Human-Centric Computer Vision Research

Recent Developments & Outlook
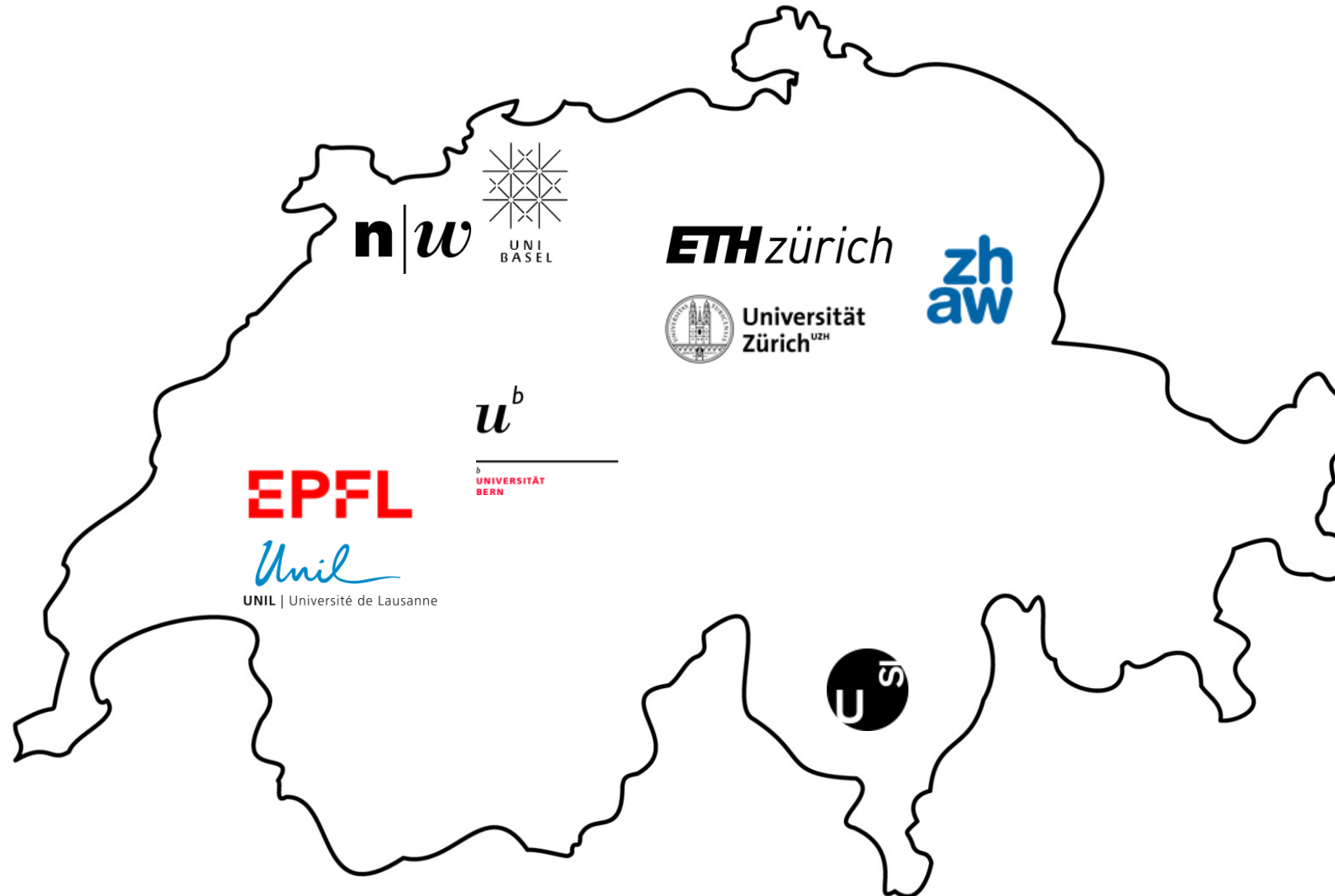
Dr. sc. ETH Manuel Kaufmann
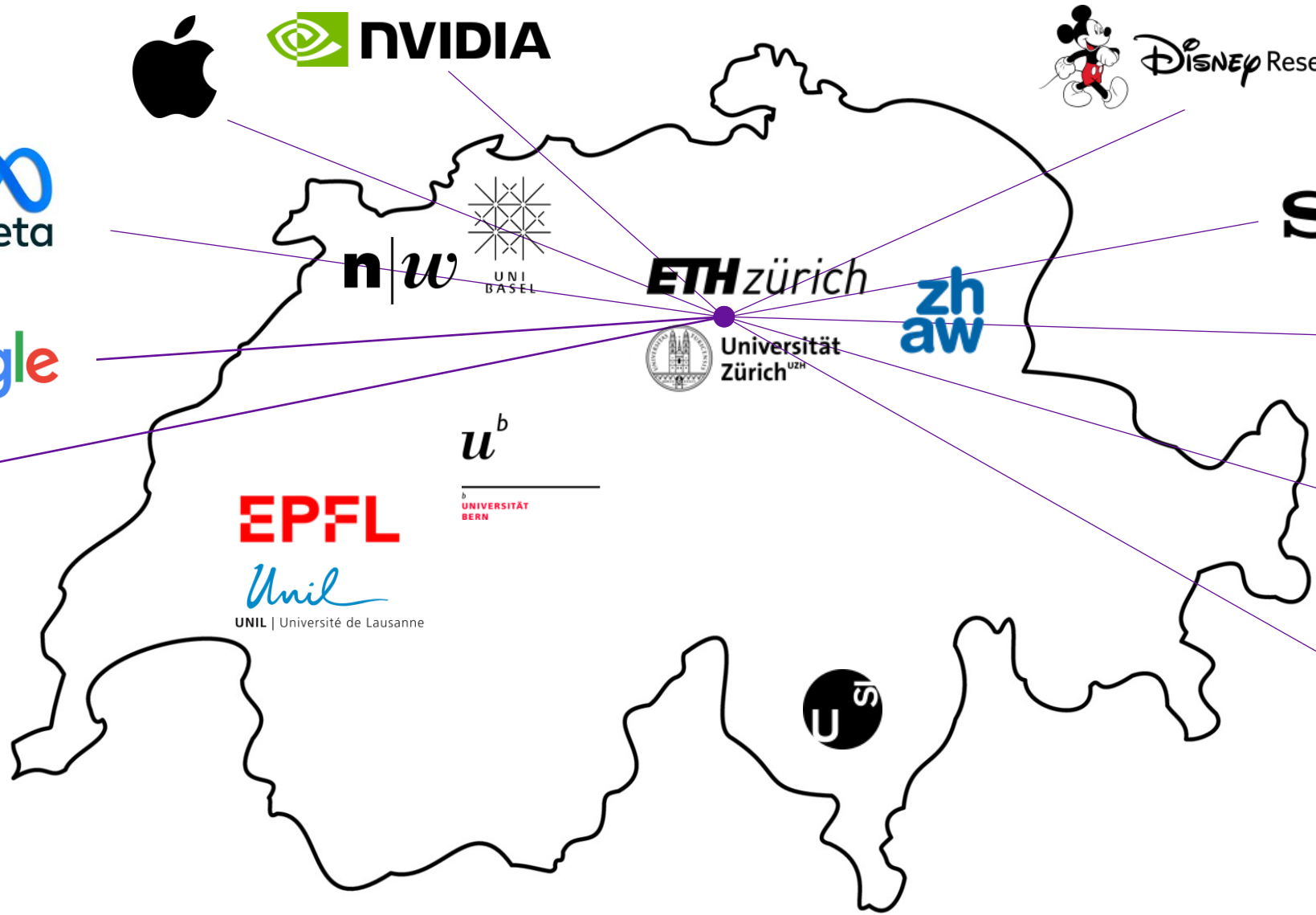
ETH AI CENTER

