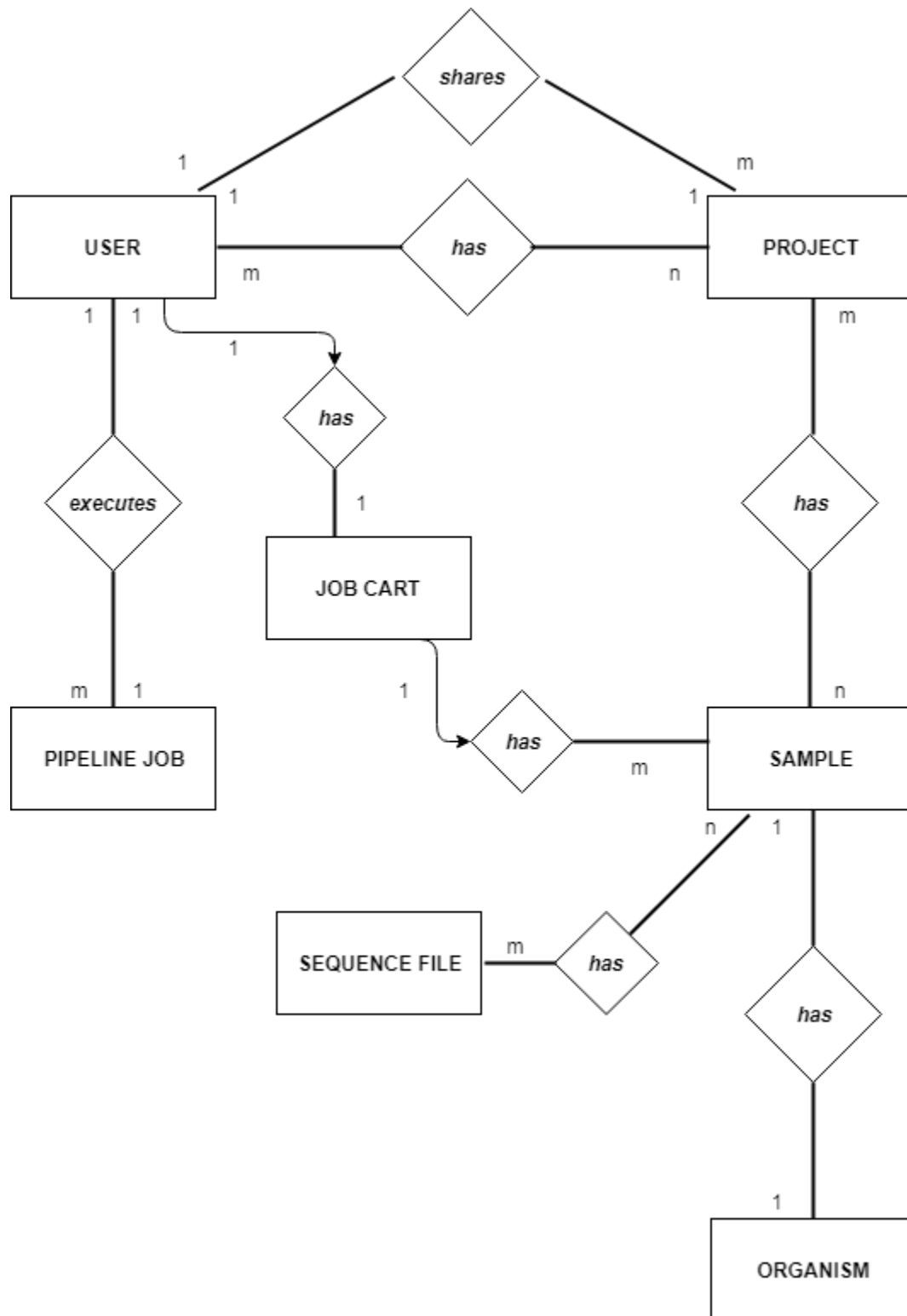


## PGDB : Database Design Document

### Assumptions

- An **Organism** associates with only 1 sample.
- A **Sample** associates with only 1 organism.
- A **Project** can have many samples.
- A **Sample** can belong to one/many projects.
- A project can have many users.
- A single **User** can belong to one/many projects.
- A user can execute many **Pipeline Jobs**.
- A pipeline job can be executed by a user.
- At a given time user can have only 1 **Job Cart**.
- A job cart can have many samples.
- A user can **Share** many projects.
- A given project can be shared only by a single user.
- A sample can have many **Sequence Files**.
- A sequence file can belong to many samples.

## ER Diagram



## **Table Structure**

Database: pgdb

MySQL credentials: user name: pgdb / password: pgdb\_0517

TABLE NAME: **organisms**

DESCRIPTION: *Stores organism related data*

<b>PK</b>	id	INT (6)
	name	VARCHAR

TABLE NAME: **projects**

DESCRIPTION: *Stores project related data*

<b>PK</b>	id	INT (6)
	name	VARCHAR
	description	VARCHAR

TABLE NAME: **samples**

DESCRIPTION: *Stores sample related data*

<b>PK</b>	id	VARCHAR (7)
	name	VARCHAR
FK	organism_id	VARCHAR
	create_date	DATETIME

TABLE NAME: **users**

DESCRIPTION: *Stores user related data*

<b>PK</b>	id	INT (6)
	name	VARCHAR
	email	VARCHAR
	user_name	VARCHAR
	encrypted_password	BLOB
	salt	BLOB
	is_admin	BOOLEAN
	dummy_password	VARCHAR
	is_pw_reset	BOOLEAN

TABLE NAME: **sequence\_files**

DESCRIPTION: *Stores uploaded fastq file related data*

<b>PK</b>	id	INT (6)
	file_path	VARCHAR
	create_date	DATETIME

TABLE NAME: **user\_projects**

DESCRIPTION: *Records relationship between users and projects*

<b>PK</b>	id	INT (6)
FK	user_id	INT (6)
FK	project_id	INT (6)

TABLE NAME: **project\_samples**

DESCRIPTION: *Records relationship between projects and samples*

<b>PK</b>	id	INT (6)
FK	project_id	INT (6)
FK	sample_id	VARCHAR (7)

TABLE NAME: **sample\_sequence\_files**

DESCRIPTION: *Records relationship between samples and sequence files*

<b>PK</b>	id	INT (6)
FK	sample_id	VARCHAR (7)
FK	sequence_file_id	INT (6)

TABLE NAME: **shared\_user\_projects**

DESCRIPTION: *Records project sharing data with the project owner and the shared user*

<b>PK</b>	id	INT (6)
FK	shared_project_id	INT (6)
FK	project_owner_id	INT (6)
FK	shared_user_id	INT (6)

TABLE NAME: **jobs\_cart**

DESCRIPTION: *Stores cart item data*

<b>PK</b>	id	INT (6)
FK	sample_id	VARCHAR (7)
FK	user_id	INT (6)
	project_name	VARCHAR

TABLE NAME: **pipeline\_jobs**

DESCRIPTION: *Stores data related to pipeline execution*

<b>PK</b>	id	INT (6)
FK	user_id	INT (6)
	name	VARCHAR
	type	VARCHAR
	file_name	VARCHAR
	file_path	VARCHAR
	input_path	VARCHAR
	output_path	VARCHAR
	status	VARCHAR
	created_date	DATETIME

Table Name: **pipeline\_jobs\_history**

DESCRIPTION: *Stores records deleted from the pipeline\_jobs table*

<b>PK</b>	id	INT (6)
	name	VARCHAR
	type	VARCHAR
	created_date	DATETIME