

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CAMPUS DOIS VIZINHOS
CURSO DE ESPECIALIZAÇÃO EM CIÊNCIA DE DADOS

LUIZ FERNANDO GIOLO ALVES

**DESENVOLVIMENTO DE UM CADERNO ELETRÔNICO PARA
GERENCIAMENTO DE EXPERIMENTOS EM UM LABORATÓRIO
DE BIOCÊNCIAS**

TRABALHO DE CONCLUSÃO DE CURSO DE ESPECIALIZAÇÃO

DOIS VIZINHOS
2022

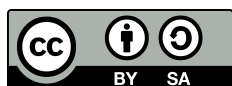
LUIZ FERNANDO GIOLO ALVES

DESENVOLVIMENTO DE UM CADERNO ELETRÔNICO PARA GERENCIAMENTO DE EXPERIMENTOS EM UM LABORATÓRIO DE BIOCÊNCIAS

Trabalho de Conclusão de Curso de Especialização apresentado ao Curso de Especialização em Ciência de Dados da Universidade Tecnológica Federal do Paraná, como requisito para a obtenção do título de Especialista em Ciência de Dados.

Orientador: Prof.^a Dra. Fernanda Paula Barbosa Pola
Coorientador: Prof. Dr. Ives Renê Venturini Pola

DOIS VIZINHOS
2022



4.0 Internacional

Esta licença permite remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

LUIZ FERNANDO GIOLO ALVES

**DESENVOLVIMENTO DE UM CADERNO ELETRÔNICO PARA
GERENCIAMENTO DE EXPERIMENTOS EM UM LABORATÓRIO
DE BIOCÊNCIAS**

Trabalho de Conclusão de Curso de Especialização
apresentado ao Curso de Especialização em Ciência de
Dados da Universidade Tecnológica Federal do Paraná, como
requisito para a obtenção do título de Especialista em Ciência
de Dados.

Data de aprovação: 10/novembro/2022

Fernanda Paula Barbosa Pola
Doutorado
Universidade Tecnológica Federal do Paraná - Câmpus Dois Vizinhos

Rosangela de Fátima Pereira Marquesone
Doutorado
Universidade Tecnológica Federal do Paraná - Câmpus Cornélio Procópio

Francisco Pereira Junior
Mestrado
Universidade Tecnológica Federal do Paraná - Câmpus Cornélio Procópio

DOIS VIZINHOS
2022

AGRADECIMENTOS

Agradeço primeiramente pelo apoio, incentivo, amor e carinho da minha esposa e companheira Bruna, sem a qual este período de estudos, trabalho e pandemia seriam insuportáveis. Agradeço também à minha mãe Regina pelas oportunidades e suporte emocional e educacional durante a minha vida. À Dra Fernanda Pola e o Dr. Ives Pola pela orientação e prontidão. Ao Dr. Paulo Lopes, pelo apoio, infraestrutura e dicas. À Dra Daniela Trivella pela chance e acolhimento ao me juntar ao time do LNBio-CNPEM. Aos cientistas, funcionários e bolsistas do LNBio por me ajudarem com opiniões e avaliações do projeto desenvolvido. E, por fim, agradeço aos professores da UTFPR pela dedicação e apreço ao curso.

Obrigado.

RESUMO

Este trabalho visa o desenvolvimento de um caderno de laboratório eletrônico (Eletronic Lab Notebook — ELN) em um centro de pesquisa em biologia, o Laboratório Nacional de Biociências (LNBio) que fica no *campus* do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM) em Campinas, SP. Para suprir as necessidades dos cientistas, que envolvem gerenciar grandes volumes de dados heterogêneos com relacionamentos entre si, utilizou-se um caderno eletrônico para consolidar, validar e relacionar os vários tipos de dados gerados em experimentos científicos distintos. Para esta implementação, utilizou-se um *software* gratuito e de código aberto (eLabFTW) com capacidade para armazenar e gerenciar dados brutos de experimentos, criar modelos e protocolos, facilitar a elaboração de relatórios e gerenciar itens e experimentos de forma colaborativa, fácil e segura. Um experimento de teste de usuários foi realizado na implementação da consolidação dos dados no caderno eletrônico, por grupos de voluntários no laboratório. Os resultados mostraram que a maioria dos usuários prefere a nova solução, substituindo o tradicional caderno de laboratório pela nova ferramenta digital. O resultado deste trabalho permite a ampliação do uso deste instrumento, onde os pesquisadores, técnicos e alunos poderão partilhar suas anotações, produções e descobertas de forma digital, instantânea e interativa, facilitando sua produção, diminuindo o tempo gasto com burocracia e aumentando o tempo para o raciocínio e reflexão. Além disso, a consolidação dos dados num único sistema digital permite o compartilhamento de dados de diferentes tipos e resultados que possuem ligação entre si, para análise conjunta pelos cientistas.

Palavras-chave: consolidação de dados heterogêneos; caderno de laboratório eletrônico; gestão de dados.

ABSTRACT

This paper aims to develop an electronic lab notebook (ELN) at a biology research center, the Brazilian Biosciences National Laboratory (LNBio) which is located at the *campus* of the Brazilian Center for Research in Energy and Materials (CNPEM), in the city of Campinas, state of São Paulo. To meet the needs of scientists, which involve managing large volumes of heterogeneous data with relationships among themselves, an electronic lab notebook was used to consolidate, validate and relate the various types of data generated in distinct scientific experiments. For this implementation, a free and open source software (eLabFTW) was used with the ability to store and manage raw data from experiments, create models and protocols, facilitate reporting, and manage items and experiments in a collaborative, easy and secure way. A user-testing experiment was conducted on the implementation of data consolidation in the electronic lab notebook by groups of volunteers in the laboratory. The results showed that most users prefer the new solution, replacing the traditional lab notebook with the electronic tool. The result of this work allows the expansion of the use of this tool, where researchers, technicians and students will be able to share their notes, productions and discoveries in a digital, instantaneous and interactive way, facilitating their production, reducing the time spent on bureaucracy and increasing the time for reasoning and reflection. Moreover, the consolidation of data in a single digital system allows the sharing of data of different types and results that are linked together, for joint analysis by scientists.

Keywords: heterogeneous data consolidation; electronic lab notebook; data management.

