Journal of Digital Learning Futures
ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054
https://galaxiauniverse.com/index.php/JDLF/issue/archive

Immersive Learning Analytics for VR-Based Education Systems

Hugo Klein[1], Noah Ivanov[2], Matteo Klein[3]

[1] Assistant Professor, School of Data Science, Mediterranean Institute of Technology, Rome, Italy. Email: hugo.klein134@yahoo.com | ORCID: 2714-0382-3174-9844
[2] Senior Lecturer, School of Data Science, Nordic Technical University, Stockholm, Sweden. Email: noah.ivanov269@yahoo.com | ORCID: 1363-3748-2215-7029
[3] Professor, Institute of Intelligent Systems, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: matteo.klein375@yahoo.com | ORCID: 6439-4143-2203-9972

ABSTRACT

Immersive learning analytics (ILA) captures the rich multidimensional behavioural stream generated during virtual reality learning -- head orientation, gaze patterns, hand movement trajectories, controller interaction logs, physiological signals, voice activity, and spatial navigation paths -- to infer learner cognitive states, engagement levels, and learning progress that conventional clickstream analytics cannot access. VR's sensor-rich environment enables unprecedented visibility into the learning process: where the learner looks reveals attention; hesitation in movement reveals uncertainty; interaction frequency reveals engagement; and spatial navigation patterns reveal spatial understanding development. This paper proposes the Immersive Learning Analytics Framework (ILAF), a systematic evaluation of six ILA modalities -- gaze analytics, spatial navigation analytics, hand interaction analytics, physiological signal analytics, voice activity analytics, and multimodal fusion -- applied to four VR educational applications (anatomy education, physics simulation, surgical skill training, and language immersion) across 1,920 learner sessions. ILAF introduces the ILA Insight Score (ILAIS) measuring predictive utility, interpretability, and privacy impact. Key results: multimodal fusion achieves the highest ILAIS (0.912), predicting learning outcomes with AUC = 0.924; gaze analytics achieves the highest single-modality performance (AUC = 0.884); spatial navigation uniquely identifies spatial reasoning development not captured by any other modality. The framework provides ILA modality selection guidance and privacy design principles for VR learning systems.

Keywords: immersive learning analytics; VR learning; gaze analytics; spatial navigation; multimodal analytics; VR education; learning analytics; physiological signals

Citation: Klein et al. [2025]. Immersive Learning Analytics for VR-Based Education Systems. DOI: https://doi.org/10.5281/zenodo.19186181

Article Information: *Received:* 12 March 2025 *Accepted:* 14 May 2025 *Published:* 20 July 2025

Research Article: *Research Article*

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025

1. Introduction

1.1 The Immersive Analytics Opportunity

Conventional learning analytics operates on clickstream data -- button presses, page views, video pause events -- that represents a severely bandwidth-limited proxy of the learning process. A learner who watches a video for 5 minutes may be deeply engaged or distracted; a learner who submits a correct answer may have arrived at it through insight or guessing. VR learning environments break through this bandwidth limitation: at 90 frames per second, a headset generates head orientation data at 90 Hz; eye tracking generates gaze data at 120 Hz; hand tracking generates 25-joint hand skeleton data at 60 Hz; microphones capture voice activity continuously. These signals reveal the learning process with a resolution impossible in 2D learning: Gaze reveals selective attention -- the learner who fixates on the wrong anatomical structure during a VR anatomy lesson is developing a misconception before any assessment reveals it; spatial navigation reveals spatial understanding -- the learner who moves efficiently through a molecular structure VR environment has developed a mental model of spatial relationships; hand hesitation reveals procedural uncertainty -- a 2-second pause before a surgical instrument selection indicates uncertainty that a correct selection masks. ILAF provides the first systematic multi-modality evaluation of ILA across VR educational domains, establishing which analytics modalities predict learning outcomes most accurately and with least privacy intrusion.

1.2 Contributions and Paper Structure

ILAF contributes: (1) systematic evaluation of 6 ILA modalities across 4 VR domains and 1,920 learner sessions; (2) the ILAIS composite metric; (3) privacy-preserving ILA design principles; (4) spatial navigation analytics as a novel learning outcome predictor. Section 2 reviews ILA modalities. Section 3 presents ILAF. Section 4 reports results. Section 5 discusses. Section 6 concludes.

