

Universidade Federal do Rio Grande do Norte

Curso: Banco de Dados - DIM0125

Relatório do Projeto de Sistemas de Banco de Dados

Aluno: Douglas Felipe de Lima Silva

Matrícula: 20220054131

Aluno: Elildes Fortaleza Santos

Matrícula: 20240078023

Aluno: Felipe Marley de Oliveira Gomes

Matrícula: 20210053408

Professor: Marcel Vinicius Medeiros Oliveira

31 de março de 2025

Sumário

1	Concepção e Definição dos Projetos	2
2	Descrição da Entrega Inicial do Projeto	3
2.1	Descrição da Visão do Produto	3
2.2	Descrição dos Atores Envolvidos e dos Usuários Finais	3
2.3	Descrição do Ambiente do Usuário	4
2.4	Principais Necessidades dos Usuários e Envolvidos	6
2.5	Alternativas e Concorrência	8
2.6	Visão Geral e Recursos do Produto	9
2.7	Outros Requisitos do Sistema	13
2.8	Cronograma do Projeto	15

1 Conceção e Definição dos Projetos

Este projeto tem como principal objetivo promover a aplicação prática dos conhecimentos adquiridos em modelagem e desenvolvimento de sistemas de banco de dados, a partir de um conjunto de dados reais provenientes de um experimento microbiológico. A base de dados utilizada contempla diferentes amostras isoladas, seus atributos morfológicos, genéticos e funcionais, bem como o registro de ensaios bioquímicos e ambientais. O cenário proposto oferece uma oportunidade concreta de lidar com dados complexos e heterogêneos, que demandam organização estruturada, normalização e tratamento semântico para garantir a integridade, escalabilidade e facilidade de consulta posterior.

Diagnóstico da Situação Atual (as-is)

Atualmente, os dados estão armazenados em planilhas independentes, com estruturas não normalizadas, campos sobrepostos, nomenclaturas variadas e ausência de chaves relacionais. Essa abordagem manual, apesar de funcional em pequena escala, dificulta:

- A manutenção e atualização consistente das informações;
- A realização de análises comparativas e agregadas;
- O versionamento e o controle histórico dos dados;
- A integração com ferramentas de análise estatística ou plataformas web.

Proposta de Situação Futura (to-be)

O projeto propõe a modelagem e implementação de um sistema de banco de dados relacional, centrado em entidades como: **Amostra**, **Isolado**, **Consórcio**, **TesteFuncional**, **Sequenciamento**, **Morfologia**, **EnsaioGenômico**, entre outras. A nova estrutura:

- Assegura integridade referencial entre os dados;
- Facilita consultas específicas e cruzamentos complexos;
- Permite o crescimento futuro da base com dados adicionais;
- Suporta análises automatizadas e construção de dashboards.

2 Descrição da Entrega Inicial do Projeto

2.1 Descrição da Visão do Produto

O objetivo deste projeto é desenvolver um sistema de banco de dados relacional voltado para a organização e gerenciamento de dados microbiológicos experimentais, com foco em amostras isoladas, consórcios microbianos, características genômicas e resultados de ensaios laboratoriais.

O sistema proposto busca atender às necessidades específicas de pesquisadores e profissionais da área de biotecnologia e microbiologia ambiental, fornecendo uma plataforma confiável para armazenar, consultar e interpretar grandes volumes de dados complexos.

Modelagem de Processo de Negócio (BPM): será elaborada uma modelagem utilizando notação BPMN com o auxílio da ferramenta PlantUML, representando o fluxo de cadastro, integração de resultados e geração de relatórios científicos.

2.2 Descrição dos Atores Envolvidos e dos Usuários Finais

Nesta seção são apresentados os principais envolvidos no projeto, incluindo tanto os **usuários finais** que utilizarão o sistema diretamente, quanto os **envolvidos** que influenciam nos requisitos, mas não necessariamente interagem com o sistema. Essa identificação é essencial para garantir que todas as necessidades sejam compreendidas e contempladas na modelagem do banco de dados.

