Product Requirements Document (PRD)

1. Problem Description

The elements, processes, interactions, and dynamics of biological systems are often too opaque or too complex to study in real life. Questions regarding the origins of life, the evolution of multicellular organisms, and population dynamics remain unanswered and, in some cases, unanswerable by conventional means of study. Experimentation via the isolation of variables is almost impossible, and much of the historical record in which these answers lie is unavailable.

Existing artificial life simulations, which seek to remedy some of the above problems, often simplify and abstract biological phenomena to the point of providing few significant insights. On the other hand, some simulations focus on realism and detail so much that there is no opportunity for unexpected discoveries. Between these two extremes, there is an under-developed middle-ground in which novel simulations can accurately model cellular biology while having a level of abstraction that allows for emergent behavior at scale.

1.1. Scope

The scope of this project is to develop an application that simulates realistic cellular life forms in a top-down-view two-dimensional environment. The user should be able to create, run, save, load, configure, and analyze simulations through a graphical user interface (GUI). The simulation should demonstrate energy metabolism, evolution (genomics, speciation, reproduction, etc.), and several features not present in existing simulations. Novel features may include regulatory gene sequences, high-fidelity chemical metabolism, intra- and extracellular chemical signals, and intercellular resource sharing.

1.2. Use Cases

Our artificial life simulation's primary objective is to provide a comprehensive and modifiable simulation environment which represents a realistic and dynamic ecosystem as it develops across generations.

Primary Use Cases:

- Scientists & researchers will use the digital simulation environment to test hypotheses or observe complex biological interactions and processes.
- Educators and students will use the platform as an academic tool to understand or represent biological processes like evolution and competition.
- Hobbyists and general enthusiasts of technology or simulations will use the platform as a source of recreation, taking interest in the complex visuals or biological interactions occurring.

2. Purpose and Vision (Background)

Our purpose is to create a high-fidelity simulation of microorganisms that demonstrates known biological phenomena, facilitates understanding of the basic elements of life, and invites observation and experimentation. Ultimately, the application should be a virtual "ant farm," allowing users to set conditions and see how they unfold.

Current artificial life simulations achieve parts of this, but none provide the same balance between abstraction and detail that this application is intended to have. Our application seeks to demonstrate population-scale behavior, such as evolution, while also simulating internal cellular processes. The idea is to give "no free lunches" to the simulated organisms but still show realistic emergent properties.

3. Stakeholders

- Evaluators
    - William Pfeil
        - Project partner and end user
        - Key decision maker regarding requirements and acceptance criteria
        - Requires at least biweekly progress reports
    - Ananya Sundararajan
        - Team management evaluator
        - Key decision maker regarding team management materials and project success
        - Requires weekly progress reports
- Engineering Team
    - Project participants
    - Key decision makers regarding design and implementation
    - Require weekly progress reports
- End Users
    - Key decision makers regarding usability and acceptance criteria
    - Require documentation and periodic release versions

4. Preliminary Context

4.1. Assumptions

- A majority of users will be inexperienced with artificial life simulations so there needs to be easy-to-understand and modifiable controls to run, save, and skip through simulations.
- We can assume researchers and those experienced in the field who will use the simulation would like access to more specific information on the generations, species,

and genomes they are seeing. They may also like more specific tools for viewing and modifying the simulation.

- A majority of users will be inexperienced with the data surrounding genomes or do not care about the emerging biological phenomena occurring, and instead will be more interested in the visually interesting nature of artificial life simulations.

4.2. Constraints

- The product will be developed and tested mostly on Windows computers. The product must be able to perform adequately on common hardware.
- We must complete a minimum viable product in three months and the entire project in nine months.
- Our team cannot be committed full-time to this project. Our team will mostly work asynchronously.
- The project partner may only be available for our weekly meeting time.

4.3. Dependencies

- We're dependent on the creation of an initial genome and simulation environment before the UI and database may be created. This is also required for the integration of modifiable features of the genome.
- We're dependent on a robust simulation engine (Godot) which allows us to separate our simulation functionality from our graphic frontend to maintain efficiency.
- We're dependent on a range of machines to test our simulation's efficiency with varying computational power.

5. Market Assessment and Competition Analysis

Alternatives:

- **The Life Engine** (https://thelifeengine.net/), **Conway's Game of Life** (https://conwaylife.com/), **Biomarker CA** (Information: https://www.youtube.com/watch?v=I2gkH1d1K5g): These simulations use pixel-sized cells as an indivisible unit and only model simple metabolisms (the presence or absence of food or neighbors). Internal cellular dynamics, an important element of the life of microorganisms, is abstracted away.

- **OpenWorm** (https://openworm.org/): This is an extremely high-fidelity simulation of a single, simple multicellular organism. While it is highly accurate, it lacks variety and the opportunity for discovery through unexpected emergent behavior.

- **Creatures** (Information: https://www.youtube.com/watch?v=Y-6DzI-krUQ): This game models metabolism, genetics, simple intelligence, and learning in detail, but it does so through a limited and outdated framework. Modern computers can allow for a more fluid structure with a finer granularity. The progression of the simulation also requires user input.

- **The Bibites** (Information: https://www.youtube.com/@TheBibitesDigitalLife), **Dylan Cope's Simulation** (Information: https://www.youtube.com/watch?v=fEDqdvKO5Y0): These simulation model organisms with a level of detail close to the intended level of this project. For example, they simulate genetics, metabolism, reproduction, and chemical signals. However, these simulations model organisms with pre-set structures such as stomachs and brains. This project intends to make organisms more flexible and diverse.

The above alternatives can all provide inspiration for the features and implementation of this project, but differences between each and this project's vision are too great to adapt between.