LISTA DE FIGURAS

Figura 1 – Ciclo de vida de um dado científico	10
Figura 2 – Sugestão do fluxo de dados para uso do eLabFTW	15
Figura 3 – Criação de um tipo de item no banco de dados. (a) Aba “Tipos de itens” no painel de administrador. (b) Configurações principais do item. Nome, cor, se é ou não agendável, visibilidade e edição padrão. (c) Caixa de edição para anotações com elementos de texto rico (pode ser codificado em html ou <i>markdown</i>). (d) Etapas. Pode-se usar para orientar o usuário ao criar o item ou mesmo gerar um passo a passo para nortear um protocolo. (e) Aba “banco de dados” no painel principal onde pode ser gerar o item recém criado.	16
Figura 4 – Grafo da organização descrita. Partindo dos nós que representam os projetos é possível chegar em todos objetos relacionados, esta conexão ajuda muito todos os pesquisadores envolvidos a buscar itens de interesse.	18
Figura 5 – Exemplo de um item do tipo “Projeto” no banco de dados. (a) outros itens de bancos de dados diretamente ligados a ele. (b) Experimentos diretamente ligados ao projeto. Ao clicar em qualquer um destes, se usuário tiver permissão de acesso, o destino será exibido.	19
Figura 6 – Diferença entre apresentação de conteúdo em sua forma padrão (a) e na forma tabular (b).	20
Figura 7 – Pré-visualização de alguns objetos anexados em um experimento	21
Figura 8 – Distribuição de uso de disco por usuário. Neste gráfico foram contabilizados apenas os usuários com mais de 1 <i>megabyte</i> em uso. JVPdG é a mesma pessoa citada na Seção 3.2. CT não aparece, pois usa mais o caderno acompanhando as inserções e acaba utilizando pouquíssimo espaço em disco.	21

LISTA DE ABREVIATURAS E SIGLAS

AD	<i>Active Directory</i>
API	Interface de Programação de Aplicação, do inglês <i>Application Programming Interface</i>
CNPEM	Centro Nacional de Pesquisa em Energia e Materiais
ELN	Caderno de Laboratório Eletrônico, do inglês <i>Electronic Lab Notebook</i>
FAIR	Fácil de encontrar Acessível Interoperável Reutilizável, em tradução livre do inglês <i>Findable Accessible Interoperable Reusable</i>
LBC	Laboratório de Biologia Computacional
LNBio	Laboratório Nacional de Biociências
LIMS	<i>Laboratory Information Management System</i>
PCR	<i>Polymerase Chain Reaction</i>
PDB	<i>Protein Data Bank</i>
PI	Cientista Chefe, do inglês <i>Principal Investigator</i>
POP	Protocolo Operacional Padrão
VM	Máquina Virtual, do inglês <i>Virtual Machine</i>

SUMÁRIO

1	INTRODUÇÃO	9
1.1	Introdução e Revisão Bibliográfica	9
1.2	Problema de Pesquisa	11
1.3	Justificativa	11
1.4	Objetivos	11
1.5	Materiais e Métodos	12
2	MATERIAIS E MÉTODOS	13
2.1	Caderno Eletrônico	13
2.2	Instalação e Disponibilidade	13
2.3	Usuários e Testes	13
3	RESULTADOS	14
3.1	Seleção do caderno eletrônico	14
3.1.1	Sugestão de uso	14
3.1.2	Fluxo de dados e operações	17
3.2	Teste de usuários	19
3.2.1	Sobre o leiaute de exibição	19
3.2.2	Sobre o compartilhamento de dados	20
3.2.3	Outras observações e potenciais	21
4	CONCLUSÃO	22
4.1	Limitações	22
4.2	Trabalhos Futuros	22
	REFERÊNCIAS	24

1 INTRODUÇÃO

1.1 Introdução e Revisão Bibliográfica

No livro “O Mundo Assombrado por Demônios”, Carl Sagan descreve a ciência como uma vela no escuro para iluminar o longo caminho de descobertas. No capítulo “A arte refinada de detectar mentiras” ele lista nove regras para o pensamento cético, sendo a nona delas a seguinte:

Devemos sempre perguntar se a hipótese pode ser, pelo menos em princípio, falseada. As proposições que não podem ser testadas ou falseadas não valem grande coisa. [...] Devemos poder verificar as afirmativas. Os céticos inveterados devem ter a oportunidade de seguir o nosso raciocínio, copiar os nossos experimentos e ver se chegam ao mesmo resultado. (SAGAN, 1996).

No método científico, para que as hipóteses sejam falseáveis e os experimentos possam ser reproduzidos, os cientistas devem registrar os dados experimentais (condições de controle, metodologia e resultados brutos e/ou analisados). O registro de dados é uma prática essencial em laboratórios e, desde o século XVIII (ARGENTO, 2020), e desde então este registro vem sendo realizado em papel. Apesar deste método ter sido muito importante até os dias atuais, por ser uma ferramenta simples de usar, acessível e barata, é possível listar algumas características que indicam sua fragilidade, como:

- O papel pode ser esquecido, perdido ou roubado;
- A reprodução de cópias pode ser um processo lento;
- O compartilhamento pode exigir muita confiança, uma vez que não é possível rastrear quem teve acesso ao dado;
- O compartilhamento específico de uma fração do caderno só é viável via fotos ou foto-cópias;
- O acesso simultâneo não é possível, a menos que todos estejam no mesmo espaço físico;
- Sua estrutura organizacional é totalmente dependente do dono do caderno (difícil padronização dos dados reportados);
- Dificilmente o dado é indexado;
- Baixa resistência a intemperes.

Com o avanço da tecnologia e do acesso à ela, novas ferramentas foram desenvolvidas e disponibilizadas para que a comunidade científica pudesse atualizar suas formas de registro. Dentre estas ferramentas, o caderno eletrônico se apresenta como uma alternativa para anotações e armazenamento de dados, visto que, além de atenuar os problemas listados, esse recurso digital apresenta grande potencial para gerar entradas em algoritmos que ajudam a agrupar as informações contidas nas anotações, inclusive associando cadernos de diferentes autores, algo impraticável em um conjunto de cadernos físicos.

A sistematização e o gerenciamento de dados experimentais no meio acadêmico ainda é um desafio em aberto (ARGENTO, 2020).

