Mini-Case - Please select the best answer for the questions listed below based on the narrative, swimlane diagram and ERD that follows these questions:

1 .What would be the best name for the actor labeled "A"?

a. Lab technician
b. Data Analysis Software
c. Volunteer
d. Secondary assistant
e. None of the above

2. What would be the most appropriate question for the decision labeled "B"?

a. Professor available?
b. Worthy of filing?
c. Any anomalous results?
d. New discovery?
e. None of the above

3. What would be the most appropriate description of the process labeled "C"?

a. Produces sequences
b. Extracts genes
c. Removes liquid
d. Replicates genes
e. None of the above

4. What would be the most appropriate description of the process labeled "D"?

a. Runs algorithm
b. Conducts data analysis
c. Updates file server
d. Sends alert to assistant
e. None of the above

5. What would be the most appropriate question for the decision labeled "E"?

a. Sequencing successful?
b. Shift over?
c. Any more sequencing to be done?
d. Data accurate?
e. None of the above

6. What would be the best name for the entity labeled "A"?

a. Lab technician
b. Supercomputer
c. Sequencing machine
d. File server
e. None of the above

7. What would be the most appropriate description for the relationship labeled "B"?

a. Analyzes
b. Presents
c. Files
d. Organizes
e. None of the above

8. What would be the most appropriate description for the relationship labeled "C"?

a. Organized into
b. Processed by
c. Cleaned by
d. Converted into
e. None of the above

9. What would be the most appropriate description for the attribute labeled "D"?

a. Expiration date
b. DNA Sequences
c. Species
d. Source
e. None of the above

10. What would be the most appropriate description for the relationship labeled "E"?

a. Uploads to cloud
b. Extracts
c. Cleans and trims
d. Annotates
e. None of the above

Mini-Case - Read the following narrative carefully

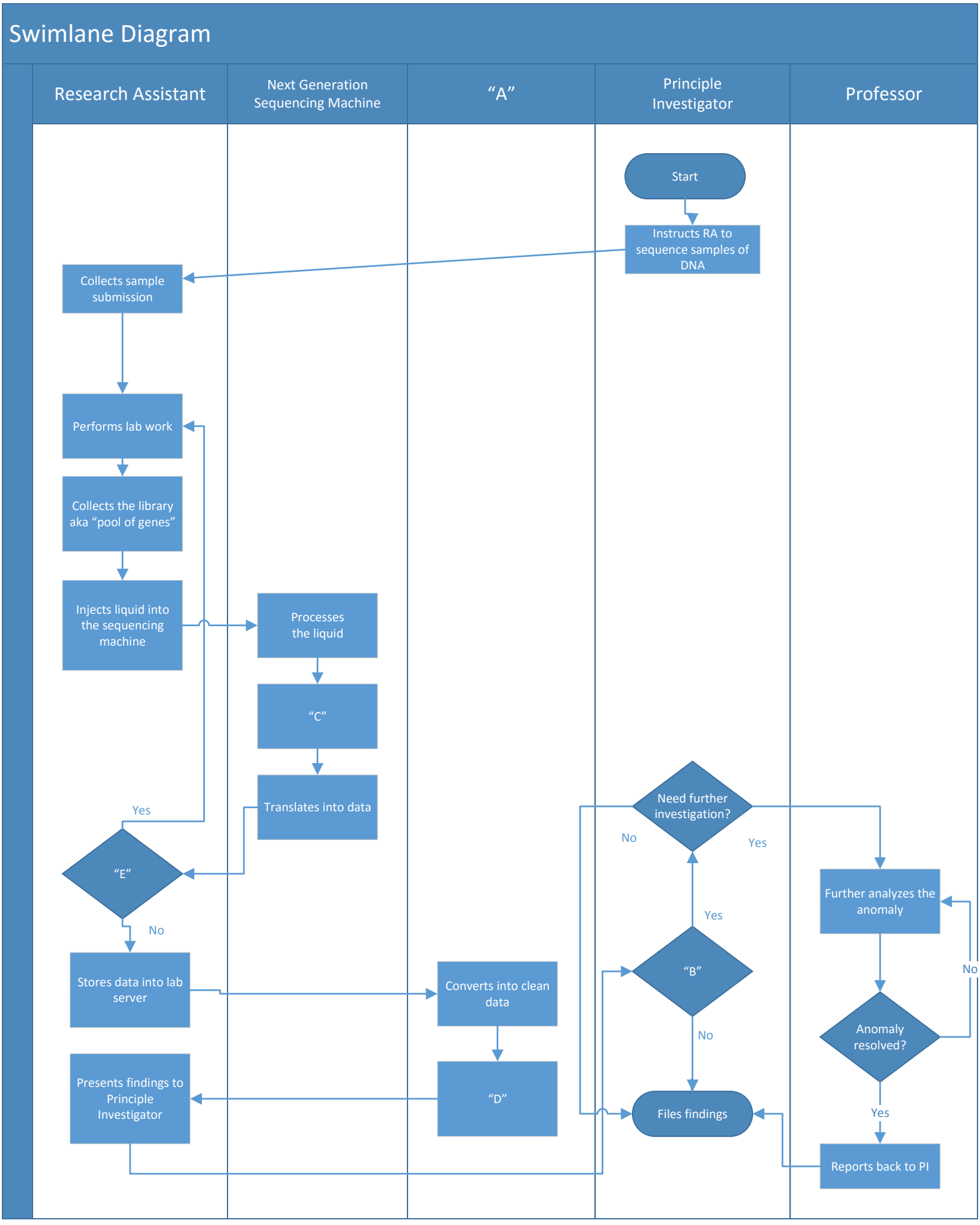
Catherine is a research assistant at University of Pennsylvania. She and her peers operate a wide variety of instruments that perform DNA sequencing, genotyping, gene expression analysis, DNA quantification and quality control. Katherine is primarily focuses on DNA sequencing core activities.

