**Long Term Learning dataset description**

**Experimental protocol:**

Twenty-one surgical residents from the Soroka Medical Center participated in the study. Once a month, each resident participated in three training sessions surrounding a single 26-hour hospital shift: before the shift (∼7 am), during the shift (∼3 pm the same day), and after the shift (∼9 am the next morning). Each resident participated in these three sessions for the duration of six months, leading to a total of 18 sessions per resident. The participants used the da Vinci Research Kit (dVRK) to perform three surgical training tasks:

* Ring transfer: the participant transferred the rings from the top towers to the side towers, using each of the right and left tools in turn.
* Knot tying: the participant tied a surgical knot to attach the two eyelets of the towers.
* Suturing: the participant sutured a wound in an artificial tissue through marked targets.

Participants who withdrew from the experiment:

* S,R – withdrew after one round.
* H - withdrew after 3 rounds.

**Coding**

* Each participant has one uppercase letter unique ID: **A-U** (ParticipantID)
* The number of the round (shift) is marked with a number: **1-6** (RoundNumber)
* The type of session is marked with a lowercase letter (SessionType):
  + **a-** before the shift
  + **b**- during the shift
  + **c**- after the shift
* The tasks are marked with their names (TaskName):
  + **Ringtowertransfer**
  + **Knottying**
  + **Suturing**

**Data in Excel files (LongTermLearningProject\Data\ExcelDataFiles):**

* **QuestionnairesData.xlsx**- questionnaires that were filled out by the participants.
  + One sheet contains the answers to the general questionnaire that each participant filled out at the beginning of the data collection.
  + The other sheets are for the questionnaires the participants filled out at the beginning of each session. columns B-G are for each of the six rounds.
* **PostSessionsData.xlsx**- data that the experimenters wrote during the sessions.
  + Each participant has one sheet.
  + Each column is one session.

**Raw data:**

The dataset consists of a folder for each participant. For each participant, there is a folder with the participant ID (A-U)

Each folder contains:

* **subj\_ParticipantID\_RoundNumber\_SessionType.txt** (18 files, one file for each session) – a text file with information regarding the recording of the session (start time, etc.)
* **subj\_ParticipantID\_RoundNumber\_SessionType.tif** (18 files, one file for each session) – a suturing tissue scan of the session.
* **subj\_ParticipantID\_RoundNumber\_SessionType\_TaskName\_Segmentation.mat** – a table with the timestamps (and video frame number) that indicate the beginning of each segment for of the trial.
* **subj\_ParticipantID\_RoundNumber\_SessionType\_TaskName.avi** (54 files, one file for each trial: 3 tasks\*18 repetitions) – the video recording of the trial.
* **subj\_ParticipantID\_RoundNumber\_SessionType\_TaskName.mat** (54 files, one file for each trial: 3 tasks\*18 repetitions) – the dVRK recordings and the timestamps of the video:
  + coag
  + MTMR, MTML, PSM1 (right tool) and PSM2 (left tool):
    - position
      * current
      * desired
    - orientation (quaternion)
      * current
      * desired
    - velocity\_current
    - wrench\_current
    - gripper (for MTMR, MTML)
    - jaw (for PSM1,PSM2)
    - VideoTimeStamps – a vector with the timestamps of the video frames. the ith element of the vector is the timestamp of the ith video frame.