Os dados científicos possuem um ciclo de vida diferenciado (CORTI et al., 2011) (Figura 1), pois o tempo de vida do dado pode ser maior que o tempo do projeto que os geraram. Uma pesquisa pode dar origem a outras e suas descobertas podem ser reutilizadas em novos experimentos. Por isso, as estratégias de armazenamento, acesso e compartilhamento devem ser feitas de modo a facilitar estes passos.

Figura 1 – Ciclo de vida de um dado científico



Fonte: Adaptada de (CORTI et al., 2011)

Atualmente, um dos princípios mais utilizados para guiar os cientistas no registro e organização de seus dados é o princípio **FAIR** (*Findable, Accessible, Interoperable, Reusable*). Sendo: **F**acilidade de acesso; **A**cessibilidade, com controle de sigilo; **I**nteroperabilidade para aumentar a colaboração e sinergia dos pesquisadores e **R**eutilização para diminuir o retrabalho, economizando recursos públicos e tempo de colegas. Essas funcionalidades contribuem para auxiliar o registro destes dados, transformando-os em dados FAIR (WILKINSON et al., 2016).

Desde a década de 1980 (HIGGINS; NOGIWA-VALDEZ; STEVENS, 2022), os ELNs e os LIMSs se apresentam como ferramentas potenciais, mas ainda hoje é possível encontrar relatos sobre os desafios da implantação e acolhimento destas tecnologias. Um dos fatores que atrasam este tipo de adoção são os custos atrelados a implantação (ROCHA; SALES; SAYÃO, 2018), altos custos financeiros com licenças de *software*, infraestrutura, gestão de *backups* e recurso humano para gerenciar a implantação; além desta oneração monetária também tem-se o custo de tempo para adoção, adaptação e treinamento de pessoal.

1.2 Problema de Pesquisa

Em um dos laboratórios nacionais do Centro Nacional de Pesquisas em Energia e Materiais (CNPEM), localizado em Campinas-SP, a necessidade de implementação de um sistema único de organização de dados se apresentou necessária. Durante a fase mais restrita de isolamento social da pandemia de COVID-19, os pesquisadores precisaram aprender a usar ferramentas da Microsoft que já eram disponibilizadas pelo Centro há algum tempo, mas que até então eram pouco utilizadas. Ferramentas como o *Teams* e o pacote *Office* em sua versão *online* foram adotadas para anotações, armazenamento e compartilhamento de dados entre participantes do mesmo grupo e de grupos diferentes. Os caderno perderam espaço para o pacote *Office*, e contrariando o que parecia mais lógico, o *PowerPoint* foi selecionado como um dos melhores correspondentes em detrimento do *Word*, pois nele criavam-se resumos objetivos do que se gostaria de compartilhar. Estes métodos atenderam a demanda momentânea, mas não foram os mais eficientes, pois muitas vezes os *links* de compartilhamento expiravam, o acesso dos arquivos se perdiam e nem sempre ficavam centralizado. Além disso, os resultados experimentais gerados em um equipamento podem vir em formatos de arquivo inadequados para ser anexado a um *slide* ou página do editor de texto.

1.3 Justificativa

A recente pandemia de COVID-19, mostrou que a tecnologia atual tem condições de ajudar um trabalho a ser construído a muitas mãos, mesmo que estas estejam separadas por cidades ou países diferentes. O Laboratório Nacional de Biociências (LNBio-CNPEM) possui diversos grupos de pesquisa que realizam seus estudos de forma interdisciplinar tanto internamente como também entre grupos de outros laboratórios e instituições. A experiência recente possibilitou a percepção das potencialidades de formas eletrônicas de anotação e compartilhamento, porém também apontou que a ferramenta mais adequada para esse registro ainda não estava dentre as já existentes e acessíveis no *campus*. Dessa forma, é necessária a busca por uma ferramenta que melhor atenda às demandas do público.

A tarefa de introduzir uma nova tecnologia vem geralmente acompanhada de alguma resistência, um ciclo de adaptações e aprendizado e até uma mudança cultural na maneira de como o grupo organiza seu trabalho. Avaliar e auxiliar na implantação de uma nova ferramenta é primordial para minimizar os atritos e desistências, assim um plano de acompanhamento também é desejável para o sucesso da adoção. Por fim, com dados e anotações mais estruturados, será possível fomentar a construção de um *datawarehouse* e análises de *BigData*.

1.4 Objetivos

O objetivo deste trabalho foi realizar a implantação de um caderno eletrônico no LNBio-CNPEM para possibilitar a organização de dados experimentais de forma mais eficiente

e segura em um sistema digital.

Esta implantação passa por uma etapa de seleção da ferramenta mais adequada, implementação e disponibilização do software escolhido e avaliação do resultado obtido nos primeiros meses de uso.

1.5 Materiais e Métodos

Na tentativa de atenuar os problemas na organização e disponibilização de dados de uma pesquisa científica, o caderno eletrônico **eLabFTW** (CARPI; MINGES; PIEL, 2017) foi selecionado. Após a seleção, o caderno foi apresentado e inicialmente implementado em alguns usuários para teste. Durante o primeiro mês de uso, os usuários foram atendidos para resolução de dúvidas e este *feedback* foi importante para antecipar as questões dos futuros utilizadores.

2 MATERIAIS E MÉTODOS

Para disponibilizar e testar o caderno eletrônico capaz de auxiliar a combinar os dados das várias fontes que os cientistas têm a sua disposição, foram organizados os recursos abaixo descritos.

2.1 Caderno Eletrônico

O eLabFTW é um caderno eletrônico de código aberto mantido pela empresa francesa Deltablot ([DELTABLOT...](#), 2022) sob licença AGPL-3.0, dando liberdade para a instituição usar sem ser onerada e, caso necessário, modificar o código. Sua interface web é escrita em PHP em cima de uma base de dados MySQL. É instalado via *docker* e acessado por navegadores web, sejam eles em computadores de mesa ou dispositivos móveis (sua interface é responsiva). Caso essa instalação seja feita em um servidor em rede local, a aplicação fica acessível para todos os usuários daquela rede mediante login. A autenticação pode ser sincronizada com o servidor Active Directory (AD), facilitando a adoção do sistema pelos usuários da instituição. Além disso, este caderno eletrônico possui uma API HTTP do tipo REST e uma biblioteca python para facilitar o uso desta através de *scripts*. Com isso, é possível automatizar inserções massivas, análises sobre do nível de adoção e uso da ferramenta e coletas para uma futura montagem de um *datawarehouse*.