ETH *zürich*

Switzerland's Strong R&D Background

Switzerland's Strong R&D Background

Prof. Markus Gross
Director Disney Research Studios

Prof. Marc Pollefeys
Director Mixed Reality and AI
Zurich Lab at Microsoft

Global Innovation and Talent Leader

#1 Global Innovation Index[1]

#1 European Innovation Scoreboard[2]

#1 Global Talent Competitiveness Index[3]

#1 World Talent Ranking[4]

#1 ETH Zurich named best university outside of the US and UK[5]

[1] WIPO & Portulans Institute, 2023
[2] European Comission, 2023
[3] INSEAD, Human Capital Leadership Institute & Descartes Institute for the Future, 2023
[4] IMD, 2023
[5] Times Higher Education, Word University Ranking 2024

ETH Among Top CV Research Institutions

Between 2014 - 2024 ETH was **top 7** CV contributor **world-wide** and the **leading** institution in **Europe**.[1]

IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR) is the **premiere Computer Science** venue.
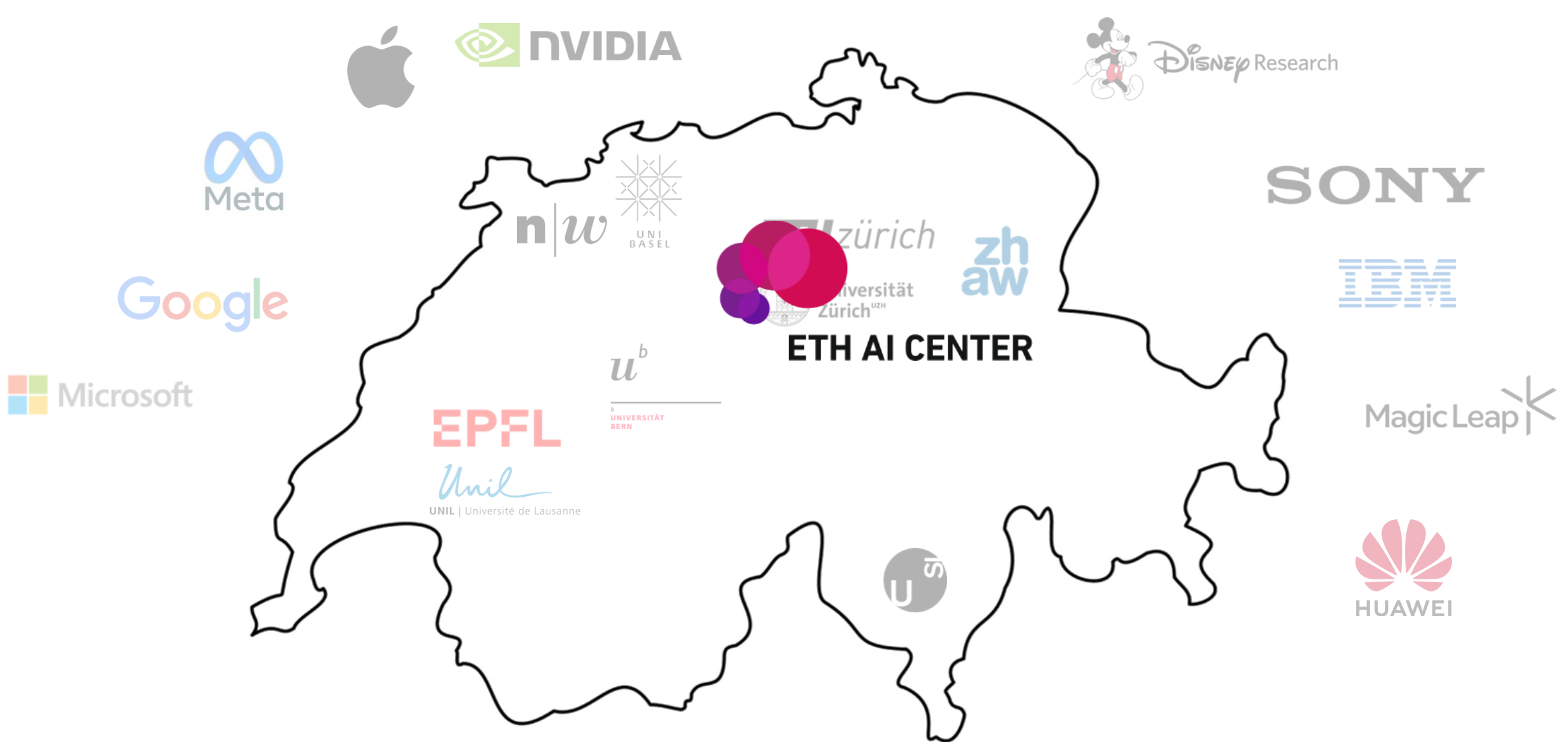

At CVPR 2022 ETH was **top 3** contributor and the **leading** institution in **Europe/America**.[2]

