Kristianstad University

WATER-BACTERIA PROJECT

Group 1

Alm Robert, Lavdim Imeri, Singh Vipin

REVISION HISTORY

Date	Version	Description	Author
30/03/2022	0.1	Initiating the project artefact. Wrote a draft of Vision and Introduction.	Alm Robert, Lavdim Imeri, Singh Vipin
11/04/2022	0.2	Updating Introduction. Initiating Requirements, Risks and Tests.	Alm Robert, Lavdim Imeri, Singh Vipin

Kristianstad University

Contents

Vision

(Write shortly what this document includes – as an abstract)

Our vision is to offer a cheap, automatic, safe, and ongoing improved method of analyzing water, using flow cytometry, producing in that way an output that can be used to draw useful conclusions about the safeness or the purity of the water and if it is possible to be used for further research, (other methods of analysis).

As a deeper goal we aim to the possibility to isolate characteristics that can indicate things like hazardless of a bacteria or clear indicators if the water is contaminated or not, to be able to analyze the water faster and more efficiently with a lower computational cost.

Introduction

Water is one of the most essential elements for the preservation of life, the survival of humankind and the development of human civilization. In modern era, while water still can't be considered yet as a something granted for everyone, human civilization is at ongoing effort to make access to water safer, easier, and cheaper. One way to make water safer and cheaper, improving that way the access to water, is by improving the method of analyzing the safeness of the water.

That is a goal of great importance as the access to clean water is one of the major goals of the 2030 Agenda and its 17 Sustainable Development Goals, (SDGs), with SDG 16 being called "Clean Water and Sanitation" and at least other 4 of the 17 SDGs being affected by the availability and access to clean water. For example, the access to clean water is essential to the production of food, thing that calls for SDG 2, (Zero Hunger), and it is an integrated part for the goal to build sustainable cities, a matter that is addressed by SDG 11, (Sustainable Cities and Communities).

A solution to the challenge of analyzing water efficiently comes from a process that is called "Flow Cytometry" and it is a process that contains of a water sample are beamed with a laser beam in order to take measurements from the illuminance and the color that several particles emit due to their fluorescent agents. This methodology was used in the past for the analysis of blood samples but now is introduced to the field of water analysis.

Flow cytometry

Flow cytometry is a technique for detecting and quantifying the physical and chemical properties of a population of cells or particles. A sample, including particles, is suspended in a fluid and injected into a flow cytometer equipment during this procedure. The flow cytometry method is used for identifying to assessing peripheral blood, bone marrow, and other bodily fluids in solution.

A flow cytometer is comparable to a microscope, except instead of creating a picture of the cell, it provides high-throughput, automated quantification of specified optical characteristics on a cell-by-cell basis. A single-cell suspension must first be created before solid tissue analysis can begin.