2.2 Instalação e Disponibilidade

A instalação do software foi feita à priori em um antigo servidor local usado para testes. A máquina é um servidor DELL *PowerEdge* R420 com processador Intel Xeon E5-2430 de 6 núcleos e 12 *threads*, 8 GB de memória RAM rodando sistema operacional Ubuntu *Server* 20.04. Com a aprovação da maioria dos usuários que testaram, o serviço foi migrado para uma Máquina Virtual (VM) com 4 núcleos, também 8 GB de RAM e provisórios 80 GB de espaço em disco. A VM foi disponibilizada pela equipe de tecnologia da informação do centro (TIC) para usufruir de todos os protocolos de segurança, integrar-se ao sistema de *login* e fazer parte das rotinas automatizadas de backups diários.

2.3 Usuários e Testes

Ao longo do ano, três grupos haviam procurado o Laboratório de Biologia Computacional (LBC) buscando por alguma ferramenta para organizar e compartilhar seus dados de maneira digital e descomplicada. Convidamos estes grupos para testar as capacidades do programa e verificar se os recursos do eLabFTW adéquam-se às necessidades de cada uma das equipes.

3 RESULTADOS

3.1 Seleção do caderno eletrônico

Existe uma plethora de cadernos eletrônicos disponíveis no mercado, tornando a escolha um trabalho desafiador. Alguns critérios foram estabelecidos para filtrar as opções disponíveis:

- Ser gratuito (obrigatório);
- Não ter limite de usuários ou conexões simultâneas (obrigatório);
- Integração com AD (obrigatório);
- Ser possível instalar em servidor local para controle dos dados (obrigatório);
- Ser agnóstico de sistema operacional (obrigatório);
- Desenvolvimento ativo (obrigatório, algum lançamento de versão nos último ano);
- Ser de código aberto (desejável);
- Ter tradução para o português (desejável).

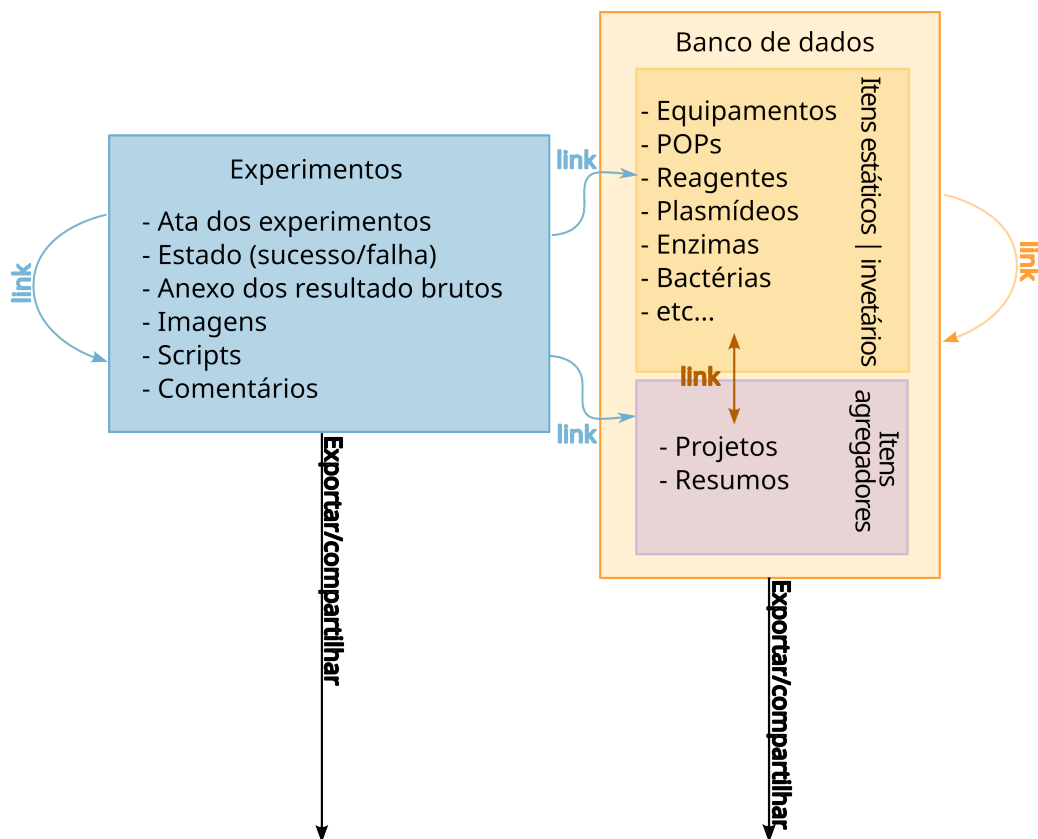
Com os critérios estabelecidos, uma análise dos cadernos eletrônicos disponíveis foi realizada para escolher o que melhor os atendessem. Durante o processo, foi utilizada uma tabela de comparação elaborada pelo grupo de trabalho de gerenciamento de dados da Harvard *Medical School* (HMS Data Management Working Group) (GROUP, 2021) com 33 opções de *softwares*. Além destes, foram adicionados como possibilidade o OpenBIS (BARILLARI et al., 2015) e o Indigo ELN (INDIGO... , 2018).

A seleção das características essenciais resultou em apenas duas opções: o OpenBIS e o eLabFTW. O SciNote (SCINOTE... , 2022) ficou entre os mais promissores especialmente por seu visual moderno de uso fácil e intuitivo, mas foi excluído por não ter integração com o AD. O desempate foi realizado pelos próprios usuários do grupo teste que puderam explorar e avaliar qual dos cadernos atendia melhor a demanda, então o eLabFTW foi escolhido por conta de sua interface mais intuitiva e amigável.

3.1.1 Sugestão de uso

O eLabFTW é uma ferramenta com poucas amarras, ele garante a liberdade do usuário criar uma maneira própria de organizar suas anotações, dados e resultados. Uma ideia de fluxo de trabalho baseada em (HEWERA et al., 2021) foi sugerida (ilustrado na Figura 2) para tentar otimizar o uso dos recursos oferecidos. O fluxo de trabalho proposto parte de dois itens principais do caderno eletrônico: o banco de dados (I) e os experimentos (II).