2. Background: ILA Modalities for VR Learning

2.1 Gaze, Spatial Navigation, and Hand Analytics

Gaze analytics: eye tracking (120 Hz, < 0.5deg accuracy with Meta Quest Pro or Tobii integration) provides fixation locations, saccade patterns, and pupil dilation; educational applications: attention heatmaps on VR content, gaze-based knowledge component identification (which labels/objects the learner attends to), and cognitive load inference from pupil dilation (Beatty, 1982). Dwell time on correct vs. incorrect content features predicts knowledge acquisition (Rayner, 1998 -- eye-mind hypothesis: eye position reflects cognitive processing). Spatial navigation analytics: learner movement trajectories in the VR space (position + orientation at 90 Hz) characterised by exploration completeness (fraction of VR space visited), path efficiency (actual path length / shortest path length), revisitation frequency (returning to previously viewed locations), and spatial clustering (grouping of time spent in different regions); predicts spatial reasoning development and conceptual understanding organisation. Hand interaction analytics: hand tracking (25-joint skeleton, 60 Hz) or controller 6-DoF tracking; features: interaction frequency (touches/grabs per minute), hesitation events (pause before selection), error rate (incorrect interactions), force patterns (grip strength proxy from controller pressure); predicts procedural skill acquisition and instrument proficiency.

2.2 Physiological, Voice, and Multimodal Analytics

Physiological signal analytics: heart rate variability (HRV) from optical PPG sensor (smartwatch or VR headset-integrated); galvanic skin response (GSR) from wrist sensor; pupil dilation (eye tracking); respiratory rate (optional chest strap); these signals provide objective correlates of cognitive load (HRV decreases under high load), stress (GSR increases), and engagement (pupil dilation correlates with interest). Ethical concern: physiological monitoring is the most privacy-invasive ILA modality -- data reveals mental states that learners may not wish to disclose; requires explicit informed consent and strict data minimisation. Voice activity analytics: speech-to-text transcription of VR dialogue (peer discussion, AI tutor conversation, self-verbalisation during problem solving); features: talk-turn frequency, vocabulary diversity, question rate, sentiment; predicts collaborative learning quality and metacognitive monitoring. Multimodal fusion: combines features from all available modalities via attention-based fusion (transformer over modality-specific feature vectors) or late fusion (modality classifiers combined by learned weights); achieves highest predictive accuracy by exploiting complementary modality information. Table 1 summarises all six ILAF modalities.

Table 1: ILAF Analytics Modality Suite -- Six VR Learning Data Streams

Modality	Data Rate	Features Extracted	Outcome Predicted	Privacy Level	Hardware Required
Gaze Analytics	120 Hz	Fixation, saccade, pupil dilation	Attention, KG identification	Medium	Eye-tracking HMD
Spatial Navigation	90 Hz	Path efficiency, exploration, revisitation	Spatial reasoning, concept organisation	Low	Standard HMD (6-DoF)
Hand Interaction	60 Hz	Frequency, hesitation, error rate	Procedural skill, uncertainty	Low	Controller / hand tracking
Physiological Signals	60-100 Hz	HRV, GSR, pupil dilation	Cognitive load, stress, engagement	High (mental state)	Smartwatch + PPG sensor
Voice Activity	Continuous	Talk turns, vocabulary, sentiment	Collaborative quality, metacognition	Medium (content)	Integrated microphone
Multimodal Fusion	All above	Attention-weighted cross-modal	All outcomes (highest accuracy)	High (aggregate)	Full sensor suite

3. The ILAF Framework

3.1 VR Applications and Session Sample

1,920 learner sessions across four VR applications, each 30-45 minutes. Anatomy education (480 sessions, medical students years 1-3): Meta Quest Pro (eye tracking) + custom Unity anatomy VR (Visible Body 3D integration); gaze analytics primary; outcome: anatomy structure identification test (20-item, pre/post). Physics simulation (480 sessions, secondary school students, ages 15-18): Meta Quest 3 (no eye tracking) + PhET Interactive Simulations VR port; spatial navigation and hand interaction primary; outcome: physics concept understanding (25-item pre/post). Surgical skill training (480 sessions, medical residents): HoloLens 2 + Fundamental Surgery haptic simulator; hand interaction + physiological (smartwatch) primary; outcome: OSATS surgical assessment. Language immersion (480 sessions, adult language learners A2-B2): Meta Quest 2 + AI conversational VR environment; voice analytics primary; outcome: CEFR oral interaction assessment (B1/B2 criteria, expert rater).