[1]According to CSRankings, https://csrankings.org/#/index?vision&world
[2]According to statistics from CVPR https://twitter.com/CSProfKGD/status/1555010601692299264/photo/2
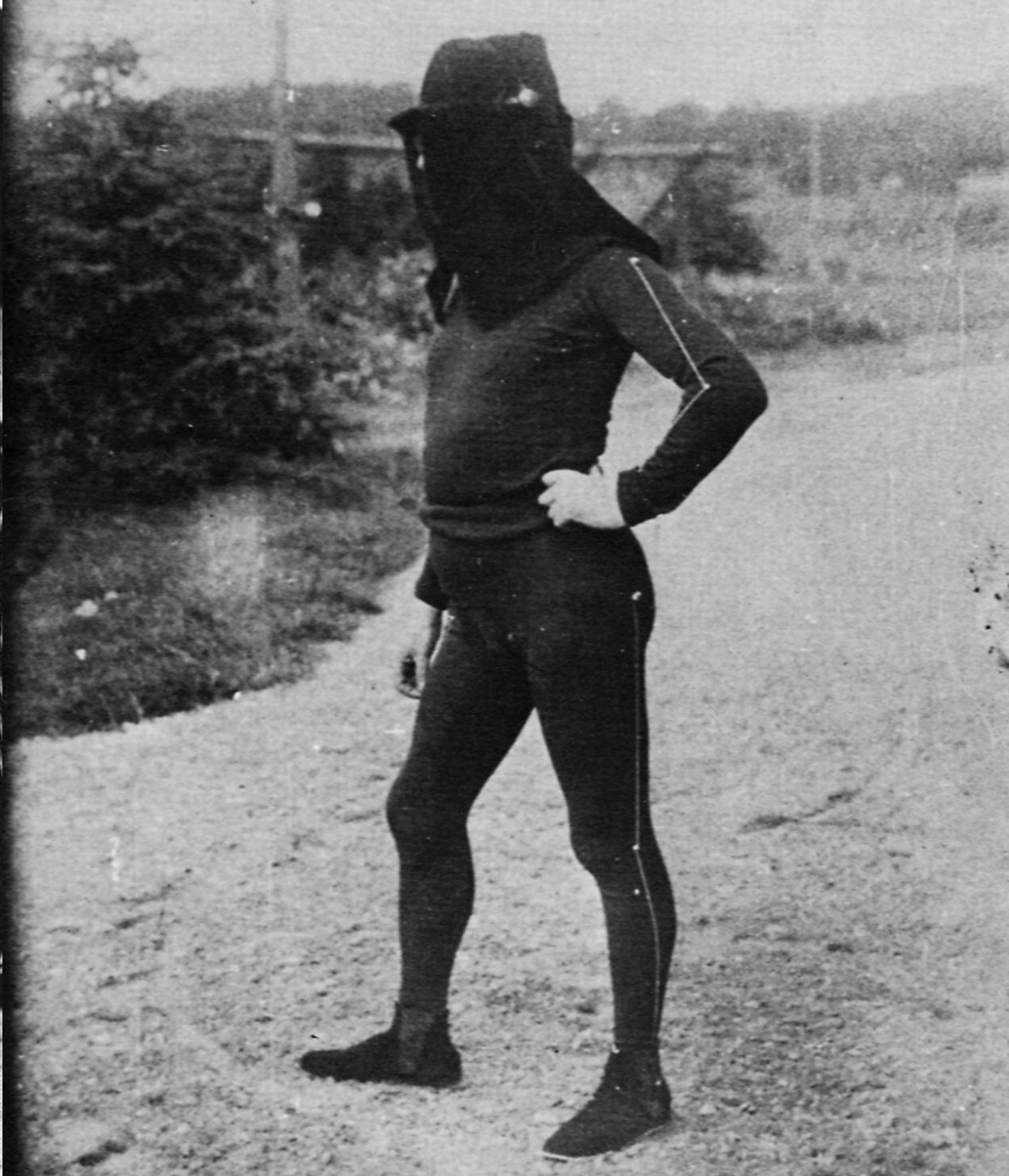