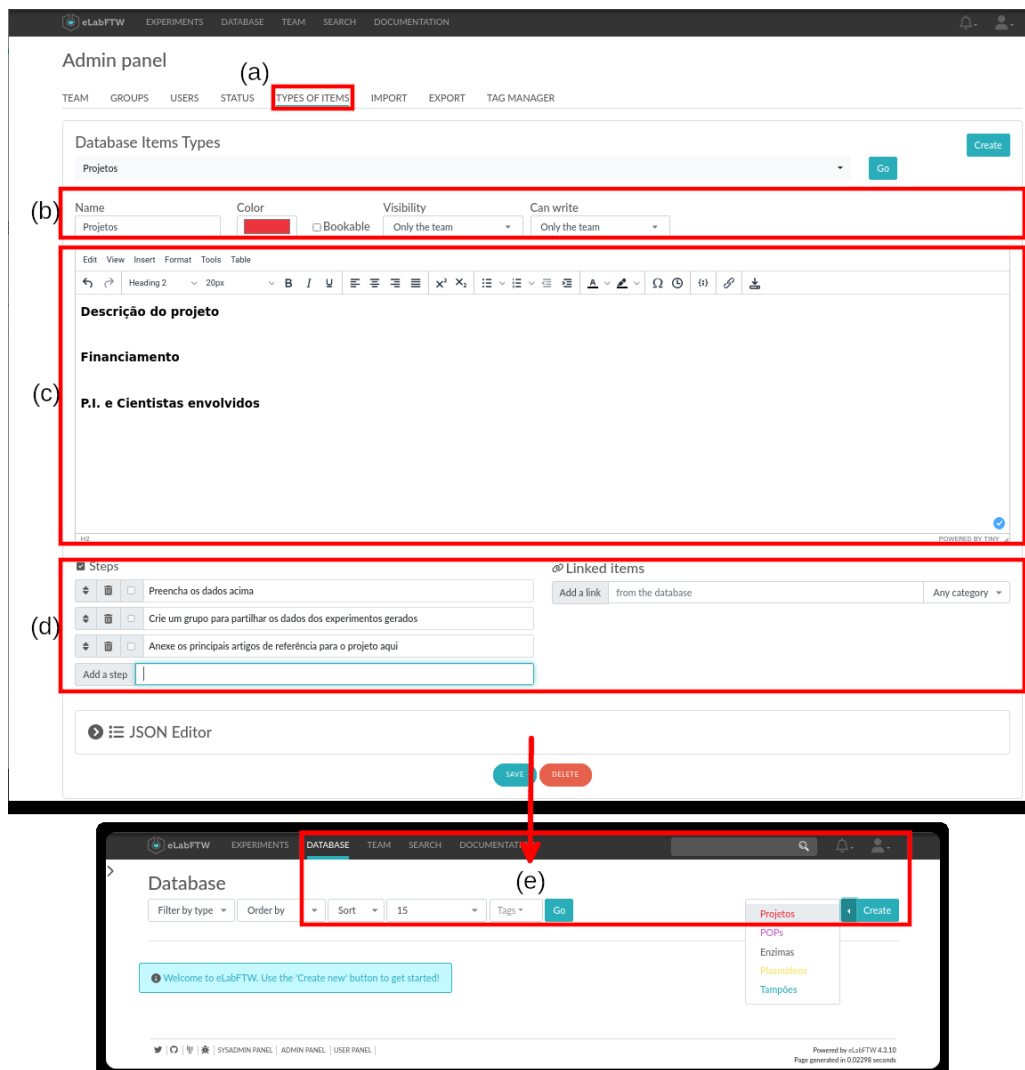
Figura 2 – Sugestão do fluxo de dados para uso do eLabFTW



Fonte: Imagem do autor

Os itens gerados no banco de dados podem ser divididos em categorias personalizadas chamadas “tipos de itens”, cada tipo tem um nome, cor representativa, e um modelo de escrita. Ao criar um item de certo tipo, o autor tem a sua disposição uma caixa de texto rico, ao qual pode anexar arquivos variados, atribuir *tags* e até colocar uma avaliação (de 1 a 5 estrelas). Além disso, também é possível adicionar uma lista de etapas, seja para o andamento de um experimento, projeto ou outra situação a qual o item se refira, ou mesmo para instruir o preenchimento do objeto criado como ilustrado na [Figura 3](#).

Figura 3 – Criação de um tipo de item no banco de dados. (a) Aba “Tipos de itens” no painel de administrador. (b) Configurações principais do item. Nome, cor, se é ou não agendável, visibilidade e edição padrão. (c) Caixa de edição para anotações com elementos de texto rico (pode ser codificado em html ou *markdown*). (d) Etapas. Pode-se usar para orientar o usuário ao criar o item ou mesmo gerar um passo a passo para nortear um protocolo. (e) Aba “banco de dados” no painel principal onde pode ser gerar o item recém criado.



Fonte: Imagem do autor

Os experimentos são muito parecidos com os itens do banco de dados, porém ao invés de terem um tipo associado, eles possuem atrelados a si um **estado**. Por padrão estes estados podem ser:

1. Em execução;
2. Sucesso;
3. Precisa ser refeito;
4. Falhou.

Uma das vantagens em se utilizar um caderno eletrônico é seu potencial em “reciclagem

de texto”. Os experimentos utilizam este potencial através da criação e utilização de modelos. Com os modelos o autor pode deixar seus principais experimentos semi preenchidos, agilizando e diminuindo uma tarefa importante, mas enfadonha, além de poder partilhar os mesmos com outros usuários, amplificando o reuso.

3.1.2 Fluxo de dados e operações

Agora que os itens de banco de dados e os experimentos foram definidos, pode-se elaborar o fluxo exibido na [Figura 2](#). Nas bases de dados, popula-se itens “estáticos” como protocolos, equipamentos, reagentes disponíveis para o grupo, POPs, etc. Nesta categoria também cabem os itens “agregadores” como os projetos. Estes itens são criados para ser referenciados pelos experimentos e aquisições feitas em seu benefício, assim fica mais fácil encontrá-los para relatórios e acompanhamentos, pois está tudo ligado a um mesmo objeto.

Em experimentos popula-se com objetos que podem ser associados a estados (em andamento, sucesso, falhou, etc) e que tenham resultados para ser anexados ou compartilhados como imagens de gel, tabela de leituras, gráficos e conteúdos similares.

