction of Knowledge from Text)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2020-10-01 – 2024-09-30
    Partner(s) Catholic University of Leuven

24. 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

25. Name **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 – 2024-12-31
    Partner(s) University of Zurich; University of Surrey;University of Applied Sciences of Special Needs Education

26. Name **STEADI** (Storytelling and first impressions in face-to-face and algorithm-powered digital interviews)
    Funding SNSF Division I
    Coordinator University of Neuchâtel
    Duration 2021-02-01 – 2025-01-31
    Partner(s) Idiap Research Institute; University of Lausanne

27. Name **SWITCH** (Learning by Switching Roles in Physical Human-Robot Collaboration)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2021-03-01 – 2024-02-29
    Partner(s) Jozef Stefan Institute

28. Name **TIPS** (Towards Integrated processing of Physiological and Speech signals)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2019-12-01 – 2023-11-30
    Partner(s) CSEM; Coaching and Moderation

4.1.2 National Innovation Projects

1. Name **ABROAD** (Development of an NLP tool for selecting the potential sources of novel antibiotic active against multiresistant microbes)
   Funding The Ark
   Coordinator Inflamalps SA
   Duration 2022-09-01 – 2024-02-29
   Partner(s) IHMA Europe Sàrl - BioArk; Idiap Research Institute; CimArk; Inflamalps SA

2. Name **BIPED** (Self-driving technology to guide blind pedestrians)
   Funding Innosuisse
   Coordinator Fusion Lab Technologies Sàrl
   Duration 2023-05-01 – 2024-10-31
   Partner(s) Idiap Research Institute

3. Name **EGUZKI** (A simulation tool for District Heating Networks (DHN) based on Artificial Intelligence (AI) to quickly assess and predict the performance of complex meshed networks)
   Funding SFOE
   Coordinator RWB Valais SA
   Duration 2020-09-01 – 2024-08-30
   Partner(s) Satom SA; Altis Groupe SA; Oiken SA; Idiap Research Institute

4. Name **EPARTNERS4ALL** (ePartners4All: a (personalized and) blended care solution with virtual buddy for child health)
   Funding Innosuisse
   Coordinator Netherlands Organisation for Applied Scientific Research
   Duration 2021-11-15 – 2024-05-15
   Partner(s) Xpert Health Cory Care; Idiap Research Institute; Delft University of Technology; Eyeware; Bern University of Applied Sciences; Organization for Fire Services, Crisis Management and Public Health; MedVision 360 BV; Leiden University Medical Center; Topicus Healthcare B.V.; Interactive Robotics B.V.; Therapieland

5. Name **HARDENING** (Heterogeneous face recognition for unified identity management)
   Funding Innosuisse
   Coordinator Idiap Research Institute
   Duration 2021-02-01 – 2023-04-30
   Partner(s) Facedapter Sàrl

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

7. Name **IVECT** (Impact of greening on the energy balance and thermal comfort of buildings and districts)
   Funding SFOE
   Coordinator HES-SO Valais
   Duration 2020-12-01 – 2023-11-30
   Partner(s) Zurich University of Applied Sciences; CREM; City of Zurich; Etat du Valais; Idiap Research Institute

8. Name **LUCIDELES** (Leveraging User-Centric Intelligent Daylight and Electric Lighting for Energy Saving)
   Funding SFOE
   Coordinator University of Fribourg
   Duration 2020-02-01 – 2023-12-31
   Partner(s) Regent ; Idiap Research Institute

9. Name **MALAT** (Machine Learning for Air Traffic)
   Funding Innosuisse
   Coordinator Idiap Research Institute
   Duration 2020-03-01 – 2023-02-28
   Partner(s) SkySoft-ATM

10. Name **PRIME** (PRoof of Identity: Multi-modalitiEs biometrics)
    Funding Innosuisse
    Coordinator Proof of Identity Labs
    Duration 2023-10-02 – 2024-02-29
    Partner(s) Idiap Research Institute

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

12. Name **SENTINEL-UK** (Securing Age Estimation and Digital IDs against Presentation and Injection Threats)
    Funding Innosuisse
    Coordinator Privately SA
    Duration 2023-11-15 – 2025-04-15
    Partner(s) Idiap Research Institute, ACCS, AVPA

13. Name **SINFONIA** (Social IntelligeNce FOr strategic aNticIpAtion)
    Funding Innosuisse
    Coordinator Bloom Suisse Sàrl
    Duration 2023-04-17 – 2025-04-16
    Partner(s) Idiap Research Institute

14. Name **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) Bern University of Applied Sciences; Idiap Research Institute

4.1.3 European and International Research Projects

1. Name **AI4MEDIA** (A European Excellence Centre for Media, Society and Democracy)
   Funding H2020
   Coordinator Centre for Research and Technology Hellas
   Duration 2020-09-01 – 2024-08-31
   Partner(s) Joanneum Research; Idiap Research Institute; Fraunhofer; Commissariat à l'Energie Atomique et aux Energies Alternatives; Catholic University of Leuven; HES-SO Valais; University of Amsterdam; Queen Mary University of London; University of Trento; Deutsche Welle; CNR; Italian Radio Television; Aristotle University of Thessaloniki; University of Florence; University of Bucarest; Globaz SA; University of Malta; Côte d'Azur University; Ircam; IBM Ireland ; Athens Technology Center; Netherlands Institute for Sound and Vision; VRT; Interdigital R&D France; F6S; Modl.ai; Imagga Technologies; Grassroots

2. Name **CRITERIA** (Comprehensive data-driven Risk and Threat Assessment Methods for the Early and Reliable Identification, Validation and Analysis of migration-related risks)
   Funding H2020
   Coordinator University of Hannover
   Duration 2021-09-01 – 2024-08-31
   Partner(s) webLyzard technology; Malta Police Force; European University of Cyprus; Estonian Police and Border Guard Board; University of Malta; Ministry of Interior Croatia; General Inspectorate of Romanian Border Police; Swedish National Police Authority; University of Groningen; Knowledge and Innovation; ARSIS Association for the Social Support of Youth; Hensoldt Analytics; Idiap Research Institute; Centre for Research and Technology Hellas

3. Name **ELIAS** (European Lighthouse of AI for Sustainability)
   Funding Horizon Europe
   Coordinator University of Trento
   Duration 2023-09-01 – 2027-08-31
   Partner(s) Jozef Stefan Institute; University of Bucarest; UVEG; UMIL; UMEA UNIVERSITET; Max Planck Society; Italian Institute of Technology; IBM Ireland ; Foundation Of The Community Of Valencia For Research, Promotion And Commercial Studies; FBK; ENG; ETHZ; BITDEFENDER; AALTO; University of Amsterdam; POLIMI; Idiap Research Institute

4. Name **GRAIL-1** (Generative Range and Altitude Identity Learning)
   Funding IARPA
   Coordinator University of Southern California
   Duration 2021-11-12 – 2023-05-11
   Partner(s) Idiap Research Institute