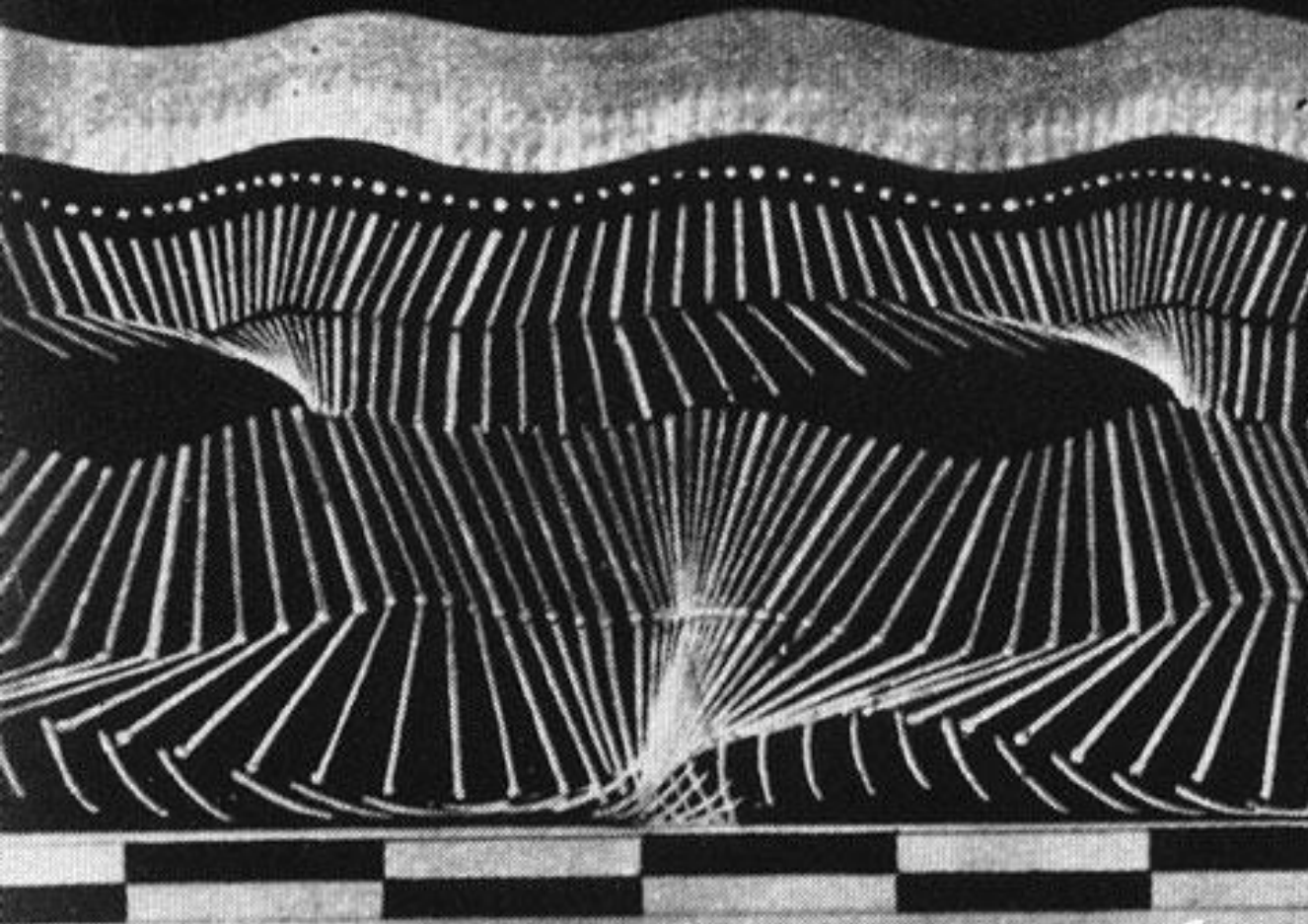
ETH AI Center's Unique Position

Human-Centric CV

Overview & Recent Developments

Performance Capture in 1883

[1] Etienne-Jules Marey. Joinville Soldier Walking

[3]

