# Health-MM 2026

## Health-MM 2026 WORKSHOP

## in Conjunction with IEEE MIPR 2026
## Bankok, Thailand

## August 9 – 11, 2026

## ABSTRACT BOOKLET

# Table of Contents

# Explainable Fundus-Based AI Glaucoma Screening: Design and Evaluation of a Healthcare System

Allan Pascual Erasmo (Yoobee Colleges)
Jaya Tiwari (Waikato Hospital)
Samaneh Chaharmiri Dokhaharani (Mae Fah Luang University)
Mohammad Norouzifard (The University of Auckland)

Glaucoma is a major cause of irreversible vision loss, and early detection remains challenging in community settings where glaucoma specialist assessment may not be readily available. This study presents a prototype fundus-based AI screening system for glaucoma detection that combines convolutional neural network (CNN) ensemble inference, Explainable AI (XAI) through Grad-CAM++ visual explanation, structured screening records, and referral-support documentation. The system uses fundus photography as a lower-cost input modality and generates structured screening support rather than providing autonomous diagnosis. Preliminary evaluation on a locked 415-image test set produced an ensemble AUC of 0.93, sensitivity of 0.85, specificity of 0.80, and negative predictive value of 0.93 at the sensitivity-first operating point. The contribution is an explainable AI-based system that links retinal image analysis to practical glaucoma screening support. An explainable fundus-based AI screening system could support earlier glaucoma detection in settings where OCT imaging or ophthalmologist assessment is not readily available, while final interpretation and decisions remain part of normal clinical review.

**Keywords:** Glaucoma screening, fundus photography, deep learning, CNN ensemble, explainable AI, Grad-CAM++, referral support.

# Ethical Implications of AI-Assisted Digital Phenotyping in Clinical Psychiatry

Deepa Shukla (IIT Jodhpur)
Nisha Daga (IIT Jodhpur)
George Kodimattam Joseph (IIT Jodhpur)

The present paper explores how AI-assisted digital phenotyping (DP) reshapes the landscape of clinical psychiatry by reassigning authority over collection and interpretation of mental health data, and interventions informed by such data. Digital phenotyping is an approach through which multimodal data streams and digital footprints of persons are collected and processed for predictive and preventive care. While the assessment in the traditional setting explicitly involved two major stakeholders, i.e., specialists and patients, the integration of artificial intelligence (AI) in the field has altered this closed relationship. This, in return, results in power asymmetry, where most of the power over the data gets transferred to AI systems, and patients become mere objects of observation and specialists lose their authority over data interpretation and clinical judgment. The paper discusses three major dimensions of authority that are challenged by this power asymmetry and proposes technical and ethical measures to safeguard legitimate human control and ownership over the collected data.

**Keywords:** Psychiatry, Multimodal, digital phenotyping, power asymmetry, authority, ethics.

## RadLens: Radiology Report Generation Using a Vision Transformer and Retrieval-Augmented Generation

**Philip Niron Nithianandan (Yoobee College of Creative Innovation)**
**Jaya Tiwari (Waikato Hospital)**
**Seyedhashem Davarpanah (Agentic AI Hub)**
**Mohammad Norouzifard (The University of Auckland)**

Automated radiology report generation systems usually produce single-section free-text reports and often lack radiologist review workflows, limiting their clinical use. This study proposed a RadLens system (Radiology Report Generation Using a Vision Transformer and Retrieval-Augmented Generation), a five-stage AI-assisted pipeline for generating structured draft reports from chest X-ray images as a case study. Each report contains five fields: clinical summary, key findings, red flags, follow-up reminders, and missing information. Three pre-trained vision transformer encoders, Bidirectional Encoder representation from Image Transformers (BEiT), Swin Transformer, and Vision Transformer (ViT), were fine-tuned on the public Indiana University Chest X-Ray dataset using contrastive loss. The fine-tuned encoder features were then used with Claude Sonnet 4.6 via Amazon Bedrock to generate structured reports. Claude Sonnet 4.6 was used with prompt engineering and Retrieval-Augmented Generation (RAG). For each test image, RAG retrieved the top three visually similar training cases using cosine similarity. RAG improved BLEU-1 scores from 0.1034 to 0.2105 for BEiT (104%), from 0.1130 to 0.2061 for Swin Transformer (82%), and from 0.0955 to 0.2140 for ViT (124%). BERTScore F1 increased from 0.8439, 0.8436, and 0.8435 to 0.8701, 0.8690, and 0.8708, respectively, suggesting improved semantic similarity. RadLens also includes a radiologist review interface for reviewing, editing, approving, or rejecting generated drafts before clinical use.

**Keywords:** Radiology report generation, vision transformer, retrieval-augmented generation, vision-language models, multimodal learning, structured output, BEiT, contrastive learning.

## Neuro-Adaptive Virtual Reality for Spider Phobia: Integrating EEG and Heart Rate Biofeedback in Closed-Loop Exposure Therapy

**Rahul Kumar Ray (FLAME University)**
**Rudra Krishna (FLAME University)**
**CHIRANJOY CHATTOPADHYAY (FLAME University)**

Virtual reality (VR) exposure therapy is increasingly recognized as an effective intervention for the treatment of specific phobias, as it enables immersive and precisely controllable environments for the graded confrontation of feared stimuli in the domain of healthcare through multimedia analytics. Arachnophobia, characterized by an intense and often debilitating fear of spiders, is a prevalent anxiety disorder that can particularly benefit from technologically mediated approaches. In this pape4r, we developed a closed-loop VR system for exposure therapy, utilizing a MAX30102 heart rate sensor and an ESP32 microcontroller to gather real-time heart rate and heart rate variability (HRV) data from users. The setup was implemented and tested using Meta Quest 3, with wireless data transmission between the hardware components. Within this framework, the VR environment is dynamically modulated as a function of the user's heart rate, allowing real-time adjustment of the number of spiders and their approach speed to individualize the intensity of stress exposure. Preliminary pilot data indicate promising feasibility and acceptability of the system. In the future, we plan to extend this closed-loop paradigm by integrating an EEG cap to continuously monitor neural markers of stress and thus enable a more comprehensive psychophysiological feedback control.

**Keywords:** Arachnophobia, Virtual Reality Exposure Therapy, EEG, Neuro-Adaptive Systems, Heart Rate Monitoring.

# Organizers

Prof. Gaurav Bhatnagar
Indian Institute of Technology Jodhpur
India
Email: goravb@iitj.ac.in

Prof. Subrahmanyam Murala
Trinity College Dublin
Ireland

Prof. Jonathan Wu
University of Windsor
Canada

Prof. Zheng Liu
University of British Columbia, Okanagan
Canada

# Partner

TIH
iHub Drishti