5. Name **GRAIL-2** (Generative Range and Altitude Identity Learning)
   Funding IARPA
   Coordinator University of Southern California
   Duration 2023-05-12 – 2024-11-11
   Partner(s) Idiap Research Institute

6. Name **ICARUS** (Innovative AppRoach to Urban Security)
   Funding H2020
   Coordinator Forum Europeen Pour La Securite Urbaine
   Duration 2020-09-01 – 2024-08-31
   Partner(s) Landeshauptstadt Stuttgart; Riga Municipal Police; City of Rotterdam; Lisbon Municipal Police; Makesense; Idiap Research Institute; University of Leeds; Greek Center for Security Studies; Salzburg University of Applied Sciences ; Ethical and Legal Plus; Erasmus University Rotterdam; Globaz SA; City of Torino; Commune de Nice; Eurocircle Association; Panteion University of Social and Political Sciences; University of Salford

7. Name **INTELLIMAN** (AI-Powered Manipulation System for Advanced Robotic Service, Manufacturing and Prosthetics)
   Funding Horizon Europe
   Coordinator University of Bologna
   Duration 2022-09-01 – 2026-02-28
   Partner(s) Technical University of Catalonia; DLR; University of Zurich; University of Erlangen-Nuremberg ; University of Genoa; Bavarian Research Alliance Gmbh; Elvez; National Institute for Industrial Accidents Insurance Inail; Eurecat Foundation; University of Campania Luigi Vanvitelli; Ocado Innovation Limited; Idiap Research Institute

8. Name **META-SPOOF** (Trustworthy AR-VR Biometric Authentication)
   Funding Meta Research
   Coordinator Idiap Research Institute
   Duration 2023-05-01 – 2024-04-30
   Partner(s) -

9. Name **MULTIPLEYE** (Enabling multilingual eye-tracking data collection for human and machine language processing research)
   Funding COST
   Coordinator University of Copenhagen
   Duration 2022-09-28 – 2026-09-27
   Partner(s) Idiap Research Institute

10. 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) Centre for Research and Technology Hellas; Idiap Research Institute; University of Bologna; University of Zaragoza; University of Ljubljana; University of Oxford; Fiat SCPA Research Centre; Institut Franco-Allemand de Recherches de Saint-Louis; Rise Research Institutes of Sweden AB; University of Latvia; Inertia Technology B.V.; Ocado Innovation Limited; Free University of Bolzano

11. Name **SOTERIA** (uSer-friendly digiTal sEcured peRsonal data and prIvacy plAtform)
    Funding H2020
    Coordinator Ariadnext
    Duration 2021-10-01 – 2024-09-30
    Partner(s) Autonomous University of Barcelona; Basque Health Service Osakidetza; ipcenter.at; Noria Onlus; Idiap Research Institute; INRIA; Catholic University of Leuven; Audencia Business School; Erdyn Atlantique; Asociatia Infocons; Scytl Election Technologies; Stelar Security Technology Law Research UG

12. Name **TRACY** (A big-data analyTics from base-stations Registrations And Cdrs e-evidence sYstem)
    Funding Horizon Europe
    Coordinator PERFORMANCE
    Duration 2023-06-01 – 2025-05-31
    Partner(s) Hellenic Police; University of Applied Sciences for Public Administration and Legal Affairs in Bavaria; TIMELEX; Idiap Research Institute; Greek Center for Security Studies; COSMOTE; POLICE ROMANIA GL; POLICE MOLDOVA

13. Name **TRESPASS-ETN** (TRaining in Secure and PrivAcy-preserving biometricS)
    Funding H2020
    Coordinator EURECOM
    Duration 2020-01-01 – 2023-12-31
    Partner(s) Catholic University of Leuven; Autonomous University of Madrid ; University of Groningen; Idiap Research Institute; Darmstadt University of Applied Sciences; Chalmers University of Technology

14. 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 – 2024-03-31
    Partner(s) -

15. Name **UNIDIVE** (Universality, diversity and idiosyncrasy in language technology)
    Funding COST
    Coordinator Université Paris-Saclay
    Duration 2022-09-23 – 2026-09-22
    Partner(s) Idiap Research Institute

16. Name **WENET** (The Internet of US)
    Funding H2020
    Coordinator University of Trento
    Duration 2019-01-01 – 2023-06-30
    Partner(s) Potosino Institute of Scientific and Technological Research; Open University of Cyprus; Ben Gurion University of the Negev; Martel; U-Hopper; London School of Economics and Political Science; Eberhard Karls University of Tübingen; National University of Mongolia; Catholic University of Asunción; Amrita University; Jilin University; University of Aalborg; Idiap Research Institute

4.1.4 Industry-Oriented Projects

Note that the following list of industry-oriented projects is not exhaustive.

1. Name **AI-SENSOR** (Sensor Fusion and Active Sensing for World-View Understanding)
   Funding Industrial
   Coordinator Idiap Research Institute
   Duration 2020-02-03 – 2024-01-31
   Partner(s) ams International

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

3. Name **S3C-CBC-2023** (Force-based assembly autonomous learning)
   Funding Swiss Cobotics Competence Center (S3C)
   Coordinator AICA
   Duration 2023-04-03 – 2023-07-26
   Partner(s) HFTM; Bota Systems; HSLU; Idiap Research Institute; Fanuc; Schunk

4. Name **SPACE** (Speech analytics for communications intelligence)
   Funding Industrial
   Coordinator armasuisse
   Duration 2022-09-01 – 2023-08-31
   Partner(s) Idiap Research Institute

5. Name **UNIPHORE** (Development of spoken language technologies for call centers)
   Funding Industrial
   Coordinator Uniphore
   Duration 2021-01-01 - ongoing
   Partner(s) Idiap Research Institute

6. Name **WAVE2-96** (H2020-SESAR-PJ.10-W2-Solution 96)
   Funding Industrial
   Coordinator Idiap Research Institute
   Duration 2020-10-01 – 2023-01-31
   Partner(s) DLR

4.2 Projects Awarded in 2023 and Starting in 2024

1. Name **COLLABCLOUD** (CollabCloud - A Collaborative Research Cloud Infrastructure powering Discovery and Exchange)
   Funding SNSF R'Equip
   Coordinator Idiap Research Institute
   Duration 2024-01-01 - 2024-12-31
   Partner(s) -

2. Name **ELOQUENCE** (Multilingual and Cross-cultural interactions for context-aware, and bias-controlled dialogue systems for safety-critical applications)
   Funding Horizon Europe
   Coordinator Telefonica
   Duration 2024-01-01 - 2026-12-31
   Partner(s) BSC CNS; FBK; CNR; Synelixis; OMILIA; Grantxpert; Transformation Lightouse; InoSens; UNS; UESSEX; Privanova SAS; European University Institute; Idiap Research Institute; Brunel University London; Brno University of Technology

3. Name **FACTCHECK** (FactCheck - Collaborative Fact Checking of Digital Media)
   Funding Hasler Foundation
   Coordinator Idiap Research Institute
   Duration 2024-04-01 - 2025-03-31
   Partner(s) -