[4]
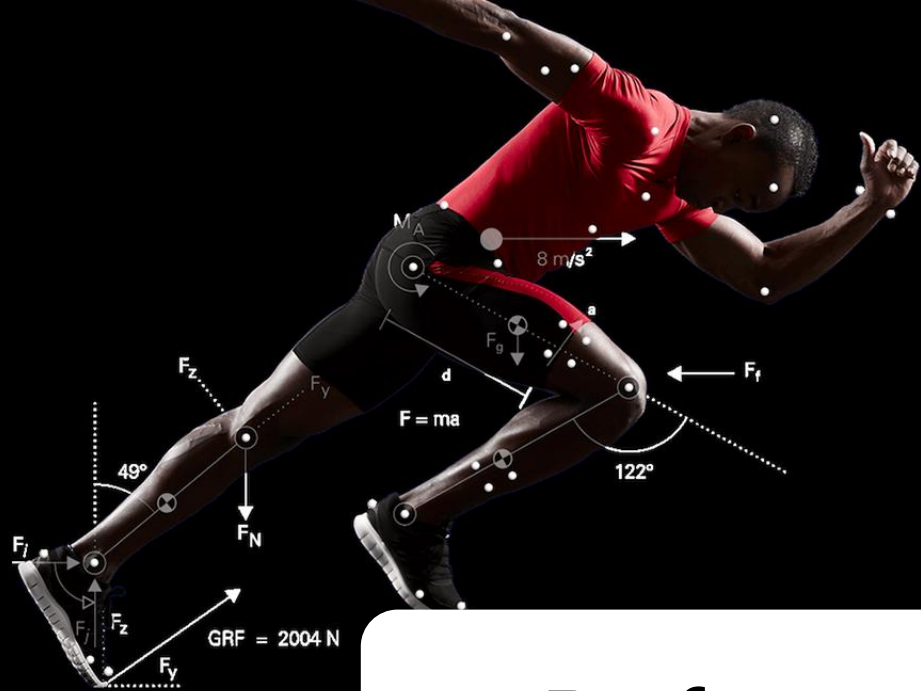
Performance Capture Today

[5]
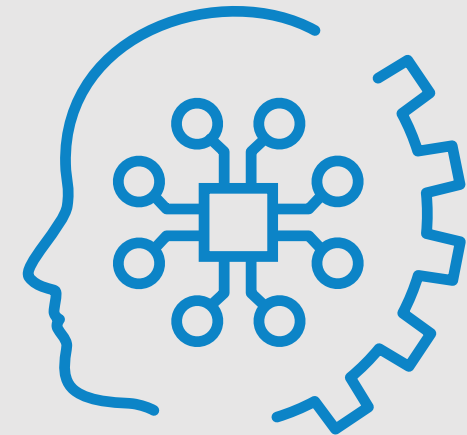
AI AGENTS

[6]

High-End Volumetric Capture

Monocular Image-Based Pose Estimation

[PyMAF, Zhang et al., ICCV 2021]
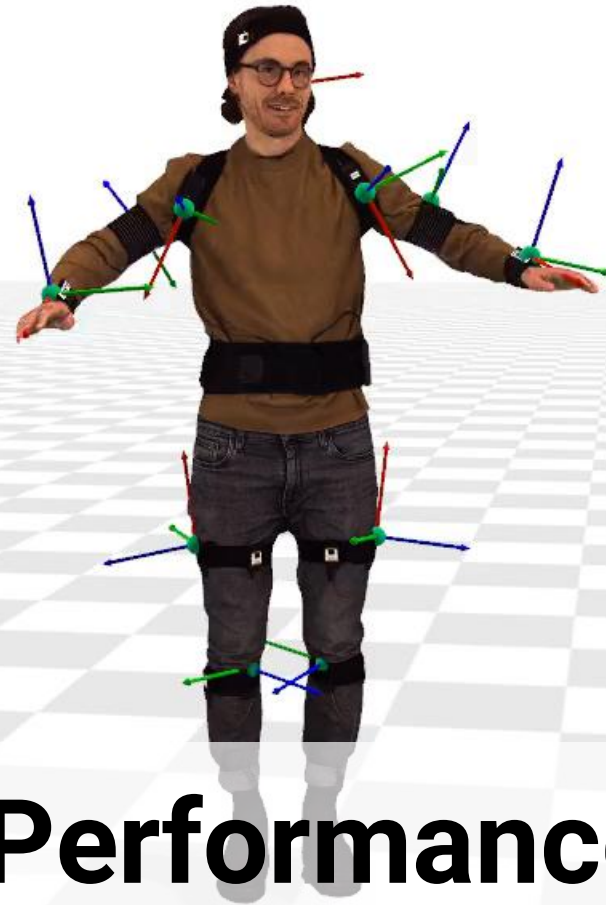
Sensor-Based Performance Capture
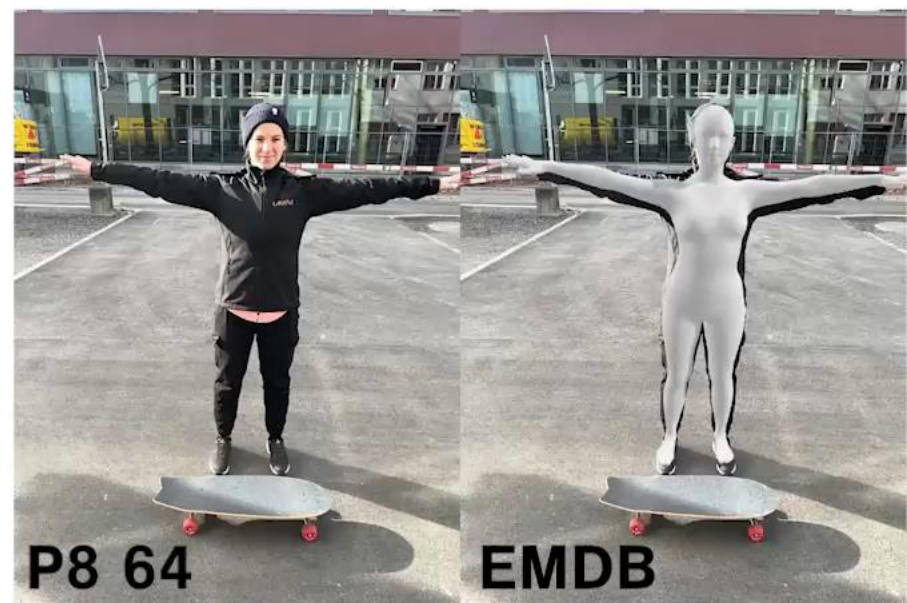here with body-worn 6-DoF Sensors

