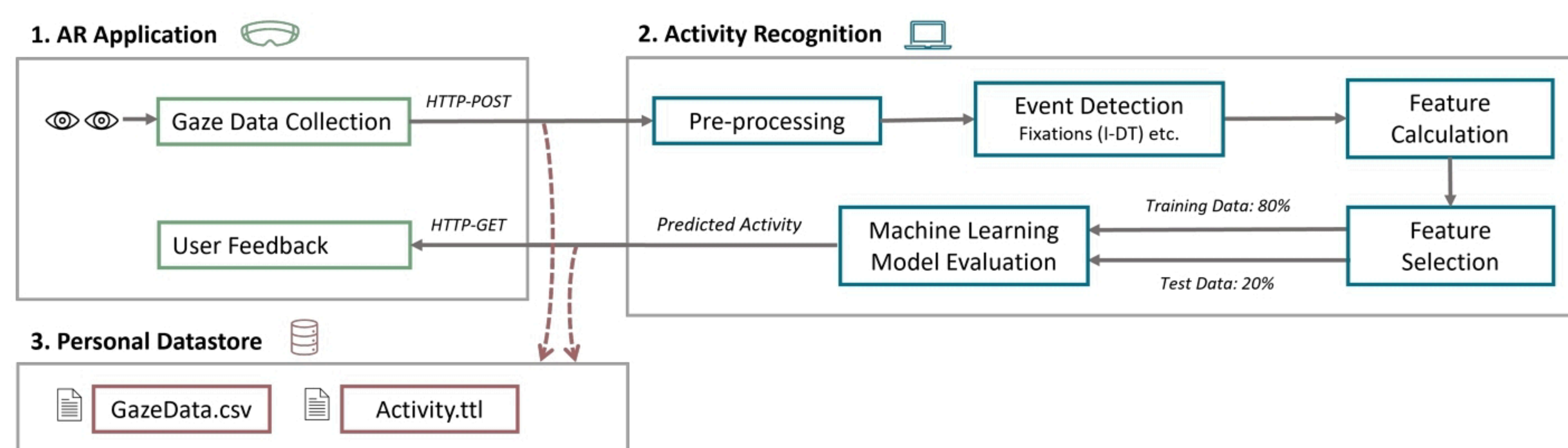


GEAR: Gaze-enabled Augmented Reality for Human Activity Recognition

Head-mounted Augmented Reality (AR) displays overlay digital information on physical objects. Through eye tracking, they allow novel interaction methods and provide insights into user attention, intentions, and activities. However, only few studies have used gaze-enabled AR displays for human activity recognition (HAR). In an experimental study, we collected gaze data from 10 users on a HoloLens 2 (HL2) while they performed three activities (i.e., read, inspect, search). We trained machine learning models (SVM, Random Forest, Extremely Randomized Trees) with extracted features and achieved an up to 98.7% activity-recognition accuracy. On the HL2, we provided users with an AR feedback that is relevant to their current activity. We present the components of our system (GEAR) including a novel solution to enable the controlled sharing of collected data. We provide the scripts and anonymized datasets which can be used as teaching material in graduate courses or for reproducing our findings. **We demonstrate accurate differentiation between three activities based on HoloLens 2 gaze data to display contextually relevant suggestions in AR.**

System Components



Collected gaze data is sent from the AR Application (1) to the Activity Recognition component (2). The recognized activity is returned to the AR application which displays appropriate feedback. As an alternative to direct HTTP-based interaction, collected gaze data and the recognized activity can be exchanged via a privacy-friendly Personal Datastore (3) that is based on Solid.

Method

Participants

- 10 participants (3 female, 7 male)
- Average age of 30.1 years
- 4 participants wear prescription glasses often or all the time
- ~30% are not familiar with MR headsets