Envolvidos (não usuários finais)

Nome	Descrição	Responsabilidade
Coordenador de Pesquisa	Responsável pela supervisão geral das atividades do laboratório.	Aprovar diretrizes para padronização dos dados e garantir integridade científica.
Professores Orientadores	Docentes que orientam os projetos de pesquisa.	Validar a modelagem dos dados e propor melhorias com base em objetivos didáticos.
Equipe de TI Acadêmica	Suporte técnico da instituição.	Apoiar a infraestrutura de hospedagem e manutenção do sistema de banco de dados.

Usuários Finais

Essa estrutura colaborativa reforça a necessidade de um sistema intuitivo, validado por diferentes perfis de usuário, e com controle de acesso adequado às funções de cada papel no processo.

Nome	Descrição	Responsabilidade	Envolvido Representante
Pesquisadores	Mestres/doutorandos que coletam e analisam dados microbiológicos.	Inserir novos registros de amostras, consultar resultados e exportar dados.	Coordenador de Pesquisa
Estudantes de Iniciação Científica	Alunos de graduação envolvidos nos projetos de pesquisa.	Auxiliar no preenchimento dos dados e revisão de informações históricas.	Professores Orientadores

2.3 Descrição do Ambiente do Usuário

O sistema será utilizado em um ambiente acadêmico de pesquisa aplicada, especificamente em um laboratório de microbiologia que conduz estudos com bactérias isoladas de ambientes contaminados, visando aplicações biotecnológicas e ambientais.

Número de Pessoas Envolvidas nas Tarefas

Atualmente, participam ativamente das atividades envolvidas neste projeto:

- 1 professora coordenadora do laboratório;
- Alunos de Iniciação Científica (IC);
- Mestrandos e doutorandos vinculados ao programa de pós-graduação.

No total, estima-se que entre **6 a 10 pessoas** estejam envolvidas diretamente nas atividades que envolvem coleta, experimentação, análise e cadastro de dados no sistema.

Duração do Ciclo de Tarefas

As tarefas relacionadas à geração de dados seguem um fluxo recorrente:

- **Coleta e isolamento das amostras:** pode levar de **4 a 5 semanas** por ciclo;
- **Realização dos testes laboratoriais:** entre **1 a 4 semanas**, dependendo do experimento;
- **Análise de dados e inserção no sistema:** geralmente feita semanalmente ou ao fim de cada ensaio.

O ciclo completo, da coleta até o registro final no sistema, pode durar entre **3 a 6 semanas**, com sobreposição entre atividades de diferentes amostras.

Restrições Existentes

O ambiente de pesquisa do laboratório apresenta desafios típicos de centros acadêmicos que lidam com grandes volumes de dados experimentais de forma descentralizada. A ausência de um sistema centralizado tem gerado limitações significativas para a gestão e evolução dos dados microbiológicos.

Entre as principais restrições observadas, destacam-se:

- **Fragmentação da informação:** os dados estão distribuídos entre diferentes arquivos, versões e dispositivos, dificultando o acesso e a consistência das informações;
- **Baixa rastreabilidade científica:** não há mecanismos eficientes para acompanhar o histórico de modificações ou identificar a autoria das edições nos registros experimentais;
- **Dependência de conhecimento informal:** a estrutura e os critérios de preenchimento das planilhas são, muitas vezes, compreendidos apenas por quem criou ou manuseia os arquivos, o que compromete a continuidade em caso de troca de bolsistas ou pesquisadores;
- **Limitações operacionais:** a inserção, recuperação e cruzamento de dados exigem esforço manual elevado, o que consome tempo dos pesquisadores e aumenta a probabilidade de erros;
- **Ausência de mecanismos de validação:** os dados não passam por nenhum tipo de verificação automatizada, o que permite inconsistências sem alerta ou correção.