In-The-Wild Recordings with Body-Worn Sensors

SPEED X2

SMPL Root Trajectory 　 Camera Trajectory

P8 64 　 EMDB

Human Performance Is More Than Pose

Skilled Hand-Object Interaction

Close Interaction with Contact Between Humans

Dynamically Deforming Clothing

Coordinated Interaction in Large Numbers

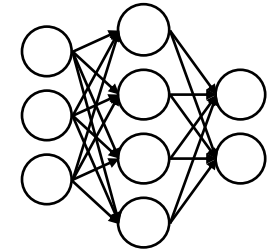Can we bring a machine to perceive the human holistically in 3D given only a single input video?

What Has Been Driving Progress?
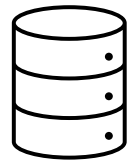
Inductive biases and statistical body models

[Loper et al., SMPL, Siggraph Asia 2015]

Advanced Deep Learning architectures and compute

Availability of data*

A Promising New Paradigm

$$\boldsymbol{x} \in \mathbb{R}^3 \qquad \rho \in [0, 1]$$

Input Image $\longrightarrow$ Neural Implicit 3D Shapes **+** Differentiable Rendering $\longrightarrow$ Reconstructed Input Image

Optimize

Compare
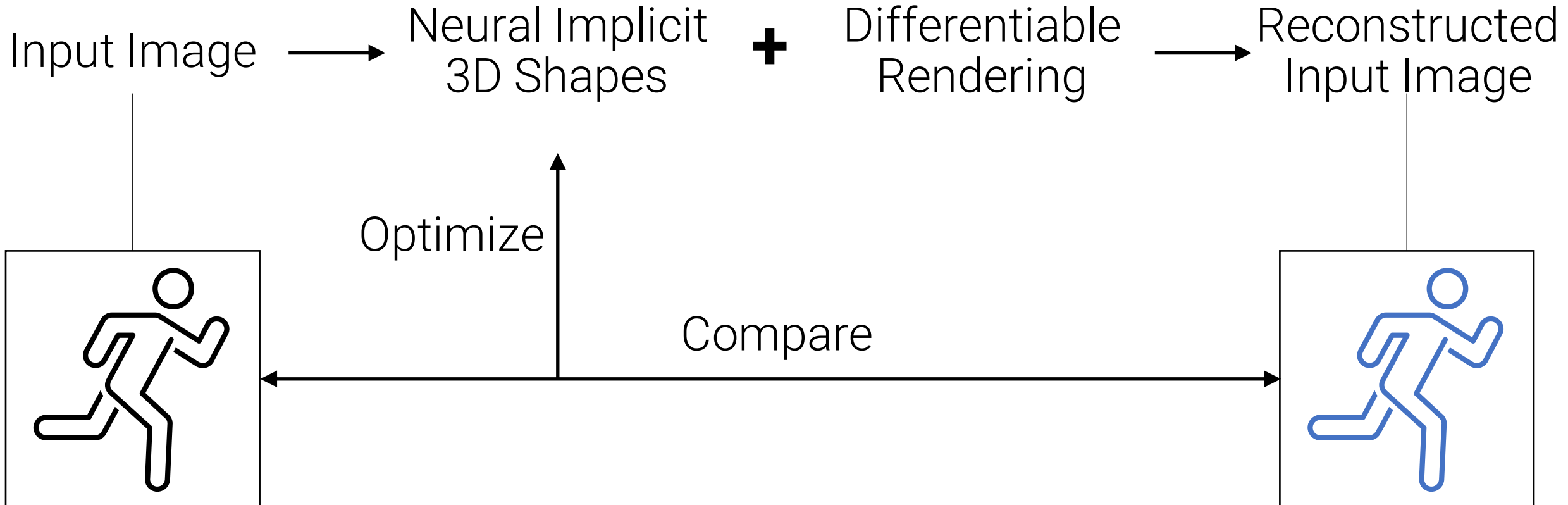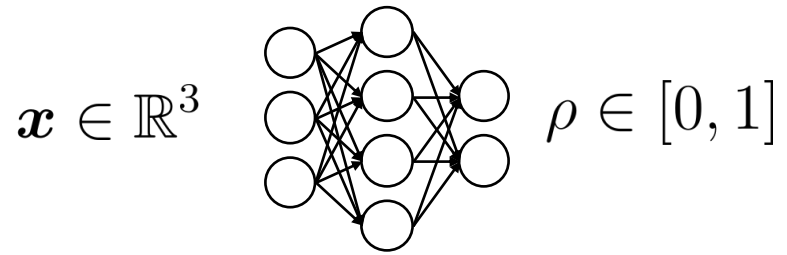
Learning Avatars from Monocular Videos

