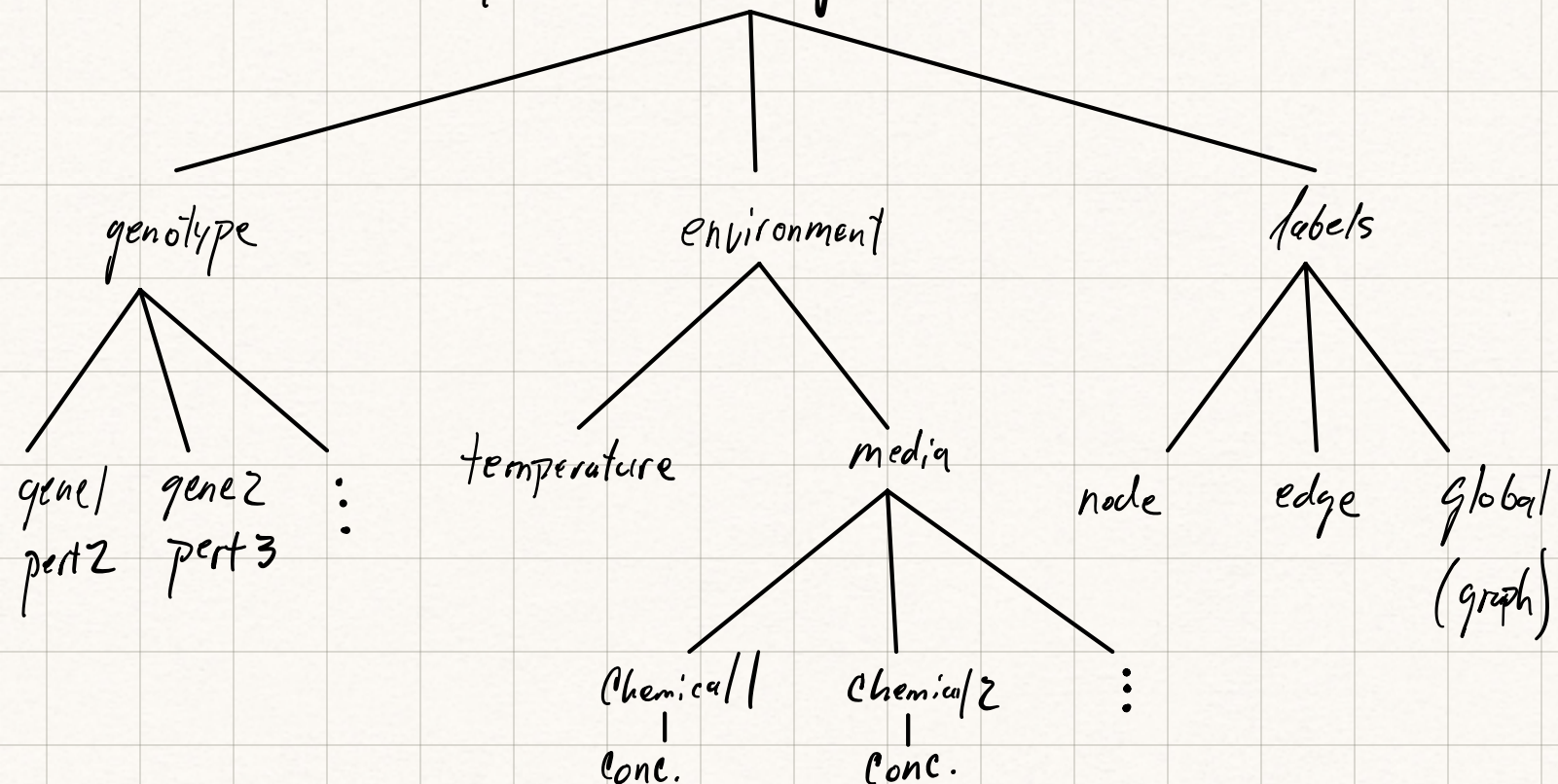


data issue with saving individual experimental data vs. saving all data.

