NWB

1. NWB bundles files together in a container. And has descriptions of columns.
2. Use of pynwb, opens , extracts into an object and searches into nwb files.
3. Its standard and helps reproducibility
4. Has dynamic tables which store vector data.
5. Uses HDF5
6. Every obj in nwb has a class.
7. Units table has single unit activity. Stores spike times.
8. Processing Modules has derived data
9. Acquisition has raw data
10. Electrical Module
11. Time Series
12. Spatial Series
13. Module -> interface -> data
14. Data in x and y, timestamps in seconds
15. .data_interfaces - data interface obj.
16. .data[] - actual numbers
17. Use of [] or .

Dataset

Excluded patients

1. 45,46,50,59 - not in NWB files as they are fmri only.
2. 53 must be ignored as they had calibration issues.
3. 44,58,60 - only completed 1 EMU run (r1).

Main Field structure with their data types

This NWB file contains the following high-level fields:
- **session_description**<class 'str'>
- identifier<class 'str'>
- **session_start_time**<class 'datetime.datetime'>
- **timestamps_reference_time**<class 'datetime.datetime'>
- file_create_date<class 'list'>
- experimenter<class 'tuple'>
- **acquisition**<class 'hdmf.utils.LabelledDict'>
- **stimulus**<class 'hdmf.utils.LabelledDict'>
- **keywords<**class 'hdmf.utils.StrDataset'>
- **processing**<class 'hdmf.utils.LabelledDict'>
- **electrodes**<class 'pynwb.ecephys.ElectrodesTable'>
- **electrode_groups**<class 'hdmf.utils.LabelledDict'>
- **devices**<class 'hdmf.utils.LabelledDict'>
- **intervals**<class 'hdmf.utils.LabelledDict'>
- subject<class 'pynwb.file.Subject'>

- **trials**<class 'pynwb.epoch.TimeIntervals'>
- **units**<class 'pynwb.misc.Units'> #look after looking at elcetrodes
- **experiment_description**<class 'str'>
- **session_id**<class 'str'>
- lab<class 'str'>
- institution<class 'str'>'''

Main Modules to focus on (inside is the type), [everything in this are the keys or an example from the data]-

1. Acquisition -
    a. Events_ttl (timeseries), experiment ids.
       Example ttl object =
       first 10 objects inside ttl events
       [4 41 42 43 44 45 46 47 48 49]

2. Processing- this is an important field, it is a container for interfaces not raw numbers, it is a module. It has clean data.
   1. Behaviour, ecephys
   1. Behaviour (processingmodule)
   Blink', 'EyeTracking', 'Fixation', 'PupilTracking', 'Saccade'
    1. **Eye tracking** (eye tracking type)
        a. Spatial_series -> SpatialSeries
        b. Example obj - [[630. 548.]
           [631. 548.]
           [631. 548.]
           [632. 548.]

    2. **Saccade**
        a. Time_series -> Timeseries
        b. dict_keys(['resolution', 'comments', 'description', 'conversion', 'offset', 'unit', 'data', 'timestamps', 'timestamps_unit', 'interval'])
        c. Description - **saccade information for R eye**; timestamps are start times; columns are ['Duration', 'StartX', 'StartY', 'EndX', 'EndY', 'Ampl', 'Pupil_vel']
        d. Duration, amplitude, velocity, start and end given already.
        e. Comments - screen_width,screen_height=1280.0,1024.0::display_w,display_h=1024.0, 768.0::display_area=128.0,128.0,1152.0,896.0
        f. Data -
           [[1.20e-02 6.25e+02 5.42e+02 6.14e+02 5.17e+02 6.00e-01 6.00e+01]
           [4.00e-02 6.16e+02 5.12e+02 3.37e+02 7.72e+02 8.28e+00 3.44e+02]

   g. Interval - 1
   h. Timestamps - [0.302 0.43 0.634 0.874 1.08 1.456 1.706 2.012 2.248 2.502]
   i. Timestamp unit is in seconds
   j. Unit - NA
   k. Resolution = -1.0
   l. Conversion - 1.0

3. **PupilTracking**
   a. Time series - ['starting_time', 'rate', 'resolution', 'comments', 'description', 'conversion', 'offset', 'unit', 'data', 'starting_time_unit']
   b. Description - pupil size (number of pixels inside the pupil contour)
   c. Seconds - unit
   d. Rate = 500.0