4. Name **LIOGPT** (LIOGPT)
   Funding Innosuisse
   Coordinator F&P Robotics
   Duration 2024-04-01 - 2024-04-30
   Partner(s) Idiap Research Institute

5. Name **MAVERICK** (Maximum evidence platform for explainable predictions of risk)
   Funding Innosuisse
   Coordinator Basinghall Analytics
   Duration 2024-03-01 - 2026-02-28
   Partner(s) Idiap Research Institute

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

7. Name **PAUSE** (PAuSE: Pathological Speech Enhancement)
   Funding SNSF Division II
   Coordinator Idiap Research Institute
   Duration 2024-01-01 - 2027-12-31

Partner(s) -

8. Name **PMPM** (Predictive Manufacturing Predictive Maintenance)
   Funding Innosuisse
   Coordinator HEI-VS - Haute école d'ingénierie
   Duration 2024-03-01 - 2025-08-31
   Partner(s) Idiap Research Institute, Constellium Valais SA, LYSR Sàrl,Eversys S.A.

9. Name **ROSALIND** (Robust anti-fraud algorithms against malicious AI-generated face images and travel documents)
   Funding Innosuisse
   Coordinator PXL Vision
   Duration 2024-02-01 - 2025-07-31
   Partner(s) Idiap Research Institute

10. Name **SPINOFF** (Navigating Spin-off Challenges)
    Funding SNSF Spark
    Coordinator Idiap Research Institute
    Duration 2024-02-01 - 2025-01-31
    Partner(s) -

5 List of Publications and Patents from 2023

5.1 PhD Theses

[1] J. Fritsch, "Novel methods for detection and analysis of atypical aspects in speech," PhD thesis, École polytechnique fédérale de Lausanne, 2023.

[2] H. Girgin, "Learning and optimization of anticipatory feedback controllers for robot manipulation," PhD thesis, École polytechnique fédérale de Lausanne, 2023.

[3] E. Hermann, "On matching data and model in lf-mmi-based dysarthric speech recognition," PhD thesis, École polytechnique fédérale de Lausanne, 2023.

[4] S. H. Kabil, "Sparse autoencoders for speech modeling and recognition," PhD thesis, École polytechnique fédérale de Lausanne, 2023.

[5] F. Mai, "Text representation learning for low cost natural language understanding," PhD thesis, École polytechnique fédérale de Lausanne, 2023.

[6] F. Marelli, "Practical computational imaging by use of spatiotemporal light modulation: From simulations to applications in biological microscopy," PhD thesis, EPFL, Oct. 2023.

[7] A. Marfurt, "Interpretable representation learning and evaluation for abstractive summarization," PhD thesis, École polytechnique fédérale de Lausanne, Jun. 2023.

[8] A. Mohammadshahi, "Modeling structured data in attention-based models," PhD thesis, EPFL, Oct. 2023.

[9] S. Sajadmanesh, "Privacy-preserving machine learning on graphs," PhD thesis, EPFL, 2023.

[10] P. T. Sivaprasad, "Confidence matters : Applications to semantic segmentation," PhD thesis, École polytechnique fédérale de Lausanne, Mar. 2023.

5.2 Articles in Scientific Journals

[1] M. Bilardo, J. H. Kämpf, and E. Fabrizio, "From zero energy to zero power buildings: A new paradigm for a sustainable transition of the building stock," *Sustainable Cities and Society*, vol. 101, p. 105 136, 2024.

[2] R. Boghetti and J. H. Kämpf, "Verification of an open-source python library for the simulation of district heating networks with complex topologies," *Energy*, vol. 290, p. 130 169, 2024.

[3] N. Ahrenhold, H. Helmke, T. Mühlhausen, O. Ohneiser, M. Kleinert, h. Ehr, L. Klamert, and Z.-G. Juan, "Validating automatic speech recognition and understanding for pre-filling radar labels-increasing safety while reducing air traffic controllers' workload," *Aerospace*, vol. 10, no. 6, p. 538, Jun. 2023.

[4] S. Aviyente, A. F. Frangi, E. Meijering, A. Muñoz-Barrutia, M. Liebling, D. Van De Ville, J.-C. Olivo-Marin, J. Kovačević, and M. Unser, "From nano to macro: An overview of the ieee bio image and signal processing technical committee," *IEEE Signal Processing Magazine*, vol. 40, no. 4, pp. 61–71, Jun. 2023.

[5] A. Bangerter, E. Mayor, S. Muralidhar, E. P. Kleinlogel, D. Gatica-Perez, and M. Schmid Mast, "Automatic identification of storytelling responses to past-behavior interview questions via machine learning," *International Journal of Selection and Assessment*, 2023.

[6] F. Battini, G. Pernigotto, F. Morandi, A. Gasparella, and J. Kämpf, "Assessment of subsidization strategies for multi-objective optimization of energy efficiency measures for building renovation at district scale," *Energies*, vol. 16, no. 15, Aug. 2023.

[7] M. Behjati and J. Henderson, "Inducing meaningful units from character sequences with dynamic capacity slot attention," *Transactions on Machine Learning Research*, 2023.

[8] C. Bilaloglu, T. Löw, and S. Calinon, "Whole-body ergodic exploration with a manipulator using diffusion," *IEEE Robotics and Automation Letters*, vol. 8, no. 12, Dec. 2023.

[9] A. Bogatu, M. Wysocka, O. Wysocki, H. Butterworth, M. Pillai, J. Allison, D. Landers, E. Kilgour, F. Thistlethwaite, and A. Freitas, "Meta-analysis informed machine learning: Supporting cytokine storm detection during car-t cell therapy," *Journal of Biomedical Informatics*, vol. 142, Jun. 2023.

[10] L. Castelo-Branco, A. Freitas, and e. al., "Learning lessons from the COVID-19 pandemic for real world evidence research in oncology–shared perspectives from an international consortia," *ESMO Open*, 2023.

[11] M. Delmas, O. Filangi, D. Christophe, P. Nils, F. Vinson, P. R. Mier, G. Franck, F. Jourdan, and C. Frainay, "Suggesting disease associations for overlooked metabolites using literature from metabolic neighbors," *GigaScience*, vol. 12, p. 13, Sep. 2023.

[12] A. Demir Dilsiz, K. E. Nweye, A. J. Wu, J. Kämpf, F. Bilijecki, and Z. Nagy, "How spatio-temporal resolution impacts urban energy calibration," *Energy and Buildings*, vol. 292, no. 113175, Aug. 2023.

[13] M. Fajcik, P. Motlicek, and P. Smrz, "Claim-dissector: An interpretable fact-checking system with joint re-ranking and veracity prediction," *Association for Computational Linguistics*, vol. Findings of the Association for Computational Linguistics (ACL), 10184–10205, Jul. 2023.

[14] A. Fornaroli and D. Gatica-Perez, "Urban crowdsourcing platforms across the world: A systematic review," *ACM Digital Government: Research and Practice*, 2023.

