

Some datasets were very easy to use.

The data was well described, complete and organised.

Sometimes code was provided to show how the data was analysed to produce the figures shown in publications.

examples include:

- [10.25493-A2KP-FKD](#)
- [10.25493-2WE9-RYF](#)
- [10.25493-VKV1-X9C](#)
- [10.25493-KXJV-K22](#)
- [10.25493-NHHM-1S5](#)
- [10.25493-ZQ9T-N4T](#)
- 

However, some datasets were harder to analyse.

Here are some of the problems I encountered with examples.

- [10.12751-g-node.pgr84e](#): datasets are not accessible, can't open them
- [10.25493-7NJQ-ANH](#): datasets are not accessible, can't open them
- [10.25493-M1AQ-3AC](#): Either missing data or not organised well because I was unable to reproduce figures with the specific cell conditions they mentioned. i.e. no cell under the conditions described could be used to reproduce the figures
- [10.25493-12X0-NBY](#): Some data was missing so unable to reproduce all figures
- [10.25493-93YQ-6QM](#): special software was required to open the files (pxp type)
- [10.25493-6C7V-HKW](#): data descriptor did not match the file so impossible to know how to analyse it
- [10.25493-VAV5-BXU](#): missing data
- [10.25493-1YN8-1J4](#): Files that are too big to analyse! Over 70 gigabytes couldn't download it on my computer
- [10.25493-REHY-EEX](#): same problem, dataset is too big
- [10.25493-XYBY-XYG](#): missing data and metadata
- [10.25493-YFQ9-MCP](#): tbk files unsure how to analyse them, tried with neo but not successful
- [10.25493-487V-4AQ.ipynb](#): missing data
- [10.25493-H38T-5C2](#): missing metadata about time when experiment started so impossible to analyse the data

Other thoughts:

- Sometimes the data available from the links in 'data availability' section in the publication was better than the data in the ebrains depository
- Also: Technologies advance very quickly and codes may not be usable; i.e. old python libraries need to be used to read certain codes so you need to download old version of libraries to be able to use the code
- The data descriptor is not very useful, often lacking numbers that can actually be used in the data analysis: for example cells 1-13 were studied under these conditions instead of describing the experiments that can already be found in the publications. Only organises the data, without explaining it
- Organisation problem: 6 different datasets for same publication

(<https://search.kg.ebrains.eu/?category=Dataset&q=The%20microcircuits%20of%20striatum%20in%20silico&experimentalApproach%5B0%5D=electrophysiology#f25ec9c9-2348-4a62-a6af-907d389f263d>,  
<https://search.kg.ebrains.eu/?category=Dataset&q=The%20microcircuits%20of%20striatum%20in%20silico&experimentalApproach%5B0%5D=electrophysiology#ba5c166a-6862-40ce-bfd7-7720b6e39015>,  
<https://search.kg.ebrains.eu/?category=Dataset&q=The%20microcircuits%20of%20striatum%20in%20silico&experimentalApproach%5B0%5D=electrophysiology#7d545473-405d-45f9-b0c4-05b627bf793a>,  
<https://search.kg.ebrains.eu/?category=Dataset&q=The%20microcircuits%20of%20striatum%20in%20silico&experimentalApproach%5B0%5D=electrophysiology#495de3d5-f9d5-4a4a-acd2-4ade0eaefb22>,  
<https://search.kg.ebrains.eu/?category=Dataset&q=The%20microcircuits%20of%20striatum%20in%20silico&experimentalApproach%5B0%5D=electrophysiology#e0cfa3fb-22d2-48c7-bb4c-a6374bfc848b>,  
<https://search.kg.ebrains.eu/?category=Dataset&q=The%20microcircuits%20of%20striatum%20in%20silico&experimentalApproach%5B0%5D=electrophysiology#02f1aec9-c709-4dd3-8929-3833569078eb>)