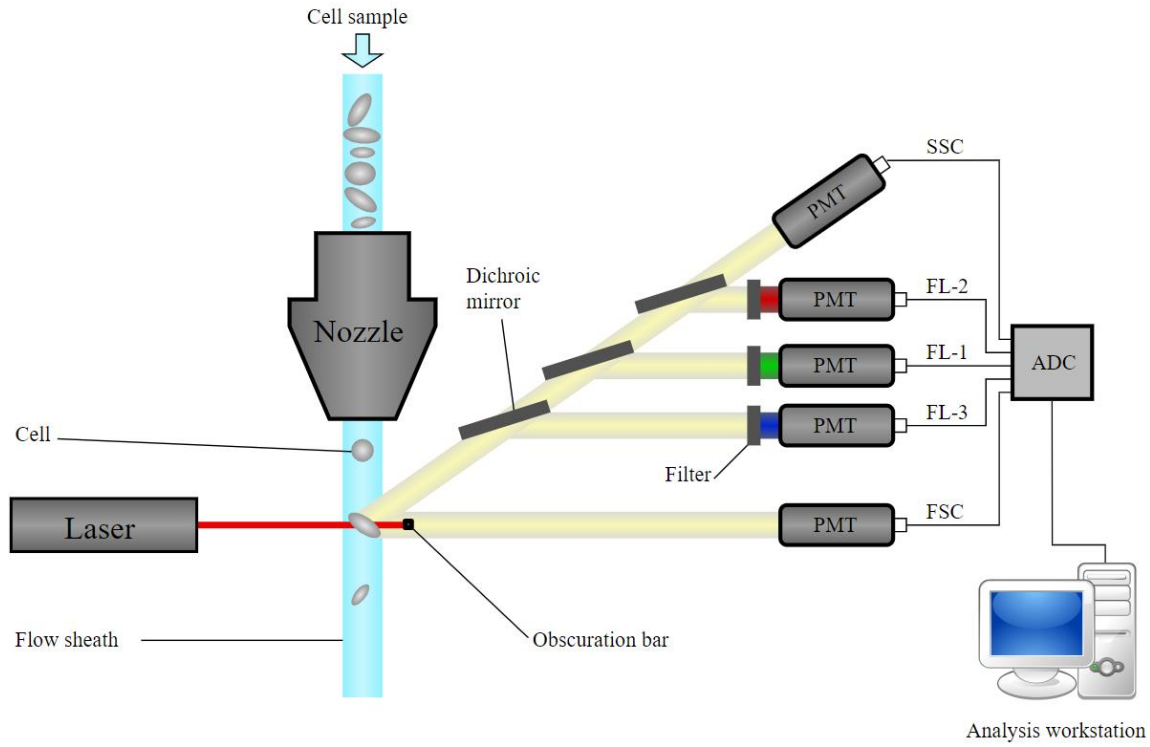
Process of flow cytometry

A cell or particle sample is suspended in a fluid and injected into a flow cytometer equipment.

The sample is concentrated such that it flows through a laser beam one cell at a time, with the light dispersed being unique to the cells and their components.

Fluorescent markers are frequently used to label cells, causing light to be absorbed and subsequently released in a spectrum of wavelengths.

Thousands of cells may be analyzed fast, and the information acquired is processed by a computer.



A flow cytometer is comprised of five key aspects:

1. Flow cell
2. Measurement system
3. Detector
4. Amplification system,
5. Computer for signal processing.

A liquid stream (sheath fluid) runs through the flow cell, carrying and aligning the cells such that they pass a single file through the light beam for sensing.

Measurement of impedance (or conductivity) and optical systems are often used in the measuring system. - mercury and xenon lamps; high-power water-cooled lasers (argon, krypton, dye laser); low-power air-cooled lasers (argon (488 nm), red-HeNe (633 nm), green-HeNe, HeCd (UV)); diode lasers (blue, green, red, violet).

Analog measurements of forward-scattered light (FSC) and side-scattered light (SSC), as well as dye-specific fluorescence signals, are converted into digital signals by the detector and analogue-to-digital conversion (ADC) system. A linear or logarithmic amplification mechanism can be used.

The word "acquisition" refers to gathering data from samples using a flow cytometer. A computer physically linked to the flow cytometer and software that handles the digital interface with the cytometer is used to facilitate the acquisition.

The program may alter settings for the sample being tested (e.g., voltage, compensation) and show initial sample information while obtaining data to check that parameters are set appropriately. Flow cytometers were first considered experimental equipment, but

technological advancements have allowed them to be widely used for therapeutic and scientific objectives.

As a result of these discoveries, a sizable industry for instruments, analytic software, and acquisition reagents such as fluorescently tagged antibodies has emerged.

Requirements

Requirement items	Priority
R1. User registration: Register new user to the system.	**High**
R2. User login: Log the user into the system.	**High**
R3. Admin login: Log the admin into the system.	**High**
R4. Upload FCS files in batch mode.	**High**
R5. Visualize the pre-processed results through web application.	**Medium**
R6. Filter out the noise in the FCS files.	**High**
R7. Store pre-processed data in cloud database.	**High**
R8. Analysing the fcs files at the server and parse .FCS to .CSV.	**High**
R9. Clustering the data by using different clustering methods.	**High**
R10. Providing a written report of the session.	**Low**

Requirement 1. User registration: Register new user to the system.