[15] T. Gentilhomme, M. Villamizar, J. Corre, and J.-M. Odobez, "Towards smart pruning: Vinet, a deep-learning approach for grapevine structure estimation," *Computers and Electronics in Agriculture*, vol. 207, p. 107 736, Apr. 2023.

[16] A. George and S. Marcel, "Attacking face recognition with t-shirts: Database, vulnerability assessment and detection," *IEEE Access*, 2023.

[17] C. Geric, Z. Z. Qin, C. M. Denkinger, S. V. Kik, B. Marais, A. Anjos, P.-M. David, F. A. Khan, and A. Trajman, "The rise of artificial intelligence reading of chest x-rays for enhanced tb diagnosis and elimination," *The International Journal of Tuberculosis and Lung Disease*, vol. 27, no. 5, pp. 367–372, May 2023.

[18] M. Hagenow, E. Senft, N. Orr, R. Radwin, M. Gleicher, B. Mutlu, D. P. Losey, and M. Zinn, "Coordinated multi-robot shared autonomy based on scheduling and demonstrations," *IEEE Robotics and Automation Letters*, vol. 8, no. 12, pp. 8335 –8342, 2023.

[19] P. Helm, A. de Götzen, L. Cernuzzi, A. Hume, S. Diwakar, S. Ruiz-Correa, and D. Gatica-Perez, "Diversity and neocolonialism in big data research: Avoiding extractivism while struggling with paternalism," *Big Data and Society*, 2023.

[20] U. Jensen, D. Rohner, O. Bornet, D. Carron, P. N. Garner, D. Loupi, and J. Antonakis, "Combating COVID-19 with charisma: Evidence on governor speeches in the United States," *The Leadership Quarterly*, May 2023.

[21] Z.-G. Juan, N. Iuliia, A. Prasad, P. Motlicek, D. Khalil, S. Madikeri, A. Tart, I. Szoke, V. Lenders, M. Rigault, and K. Choukri, "Lessons learned in transcribing 5000 h of air traffic control communications for robust automatic speech understanding," *Aerospace*, vol. 10, no. 10, Oct. 2023.

[22] Z.-G. Juan, A. Prasad, N. Iuliia, P. Motlicek, and M. Kleinert, "A virtual simulation-pilot agent for training of air traffic controllers," *Aerospace*, vol. 10, no. 5, 2023.

[23] D. Khalil, A. Prasad, P. Motlicek, Z.-G. Juan, N. Iuliia, S. Madikeri, and S. Christof, "An automatic speaker clustering pipeline for the air traffic communication domain," *Aerospace*, vol. 10, Sep. 2023.

[24] V. Krivokuca, “Exploring human biometrics: A focus on security concerns and deep neural networks,” *APSIPA Transactions on Signal and Information Processing*, 2023.

[25] T. Kulak and S. Calinon, “Combining social and intrinsically-motivated learning for multi-task robot skill acquisition,” *IEEE Transactions on Cognitive and Developmental Systems*, vol. 15, no. 2, pp. 385–394, 2023.

[26] G. Lamberti, R. Boghetti, J. Kämpf, F. Fantozzi, F. Leccese, and G. Salvadori, “Development and comparison of adaptive data-driven models for thermal comfort assessment and control,” *Total Environment Research Themes*, vol. 8, Dec. 2023.

[27] Y. Lin and D. Gatica-Perez, “Characterizing swiss alpine lakes: From wikipedia to citizen science,” *ACM Journal on Computing and Sustainable Societies*, 2023.

[28] T. Löw and S. Calinon, “Geometric algebra for optimal control with applications in manipulation tasks,” *IEEE Transactions on Robotics*, May 2023.

[29] R. Luisier, C. Andreassi, L. Fournier, and A. Riccio, “The predicted rna-binding protein regulome of axonal mrnas,” *Genome Research*, 2023.

[30] F. Marelli, A. Ernst, N. Mercader, and M. Liebling, “Paaq: Paired alternating acquisitions for virtual high frame rate multichannel cardiac fluorescence microscopy,” *Biological Imaging*, vol. 3, e20, Nov. 2023.

[31] F. Marelli and M. Liebling, “Efficient compressed sensing reconstruction for 3d fluorescence microscopy using optomechanical modulation tomography (ommt) with a 1+2d regularization,” *Optics Express*, vol. 31, no. 20, pp. 31 718–31 733, Nov. 2023.

[32] K. Matoba, N. Dimitriadis, and F. Fleuret, “Benefits of max pooling in neural networks: Theoretical and experimental evidence,” *Transactions on Machine Learning Research*, 2023.

[33] J. Meadows and A. Freitas, “Introduction to mathematical language processing: Informal proofs, word problems, and supporting tasks,” *Transactions of the ACL*, 2023.

[34] A. Miani, L. van der Plas, and A. Bangerter, “Loose and tight: Creative formation but rigid use of nominal compounds in conspiracist texts,” *The Journal of Creative Behavior*, 2023.

[35] C. Niklaus, M. Cetto, A. Freitas, and S. Handschuh, “A canonical context-preserving representation for open ie: Extracting semantically typed relational tuples from complex sentences,” *Knowledge-based Systems*, 2023.

[36] O. Ohneiser, H. Helmke, S. Shetty, M. Kleinert, h. Ehr, S. Schier-Morgenthal, S. S. Sarfjoo, P. Motlicek, Š. Murauskas, T. Pagirys, H. Usanovic, M. Meštrović, and A. Černá, “Assistant based speech recognition support for air traffic controllers in a multiple remote tower environment,” *Aerospace*, vol. 10, no. 6, Jun. 2023.

[37] H. Otroshi Shahreza and S. Marcel, “Comprehensive vulnerability evaluation of face recognition systems to template inversion attacks via 3d face reconstruction,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[38] H. Otroshi Shahreza, Y. Y. Shkel, and S. Marcel, “Measuring linkability of protected biometric templates using maximal leakage,” *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 2262 –2275, 2023.

[39] H. Otroshi Shahreza, A. Veuthey, and S. Marcel, “Toward high-resolution face image generation from coded aperture camera,” *IEEE Sensors Letters*, 2023.

[40] M. Piwecka, R. Luisier, and C. Andreassi, “Rna at a breaking point? cytoplasmic cleavage and other post-transcriptional rna processing in neurodevelopment and disease,” *Frontiers in Molecular Neuroscience*, 2023.

[41] P. Praveena, Y. Wang, E. Senft, M. Gleicher, and B. Mutlu, "Periscope: A robotic camera system to support remote physical collaboration," *Proceedings of the ACM on Human Computer Interaction*, 2023.

[42] G. Ramírez-de-la Rosa, H. Jiménez-Salazar, E. Villatoro-Tello, V. Reyes-Meza, and J. Rojas-Avila, "A lexical-availability-based framework from short communications for automatic personality identification," *Cognitive Systems Research*, vol. 79, pp. 126–137, Jan. 2023.

