Abstract: We are entering the era of big video data where cameras are ubiquitous. In particular, the amount of videos from wearable cameras and robots is explosively rising. These videos, taken from an actor's own viewpoint, are called 'first-person videos' or 'egocentric videos'. In this seminar, we discuss different types of videos robots are expected to face in their everyday operation, and overview approaches to recognize human activities from such videos. We present features and recognition algorithms necessary for activity-level understanding of first-person videos, and describe how they make recognition of human-human (and human-robot) interactions possible. Approaches for human activity 'prediction' from streaming videos will also be described, which allow the robots to infer humans' intended activities at their early stage.