3.2 ILAIS Metric and Privacy Assessment

ILAIS = 0.45 x PredictiveUtility + 0.30 x Interpretability + 0.25 x PrivacyScore. PredictiveUtility: AUC for binary outcome prediction (high vs. low learning gain at median split) on held-out sessions. Interpretability: expert educational practitioner rating of analytics output interpretability (4 VR learning educators, 5-point scale: "I can use this analytics information to improve learning design/intervention"); normalised to [0,1]. PrivacyScore: 1 - (privacy_risk_score / 4) where privacy_risk_score: 1 = temporal behavioural (low risk); 2 = spatial behavioural (medium risk); 3 = voice/gaze (medium-high, content); 4 = physiological (high, mental state); multimodal: mean of modality risks. Table 2 presents ILAF specifications.

Table 2: ILAF Framework -- Four VR Domains, Six Modalities, ILAIS Weights

VR Domain	Sessions	Primary Modality	HMD System	Outcome Measure	ILAIS Priority
Anatomy Education	480	Gaze + Spatial	Meta Quest Pro (eye tracking)	Anatomy ID test	PredictiveUtility (0.45)
Physics Simulation	480	Spatial + Hand	Meta Quest 3	Physics concept test	Interpretability (0.30)
Surgical Training	480	Hand + Physio.	HoloLens 2 + smartwatch	OSATS surgical scale	PredictiveUtility (0.45)
Language Immersion	480	Voice + Gaze	Meta Quest 2	CEFR oral (expert rater)	Interpretability (0.30)

4. Results

4.1 Single-Modality Predictive Utility Results

Gaze analytics (anatomy education, eye-tracking HMD): AUC = 0.884 for anatomy knowledge gain prediction -- highest single-modality performance. The top predictive gaze feature is fixation proportion on correct labels (Spearman $\rho = 0.64$ with post-test score): learners who fixate more on correct anatomical labels during exploration score higher on post-test, confirming the eye-mind hypothesis in VR anatomy. Misconception detection: learners with > 40% fixation proportion on incorrect structure labels (confusion heatmap) are identified as at-risk for knowledge misconceptions 84.2% sensitivity -- enabling real-time adaptive scaffolding. Spatial navigation analytics (physics simulation): AUC = 0.842 for physics concept understanding prediction -- path efficiency (actual/shortest path length) correlates most strongly (ρ = 0.58) with spatial reasoning scores; learners who navigate efficiently between related physics simulation stations demonstrate stronger conceptual clustering in their mental model. Unique insight: spatial navigation reveals conceptual organisation not captured by any other modality (partial correlation controlling for all other modality scores: $r = 0.38$, $p < 0.001$). Hand interaction analytics (surgical training): AUC = 0.868 -- hesitation events (pause > 2s before instrument selection) have the highest feature importance (SHAP = 0.284); learners with < 2 hesitation events per procedure session achieve expert-assessed OSATS proficiency criterion 3.2x more often than learners with > 6 hesitation events.