Essas restrições reforçam a importância da implementação de um sistema que não apenas organize os dados de forma relacional, mas também ofereça suporte à colaboração segura e à preservação do conhecimento gerado no laboratório ao longo do tempo.

Plataformas Utilizadas Atualmente e Futuras

Atualmente:

- Microsoft Excel (armazenamento dos dados);
- Google Drive (compartilhamento de arquivos).

Plataformas Futuras:

- Sistema de banco de dados relacional;
- Interface web para cadastro e consulta dos dados.

Integração com Outros Aplicativos Existentes

- Compatibilidade com ferramentas como **Excel** e **Google Sheets**, permitindo importação/exportação de dados;
- Exportação de dados em formatos padronizados (CSV, JSON) para submissão em bases externas.

2.4 Principais Necessidades dos Usuários e Envolvidos

A partir da análise do ambiente atual do laboratório e das interações dos usuários com os dados experimentais, foi possível identificar um conjunto de problemas centrais, suas causas e as soluções desejadas pelos envolvidos.

Problema 1: Fragmentação e Duplicidade de Dados

- **Causa:** Utilização de múltiplas planilhas e arquivos soltos, sem controle de versões nem integração entre as fontes.
- **Solução Atual:** Compartilhamento manual de arquivos via Google Drive e comunicação informal entre os membros do laboratório.
- **Solução Desejada:** Centralização dos dados em um banco relacional único, com interface de acesso controlada e registros consolidados.

Prioridade	Preocupações	Solução Atual	Solução Proposta
Alta	Perda de dados, inconsistência	Planilhas desconectadas	Banco relacional com acesso centralizado

Problema 2: Falta de Padronização no Registro de Informações

- **Causa:** Ausência de campos obrigatórios e padronizações nas planilhas utilizadas.
- **Solução Atual:** Criação de modelos manuais (templates) com instruções informais.
- **Solução Desejada:** Implementação de um sistema com campos validados, formatos consistentes e menus controlados.

Prioridade	Preocupações	Solução Atual	Solução Proposta
Alta	Dados inválidos, retrabalho, ambiguidade	Templates soltos	Validações automáticas e estrutura padronizada

Problema 3: Dificuldade na Rastreabilidade de Resultados

- **Causa:** Modificações nos dados não são registradas ou associadas a um responsável.
- **Solução Atual:** Depende da memória dos usuários ou da criação de anotações paralelas.
- **Solução Desejada:** Sistema com controle de usuários, histórico de alterações e identificação de autoria nos cadastros.

Prioridade	Preocupações	Solução Atual	Solução Proposta
Média	Falta de auditoria e rastreamento	Anotações manuais	Sistema com login e controle de versões

Problema 4: Baixa Eficiência na Recuperação de Dados

- **Causa:** Falta de mecanismos de busca estruturada e filtros para cruzar informações.
- **Solução Atual:** Busca visual e filtros manuais nas planilhas.
- **Solução Desejada:** Consultas flexíveis com critérios por atributos.

Prioridade	Preocupações	Solução Atual	Solução Proposta
Alta	Demora na análise e risco de erro	Filtros no Excel	Sistema com formulários e consultas parametrizadas

Problema 5: Escalabilidade Limitada para Novos Dados e Análises

- **Causa:** Estrutura atual não comporta grandes volumes ou novos tipos de dados sem reestruturações manuais.
- **Solução Atual:** Criação de novas abas ou planilhas quando necessário.
- **Solução Desejada:** Modelo de dados relacional, escalável e extensível.

Prioridade	Preocupações	Solução Atual	Solução Proposta
Alta	Reestruturação constante, risco de erros	Expansão manual	Banco relacional escalável com suporte a extensões

Essas necessidades foram levantadas em diálogo com os usuários reais do sistema e orientam a priorização das funcionalidades a serem implementadas, garantindo que a solução atenda de forma eficaz às demandas do laboratório.