[43] R. Saese, A. Freitas, and e. al., "Defining the role of real-world data in cancer clinical research: The position of the european organisation for research and treatment of cancer," *European Journal of Cancer*, 2023.

[44] S. Shetty, T. S. Lembono, T. Löw, and S. Calinon, "Tensor train for global optimization problems in robotics," *International Journal of Robotics Research*, p. 25, 2023.

[45] J. Taylor, A. Subramanian, A. Freitas, D. Mendes, and C. Dickinson, "What do individuals with visual impairment need and want from a dialogue-based digital assistant?" *Clinical and Experimental Optometry*, 2023.

[46] B. Ti, A. Razmjoo, Y. Gao, J. Zhao, and S. Calinon, "A geometric optimal control approach for imitation and generalization of manipulation skills," *Robotics and Autonomous Systems*, 2023.

[47] B. Tobler-Ammann, F. Schuind, L. Voillat, T. Gentilhomme, E. Vögelin, N. Murith, and B. Masserey, "Development of 3d-printed patient-specific anatomical braces (psab) for distal radius and scaphoid fractures," *Journal of wrist Surgery*, 2023.

[48] M. Wysocka, O. Wysocki, M. Zufferey, D. Landers, and A. Freitas, "A systematic review of biologically-informed deep learning models for cancer: Fundamental trends for encoding and interpreting oncology data," *BMC Bioinformatics*, vol. 24, no. 198, May 2023.

[49] O. Wysocki, M. Bykowska-Tumasz, and K. Sikorska, "Epidemiological and clinical analysis of polish short-term and long-term travelers returning from tropical countries," *Travel Medicine and Infectious Disease*, vol. 55, Sep. 2023.

[50] O. Ziff, J. Neeves, J. Mitchell, G. E. Tyzack, C. Martinez-Ruiz, R. Luisier, A. Chakrabarti, N. McGranahan, K. Litchfield, S. Boulton, A. Al-Chalabi, G. Kelly, J. Humphrey, and R. Patani, "Integrated transcriptome landscape of als identifies genome instability linked to tdp-43 pathology," *Nature Communications*, 2023.

5.3 Book Chapters

[1] S. Calinon, "Programming industrial robots from few demonstrations.," in *Human-Robot Collaboration: Unlocking the potential for industrial applications*, Z. Doulgeri and F. Dimeas, Eds., Doulgeri, Z. and Dimeas, F., Institution of Engineering and Technology (IET), 2023, pp. 9–37.

[2] S. R. Dash and E. VILLATORO-TELLO, *Intelligent Technologies: Concepts, Applications, and Future Directions, Volume 2*, S. R. Dash, H. Das, K.-C. Li, and E. VILLATORO-TELLO, Eds. Springer, Jun. 2023, vol. 1098.

[3] A. George and S. Marcel, "Robust face presentation attack detection with multi-channel neural networks," in *Handbook of Biometric Anti-Spoofing*, Springer, 2023.

[4] H. Girgin, J. Jankowski, and S. Calinon, "Reactive anticipatory robot skills with memory," in *Robotic Research*, A. G. Billard, T. Asfour, and O. Khatib, Eds., Springer, 2023, pp. 436–451.

5.4 Articles in Conference Proceedings

[1] S. Kumar, S. Madikeri, N. Iuliia, E. VILLATORO-TELLO, P. Motlicek, K. P. D. S, S. P. Dubagunta, and A. Ganapathiraju, "Multitask speech recognition and speaker change detection for unknown number of speakers," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[2] Z. Liu, W. Sun, Y. Hong, D. Teney, and S. Gould, "Bi-directional training for composed image retrieval via text prompt learning," in *IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024.

[3] N. Tarigopula, P. Garg, S. Muralidhar, S. Tornay, D. B. Jayagopi, and M. Magimai-Doss, "Content-based objective evaluation of artificially generated sign language videos," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[4] E. VILLATORO-TELLO, S. Madikeri, B. Sharma, D. Khalil, S. Kumar, N. Iuliia, P. Motlicek, and A. Ganapathiraju, "Probability-aware word-confusion-network-to-text alignment approach for intent classification," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[5] M. Agarwal, S. Agarwal, A. Anastasopoulos, L. Bentivogli, O. Bojar, C. Borg, M. Carpuat, R. Cattoni, M. Cettolo, M. Chen, W. Chen, K. Choukri, A. Chronopoulou, A. Currey, T. Declerck, Q. Dong, K. Duh, Y. Esteve, M. Federico, S. Gahbiche, B. Haddow, B. Hsu, P. M. Htut, H. Inaguma, D. Javorsky, J. Judge, Y. Kano, T. Ko, R. Kumar, P. Li, X. Ma, P. Mathur, E. Matusov, P. McNamee, J. P. McCrae, K. Murray, M. Nadejde, S. Nakamura, M. Negri, H. Nguyen, J. Niehues, X. Niu, A. K. Ojha, J. E. Ortega, P. Pal, J. Pino, L. van der Plas, P. Polak, E. Rippeth, E. Salesky, J. Shi, M. Sperber, S. Stuker, K. Sudoh, Y. Tang, B. Thompson, K. Tran, M. Turchi, A. Waibe, M. Wang, S. Watanabe, and R. Zevallos, "Findings of the IWSLT 2023 evaluation campaign," in *International Conference on Spoken Language Translation (IWSLT)*, 2023.

[6] V. Amiot, O. Jimenez-del Toro, P. Eyraud, Y. Guex-Crosier, C. Bergin, A. Anjos, F. Hoogewoud, and M. Tomasoni, "Fully automatic grading of retinal vasculitis on fluorescein angiography time-lapse from real-world data in clinical settings," in *IEEE International Symposium on Computer-Based Medical Systems*, 2023, pp. 689–693.

[7] T. Apicella, A. Xompero, E. Ragusa, R. Berta, A. Cavallaro, and P. Gastaldo, "Affordance segmentation of hand-occluded containers from exocentric images," in *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2023.

[8] K. Assi, L. B. Meegahapola, W. Droz, P. KUN, A. de Götzen, M. Bidoglia, S. Stares, G. Gaskell, A. Chagnaa, A. Ganbold, T. Zundui, C. Caprini, D. Miorandi, J. L. Zarza, A. Hume, L. Cernuzzi, I. Bison, M. R. Britez, M. Busso, R. Chenu-Abente, F. Giunchiglia, and D. Gatica-Perez, "Complex daily activities, country-level diversity, and smartphone sensing: A study in Denmark, Italy, Mongolia, Paraguay, and UK," in *ACM CHI Conference on Human Factors in Computing Systems (CHI)*, Association for Computing Machinery, 2023.

[9] S. Bagshy, P. Holthaus, G. Beraldo, E. Senft, D. Hernandez Garcia, Z. Han, S. Kumaar Jayaraman, A. Rossi, C. Esterwood, A. Andriella, and P. Pridham, "Towards improved replicability of human studies in human-robot interaction: Recommendations for formalized reporting," in *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, pp. 629–633.