Exemplo: Um grupo possui o seguinte banco de dados (ilustrado em partes na [Figura 4](#)):

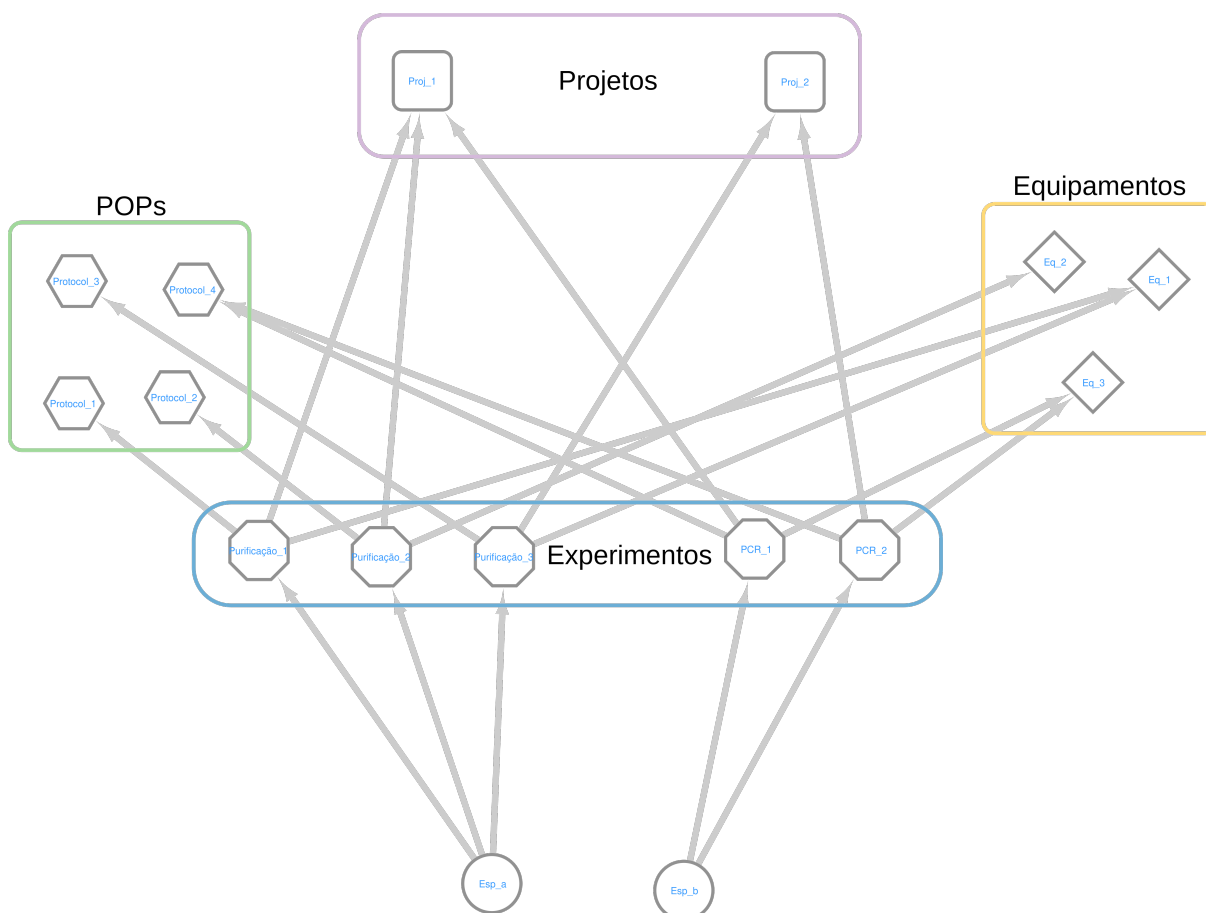
- Projetos
 - Proj₁
 - Proj₂
- POPs
 - Protocol₁
 - Protocol₂
 - ...
 - Protocol_n
- Tampões
 - T₁
 - T₂
 - ...
 - T_n
- Proteínas
 - Prot₁
 - Prot₂
 - ...
 - Prot_n
- Equipamentos
 - Eq₁
 - Eq₂
 - ...
 - Eq_n

A Especialista Esp_a realiza três experimentos de purificação de proteínas, um para o Proj₁ e dois para o Proj₂, cada purificação usa seus respectivos protocolos, materiais e equipamentos.

Outra especialista a Esp_b realiza uma *Polymerase Chain Reaction* (PCR) para confirmação de dados para cada projeto, também referenciando aos protocolos, materiais e equipamentos. E assim sucessivamente.

Desta forma, quando a PI quer acompanhar a situação de experimentos em determinado projeto, sendo “Projetos” um item agregador (Figura 5), basta clicar em um deles (Proj₁ por exemplo) e ela terá acesso a todos experimentos e itens nele ligados, assim como seus documentos e resultados anexados.

Figura 4 – Grafo da organização descrita. Partindo dos nós que representam os projetos é possível chegar em todos objetos relacionados, esta conexão ajuda muito todos os pesquisadores envolvidos a buscar itens de interesse.



Fonte: Imagem do autor

A adoção do caderno eletrônico ainda encontra-se em etapa experimental, mas o fato dele ter sido exaltado neste primeiro momento mostra uma *feedback* positivo por parte dos usuários, visto que relatos sugerem que uma tentativa anterior de implantação de outra ferramenta não obteve o mesmo sucesso em aceitação.

Figura 5 – Exemplo de um item do tipo “Projeto” no banco de dados. (a) outros itens de bancos de dados diretamente ligados a ele. (b) Experimentos diretamente ligados ao projeto. Ao clicar em qualquer um destes, se usuário tiver permissão de acesso, o destino será exibido.

The screenshot displays a web interface for a database project. At the top, there's a navigation bar with 'eLabFTW' and tabs for 'EXPERIMENTS', 'DATABASE', 'TEAM', 'SEARCH', and 'DOCUMENTATION'. The main content area is titled 'Database' and shows details for 'Project 011 - Projeto 1'. It includes a 'Considerações gerais' section with a paragraph of text, a 'Referências' section with two references, a 'Steps' section with two completed tasks, a 'Linked items' section (a) with four items, a 'Related experiments' section (b) with two items, and a 'Comments' section with one comment.