Libraries and game engines that provide GUI features already exist. This project will utilize the Godot game engine for these functionalities.

6. Target Demographics (User Persona)

- Tucker is a 20-year-old male college student majoring in Biology. He spends his free time either at home playing games on his computer or outside hiking on the nearby trails. He has experience using computers and using some other simulation software from his other science classes. Tucker also has some knowledge of microbiology from the various classes he has taken.

- Kelly is a 42-year-old female 7th-grade science teacher making a little over 50k a year in Missouri. She spends her weekends mainly at home watching TV while grading assignments or going for walks around her neighborhood. She has some experience with computers at home and in the workplace but always calls the school tech if something isn't working properly. Kelly is always looking for cool new ways to teach her students about science.

- Sarah is a 15-year-old female high school student who loves using computers. She spends her free time inside at home playing video games or learning to code. Sarah has lots of experience with computers and knows some of the basics of programming. Sarah also loves seeing other cool projects that people make.

- Roger is a 45-year-old male research professor mainly researching micro-biology. He spends the majority of his free time either playing with his kids outside or cooking meals

for his family. Roger is pretty good with computers since having to use tech over the pandemic.

7. Requirements

7.1. User Stories and Features (Functional Requirements)

User Story	Feature	Priority	GitHub Issue	Dependencies
As a researcher, I want to observe simulated organisms' statistics as they develop over generations so that I can understand how their traits and complex interactions evolve over time in response to their environment.	Simulation statistics	Must Have	Issue #15	Issue #16
As a teacher, student, or researcher, I want to observe a realistic simulation of microorganisms so that I can study real biology through this program.	Cell life, interactions, and visuals	Must Have	Issue #16	Issue #14
As a student with an interest in microbiology, I want to see and control the progression of simulated cellular organisms through time so that I can learn about the evolution and behavior of microbes.	Playback controls	Must Have	Issue #17	Issue #1
As a teacher, student, or researcher, I want to observe a realistic simulation with a dynamic environment so that I can study changing biological systems.	Simulation environment	Must Have	Issue #20	Issue #17
As a teacher I want an introduction to the program so I can use it for the first time and watch an interesting simulation of some organisms.	UI tutorial	Could Have	Issue #8	Issue #14

7.2. Non-Functional Requirements

Requirement	Metric
Simulations run smoothly, consistently, and without interruptions.	Frame rate never dips noticeably; no crashes
Simulations should demonstrate evolution	Organism genetics respond to environment factors via natural selection
Simulations should demonstrate speciation	Organisms adopt different genetic patterns for different niches

7.3. Data Requirements

The application will only need to manage simulation data while running.

- Current organisms that are present in the simulation and their information
    - Genome: the structure that manages genetic information
    - CellState: the structure that list all custom organism variables such as resource amounts and physical parameters
    - CellStructures: a list of simulated organelles and other sub-structures within a cell
    - Lineage: a history of all cells, living and dead, in the current simulation
- The user may set simulation parameters before execution.
    - Parameters of stored simulations are static and cannot be changed after initialization.

7.4. Integration Requirements

The application will be developed in and integrated with the Godot game engine. Any external interfaces will be managed by Godot.

7.5. User Interaction and Design

The user interface design will adhere to the following principles:

- Visibility: options and actions are visible to the user

- Feedback: the user receives notification of the outcome of actions
- Constraints: available actions are restricted to only those that are valid and necessary
- Consistency: the application maintains a standard pattern of looks and behavior throughout; that pattern matches the common practices of other applications
- Affordance: the appearance of elements matches how they are used

The user interface will be designed at a later stage.

8. Milestones and Timeline

- 0.1.0 Release - End of fall term (December 15th) GitHub Issue
    - The genome and simulation environment must be operational in its basic form.
    - Version 0.1.0 should be tagged and documented as the initial working version.
    - The environment should successfully initialize with a specified number of nodes.
    - Implementation of a time progression mechanism where "time" can be tracked and visualized within the simulation.
    - At least one type of specified interaction or rule must be clearly implemented and functional between the nodes.
    - This interaction should be observable and should change according to the passage of time within the simulation.
    - The system should provide a basic readout or dashboard.
    - This readout must display the total number of nodes at any given time.
    - All interactions or significant events occurring between nodes should be logged or marked in the readout in real-time or as a summary.
    - **Potential Dependencies:**
        - Select simulation framework (unless created in backend source for portability)
        - Select simulation source language
        - Testing framework (decide parameters for efficiency and emerging behaviors)

        - Version control system setup
- 0.2.0 Release - End of winter term (March)
    - Statistics generated and viewable by user, issue
    - Organisms will model realistic cells as determined by this issue
    - User can control simulation playback and save and load simulations, issue
    - Simulation can demonstrate biological phenomena as determined by this issue
- 1.0.0 Release - Mid-spring term (May 26)
    - Project is polished and ready for users to use and view
    - Organisms demonstrate divergent evolution and can adapt to face challenges

9. Goals and Success Metrics

Goal	Metric	Baseline	Target	Tracking Method
Interface Usability (GUI specific)	User Acceptance Testing	5/10 across all categories	8/10 across all categories	User study
Educational Value	User ratings	Moderately informative	Moderately to highly informative	User survey
Performance	Stress Testing, User Acceptance Testing	Simulation executes with minor stutters, crashes, errors, etc. does not largely affect simulation	Simulations execute with no stutters, crashes, errors, etc.	User survey, demo testing

10. Open Questions

- Will development continue after this year?
- What is the best set of parameters to display emergent behavior?

11. Out of Scope

There will be no networking capabilities to the application.

The application will not be supported on mobile devices.

Simulations will be two-dimensional.

Simulations will have a set of parameters that are fixed at creation.

Security will be secondary to performance.