[10] A. E. Baia, V. Poggioni, and A. Cavallaro, "Black-box attacks on image activity prediction and its natural language explanations," in *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2023.

[11] D. Alonso del Barrio and D. Gatica-Perez, "Examining european press coverage of the COVID-19 no-vax movement: An nlp framework," in *ACM International Workshop on Multimedia AI against Disinformation*, 2023.

[12] D. Alonso del Barrio and D. Gatica-Perez, “Framing the news: From human perception to large language model inferences,” in *International Conference on Multimedia Retrieval*, 2023.

[13] M. Behjati, F. Fehr, and J. Henderson, “Learning to abstract with nonparametric variational information bottleneck,” in *Conference on Empirical Methods in Natural Language Processing*, 2023.

[14] M. Bhattacharjee, P. Motlicek, N. Iuliia, H. Helmke, O. Ohneiser, M. Kleinert, and h. Ehr, “Customization of automatic speech recognition engines for rare word detection without costly model re-training,” in *Proceedings of the 13th SESAR Innovation Days*, 2023.

[15] M. Bilardo, J. Kämpf, and E. Fabrizio, “Energy assessment of a district by integrating solar thermal in district heating network: A dynamic analysis approach,” in *Journal of Physics: Conference Series*, vol. 2600, IOP Publishing Ltd, 2023.

[16] R. Boghetti and J. Kämpf, “A benchmark for the simulation of meshed district heating networks based on anonymised monitoring data,” in *Journal of Physics: Conference Series*, vol. 2600, IOP Publishing Ltd, 2023.

[17] R. Boghetti, G. Peronato, and J. Kämpf, “Verification of pydhn - a python library for the thermo-hydraulic simulation of district heating networks - through the destest,” in *Proceedings of Building Simulation 2023: 18th Conference of IBPSA*, ser. Building Simulation, IBPSA, vol. 18, IBPSA, 2023.

[18] S. Burdisso, E. Villatoro-Tello, S. Madikeri, and P. Motlicek, “Node-weighted graph convolutional network for depression detection in transcribed clinical interviews,” in *Interspeech*, 2023.

[19] D. Carvalho, G. Mercatali, Y. Zhang, and A. Freitas, “Learning disentangled representations for natural language definitions,” in *In Findings of the European chapter of Association for Computational Linguistics*, 2023.

[20] S. Chandrasekar, A. Ramesh, T. Purohit, and P. K. Ghosh, “A study on the importance of formant transitions for stop-consonant classification in vcv sequence,” in *Interspeech*, ISCA, 2023.

[21] H. Chen and P. N. Garner, “Diffusion transformer for adaptive text-to-speech,” in *Proceedings of the 12th ISCA Speech Synthesis Workshop (SSW 12)*, 2023.

[22] H. Chen, M. He, L. Coppieters de Gibson, and P. N. Garner, “The idiap speech synthesis system for the blizzard challenge 2023,” in *Proceedings of the 18th Blizzard Challenge Workshop*, 2023.

[23] L. Colbois, H. Otroshi Shahreza, and S. Marcel, “Approximating optimal morphing attacks using template inversion,” in *IEEE International Joint Conference on Biometric*, 2023.

[24] M. Ewerton, M. Villamizar, J. Jankowski, S. Calinon, and J.-M. Odobez, “A multitask and kernel approach for learning to push objects with a task-parameterized deep q-network,” in *IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2023.

[25] T. Gentilhomme, M. Villamizar, J. Corre, and J.-M. Odobez, “Efficient grapevine structure estimation in vineyards conditions,” in *International Conference on Computer Vision (ICCV)*, 2023, pp. 712–720.

[26] A. George, C. Ecabert, H. Otroshi Shahreza, K. Kotwal, and S. Marcel, “Efar 2023: Efficient face recognition competition,” in *IEEE International Joint Conference on Biometrics (IJCB 2023)*, 2023.

[27] A. George and S. Marcel, “Bridging the gap: Heterogeneous face recognition with conditional adaptive instance modulation,” in *IJCB*, 2023.

[28] A. George and S. Marcel, “The unconstrained ear recognition challenge 2023: Maximizing performance and minimizing bias,” in *IEEE International Joint Conference on Biometrics (IJCB 2023)*, 2023.

[29] M. He and P. N. Garner, “Can ChatGPT detect intent? evaluating large language models for spoken language understanding,” in *Interspeech*, 2023.

[30] M. He and P. N. Garner, “The interpreter understands your meaning: End-to-end spoken language understanding aided by speech translation,” in *Findings of EMNLP*, 2023.

[31] H. Helmke, M. Kleinert, N. Ahrenhold, h. Ehr, T. Mühlhausen, O. Ohneiser, P. Motlicek, A. Prasad, Z.-G. Juan, L. Klamert, J. Dokic, and E. Pinska Chauvin, "Automatic speech recognition and understanding for radar label maintenance support increases safety and reduces air traffic controllers' workload," in *Fifteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2023)*, Eurocontrol (Europe), FAA (U.S.), 2023.

[32] J. Henderson and F. Fehr, "A vae for transformers with nonparametric variational information bottleneck," in *The Eleventh International Conference on Learning Representations*, 2023.

[33] J. Henderson, A. Mohammadshahi, A. C. Coman, and L. Miculicich, "Transformers as graph-to-graph models," in *Big Picture Workshop at EMNLP*, 2023.

[34] E. Hermann and M. Magimai-Doss, "Few-shot dysarthric speech recognition with text-to-speech data augmentation," in *Interspeech*, 2023, pp. 156–160.

[35] N. Iuliia, S. Madikeri, E. Villatoro-Tello, P. Motlicek, Z.-G. Juan, K. P. D. S, and A. Ganapathiraju, "Implementing contextual biasing in gpu decoder for online asr," in *Interspeech*, 2023.

[36] J. Jankowski, L. Brudermuller, N. Hawes, and S. Calinon, "Vp-sto: Via-point-based stochastic trajectory optimization for reactive robot behavior," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023.

[37] M. M. Johari, C. Carta, and F. Fleuret, "Eslam: Efficient dense slam system based on hybrid representation of signed distance fields," in *Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[38] Z.-G. Juan, Z. Huang, X. Niu, S. Srinavasan, P. Mathur, B. Thompson, and M. Federico, "End-to-end single-channel speaker-turn aware conversational speech translation," in *The 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, ser. 1, vol. 1, 2023.

[39] Z.-G. Juan, A. Prasad, N. Iuliia, S. S. Sarfjoo, P. Motlicek, M. Kleinert, H. Helmke, O. Ohneiser, and Q. Zhan, "How does pre-trained wav2vec 2.0 perform on domain-shifted asr? an extensive benchmark on air traffic control communications," in *2023 IEEE Spoken Language Technology Workshop (SLT)*, ser. 1, IEEE, vol. 1, 2023.

[40] Z.-G. Juan, A. Sara, V. Danielius, and S. Cem, "Commonaccent: Exploring large acoustic pretrained models for accent classification based on common voice," in *Interspeech*, 2023.

