

AI FOR HUMAN FACTORS AND
ERGONOMICS

HANDS-ON WORKSHOP

Csongvai Roland

2 July 2026 · 14:00-16:00 · Interactive session

Sensing → Estimation → Standards mapping → Deployment decision

1

WHY THIS SESSION EXISTS

1

Industry 4.0 / 5.0 puts the human at the centre of system design.

— and the sensors to watch that human are now cheap.

2

Pose estimation, wearables and LLMs make ergonomic monitoring scalable.

— what needed a lab now runs in a browser.

3

Field-ready, practitioner-usable tooling is still rare.

— the gap is not the models — it is the audit skill.

"You will not listen to AI today. You will run it, break it, and audit it."

2

THE QUESTION WE STARTED FROM

Can AI take over parts of the ergonomist's job — and where must it not?

What is actually mature — and what is hype?

What do the reported numbers really prove?

What would a field-runnable, auditable tool look like?

This afternoon retraces that path: literature → synthesis → tool → your data.

3

LITERATURE AT AI SPEED · ASTA (Ai2)

WHAT IT IS

A scholarly research assistant combining literature understanding and data-driven discovery.

Searches 108M+ abstracts and 12M+ full-text papers.

Finds, summarises and analyses evidence — with citations.

Find papers

 — cited answers to the queries on the right

Generate a report

 — the literature summary behind Part A

Analyze data

 — your Station-1 CSV, at Station 3 or tonight

Anonymous access, no sign-up: [asta.allen.ai](#)

ASK IT WHAT WE ASKED

"State of the art in markerless pose estimation for ergonomic risk assessment?"

"Which studies validate camera-based RULA scoring against expert raters?"

"Multivariate anthropometric accommodation beyond the P5–P95 convention?"

Cited answers make stage 2 auditable — you still open the citations. That is the gate.

4

THE CORPUS AT A GLANCE

16

sources

2019–2025

publication window

Sensors

dominant venue family

Physical assessment 3

Posture & biomechanics 3

Anthropometry 4

Cognitive workload 4

Human–robot 2

Full corpus map: Fig. 1 of the companion paper.

5

THE ERGONOMIC INFERENCE CHAIN

Fig. 2 of the companion paper

1

SENSING

IMU · sEMG · EEG · RGB / depth camera

2

ESTIMATION

joint angles · posture · body dims · workload

3

STANDARDS MAPPING

RULA · REBA · OWAS · NIOSH · ISO · DHM fit

4

DEPLOYMENT DECISION

feedback · sizing · HRI · adaptive design

Error, bias, proxy labels and modelling assumptions propagate downstream.

Every station today is one link of this chain — deliberately broken.

6

1

FINDINGS I · WEARABLES + CLASSICAL ML

VALUE	WHAT IT MEASURES	WHAT IT DOES NOT PROVE
99.4% SVM	safe / unsafe posture classification in a manual-handling setting	validity of a RULA, REBA or NIOSH decision
99.94% KNN	correct / incorrect posture labels under reported study conditions	injury, fatigue or productivity outcomes
ICC 0.72 ± 0.23	repeatability of joint patterns — 10 workers, one cable plant	universal automated risk scoring
38%	reported optimisation of ergonomic conditions vs traditional approaches	clinical injury reduction — a deployment hypothesis

High accuracy in a controlled setting is evidence about the setting — audit before you transfer.

The paper explicitly REFUSES two requested intervention values (VAS Δ −0.28; low-back OR 0.53) — no primary evidence in the corpus. That discipline returns at Station 3.

7

FINDINGS II · MARKERLESS POSE ESTIMATION

WHY IT IS THE CLEAREST ROUTE

Ordinary video → landmarks → joint angles → category → RULA / REBA / DHM.

Scalable, real workplaces, no lab motion-capture.

This is exactly what the Lab you will run does.

WHERE IT BREAKS

Self-occlusion, lighting, camera view.

Muscle force and joint load remain out of reach.

A small angle error near a band edge flips the category → uncertainty-aware scoring, not one deterministic label.

The Lab's "near boundary" chip exists because of this paragraph.

8

FINDINGS III · ANTHROPOMETRY

THE UNIVARIATE TRAP

The P5-female / P95-male corner-point convention fails when body dimensions covary → multivariate accommodation is required.

Missing populations — older adults, children, people with disabilities, plus-size, non-Western — are a fairness error that propagates into design.

WHAT THE CORPUS REPORTS

89.66% vs 68.97% — SVM vs PCA-SVM clothing-size classification from 3-D scans.

52 subjects · ISO 20685 protocol · 4 measures — one validation, not universality.

DHM families: RAMSIS · JACK on ANSUR / CAESAR databases.

Try it: [dined.io.tudelft.nl](#) — 1-D percentiles vs the Ellipse tool.

9

FINDINGS IV · COGNITIVE WORKLOAD + HRI

COGNITIVE WORKLOAD

76.59% — 4-class workload, EEG + eye-tracking, 34 features.

F1 0.97 — 3-class, eye-tracking only, 7 of 20 features.

Weak link: ground truth is NASA-TLX-style proxy labels.

HUMAN-ROBOT INTERACTION

Cobot safety is bounded by the weakest upper estimate — not by the controller.

Transparent policies over black-box speed.

Construct-validity evidence — not intervention evidence.

10

FOUR WAYS THE CHAIN BREAKS

1 MEASUREMENT ERROR

Occlusion, lighting and camera view shift the angle before anything else happens.

2 PROXY LABELS

The "ground truth" is a label someone chose — not an injury outcome.

3 REPRESENTATION BIAS

The model only knows the population it was trained on.

4 STANDARDS-TRANSLATION AMBIGUITY

A probability is not a RULA decision until the mapping is auditable.

You will trigger all four before 16:00.

11

THE GAP → SO WE BUILT ONE

WHAT THE LITERATURE LEAVES OPEN

Strong technical results — thin field validation.

Black-box outputs block practitioner audit.

Credentialed, hands-on training is scarce — and practitioners ask for it.

The missing artifact is a field-runnable, auditable demonstrator. So we vibe-coded one this morning.

Part B: you run it.

12

TAKEAWAYS

1

Wearables and pose estimation are mature — generative AI is early.

— pick the mature end to deploy, the early end to watch.

2

Reported accuracies are high — and laboratory-biased.

— you broke a 95%-class tool with a desk lamp and a folder.

3

The limiting factor is not the model. It is the audit layer: you.

— explainability, representativeness, standards translation.

The mini-report you filled in is the take-home — use it on the next vendor pitch.

19

THE PIPELINE BEHIND THE LAB

Problem → literature → synthesis → vibe-coded tool → your data — built in one morning, audited at every hand-off

1

PROBLEM

Can a phone camera score working posture in the field?

your domain expertise

2

LITERATURE / SOTA

Agentic review over 108M+ abstracts — answers with citations

Asta (AI2) · Semantic Scholar

3

SYNTHESIS

16 sources → inference chain + four failure modes

Claude · the paper

4

VIBE-CODED TOOL

Prompt → working web app; RULA maths unit-tested

Claude · MediaPipe

5

YOUR DATA

Mark → CSV → analysis of your own postures

the Lab · Asta · any LLM

Every arrow is a human audit gate — hallucinated citations, mis-binned angles and proxy labels do not pass.

You just ran stages 4–5 yourselves.

20

THANK YOU

Questions — and your neighbour-test stories

Csongvai Roland

Slides, worksheet and the tool remain available after the session.

21

4