[Guo et al., Vid2Avatar, CVPR 2023]

Multiple Humans In Interaction

Input
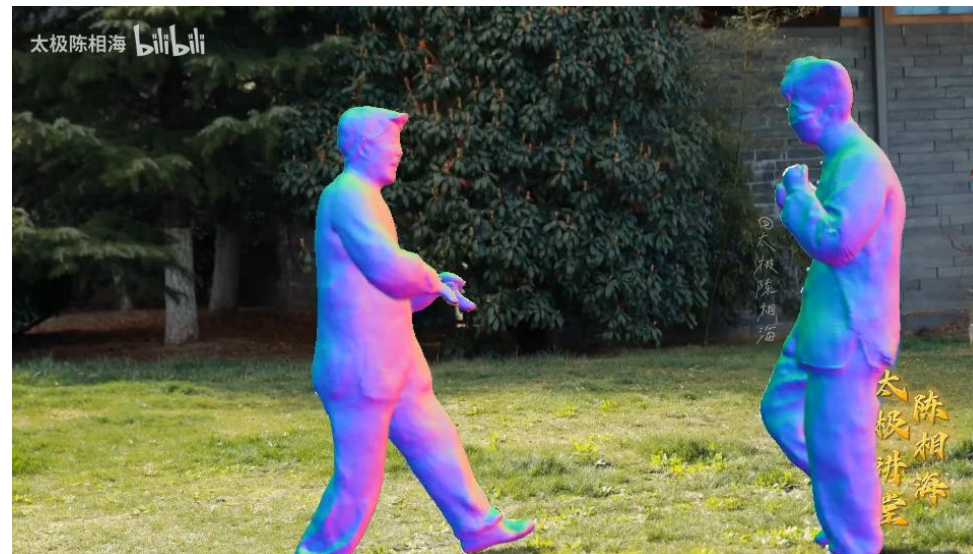
Reconstructions

3D Instances

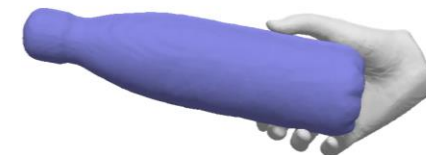3D Novel View

[Jiang and Guo et al., MultiPly, CVPR 2024]

Hands and Objects

Input Video

Reconstructed hand and object surfaces

[Fan et al., HOLD, CVPR 2024]

Hands and Objects

[Fan et al., HOLD, CVPR 2024]
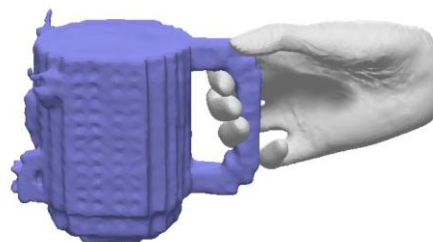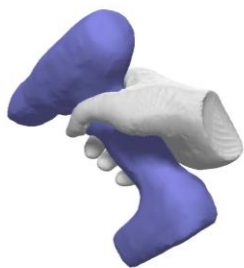
Outlook

Human-Centric AI for Health

The AI Assistant of the Future

prompt 1
prompt 2
...