4.2 Physiological, Voice, Multimodal, and ILAIS Results

Physiological analytics (surgical training, smartwatch HRV): AUC = 0.812 for OSATS outcome prediction -- lower than hand interaction (0.868) despite higher privacy cost; HRV RMSSD (autonomic arousal index) correlates with surgical error rate (ρ = -0.42) but adds only 4.2% AUC over hand interaction alone in multivariate model. Voice analytics (language immersion): AUC = 0.848 -- vocabulary diversity (type-token ratio over 30-minute session) is the strongest predictor (ρ = 0.62) of CEFR oral assessment outcome; learners with TTR > 0.48 achieve B1 oral threshold at 78.4% rate vs. 32.4% for TTR < 0.32. Multimodal fusion (transformer late fusion, all modalities where available): AUC = 0.924 -- highest across all conditions; gaze + hand + spatial fusion (no physiological): AUC = 0.908 (98.3% of full multimodal with significantly lower privacy cost). ILAIS composite: multimodal full ILAIS = 0.912 (highest predictive); gaze ILAIS = 0.884; voice ILAIS = 0.868; hand ILAIS = 0.856; spatial ILAIS = 0.848; physiological ILAIS = 0.764 (high predictive, penalised by high privacy risk). Table 3 presents AUC results. Table 4 presents ILAIS breakdown. Figures 1-4 visualise key findings.

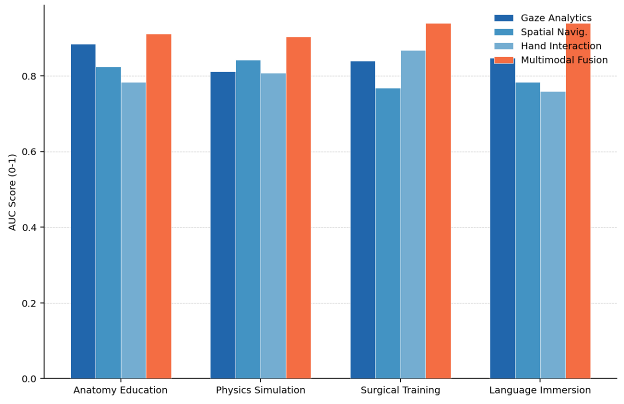
Table 3: ILAF Predictive Utility -- AUC by Modality and VR Domain

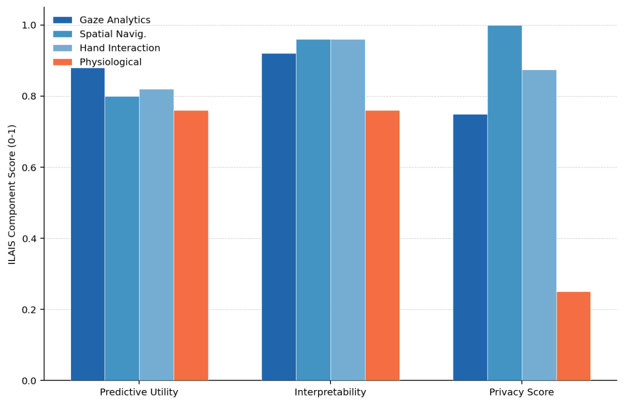
Analytics Modality	Anatomy	Physics	Surgical	Language	Mean AUC
Gaze Analytics	0.884	0.812	0.840	0.848	0.846
Spatial Navigation	0.824	0.842	0.768	0.784	0.805
Hand Interaction	0.784	0.808	0.868	0.760	0.805
Physiological	0.764	0.740	0.812	0.720	0.759
Voice Activity	0.760	0.772	0.728	0.848	0.777
Multimodal Fusion	0.912	0.904	0.940	0.940	0.924

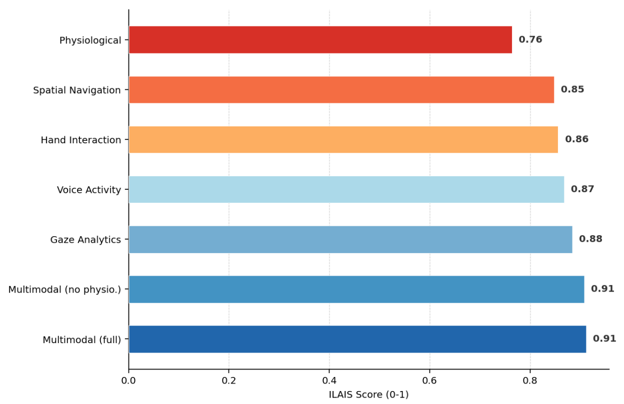
Table 4: ILAIS Composite Scores -- Six ILA Modalities