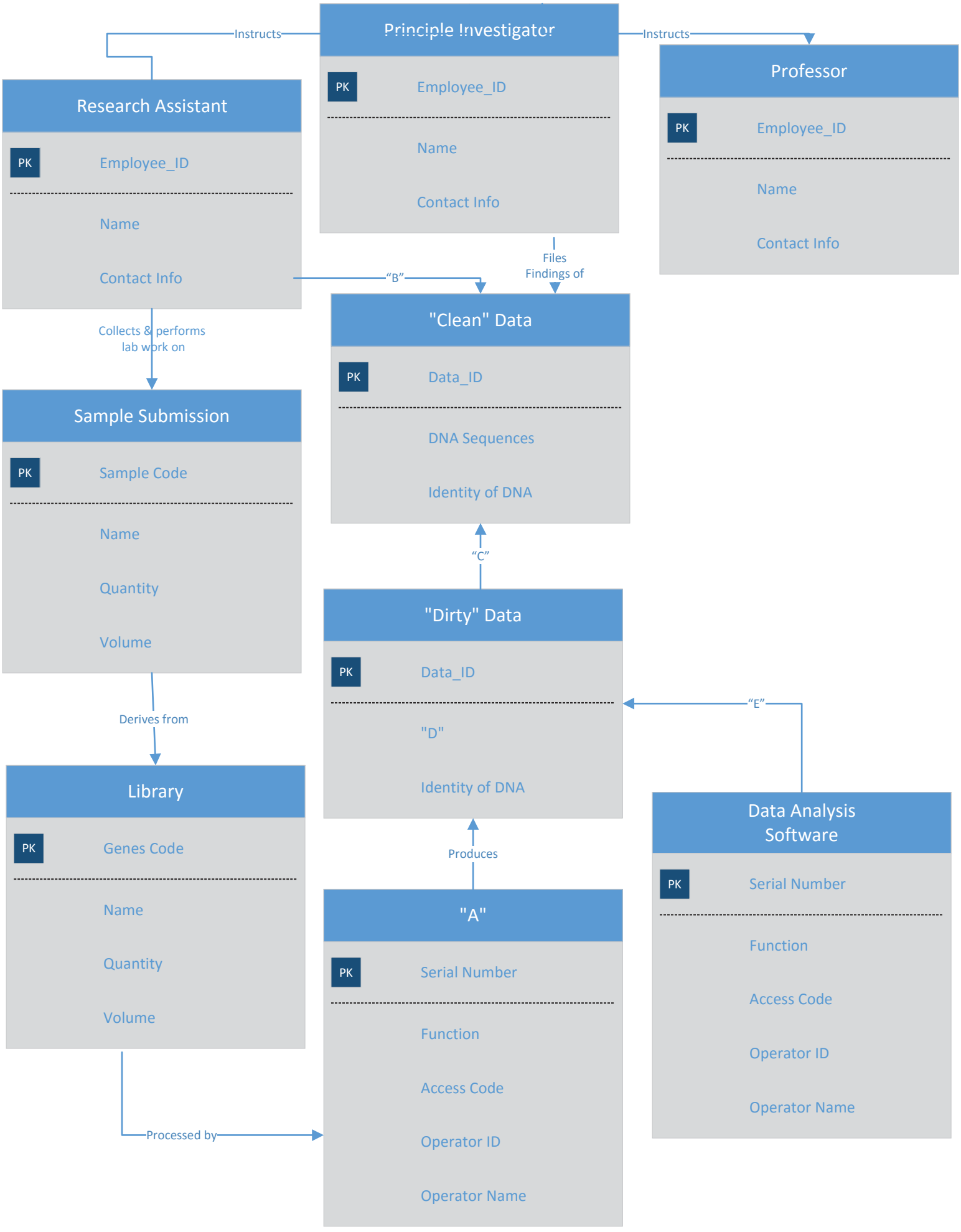
The DNA sequencing core process starts when she receives an instruction from her Principle Investigator to sequence a collection of samples of DNA. She will start the first stage of the DNA sequencing core, that is, library preparation. Library preparation entails a few steps, all done by Catherine. In the morning, she starts by collecting any sample submissions that are to be sequenced. Then, she conducts various lab work, at the end of which she derives the "library" - a liquid consisting of a pool of genes. This concentrated pool of genes is very important, in that it contains the DNA strands that are to be analyzed and sequenced.

