Source	Entity List	How captured
Metamath	1029 Terms 27688 Propositions 28698 Formal proofs (All knowledge from metamath's set.mm and iset.mm datasets at the time of extraction.) 7 new PGMS (Metamath didn't have an automated proof-generating scheme.)	JSON files for this dataset were generated from metamath's primary proof assistant to extract all knowledge from metamath's primary logics (*set.mm* and *iset.mm*) Sophize's data extraction code in Metamath's primary **Java**-based proof assistant
Planetmath	1879 Propositions 4771 Terms (mainly definitions) 388 Proofs 1675 Articles	The JSON files for this dataset were extracted using python based crawlers. **Python** code Repository for data extraction
2 research papers	28 propositions 20 arguments (subproofs) 24 Definitions	**Manually** curated JSON files
Wikipedia	3685 Definitions	The JSON files for this dataset were extracted from wikipedia math glossaries. **Python** code Repository for data extraction