PREVENTION &
EARLY
INTERVENTION

from on-screen,
prompt-based AI

to **always-on**, **contextualized**
co-pilots that **amplify** human
potential

to **assist** healthcare
professionals and individuals in
personalized health
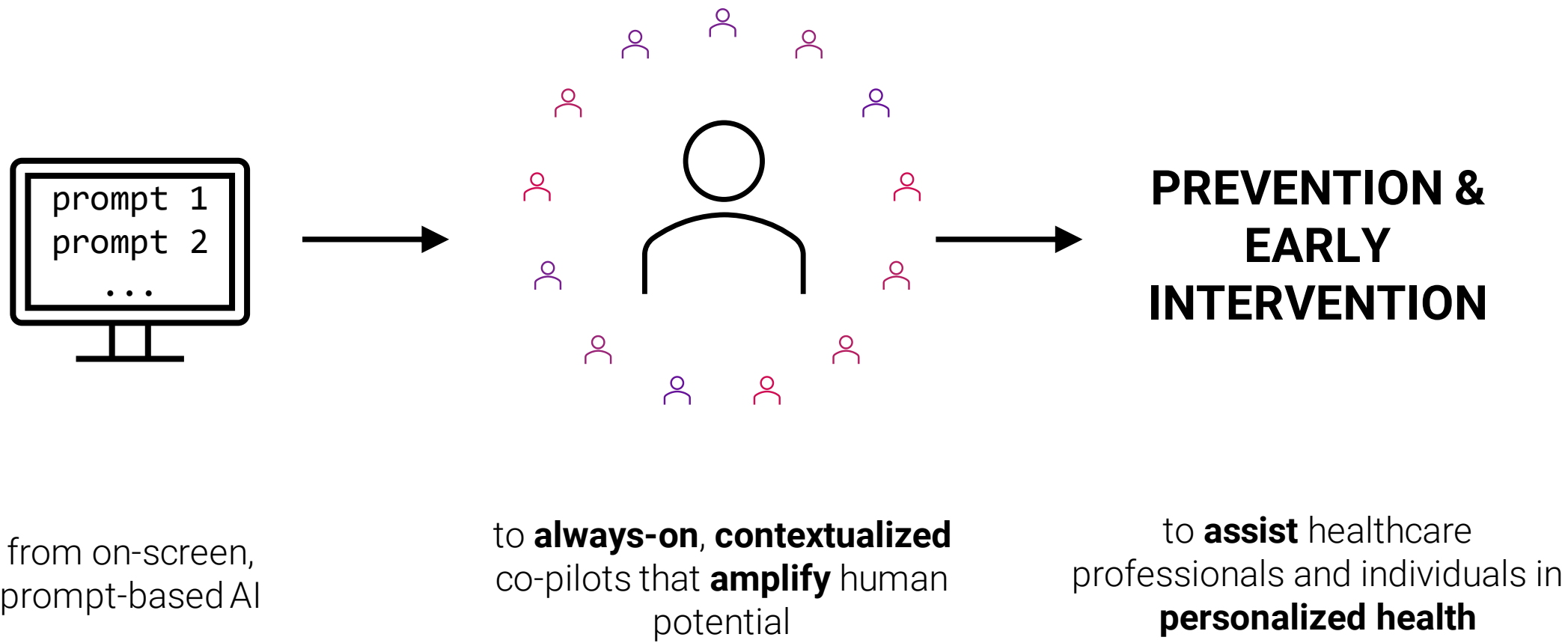
Enabled By Contextual Awareness

MULTIPLE DEVICES & USERS   ✕   **TRACKING ENVIRONMENT**   ✕   **EGOCENTRIC FOUNDATION MODELS**

In The Wild

Lab-Based

Simulated

Topic 1
Egocentric AI for Healthcare Professionals to Support Error Prevention
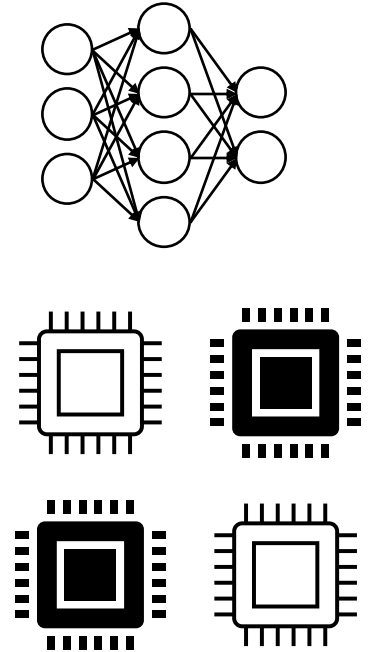
Multimodal digitization of realistic surgeries in living lab conditions:

- multi-person tracking
- object tracking
- exo- and egocentric views
- eye tracking
- audio including commentary

Strategic partnership between OR-X and the ETH AI Center to research the future of surgery.
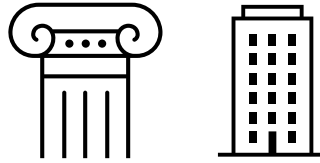
Topic 2
Advanced Preventive Care In the Wild

Move preventive medicine out of the hospital into people's homes.

- Establish new correlations between diseases and wearable health data.
- Provide expert information based on long-term tracking.
- E.g., continued gait analysis from few sensors.
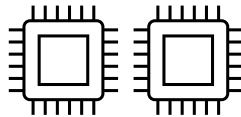- E.g., standardize and simplify clinical tests

ETH AI Center

120 associated faculty, 30 related to health, machine learning, computer vision, > 60 affiliated industry partners

Strategic Research Platform
Human-Centric Data Acquisition

Various recording devices and data capture platforms

Large-Scale Compute via Swiss AI Initiative, ALPS Super-Computer at CSCS

References

[1] Etienne-Jules Marey. Joinville Soldier Walking. Geometric chronophotograph, Paris College de France, via https://www.khanacademy.org/humanities/becoming-modern/earlyphoto/early-photo-france/a/marey-joinville-soldier-walking

[2] Eadward Muybridge. The Horse in Motion. Provided directly by Library of Congress Prints and Photographs Division, Public Domain, https://commons.wikimedia.org/w/index.php?curid=57260211

[3] https://www.patcash.co.uk/2011/09/biomechanics-what-it-is-and-why-it-is-important-for-tennis/

[4] https://www.ign.com/articles/meta-quest-3-review

[5] https://community.robotshop.com/blog/show/robots-are-transforming-the-healthcare-industry

[6] https://iis.inf.ethz.ch/

[7] Switzerland Map Vectors by Vecteezy