2.5 Alternativas e Concorrência

Durante a análise inicial do projeto, foram consideradas diferentes alternativas para solucionar os problemas enfrentados atualmente no gerenciamento de dados do laboratório. Abaixo estão listadas as principais opções identificadas, incluindo soluções concorrentes, abordagens locais e a manutenção do cenário atual.

Opção 1: Manutenção do Status Quo

Descrição: Continuar utilizando planilhas do Excel/Google Sheets para organizar os dados de forma manual e descentralizada.

Pontos Fortes:

- Ferramentas conhecidas por todos os usuários;
- Sem necessidade de instalação ou desenvolvimento adicional;
- Fácil compartilhamento via Google Drive.

Pontos Fracos:

- Alta propensão a erros e inconsistências;
- Nenhum controle de versões ou histórico de alterações;
- Dificuldade de escalar para novos dados e análises;
- Trabalho repetitivo e baixa rastreabilidade.

Opção 2: Desenvolvimento de um Sistema Local com Interface Web e Banco de Dados

Descrição: Implementar um sistema customizado com banco de dados relacional, interface web para cadastro e consulta, e estrutura normalizada de dados.

Pontos Fortes:

- Atende exatamente às necessidades do laboratório;
- Permite controle de usuários, histórico e validação dos dados;
- Estrutura escalável e de longo prazo;
- Facilita análises estatísticas e exportações.

Pontos Fracos:

- Requer tempo e conhecimento técnico para o desenvolvimento;
- Necessita de manutenção contínua e apoio institucional;
- Depende da familiaridade dos usuários com a nova interface.

Opção 3: Uso de Softwares de LIMS (Laboratory Information Management Systems)

Descrição: Adotar uma plataforma pronta para gestão de laboratórios, como Benchling, LabWare ou openLIMS.

Pontos Fortes:

- Soluções profissionais, prontas para uso;
- Recursos robustos de rastreabilidade e integração com workflows;
- Suporte a anexos, resultados e documentos experimentais.

Pontos Fracos:

- Custo elevado (a maioria das versões completas é paga);
- Pouca flexibilidade para adaptar à estrutura de dados do laboratório;
- Necessidade de treinamento e possível resistência dos usuários;
- Alguns dados podem precisar ser convertidos ou retrabalhados.

Escolha Estratégica

Optou-se pelo **desenvolvimento de um sistema local customizado**, pois essa abordagem:

- Equilibra custo, flexibilidade e controle;
- É compatível com a realidade do laboratório e com os objetivos da disciplina;
- Proporciona aprendizado prático aos alunos na concepção de um sistema real;
- Garante aderência ao domínio dos dados e às necessidades específicas da pesquisa.

2.6 Visão Geral e Recursos do Produto

Perspectiva do Produto

O sistema a ser desenvolvido será **independente**, com uso exclusivo interno pelo laboratório de microbiologia. Seu propósito é substituir o atual modelo baseado em planilhas dispersas, oferecendo uma solução centralizada, organizada e confiável para o controle de amostras e resultados experimentais.

A interação com o sistema se dará por meio de uma **interface web**, desenvolvida em **Python**, com funcionalidades acessíveis via navegador por diferentes perfis de usuários.

O sistema incluirá módulos para cadastro, atualização, consulta e geração de relatórios científicos, além de funcionalidades para análise e rastreamento de dados ao longo do tempo.

Componentes principais do sistema:

- **Módulo de Cadastro de Amostras:** registro de novas bactérias, consórcios e isolados;
- **Módulo de Testes e Ensaios:** inserção dos resultados funcionais, morfológicos e bioquímicos;
- **Módulo de Análise Genômica:** organização das informações de sequenciamento e dados de bioinformática;
- **Módulo de Relatórios:** geração de relatórios por filtros e exportações;
- **Controle de Usuários:** autenticação, permissões e rastreamento de modificações;
- **Banco de Dados:** armazenamento relacional estruturado e normalizado.