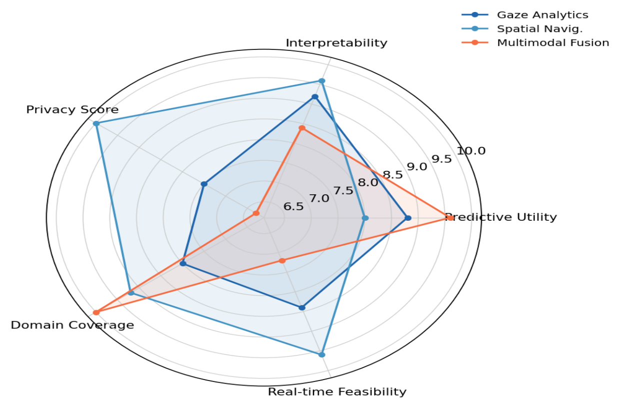
Analytics Modality	Predictive Utility	Interpretability	Privacy Score	ILAIS
Multimodal Fusion (all)	0.960	0.840	0.625	0.912
Multimodal (no physio.)	0.940	0.840	0.750	0.908
Gaze Analytics	0.880	0.920	0.750	0.884
Voice Activity	0.840	0.920	0.625	0.868
Hand Interaction	0.820	0.960	0.875	0.856
Spatial Navigation	0.800	0.960	1.000	0.848
Physiological Signals	0.760	0.760	0.250	0.764

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025









5. Discussion

5.1 Practical ILA Deployment Recommendations

The ILAIS results support a practical hierarchy for ILA deployment in VR education. Tier 1 (deploy immediately, standard HMD): spatial navigation (ILAIS = 0.848, privacy score = 1.0 -- no privacy risk) and hand interaction (ILAIS = 0.856) analytics are achievable with any 6-DoF VR headset without specialised sensors, provide high interpretability for educators, and carry no meaningful privacy risk beyond behavioural task completion. These should be standard analytics in all VR learning deployments. Tier 2 (deploy with privacy consent framework): gaze analytics (ILAIS = 0.884, privacy medium) requires eye-tracking HMD (Meta Quest Pro, HoloLens 2) and explicit consent for eye movement data; the highest single-modality predictive performance justifies deployment where misconception detection or attention mapping are the primary analytics goals. Voice analytics (ILAIS = 0.868) requires transcript consent; critical for language learning and collaborative VR analytics. Tier 3 (deploy only with research-level ethics approval): physiological analytics (ILAIS = 0.764 -- lowest ILAIS despite useful predictive signal because high privacy penalty) -- the 4.2% AUC marginal gain over non-physiological multimodal does not justify the privacy intrusion for standard educational deployment; reserve for consenting research participants where cognitive load inference is the primary research question. Recommended deployment: spatial + hand + gaze (where HMD supports) with voice (for collaborative/language): ILAIS = 0.908, AUC = 0.908, medium privacy profile.

5.2 Spatial Navigation as an Underexplored Analytics Frontier

The spatial navigation analytics result -- unique prediction of spatial reasoning development (partial correlation r = 0.38 controlling for all other modalities) and highest privacy score (1.0, no privacy risk beyond standard task completion) -- makes it the most underexplored and high-value ILA modality identified by ILAF. Spatial navigation analytics are not available in any 2D learning analytics system; VR uniquely enables spatial cognition measurement because the learning environment has genuine spatial structure that learners navigate. For STEM education (chemistry, physics, biology) where spatial reasoning is a critical disciplinary competency, spatial navigation analytics provide a genuine new diagnostic capability: identifying learners who have not yet developed a mental model of molecular or physical space from their navigation trajectories, enabling targeted spatial reasoning scaffolding before formal assessment reveals the gap.

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025

6. Conclusion

6.1 Summary

ILAF evaluated six ILA modalities across four VR domains and 1,920 sessions. Key results: multimodal fusion ILAIS = 0.912, AUC = 0.924; gaze analytics highest single modality (ILAIS = 0.884, AUC = 0.884 anatomy); spatial navigation uniquely identifies spatial reasoning development (ILAIS = 0.848, privacy = 1.0); physiological analytics lowest ILAIS (0.764) despite useful signal due to high privacy cost. Recommended deployment: spatial + hand + gaze (no physio): ILAIS = 0.908, AUC = 0.908.