[41] Z.-G. Juan, S. S. Sarfjoo, A. Prasad, N. Iuliia, P. Motlicek, K. Ondřej, and O. Ohneiser, "Bertraffic: Bert-based joint speaker role and speaker change detection for air traffic control communications," in *2023 IEEE Spoken Language Technology Workshop (SLT)*, ser. 1, IEEE, vol. 1, 2023.

[42] M. Jullien, M. Valentino, H. Frost, P. O'Reagan, D. Landers, and A. Freitas, "Nli4ct: Multi-evidence natural language inference for clinical trial reports," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)*, 2023.

[43] M. Jullien, M. Valentino, H. Frost, P. O'Reagan, D. Landers, and A. Freitas, "Semeval-2023 task 7: Multi-evidence natural language inference for clinical trial data," in *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)*, Association for Computational Linguistics, 2023.

[44] N. Kammoun, L. B. Meegahapola, and D. Gatica-Perez, "Understanding the social context of eating with multimodal smartphone sensing: The role of country diversity," in *25th ACM International Conference on Multimodal Interaction*, 2023.

[45] T. Kew, A. Chi, L. Vásquez-Rodríguez, S. Agrawal, D. Aumiller, F. Alva-Manchego, and M. Shardlow, "Bless: Benchmarking large language models on sentence simplification," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2023.

[46] H. Kim and D. Gatica-Perez, "Referencing in youtube knowledge communication videos," in *ACM International Conference on Interactive Media Experiences (IMX '23), June 2023, Nantes, France*, 2023.

[47] A. Komaty, V. Krivokuca, C. Ecabert, and S. Marcel, “Can personalised hygienic masks be used to attack face recognition systems?” In *Proceedings of IEEE International Joint Conference on Biometrics (IJCB2023)*, 2023.

[48] A. Konstant, N. Orr, M. Hagenow, E. Senft, I. Gundrum, B. Mutlu, M. Zinn, M. Gleicher, and R. Radwin, “Human-robot collaboration in a sanding task,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2023.

[49] P. Korshunov, H. Chen, P. N. Garner, and S. Marcel, “Vulnerability of automatic identity recognition to audio-visual deepfakes,” in *IEEE International Joint Conference on Biometrics*, 2023.

[50] C. Layfield, R. Zandbergen, L. F. Davis, J. Abela, C. Bowern, M. Rosner, and L. van der Plas, “International conference on the Voynich manuscript 2022,” in *Proceedings of the International Conference on Historical Cryptology*, 2023.

[51] Y. Li, Y. Zhang, A. Razmjoo, and S. Calinon, “Representing robot geometry as distance fields: Applications to whole-body manipulation,” in *Idiap-Internal-RR-39-2023*, 2023.

[52] J. Liu, Z. Li, S. Calinon, and F. Chen, “Softgpt: Learn goal-oriented soft object manipulation skills by generative pre-trained heterogeneous graph transformer,” in *IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2023.

[53] F. Mai, Z.-G. Juan, T. Parcollet, and P. Motlicek, “Hyperconformer: Multi-head hypermixer for efficient speech recognition,” in *Interspeech*, ser. 1, vol. 1, 2023.

[54] F. Mai, A. Pannatier, F. Fehr, H. Chen, F. Marelli, F. Fleuret, and J. Henderson, “Hypermixer: An mlp-based low cost alternative to transformers,” in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, 2023, pp. 15 632–15 654.

[55] L. B. Meegahapola, M. Constantinides, Z. Radivojevic, H. Li, D. Quercia, and M. S. Eggleston, “Quantified canine: Inferring dog personality from wearables,” in *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI) 2023*, Association for Computing Machinery, 2023.

[56] P. Melzi, H. Otroshi Shahreza, C. Rathgeb, R. Tolosana, R. Vera-Rodriguez, J. Fierrez, S. Marcel, and C. Busch, “Multi-ive: Privacy enhancement of multiple soft-biometrics in face embeddings,” in *IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023.

[57] A. Mohammadshahi and J. Henderson, “Syntax-aware graph-to-graph transformer for semantic role labelling,” in *Procceedings of 8th Workshop on Representation Learning for NLP*, 2023.

[58] A. Mohammadshahi, T. Scialom, M. Yazdani, P. Yanki, A. Fan, J. Henderson, and M. Saeidi, “Rquge: Reference-free metric for evaluating question generation by answering the question,” in *Association for Computational Linguistics*, 2023.

[59] A. Montazeri, J. Kämpf, and G. Mutani, “Data-driven urban building energy modeling with machine learning in satom (ch),” in *6th International IEEE Conference AND Workshop in Óbuda on Electrical and Power Engineering*, 2023.

[60] A. Nanchen, L. B. Meegahapola, W. Droz, and D. Gatica-Perez, “Keep sensors in check: Disentangling country-level generalization issues in mobile sensor-based models with diversity scores,” in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2023.

[61] B. Nemec, M. Hrovat, M. Simonič, S. Shetty, S. Calinon, and A. Ude, “Robust execution of assembly policies using a pose invariant task representation,” in *20th International Conference on Ubiquitous Robots (UR)*, IEEE, 2023.

[62] A. M. Nicolicioiu, A. L. Nicolicioiu, B. Alexe, and D. Teney, “Learning diverse features in vision transformers for improved generalization,” in *ICML Workshop on Spurious Correlations, Invariance and Stability*, 2023.

[63] H. Otroshi Shahreza, A. Bassit, S. Marcel, and R. Veldhuis, “Remote cancelable biometric system for verification and identification applications,” in *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2023.

[64] H. Otroshi Shahreza, A. George, and S. Marcel, “Synthdistill: Face recognition with knowledge distillation from synthetic data,” in *IEEE International Joint Conference on Biometrics*, 2023.

[65] H. Otroshi Shahreza, V. Krivokuca, and S. Marcel, “Mlp-hash: Protecting face templates via hashing of randomized multi-layer perceptron,” in *Proceedings of the European Signal Processing Conference*, 2023.

[66] H. Otroshi Shahreza and S. Marcel, “Blackbox face reconstruction from deep facial embeddings using a different face recognition model,” in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2023.

[67] H. Otroshi Shahreza and S. Marcel, “Face reconstruction from facial templates by learning latent space of a generator network,” in *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

[68] H. Otroshi Shahreza and S. Marcel, “Inversion of deep facial templates using synthetic data,” in *Proceedings of the IEEE International Joint Conference on Biometric*, 2023.

[69] H. Otroshi Shahreza and S. Marcel, “Template inversion attack against face recognition systems using 3d face reconstruction,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023, pp. 19 662–19 672.

[70] M. Papinutto, M. Colombo, R. Boghetti, C. Basurto, K. Reutter, D. Lalanne, J. Kämpf, and J. Nembrini, “Enhancing user acceptance in automated systems with human-centric lighting: The role of visual comfort, personality, and preference,” in *Journal of Physics: Conference Series*, vol. 2600, IOP Publishing Ltd, 2023.