- All data cannot be stored in common GPU memory
- 1.pt for each leads to IO bottleneck

Solutions:

- Optimize IO
- Optimize data transfer Globus - haven't tried tar before transfer
- Smart batching on save - this will avoid issues for subsequent joins
  - maybe can get around this during cell process writing



2023.09.05

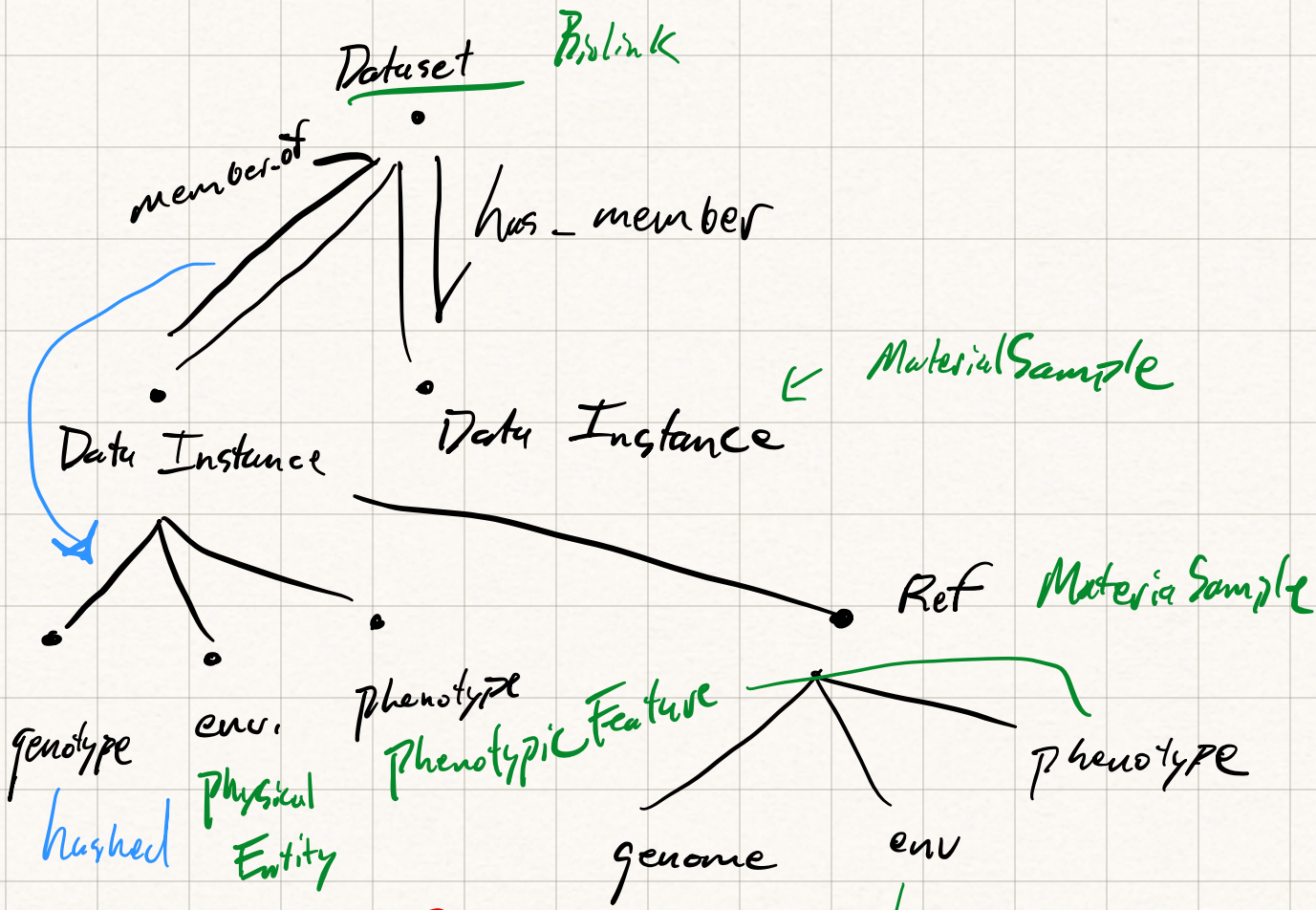
- First pass on using edge info will use fixed graphs. I want to do something like ms-gen but my graphs are sparse, at least PPI and Reg.
- I think there are a few ways to conceptualize this one with GNNs one with transformers.

2023.12.16

- Biolink Ontology

predicates : Concept level : has\_member  
member\_of

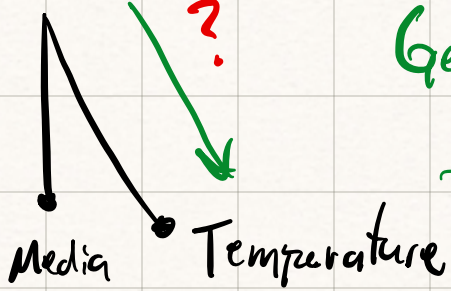
Instance level : has\_phenotype





Genotype

Genome



Chemical/Entity