Apparatus

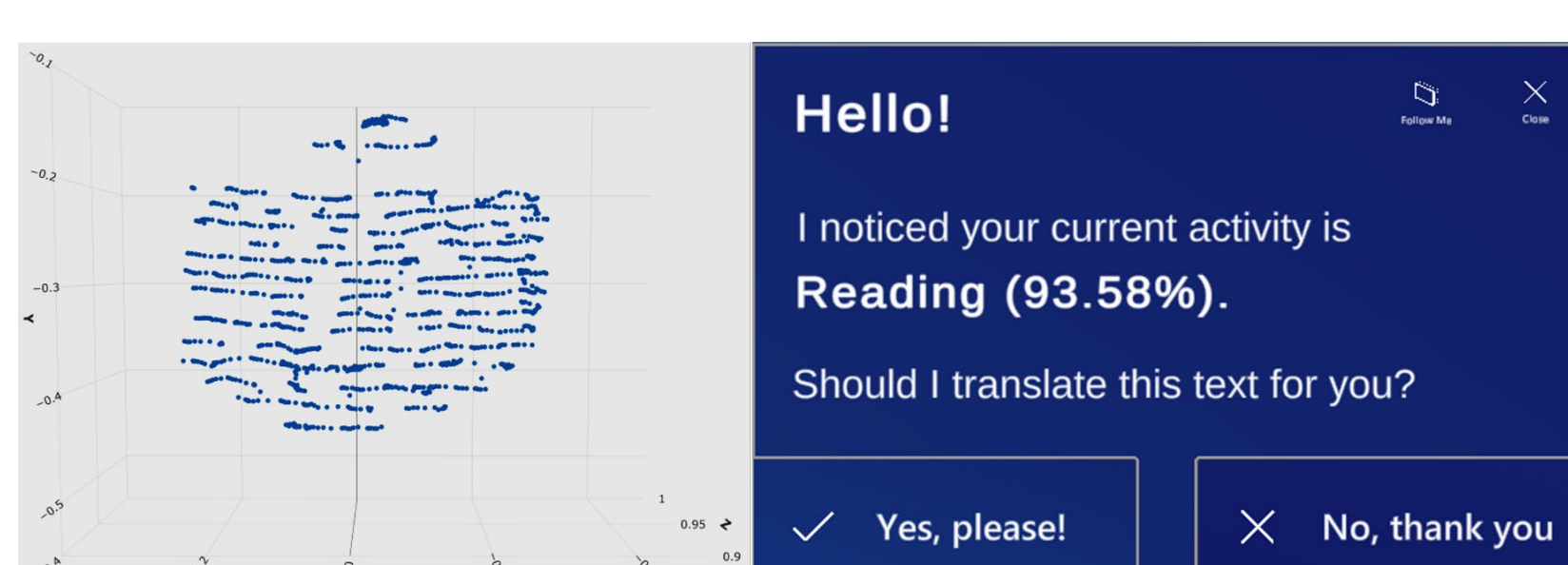
- Microsoft HoloLens 2 (Eye tracking sampling rate: 30Hz; Accuracy: ~1.5°)

Procedure

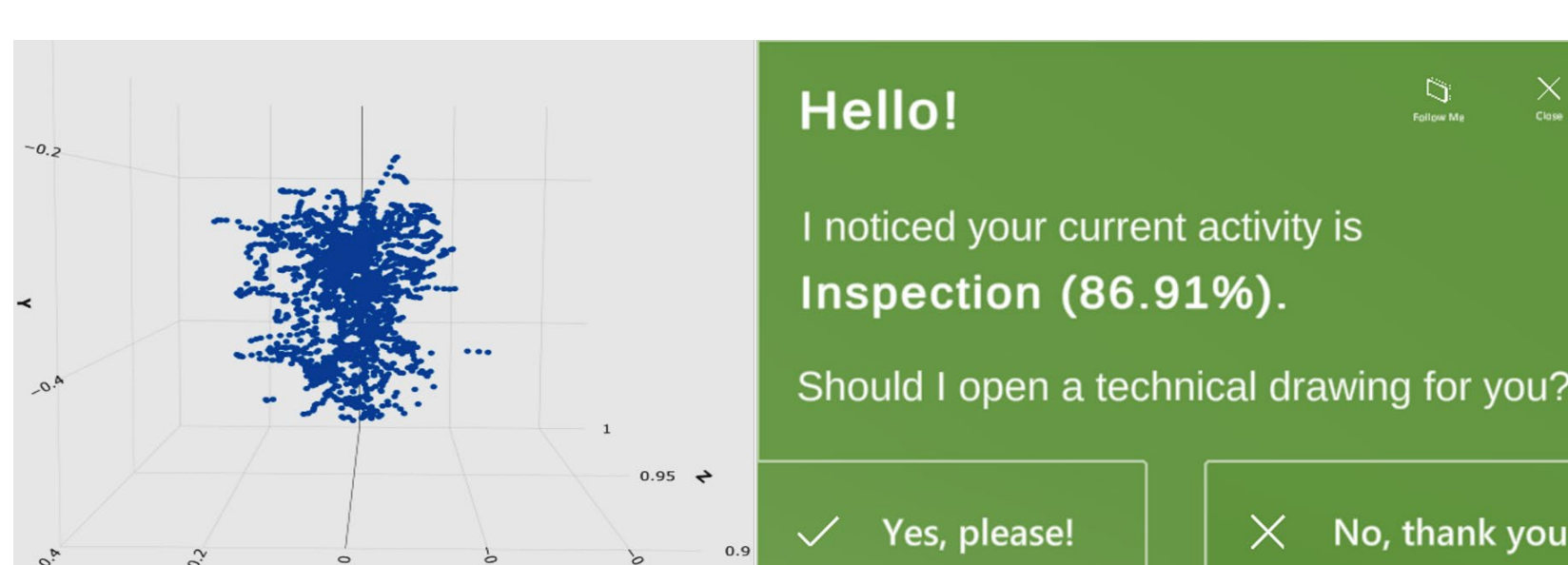
- First, the participants were asked to *read* instructions on how to inspect a robot.
- Then they *inspected* the robot.
- Finally, they *searched* for a missing piece of the robot in a shelf.