Description: It is the action of registering a new user to the system. Logically they will be 2 different types of registration. The one that is made by the users themselves for themselves, and a secondary that is applicable from the panel of the admin, (as a “Create a User” action). When the registration happens through the admin panel, the admin can give the Admin property to the user with the privileges that an Admin has.

Requirement 2. User login: Log the user into the system.

Description: It is the act when the user, (or the admin), provides the correct credentials to the system and gains access to the system. The login credentials are expected to be the email and the password that the user has defined for their account.

Requirement 3. Admin login: Log the admin into the system.

Description: It is the act when the admin, (admins are actually users with more privileges), provides the correct credentials to the system and gains access to the system. The login credentials are expected to be the email and the password that the user has defined for their account.

Requirement 4. Upload FCS files in batch mode.

Description: FCS files need to be uploaded from the user interface to a google storage service.In this process the user chooses the files that needs to be uploaded. While uploading, a status bar will show the progress of upload for each file.

Requirement 5. Visualize the pre-processed results through web application.

Description: The user chooses the file that needs to be shown. A

scatter plot will show the cells in the sample using two values FL1-A, Fl2-A as the two coordinates.

Requirement 6. Filter out the noise in the FCS files.

Description: The uploaded by the user data is expected to contain unnecessary information and impurities that can affect the results of the analysis. A user defined gating process will be used define the useful data and omit the data that does not contribute to the analysis.

Requirement 7. Store pre-processed data in cloud database.

Description: The collected and filtered data, is useful to be stored into a database so it will be available for processing or analysis. Understandably the data on the database needs to be mapped with the corresponding FCS file.

Requirement 8. Analyzing the fcs files at the server and parse .FCS to .CSV.

Description: It is necessary for our system and for any further analysis to have a conversion from FCS to CSV especially when the data is processed and analyzed properly.

Requirement 9. Clustering the data by using different clustering methods.

Description: Clustering is an important analysis tool that our system wish to utilize. To make the system more effective, the use should have the option to select between different clustering algorithms.

Requirement 10. Providing a written report of the session.

Description: It is expected that the user will need to have the analyzed data in a human readable report, so the system is expected to be able to deliver a structured documentation of the analysis and its results.

Supplementary Requirements

Supplementary requirements	Priority
SR1. Availability	High
SR2. Robustness	High
SR3. Graciousness	Medium
SR4. Scalability	High
SR5. User Friendliness	Medium
SR6. Efficiency	High

Supplementary Requirement 1. Availability
Description: The system needs to be available to the users 24/7.

Supplementary Requirement 2. Robustness
Description: The code needs to be well written and documented with as few bugs or errors as possible.

Supplementary Requirement 3. Graciousness
Description: The system needs to have a smooth progress without sudden activity spikes or unexplained behavior.

Supplementary Requirement 4. Scalability
Description: The system needs to be able to handle large amounts of workload.

Supplementary Requirement 5. User Friendliness
Description: The user interface needs to be intuitive and easy to use.

Supplementary Requirement 6. Efficiency
Description: The system needs to perform as efficiently as possible and to consume as less resources as possible.

Risks

Risk items	Priority
R1. Excessive Workload	**Essential**
R2. Compromised Personal Productivity	**Desirable**
R3. Git Related Issues	**Essential**
R4. New Technologies	**Essential**
R5. External Factors	**Essential**
R6. Third Party Services	**Essential**
R7. Bugs and Vulnerabilities	**Essential**
R8. Compromised Performance	**Essential**

Risk 1. Excessive Workload

Description: Even if we are confident that we can proceed to the completion of the project without difficulties, it is clear to us that it is a complicated project that requires high standards of quality, and a lot of effort. The possibility of assign to ourselves too many tasks and an amount of workload that we can't deliver, it is a factor that we should be very aware of.