4. **Fixation**
   a. Same keys as saccade
   b. Description - fixation information for R eye; timestamps are start times; columns are ['Duration', 'FixationX', 'FixationY', 'Pupil_size_avg']
   c. Data -
      [[1.16e-01 6.15e+02 5.10e+02 2.93e+02]
       [1.64e-01 3.15e+02 8.03e+02 2.99e+02]
   d. Pupil size is given

5. **Blink**
   a. blink information for R eye; timestamps are start times; data are 'Duration'

2. Ecephys (processing module)
   1. LFP_micro and LFP_macro (both are LFP)
   2. Electrical series inside both

3. Stimulus
   1. Movie frame time (misc.AnnotationSeries)
   First 10 objects inside movieframe_time
   ['0' '1' '2' '3' '4' '5' '6' '7' '8' '9']

4. Session description - Movie watching and new/old recognition task for session: P41CS_R1

5. Session start time - 2016-08-01 00:00:00-07:00
6. time reference time is 2016-08-01 00:00:00-07:00
7. Keywords - <StrDataset for HDF5 dataset "keywords": shape (7,), type "|O">
8. Devices - Fields:

description: CS - Neuralynx-Atlas}   From paper - EyeLink 1000,

9. Intervals
    a. Trials -> dict_keys(['description', 'id', 'colnames', 'columns']
    b. Description - experimental trials
    c. ('start_time', 'stop_time', 'stim_phase', 'stimulus_file', 'response_correct', 'response_confidence', 'actual_response', 'response_time')
    d. Stim_phase is encoding or recognition
    e. Response correct is either 0,1 or NaN
    f. Stimulus file is the picture shown
    g. Confidence =
    h.
10. Fields for each electrode - description, location, microwire group present
11. Electrode_groups has multiple bundles of microwires and their metadata.
    Fields: comp address of electrode group : description: electrode wires physical type, device address on computer, fields: description of device , location of wires physically in brain,

12. Electrodes
    a. ['description', 'id', 'colnames', 'columns'])
    b. Description = metadata about extracellular electrodes
    c. Colnames = ('location', 'group', 'group_name', 'origchannel', 'pairwise_distances', 'origchannel_name', 'x', 'y', 'z')
    d. Group_name = ['Neuralynx-Atlas-microwire-1' 'Neuralynx-Atlas-microwire-2']
    e. Origchannel = ['micro-1' 'micro-2' 'micro-3' 'micro-4' 'micro-5' 'micro-6' 'micro-7'
    f. 'micro-8' 'micro-9' 'micro-10']
    g. Paiwise_distances = ['NA' 'NA' 'NA' 'NA' 'NA' 'NA' 'NA' 'NA' 'NA' 'NA']
    h. Origchannel_name = ['LACC1' 'LACC2' 'LACC3' 'LACC4' 'LACC5' 'LACC6' 'LACC7' 'LACC8' 'LSMA1'
    i. 'LSMA2']
    j. X = [0.18 0.18 0.18 0.18 0.18 0.18 0.18 0.18 -0.04 -0.04]

13. Units - terms given in technical validation section of paper:
    a. dict_keys(['description', 'id', 'colnames', 'columns', 'waveform_unit'])
    b. Description - autogen by NWBfile
    c. ('unit_id', 'unit_id_session', 'electrode_id', 'electrodegroup_label', 'origcluster_id', 'waveform_mean_encoding', 'waveform_mean_recognition', 'isolationdist', 'meanSNR', 'peakSNR', 'isibelow', 'cv2', 'waveform_mean_sampling_rate', 'spike_times', 'electrodes')
    d. waveform_mean_encoding
    e. waveform_mean_recognition
    f. Isolationdis = how far is a single neuron’s cluster away from the noise
    g. meanSNR = ratio between the mean amplitude of the entire waveform and the standard deviation of the noise.

[1.39248713 1.17005618 1.25066934 1.33258502 1.28024761 2.81413472 2.06640232]

h. peakSNR = the ratio between peak amplitude of the mean waveform of each cluster and the standard deviation of the noise
i. Isibelow = interspike intervals (ISIs) below 3 ms to show that neurons were well isolated without refractory period violations
j. cv2
k. Waveform_mean_sampling_rate: frequency (100,000 Hz) at which the voltage was sampled to create the mean waveforms
l. Spike_times
m. Electrodes
n. the pairwise projection: distance in clustering space between all neurons isolated on the same wire