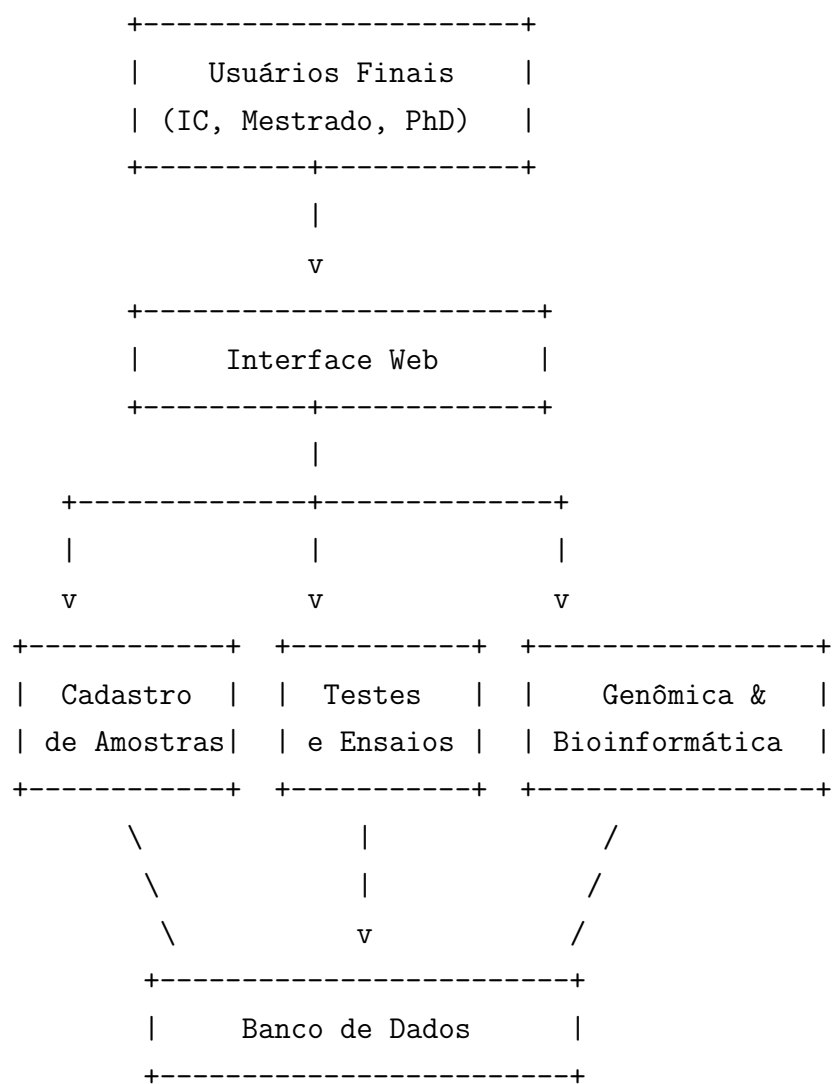
Premissas de Infraestrutura

- O sistema será hospedado nos servidores da faculdade, exigindo suporte a:
 - Banco de dados relacional (PostgreSQL ou MySQL);
 - Servidor web com suporte a Python (Flask ou Django);
 - Armazenamento seguro para dados e backups periódicos;
 - Acesso em rede local, com possibilidade de expansão para acesso remoto.

Suposições

- Os dados laboratoriais seguirão o padrão atual de estrutura (refletido nas planilhas fornecidas);
- Novos experimentos poderão ser integrados futuramente com a adição de novas tabelas e funcionalidades;
- O sistema será utilizado além da disciplina, exigindo documentação completa, interface amigável e possibilidade de manutenção evolutiva.

Diagrama de Blocos do Sistema



Funcionalidades do Sistema (Requisitos Funcionais)

O sistema será estruturado para garantir a organização eficiente dos dados microbiológicos, permitindo um fluxo de trabalho claro para os usuários. A seguir, os requisitos funcionais são apresentados por módulo, com descrição, prioridade e análise de estabilidade, benefício, esforço e risco.