Fonte: Imagem do autor

3.2 Teste de usuários

Em uma breve conversa com JVPdG e CT, respectivamente pesquisador e PI de uma das equipes que testaram o caderno, as seguintes impressões foram relatadas após o primeiro mês de uso.

3.2.1 Sobre o leiaute de exibição

O visual, em si, é agradável e fácil de usar, porém, o estilo padrão de apresentação tende a ser confuso. Este leiaute exhibe pequenos cartões para representar os objetos. Por sorte, no painel de configurações pessoais é possível escolher uma formatação tabular que é muito mais agradável (Figura 6).

Figura 6 – Diferença entre apresentação de conteúdo em sua forma padrão (a) e na forma tabular (b).

(a) Padrão

DATE	TITLE	NEXT STEP	CATEGORY	TAGS	ACTIONS	RATING	OWNER
2022-10-13	PCR	Limpar bancada	Running	PCR	🔍		👤
2022-09-26	PCR	Limpar bancada	Running	PCR	🔍		👤
2022-09-23	Conjugação de MAb		Running		🔍		👤
2022-09-22	PCR	Verificar se está compartilhado com as pessoas necessárias	Success	PCR Projeto A Lote1 Purificacao	🔍		👤

(b) Tabela

Fonte: Imagem do autor

3.2.2 Sobre o compartilhamento de dados

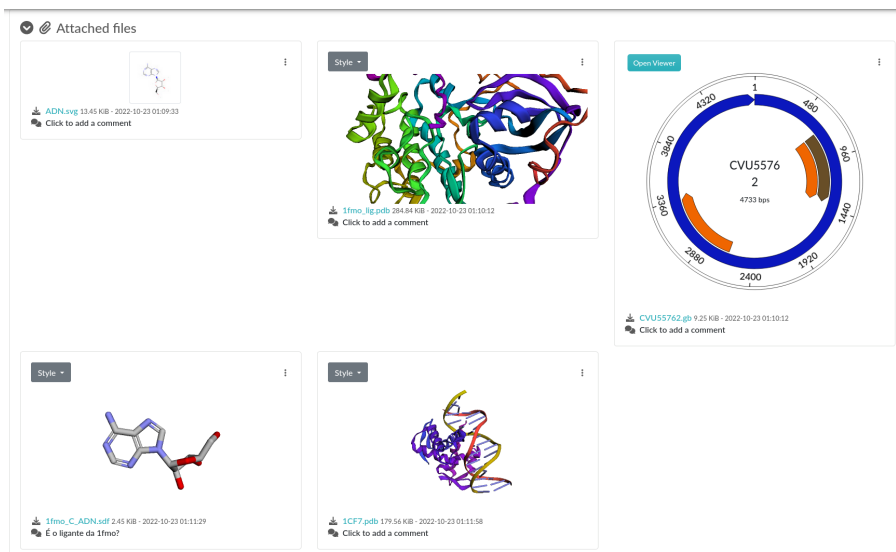
JVPdG tem usado muito o eLabFTW para compartilhar relatórios e resultados com um pesquisador de outro grupo.

O relato é que partilhar os dados através do caderno eletrônico ajuda a centralizar acesso à informação, evitando a fragmentação dos dados que antes ficavam numa combinação de e-mail, *Sharepoint*, *OneDrive* e *Teams*.

Os links entre os objetos ajudam os envolvidos a rastrear os dados de maneira rápida e eficiente e o uso de etiquetas resume o trabalho de agrupamentos e filtros a um clique. Caso seja preciso uma busca mais refinada é possível construir *queries* para buscar por autor, *tag*, conteúdo no corpo do texto, data, situação... incluindo combinações. Usuários que não têm familiaridade com este tipo de redação podem contar com um construtor de *queries* que o ajudará a chegar no conteúdo buscado.

Conseguiu-se agregar muitos dados heterogêneos, superando o desafio dos cadernos físicos e das ferramentas anteriores, agora é possível concentrar dados de texto, imagem, proteínas (.pdb), moléculas, arquivos de sequenciamento ou qualquer outro formato querido. Com a vantagem de que vários destes podem ser pré-visualizados (Figura 7), agilizando para o cientista que muitas vezes precisa visualizar rapidamente tais arquivos, mas antes é preciso baixá-lo e abri-lo no *software* adequado.

Figura 7 – Pré-visualização de alguns objetos anexados em um experimento

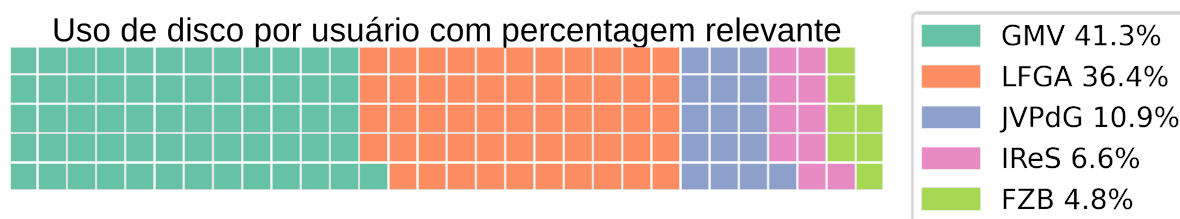


Fonte: Imagem do autor

3.2.3 Outras observações e potenciais

Nestes dois meses de usos tem-se, por enquanto, 13 usuários ativos com 71 experimentos registrados e 116 itens de banco de dados. Estes usuários consumiram cerca de 147,6 megabytes sendo que 41,3% pertence a um único usuário.

Figura 8 – Distribuição de uso de disco por usuário. Neste gráfico foram contabilizados apenas os usuários com mais de 1 megabyte em uso. JVPdG é a mesma pessoa citada na Seção 3.2. CT não aparece, pois usa mais o caderno acompanhando as inserções e acaba utilizando pouquíssimo espaço em disco.



Fonte: Imagem do autor

Até o momento a solução implementada foi de grande aceitação, pois facilita o gerenciamento de dados gerados diariamente, principalmente ao repetir experimentos parecidos.

Outro ponto positivo foi observado na rotina laboratorial. É comum a realização vários experimentos simultâneos e anotá-los no caderno é um incômodo, pois nem sempre o espaço ocupado nas anotações é equivalente ao espaço reservado. Usando um navegador para tomar estas notas, bastas usar uma aba para cada experimento, ou até mesmo fechar e abrir cada arquivo à medida que for necessário.

