

DBLP Parsing and Loading into Database

Do-ables

- Download DBLP dataset available in XML form
- Design an ER Diagram to load entire data
 - Current records in the dataset are 4.7 million Publications (Oct 14, 2019)
- Issues to solve/learn
 - how to smartly read and process data for the proposed ERD using your PC (imagine your PC has very low RAM and Processing Power)
- Deliverables: A report, in MS Word form, containing
 - Proposed ER Model
 - Programming Code