Katherine then injects the library liquid into a next generation sequencing machine. This sequencing machine is able to conduct advanced sequencing and processing functions that humans simply cannot do on their own. The library is processed by the machine, and then the machine produces sequences which are subsequently translated into data.

Following the machine's work, Katherine checks if there is any more sequencing to be done by her. If there is, she performs her lab work again to collect another set of library. Otherwise, she can move on to the second phase of the process which is data analysis. First, she will retrieve the new data and store it into a database in her lab server – which stores huge data consisting of millions of sequences. After that, she proceeds to transform the data using a software so that it can be used and sold – the analysis software cleans and trims the dirty data. The data is converted from a raw and unstructured form to a clean and organized form. After the data has been cleaned, the software conducts data analysis to find out what sequences are in the original samples.

Sometimes there will be some unusual findings in the sequence, like gene defects, mutation, deletion in the genes, etc. Regardless, she will present the cleaned data to her Principle Investigator (PI). If there are no anomalous results, the PI will file the findings and the process ends. If there are any anomalous results, the PI will then determine if it is necessary to further investigate the anomalous results. Sometimes, the PI will decides that the anomalies are inconsequential and not worthy of further investigation, so the process ends with him filing the findings. If the PI deems it necessary to further investigate the anomaly, he will delegate the task to professors. Upon analysis, if the professor solves the anomaly, he will report back to the PI and the latter will file the findings. Otherwise, the professor will do further analysis and this will continue until anomalies are resolved and the findings can be filed.


Swimlane Diagram
Research Assistant
Next Generation Sequencing Machine
"A"
Principle Investigator
Professor
Start
Instructs RA to sequence samples of DNA
Collects sample submission
Performs lab work
Collects the library aka "pool of genes"
Injects liquid into the sequencing machine
Processes the liquid
"C"
Translates into data
Yes
"E"
No
Stores data into lab server
Converts into clean data
"D"
Presents findings to Principle Investigator
Need further investigation?
No
Yes
Yes
"B"
No
Files findings
Further analyzes the anomaly
Anomaly resolved?
No
Yes
Reports back to PI


Principle Investigator
PK
Employee_ID
Name
Contact Info
Instructs
Instructs
Research Assistant
PK
Employee_ID
Name
Contact Info
Professor
PK
Employee_ID
Name
Contact Info
"B"
Files
Findings of
Collects & performs
lab work on
"Clean" Data
PK
Data_ID
DNA Sequences
Identity of DNA
Sample Submission
PK
Sample Code
Name
Quantity
Volume
"C"
"Dirty" Data
PK
Data_ID
"D"
Identity of DNA
"E"
Derives from
Library
PK
Genes Code
Name
Quantity
Volume
Produces
Data Analysis
Software
PK
Serial Number
Function
Access Code
Operator ID
Operator Name
"A"
PK
Serial Number
Function
Access Code
Operator ID
Operator Name
Processed by