4 CONCLUSÃO

O caderno eletrônico eLabFTW tem grande potencial para facilitar a gerência dos dados relacionados a experimentos e melhorar o controle dos projetos que os contém. Mesmo baseado em um banco de dados relacional, foi pensado e implementado de modo a deixar o usuário bem à vontade para organizar seu fluxo de trabalho e gerência sem muitas amarras. Os dados brutos, fotos e arquivos correlatos aos experimentos podem ser facilmente transportados e compartilhados com pouco cliques, liberando grupos específicos a visualizar ou editar relatos, modelos, itens ou documentos.

4.1 Limitações

Apesar das facilidades e benefícios que o software oferece, existe margem para algumas melhorias.

Uma ferramenta que ajuda o usuário a compartilhar seus trabalhos com pessoas específicas, é a ferramenta de criação de grupos. O usuário pode criar vários grupos com qualquer pessoa cadastrada na plataforma, porém identificamos dois problemas com os grupos.

O primeiro é o gerenciamento, apenas o criador do grupo pode modificá-lo, isso dificulta a adição e gerenciamento de grupos maiores. Além disso, como não é possível mudar o dono do grupo, isso pode gerar um problema no alcance de alguns experimentos no desligamento de algum pesquisador.

O segundo é o fato de não ser possível unir grupos para dar as devidas permissões de visualização ou escrita, sempre que for necessário um arranjo diferente de pessoas, um grupo novo deve ser criado para esse papel.

Indo afundo nas questões administrativas, usuários colocados como administradores de equipe, têm privilégios administrativos em qualquer equipe que pertecerem, isso pode ser visto como encorajamento para que cada usuário pertença a uma equipe só, mas dado que é possível participar de várias, talvez seja interessante uma granularidade maior nas definições de papéis.

E já que os papéis e permissões vieram à tona, falta um papel de "supervisor geral", pois não existe nenhum tipo de usuário que tenha acesso a todo conteúdo dos bancos, algo que pretende-se solucionar futuramente a partir de uma varredura através da API, dado que tal permissão não está nos planos dos desenvolvedores, segundo resposta dada a pergunta no *GitHub*.

4.2 Trabalhos Futuros

Um laboratório de pesquisa é uma fonte constante de dados. São experimentos, microscópios, biotérios e vários outros equipamentos e pessoas gerando e analisando informações.

Centralizar estes dados em um servidor de arquivos, automatizar a organização, coletar e cruzar metadados e os relatórios gerados nos cadernos em um *datawarehouse* fazem parte do caminho natural deste trabalho.

Com o caderno eletrônico muitos dados heterogêneos estão sendo integrados, possibilitando futuras pesquisas envolvendo uma mineração mais profunda dos mesmos.

Com estas informações organizadas e relacionadas, novas descobertas podem surgir tanto no campo comportamental, traçando o perfil de trabalho do laboratório, quanto no campo da pesquisa, deixando a máquina procurar por denominadores comuns escondidos dos olhos humanos.

Referências

- ARGENTO, N. Institutional eln/lims deployment. **EMBO reports**, v. 21, n. 3, p. e49862, 2020. Disponível em: <<https://www.embopress.org/doi/abs/10.15252/embr.201949862>>. Citado na página 9.
- BARILLARI, C. et al. openBIS ELN-LIMS: an open-source database for academic laboratories. **Bioinformatics**, v. 32, n. 4, p. 638–640, 10 2015. ISSN 1367-4803. Disponível em: <<https://doi.org/10.1093/bioinformatics/btv606>>. Citado na página 14.
- CARPI, N.; MINGES, A.; PIEL, M. elabftw: An open source laboratory notebook for research labs. **Journal of Open Source Software**, The Open Journal, v. 2, n. 12, p. 146, 2017. Disponível em: <<https://doi.org/10.21105/joss.00146>>. Citado na página 12.
- CORTI, L. et al. **Managing and Sharing Data**. Los Angeles, Calif.: Sage, 2011. ISBN 1-904059-82-1. Citado na página 10.
- DELTABLOT, Open Source software for research labs. 2022. Disponível em: <<https://www.deltablot.com/>>. Citado na página 13.
- GROUP, H. L. M. A. R. D. M. W. **Electronic Lab Notebook Comparison Matrix**. Zenodo, 2021. Additional website content archived through Archive-It. Disponível em: <<https://doi.org/10.5281/zenodo.4723753>>. Citado na página 14.
- HEWERA, M. et al. elabftw as an open science tool to improve the quality and translation of preclinical research. 2021. Disponível em: <<https://doi.org/10.12688/f1000research.52157.3>>. Citado na página 14.
- HIGGINS, S. G.; NOGIWA-VALDEZ, A. A.; STEVENS, M. M. Considerations for implementing electronic laboratory notebooks in an academic research environment. **Nature Protocols**, v. 17, n. 2, p. 179–189, Feb 2022. ISSN 1750-2799. Disponível em: <<https://doi.org/10.1038/s41596-021-00645-8>>. Citado na página 10.
- INDIGO ELN v2.0: The open-source electronic lab notebook. 2018. Disponível em: <<https://lifescience.opensource.epam.com/indigo/eln/index.html>>. Citado na página 14.
- ROCHA, L. d. L.; SALES, L. F.; SAYÃO, L. F. Uso de cadernos eletrônicos de laboratório para as práticas de ciência aberta e preservação de dados de pesquisa. **PontodeAcesso**, v. 11, n. 3, p. 2–16, jan. 2018. Disponível em: <<https://periodicos.ufba.br/index.php/revistaici/article/view/24945>>. Citado na página 10.
- SAGAN, C. **O Mundo Assombrado Pelos Demônios**. São Paulo: Companhia de Bolso, 1996. Citado na página 9.
- SCINOTE ELN: Electronic Lab Notebook; Inventory Management. 2022. Disponível em: <<https://www.scinote.net/>>. Citado na página 14.
- WILKINSON, M. D. et al. The fair guiding principles for scientific data management and stewardship. **Scientific Data**, v. 3, n. 1, p. 160018, Mar 2016. ISSN 2052-4463. Disponível em: <<https://doi.org/10.1038/sdata.2016.18>>. Citado na página 10.