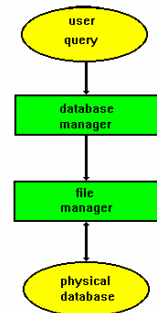


BINF 630: Introduction to Bioinformatics

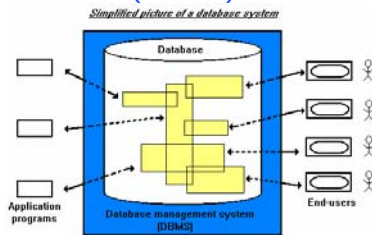
Iosif Vaisman

Email: ivaisman@gmu.edu

Database Organization



Database Management System (DBMS)



Four major components of DBMS:

Data * Hardware * Software * Users

Data Model

- A named logical unit (record type, data item)
- Relationships among logical units

Relationships among logical units

- one to one
- one to many
- many to one

Relational Database Model

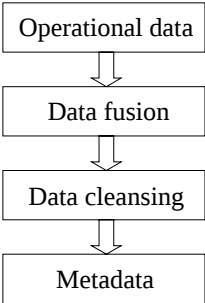
protab1			
Protein-code	Protein-name	Length	Species-origin
P1001	Hemoglobin	145	Bovine
P1002	Hemoglobin	136	Ovine
P1003	Eye Lens Protein	234	Human
.....			

protab2	
Protein-code	Protein-sequence
P1001	MDRTTHGFDLKLSPRTVNQWLMLALFFGHS...
P1002	MDRTSHGFELKLLTPKKLQQWLMIAIYFGHT...
P1003	SRTHEEGKLMQWPPRPLYALFTEPPYP...
.....	

Database administration

- Redundancy eliminated
- Inconsistency avoided
- Data shared
- Standards enforced
- Security applied
- Integrity maintained
- Requirements balanced

Data Warehouse



InterPro Database

