BASIC OUTLINE and NOTES

Define the problem

There is no consistent format or template for medical records

- often electronic medical records (EMR) are stored in various forms and formats
  - e.g. csv, xls, sas, etc...
- there was a big push to have all hospitals use the same systems
  - has not been done and does not look like it will happen soon
  - resistance from users, administrators, stakeholders

Offer solution: Introduce Pau and Tree Notation

- define a grammar for your tree language with respect to each EMR system
- NLP tools, possibly powered by machine/deep learning, can help to quickly define grammars and languages for EMR systems

Advantage of Pau

- hospitals and healthcare systems can continue their current workflow
  - charting and EMR entry can continue as usual
  - clinicians do not have to learn a new tool
- Tree Notation Works with paper offline as well.
- grammars, defined for each EMR system, can convert everything into a consistant tree notation format
  - efficiently aggrigate/curate data
  - medical field has entered the big data era
    - data consistency is important for big data analysis
      - e.g. machine/deep learning model construction
- Grammars can be Concatenated by simply concatenating 2 grammar files
  - systems can use only the grammars for the target domain
  - multiomics systems can use all grammars at once.
- include code snipits and examples of Pau
  - or include example of tree

        * the toCSV() example in the sandbox is very practicle and applicable to EMR
- Omnifix notation
- Grammars can be used to generate mock data
  - Code can be written and tested against mock data, and then run on data held in a black box to safeguard privacy.
- Internationalization
  - i18n'izing each grammar is as simple as adding 1 word per concept
- Wikipedia like crowdsourcing of a global grammar for EMR

A big prediction

- Our hypothesis: This is the simplest possible (as in, fewest symbols needed starting from 0 to 1) complete medical records system that can be devised. A simpler system will not be found.

On going and future work

- grammars for particular EMR systems are being defined
  - tools and techniques to optimize this process being worked on
  - a person/expert is needed to QA/QC the grammars etc
- potential use cases and example of how it optimizes analysis pipelines
- how Pau fits in the grand scheme or vision of tree notation