Mitigation Strategy: we need to be aware about the size of workload, and how long time it will take us to complete each task

Risk 2. Compromised Personal Productivity

Description: We are really confident that everyone in the team will do their very best while working on the project, but we should always be aware of the unforeseen factors that can affect someone's productivity, like sickness, personal life problems, exhaustion, external factors, etc.

Mitigation Strategy: Each task should involve the three of us, with one of us doing the task the two serving as consultants. That tactic increases our knowledge pool, and ensures that if someone of us is unable to complete the task, someone else will jump and take over from the spot the other person stopped.

Risk 3. Git Related Issues

Description: Git is an important part of our programming culture, and we are highly trained on the use of Git/GitHub, but factors like, external promoted strategies, third party integrations, and

the typical "human mistake" factor can lead to conflicts, that are not going to compromise our project but they are capable to delay us, sometimes even for days.

Mitigation Strategy: Ownership of our own strategy, even if our strategy will have to conform to other strategies. We have the be the rulers of our repositories at all times.

Risk 4. New Technologies

Description: New technologies is a common point of failure, as we are not fully educated them about them, we are not fully aware of all the issues that may be hidden behind the new technology, and of course there is a whole set of unforeseen factor behind each new technology.

Mitigation Strategy: Avoid new technologies. Study extensively the documentation behind each new technology that we tend to use.

Risk 5. External Factors

Description: From Dramatic global events, like pandemics or wars to more subtle events like incosistencies on the past research or change of the requirements, can affect our project in various ways.

Mitigation Strategy: A good documentation and good cooperation between the team members can assure that whatever happens the team will be consistent towards the completion of the project.

Risk 6. Third Party Services

Description: Depedency on third party products or services is always something negative because of the depedency to their contribution. If a third party service change or cease to exist, it is expected that they will not take into consideration the needs of our project.

Mitigation Strategy: Avoidance of third party services if possible. Usage of the most reliable third part services. Study throughly the documentation of the use third party services, (or products).

Risk 7. Bugs and Vulnerabilities

Description: Software Bugs can compromise the performance of a system dramatically or even damage the system and its reliability. System vulnerabilities, when not detected or fixed can compromize the security of the system.

Mitigation Strategy: Extensive Testing, andlysis of the results, patching, and of course peer review.