1. Módulo de Cadastro de Amostras

- **Descrição:** Permite o registro e organização das amostras microbiológicas.
- **Funcionalidades:**
 - Cadastro, edição e atualização de amostras;
 - Associação a ensaios e análises genômicas;
 - Importação via CSV ou API.
- **Atributos:** Estabilidade: Alta, Benefício: Essencial, Esforço: Médio, Risco: Baixo.

2. Módulo de Testes e Ensaios

- **Descrição:** Inserção e consulta de dados laboratoriais.
- **Funcionalidades:**
 - Registro de dados funcionais, morfológicos, bioquímicos;
 - Histórico de alterações;
 - Exportação para análise estatística.
- **Atributos:** Estabilidade: Média, Benefício: Alto, Esforço: Alto, Risco: Médio.

3. Módulo de Análise Genômica

- **Descrição:** Organização de dados de sequenciamento e bioinformática.
- **Funcionalidades:**
 - Armazenamento de metadados e associações;
 - Consulta por características funcionais.
- **Atributos:** Estabilidade: Média, Benefício: Alto, Esforço: Médio, Risco: Alto.

4. Módulo de Relatórios

- **Descrição:** Geração de relatórios personalizáveis.
- **Funcionalidades:**
 - Filtros específicos (por espécie, origem, etc.);
 - Exportação em CSV e JSON;
 - Compartilhamento interno.
- **Atributos:** Estabilidade: Alta, Benefício: Alto, Esforço: Médio, Risco: Baixo.

5. Controle de Usuários

- **Descrição:** Gerenciamento de acessos e rastreamento de ações.
- **Funcionalidades:**
 - Cadastro de usuários por perfil;
 - Login seguro e logs de atividade;
 - Recuperação de senha e controle de sessões.
- **Atributos:** Estabilidade: Alta, Benefício: Alto, Esforço: Médio, Risco: Médio.

6. Banco de Dados

- **Descrição:** Armazenamento estruturado e normalizado dos dados.
- **Funcionalidades:**
 - Relacionamento entre todas as entidades;
 - Otimização para consultas eficientes;
 - Backup e políticas de recuperação.
- **Atributos:** Estabilidade: Alta, Benefício: Crítico, Esforço: Alto, Risco: Médio.

Este conjunto de requisitos funcionais garante que o sistema atenda às necessidades dos usuários, fornecendo uma plataforma **robusta, segura e eficiente** para o gerenciamento de dados microbiológicos.

2.7 Outros Requisitos do Sistema

Esta seção contempla os **requisitos não funcionais** que devem ser considerados para garantir a confiabilidade, escalabilidade e usabilidade do sistema ao longo do tempo. Tais requisitos são fundamentais para assegurar que a solução possa ser adotada de forma eficiente no ambiente de pesquisa e mantida de maneira sustentável.

Padrões Aplicáveis

- Seguir boas práticas de desenvolvimento web e banco de dados (ex: RESTful APIs, MVC, normalização 3NF);
- Utilização de convenções de nomenclatura claras e padronizadas para tabelas, campos e variáveis;
- Adoção de padrões abertos para exportação de dados: CSV, JSON e SQL.

Hardware / Plataforma

- O sistema será hospedado em **servidores da instituição de ensino** (ou ambiente em nuvem, se necessário);
- A aplicação será acessada via **navegador web**, sem necessidade de instalação local;
- O servidor deve possuir requisitos mínimos para rodar aplicações em **Python** (Flask ou Django) com **PostgreSQL** ou **MySQL**, além de armazenamento seguro e backups periódicos.

Desempenho

- O sistema deverá ser capaz de realizar **consultas em menos de 2 segundos**, mesmo com aumento do volume de dados;
- Suportar **múltiplos acessos simultâneos** sem degradação perceptível no desempenho;
- A estrutura do banco deve ser otimizada com **índices**, relacionamentos eficientes e consultas bem estruturadas.