Activities

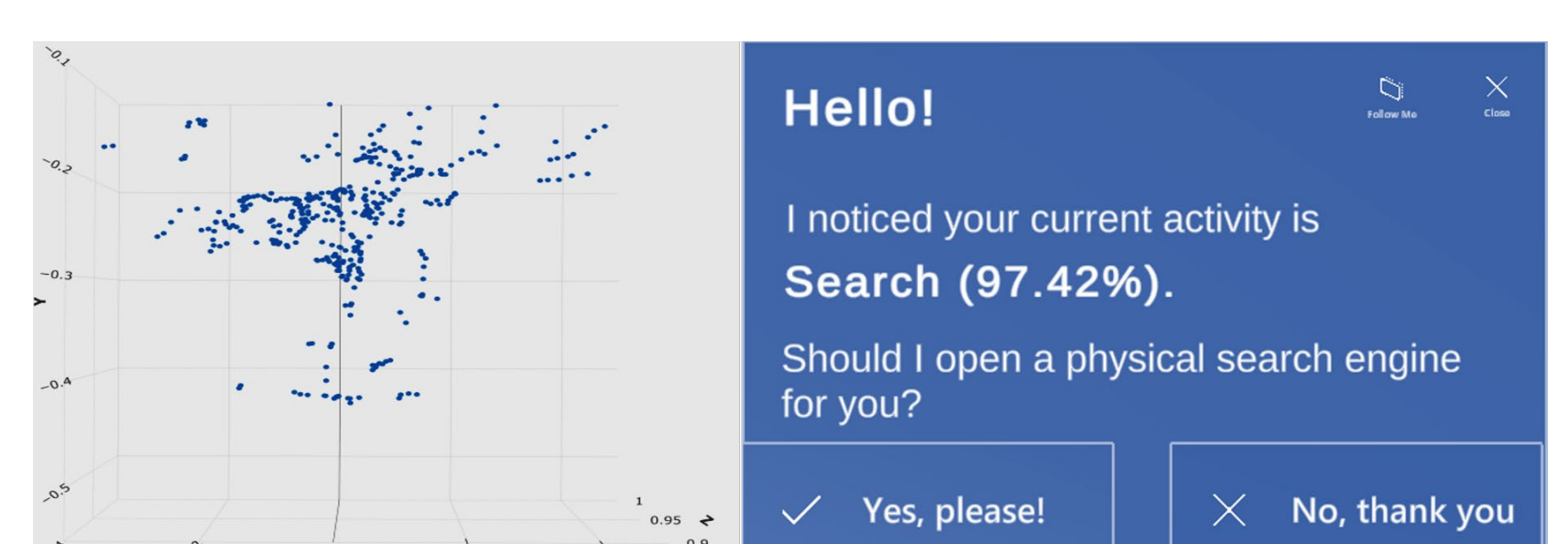
1 Reading instructions



2 Inspecting a robot



3 Searching for a missing piece



Example 3D-Plots of the normalized gaze direction data points (x/y/z) for the three activities *reading*, *inspecting*, and *searching* (left). After an activity is recognized, GEAR displays contextually relevant feedback and suggestions in AR on the HoloLens 2 (right).

Model Performance

	5s	10s	15s	20s
Support Vector Machine	78.7%	93.3%	85.0%	100%
Random Forest	93.2%	96.6%	94.4%	94.9%
Extremely Randomized Trees	89.3%	98.7%	96.2%	96.6%

Performance benchmark of the three models that are trained with 19 features. The number in each cell presents the accuracy (%) of the model for a 5, 10, 15, and 20 seconds time window.



Link to the paper



kenan.bektas@unisg.ch



interactions.ics.unisg.ch

Link to the data, code
and demo video