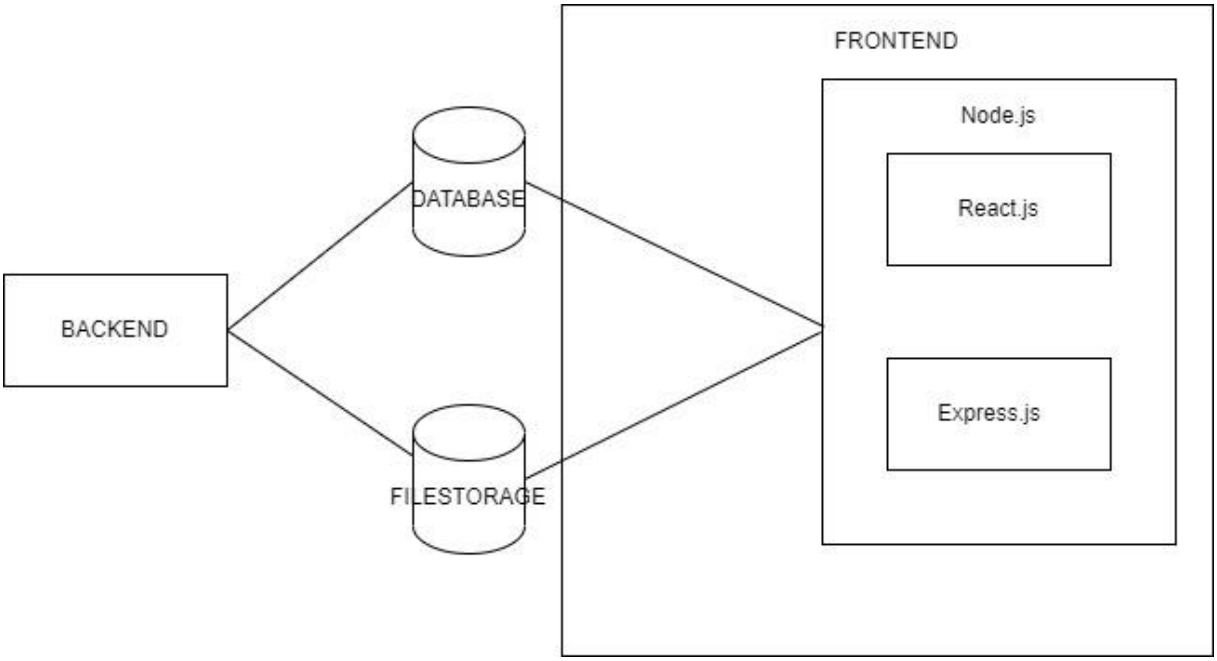
Risk 8. Compromised Performance

Description: There are many reasons why a performance of a system is compromized. Sometimes it can be a software bug, a badly designed component of the system, faulty hardware, or simply inadquate design.

Mitigation Strategy: The understanding of the reasons why systems perform badly sometimes can help us having ways to mitigate the issue. Access to several experts and various technical support services can mitigate this risk drastically. Algorithmic analysis and data structures need to be taken into serious consideration as they are usually the number one reason for bad perfomance.

Design

Design items	Priority
D1. Front-End	**Essential**
D2. Back-End	**Essential**
D3. File Storage	**Essential**
D4. Database	**Essential**



The basic design idea

Text

Design Item 1. Front-End

Description: It is clear that the front end will be On React.js Instance, and a special attention will be given to the responsive design, possibly with the use of responsive design.

Design Item 2. Back-End

Description: Back End is expected to be A Python instance with utilization of the Django Framework, in order to be able to serve as a standalone backend server with the ability to receive, process and return large amounts of data in different forms.

It is expected that the Python instance will be formatted inside a GitHub repository, so it will be able to be cloned inside a VM, (Linux probably), and with the support of Server Software that is not specified yet.

Design Item 3. File Storage

Description: While, for the file storage, we will not have to implement a special instance, The google cloud storage instance has to be set up correctly, and 2 different interfaces, (backend and frontend), will have to have specific functionality implemented to be able to communicate and interact with the google cloud storage image.

Design Item 4. Database

Description: Database has to have 2 different functionalities, namely the storage of the processed data, and the storage of user credentials. Because the different functionalities have different requirements for reliability and performance, the use of two different databases may be recommended. The Database component need to have interface on both front- and backend.

Tests

Tests	Passed/Failed
T1. User registration	**Not Tested Yet**
T2. User login	**Not Tested Yet**
T3. Admin login	**Not Tested Yet**
T4. Test FCS Files Uploading in Batch Mode.	**Not Tested Yet**
T5. Test Visualization of pre-processed results through a web application.	**Not Tested Yet**
T6. Test Noise Filtering in the FCS files.	**Not Tested Yet**
T7. Test the storing of pre-processed data in cloud database.	**Not Tested Yet**
T8. Test the Analysis of fcs files and the parsing from .FCS to .CSV.	**Not Tested Yet**
T9. Test the Clustering of the data by using different clustering methods.	**Not Tested Yet**
T10. Test the generation of a result report.	**Not Tested Yet**

Test 1. User registration

Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 2. User login

Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 3. Admin login

Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 4. Test FCS Files Uploading in Batch Mode.
Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 5. Test Visualization of pre-processed results through a web application.
Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 6. Test Noise Filtering in the FCS files.
Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 7. Test the storing of pre-processed data in cloud database.
Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 8. Test the Analysis of fcs files and the parsing from .FCS to .CSV.
Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 9. Test the Clustering of the data by using different clustering methods.
Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Test 10. Test the generation of a result report.
Testing Procedure: Not Defined Yet

Expected Results: Not Defined Yet

Evaluation or Analysis of Test Results

Conclusion

References

(Use IEEE reference system. It is in-built in Word.)

(Books)

(Articles and Journals)

(WWW)