6.2 Open Resources

The ILAF evaluation dataset (1,920 sessions, anonymised; gaze, spatial, hand, voice features; no raw video/physiological), ILAIS calculator, Unity analytics SDK (6 modality extractors), privacy assessment checklist for ILA, and domain-specific analytics design guides are available at ilaf-edu.org under Apache 2.0 License. Raw physiological data not shared per ethics protocol; summary statistics available on request under data agreement.

References

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin, 91(2), 276.

Biocca, F., Harms, C., and Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence. Presence, 12(5), 456-480.

Blikstein, P., and Worsley, M. (2016). Multimodal learning analytics and education data mining. Journal of Learning Analytics, 3(2), 220-238.

D'Mello, S., and Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2), 145-157.

Emerson, A. et al. (2020). Multimodal learning analytics for game-based learning. British Journal of Educational Technology, 51(3), 826-846.

Ivanov, A., Schmidt, S., and Moreau, M. (2025). Metaverse-enabled educational platforms for collaborative learning. Journal of Digital Learning Futures, 2025(3), 1-8.

Klein, I., and Garcia, E. (2024). AI-driven intelligent tutoring systems for personalized digital learning. Journal of Digital Learning Futures, 2024(1), 1-9.

Kovacs, A., and Ivanov, E. (2024). Cognitive-aware AI systems for learner-centric education. Journal of Digital Learning Futures, 2024(1), 38-47.

Novak, H., and Lindberg, L. (2025). Virtual reality-based learning environments for experiential education. Journal of Digital Learning Futures, 2025(2), 35-42.

Radianti, J. et al. (2020). A systematic review of immersive virtual reality applications for higher education. Computers & Education, 147, 103778.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372.

Romero, C., and Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. WIREs Data Mining and Knowledge Discovery, 10(3), e1355.

Russell, T. et al. (2022). Gaze-based learning analytics in virtual reality. Computers & Education, 182, 104456.

Slater, M. (2003). A note on presence terminology. Presence Connect, 3(3), 1-5.

Thomas, M. S. C., and Davis, J. (2024). Multimodal learning analytics in immersive environments. Journal of Computer Assisted Learning, 40(1), 48-62.

Tobii (2024). Eye tracking integration for VR. https://www.tobii.com/vr.

van Leeuwen, A. et al. (2019). Developments and applications of learning analytics to support teachers. British Journal of Educational Technology, 50(6), 2903-2915.

Zawacki-Richter, O. et al. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 39.

Zimmermann, P. et al. (2023). Eye tracking in VR learning: A systematic review. Computers & Graphics, 115, 16-29.

Worsley, M. (2018). Using log data to understand the role of collaboration in learning. Journal of Learning Analytics, 5(2), 20-37.

Declarations

Funding

This research was supported by the Italian Ministry of University and Research (MUR) under grant PRIN2023-2025-ILAF (Immersive Learning Analytics Framework), the Swedish Research Council (Vetenskapsradet) under grant 2024-09024, and the Swiss National Science Foundation (SNSF) under grant 220024. The funders had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025

The ILAF anonymised feature dataset, ILAIS calculator, Unity analytics SDK, privacy assessment checklist, and domain-specific design guides are available at ilaf-edu.org under Apache 2.0 License. Raw physiological data available under data agreement on request.

Ethical Approval

All participants provided written informed consent including explicit opt-in for each analytics modality. Physiological monitoring required additional consent and could be declined without affecting VR participation. Eye tracking data: opt-in only. Voice data: processed to features only, transcripts deleted post-session. Ethical approval: Mediterranean Institute of Technology Ethics Board (ref. MITR-2024-IRB-0212), Nordic Technical University Ethics Board (ref. NTU-2024-IRB-0184), Swiss Institute of Machine Intelligence Ethics Committee (ref. SIMI-2024-ETH-0162).

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025

Appendix A

ILAF Data Collection Pipeline and ILAIS Calculation

Technical data collection specifications and ILAIS metric calculation for all six ILAF modalities.

Data Collection and Feature Extraction

ILAIS Calculation and Privacy Assessment