[71] G. Peronato and J. Kämpf, “Potential for district heating networks from waste heat: An assessment tool and its application to sewage treatment plants in the canton of zurich,” in *Journal of Physics: Conference Series*, vol. 2600, IOP Publishing Ltd, 2023.

[72] M. R. Petersen and L. van der Plas, “Can language models learn analogical reasoning? investigating training objectives and comparisons to human performance,” in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 2023.

[73] F. Piras, F. Marelli, E. De Moura Presa, P. Wellig, and M. Liebling, “Multi-image deconvolution of thermal images with a boundary condition weighting scheme,” in *Target and Background Signatures IX*, International Society for Optics and Photonics, vol. 12736, SPIE, 2023, pp. 149–158.

[74] T. Piton, E. Hermann, A. Pasqualotto, M. Cohen, M. Magimai-Doss, and D. Bavelier, “Using commercial asr solutions to assess reading skills in children: A case report,” in *Interspeech*, 2023, pp. 4573–4577.

[75] T. Purohit, B. Vlasenko, and M. Magimai-Doss, “Implicit phonetic information modeling for speech emotion recognition,” in *Interspeech*, ISCA, 2023.

[76] T. Purohit, S. Yadav, B. Vlasenko, S. P. Dubagunta, and M. Magimai-Doss, “Towards learning emotion information from short segments of speech,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, IEEE, 2023.

[77] P. Rahimi, C. Ecabert, and S. Marcel, “Toward responsible face datasets: Modeling the distribution of a disentangled latent space for sampling face images from demographic groups,” in *IEEE International Joint Conference on Biometrics*, 2023.

[78] M. Ramirez, A. Bogatu, N. Paton, and A. Freitas, “Transformers, tables and frame semantics,” in *International Conference on Semantic Computing*, 2023.

[79] A. Razmjoo, T. Brecelj, K. Savevska, A. Ude, T. Petric, and S. Calinon, “Learning joint space reference manifold for reliable physical assistance,” in *IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2023.

[80] J. Rozanova, M. Valentino, L. Cordeiro, and A. Freitas, “On interventional probing in high dimensions: An nli case study,” in *Findings of the 17th European Chapter of the Association for Computational Linguistics*, 2023.

[81] S. Sajadmanesh, A. S. Shamsabadi, A. Bellet, and D. Gatica-Perez, “Gap: Differentially private graph neural networks with aggregation perturbation,” in *32nd USENIX Security Symposium (USENIX Security 23)*, 2023.

[82] E. Sarkar and M. Magimai-Doss, “Can self-supervised neural representations pre-trained on human speech distinguish animal callers?” In *Interspeech*, 2023.

[83] G. Schu, P. Janbakhshi, and I. Kodrasi, “On using the UA-Speech and TORGO databases to validate automatic dysarthric speech classification approaches,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2023.

[84] L. Scimeca, A. Rubinstein, A. M. Nicolicioiu, D. Teney, and Y. Bengio, “Leveraging diffusion disentangled representations to mitigate shortcuts in underspecified visual tasks,” arXiv preprint, Under review, 2023.

[85] L. Stegner, E. Senft, and B. Mutlu, “Situated participatory design: A method for in situ design of robotic interaction with older adults,” in *CHI ’23: Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2023.

[86] S. Tafasca, A. Gupta, N. Kojovic, M. Gelsomini, T. Maillart, M. Papandrea, M. Schaer, and J.-M. Odobez, “The ai4autism project: A multimodal and interdisciplinary approach to autism diagnosis and stratification,” in *Companion Publication of the 25th International Conference on Multimodal Interaction*, ser. ICMI ’23 Companion, Association for Computing Machinery, 2023, 414–425.

[87] S. Tafasca, A. Gupta, and J.-M. Odobez, “Childplay: A new benchmark for understanding children’s gaze behaviour,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023.

[88] D. Teney, Y. Lin, S. J. Oh, and E. Abbasnejad, “ID and OOD performance are sometimes inversely correlated on real-world datasets,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

[89] M. Tognoli, G. Peronato, and J. Kämpf, “A machine learning model for the prediction of building hourly heating demand from citygml files: Training workflow and deployment as an api,” in *Proceedings of Building Simulation 2023: 18th Conference of IBPSA*, 2023, pp. 2932 –2939.

[90] D. Vallejo-Aldana, A. P. López-Monroy, and E. Villatoro-Tello, “Enhancing multi-modal classification of violent events using image captioning,” in *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2023), CEUR Workshop Proceedings*, vol. 3496, 2023.

[91] G. Vanderreydt, A. Prasad, D. Khalil, S. Madikeri, K. Demuynck, and P. Motlicek, “Parameter-efficient tuning with adaptive bottlenecks for automatic speech recognition,” in *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, 2023.

[92] L. Vásquez-Rodríguez, M. Shardlow, P. Przybyla, and S. Ananiadou, “Document-level text simplification with coherence evaluation,” in *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, 2023.

[93] E. VILLATORO-TELLO, S. Madikeri, Z.-G. Juan, B. Sharma, S. S. Sarfjoo, N. Iuliia, P. Motlicek, A. V. Ivanov, and A. Ganapathiraju, “Effectiveness of text, acoustic, and lattice-based representations in spoken language understanding tasks,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2023.

[94] T. Xue, H. Girgin, T. S. Lembono, and S. Calinon, “Demonstration-guided optimal control for long-term non-prehensile planar manipulation,” in *IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2023.

5.5 Patents

- Granted patents

IDIAP-1 [US 9,689,959] A. Asaei, H. Bourlard, V. Cevher, “Method, apparatus and computer program product for determining the location of a plurality of speech sources”

IDIAP-2 [US 9,058,541] C. Dubout, F. Fleuret, “Object detection method, object detector and object detection computer program”

IDIAP-5 [EP3154407] K. A. Funes Mora, J-M. Odobez, “A gaze estimation method and apparatus”

IDIAP-8 [US 9,973,503] S. Marcel, A. Anjos, P. Abbet, “Method and internet-connected server for reviewing a computer-executable experiment”

IDIAP-10 [EP 3 691 258] M. Liebling, C. Jacques, “System and method for acquiring images”

IDIAP-11 [US 11,514,918] | [EP 3 719 679 B1] S. Marcel, V. Krivokuca, “A method for protecting biometric templates, and a system and method for verifying a speaker’s identity”

- Submitted applications

globID-1 [WO 2019/150254 A1] L. Sonna Momo, L. Cerqueira Torres, S. Marcel, A. Anjos, M. Liebling, A. Shajkofci, S. Amoos, A. Woeffray, A. Sierro, “Method and device for biometric vascular recognition and/or identification”

IDIAP-USC-1 [US 2022 0398820 A1] W. Abd-Almageed, L. Spinoulas, M. E. Hussein, D. Geissbühler, S. Marcel, “Multispectral biometrics system”

IDIAP-D-10 [EP 4250249A1, US 2023 306732A1] A. George, S. Marcel, “Heterogenous face recognition system and method”

idiap
RESEARCH INSTITUTE