Tolerância a Erros

- Implementar **tratamento de exceções** para erros de entrada, falhas de conexão e operações inválidas;
- Fornecer **mensagens de validação claras** ao usuário em caso de erro;
- Manter uma política de **backup automático** e rotinas de recuperação de dados para evitar perdas.

Usabilidade

- Interface acessível e responsiva, com menus e campos organizados de forma intuitiva;
- Campos com instruções ou *placeholders* para facilitar o preenchimento;
- Navegação fluida e com número reduzido de cliques para ações frequentes.

Documentação

- Documentação técnica incluindo:
 - Instruções de instalação e configuração;
 - Modelo Entidade-Relacionamento (ER), tabelas e relacionamentos;
 - Guia de uso da interface.
- Instruções claras para manutenção e extensão futura por outros alunos ou técnicos da instituição.

Restrições Externas

- A implantação depende da disponibilidade de **infraestrutura de TI da faculdade** (acesso a servidores, permissões, rede);
- A entrada de dados depende da **continuidade dos experimentos** e do **comprometimento dos usuários** em manter o sistema atualizado.

Esses requisitos não funcionais são essenciais para garantir que o sistema não apenas atenda às necessidades imediatas do laboratório, mas também seja **sustentável, seguro e utilizável ao longo do tempo** por diferentes gerações de pesquisadores.

2.8 Cronograma do Projeto

O cronograma a seguir foi estruturado com base nas etapas naturais de desenvolvimento de um sistema de banco de dados, considerando também os prazos da disciplina e os marcos esperados no contexto do laboratório. Ele está dividido em quatro partes: Estrutura Analítica do Projeto (EAP), entregáveis intermediários, previsão de conclusão e a lista de documentos a serem gerados.

Estrutura Analítica do Projeto (EAP)

1. Planejamento Inicial

- Levantamento dos requisitos e entendimento do problema;

- Análise exploratória das planilhas de dados;
- Escolha de tecnologias (Python, Flask/Django, banco relacional).

2. Modelagem

- Criação do Modelo Entidade-Relacionamento (MER);
- Normalização e estruturação lógica;
- Definição de tabelas e relacionamentos.

3. Desenvolvimento

- Implementação do banco de dados relacional;
- Construção da interface web;
- Integração com os módulos de cadastro, testes e genômica.

4. Validação e Testes

- Testes funcionais com usuários internos;
- Validação de consistência dos dados;
- Ajustes com base em feedback dos pesquisadores.

5. Documentação e Entrega Final

- Finalização do relatório técnico e manual do sistema;
- Preparação de apresentação final;
- Publicação do repositório e entrega oficial.

Entregáveis Intermediários

- Documento de Visão do Produto;
- Modelo Entidade-Relacionamento (MER);
- Modelo Relacional com tabelas normalizadas;
- Scripts SQL de criação do banco;
- Protótipo da Interface Web;
- Código-fonte no GitHub;
- Relatórios de testes e validações;
- Manual do usuário e instruções de uso.

Etapa	Período Estimado
Planejamento Inicial	Semanas 1 a 2
Modelagem	Semanas 3 a 4
Desenvolvimento	Semanas 5 a 8
Validação e Testes	Semanas 9 a 10
Documentação e Entrega Final	Semanas 11 a 12

Tabela 1: Cronograma de execução do projeto

Previsão de Conclusão

Obs.: O cronograma poderá ser ajustado conforme o andamento do semestre e a disponibilidade da equipe.

Lista de Documentos Mencionados

Título	Nº Documento	Data Estimada	Organização
Documento de Visão do Produto	—	—	Laboratório de Microbiologia / Professor
Modelo Entidade-Relacionamento (MER)	—	—	Projeto de Banco de Dados
Relatório Final do Projeto	—	—	Professor
Scripts de Criação do Banco (SQL)	—	—	Equipe de Desenvolvimento
Manual do Sistema	—	—	Disciplina de Projeto de BD

Tabela 2: Documentos gerados e utilizados no projeto