

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
INSTITUTO DE MATEMÁTICA  
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

MAYARA MORAIS DOS SANTOS  
MARCOS VINICIUS MACHADO RODRIGUES

CRF4FAIR: MÓDULO PARA VERSIONAMENTO DE QUESTIONÁRIOS PARA  
APLICAÇÃO VODAN BR

RIO DE JANEIRO  
2022

MAYARA MORAIS DOS SANTOS  
MARCOS VINICIUS MACHADO RODRIGUES

CRF4FAIR: MÓDULO PARA VERSIONAMENTO DE QUESTIONÁRIOS PARA  
APLICAÇÃO VODAN BR

Trabalho de conclusão de curso de graduação  
apresentado ao Departamento de Ciência da  
Computação da Universidade Federal do Rio  
de Janeiro como parte dos requisitos para ob-  
tenção do grau de Bacharel em Ciência da  
Computação.

Orientador: Profa. Maria Luiza Machado Campos  
Co-orientador: Vânia Jesus de Araujo Soares Borges

RIO DE JANEIRO  
2022

### CIP - Catalogação na Publicação

R484t      Ribeiro, Tatiana de Sousa  
              Titulo / Tatiana de Sousa Ribeiro. -- Rio de  
              Janeiro, 2018.  
              44 f.

              Orientador: Maria da Silva.  
              Trabalho de conclusão de curso (graduação) -  
Universidade Federal do Rio de Janeiro, Instituto  
de Matemática, Bacharel em Ciência da Computação,  
2018.

              1. Assunto 1. 2. Assunto 2. I. Silva, Maria da,  
orient. II. Titulo.

MAYARA MORAIS DOS SANTOS  
MARCOS VINICIUS MACHADO RODRIGUES

CRF4FAIR: MÓDULO PARA VERSIONAMENTO DE QUESTIONÁRIOS PARA  
APLICAÇÃO VODAN BR

Trabalho de conclusão de curso de graduação  
apresentado ao Departamento de Ciência da  
Computação da Universidade Federal do Rio  
de Janeiro como parte dos requisitos para ob-  
tenção do grau de Bacharel em Ciência da  
Computação.

Aprovado em \_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_

BANCA EXAMINADORA:

---

Nome do Professor Orientador  
Titulação (Instituição)

---

Nome do Professor1  
Titulação (Instituição)

---

Nome do Professor2  
Titulação (Instituição)

Às famílias e amigos, que estiveram ao nosso lado dando suporte e nos apoiando em todos os momentos nessa jornada das nossas vidas

## **AGRADECIMENTOS**

Este trabalho como parte de um todo, a iniciativa VODAN BR, tem sido realizado por conta de muitos esforços. Nossos agradecimentos à professora Maria Luíza Machado Campos por nos orientar ao longo deste trabalho, aos docentes que estão trabalhando na evolução da base de dados do projeto, em especial a professora Vânia Borges e sua co-orientação. Agradecemos também a equipe VODAN-Br, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - BRASIL (CAPES), ao CNPq - Código de Financiamento 3153/2018-0 , à FAPERJ, ao hospital Federal Gaffrée Guinle, do Rio de Janeiro e ao Hospital Municipal São José, de Duque de Caxias.

"Procure fazer as coisas da forma mais simples que você puder".

**Albert Einstein**

## RESUMO

Este trabalho nasce como um dos trabalhos que surgiram a partir da iniciativa **Virus Outbreak Data Network(VODAN) Implementation Network<sup>1</sup>**, a **Rede VODAN**, um esforço conjunto do CODATA<sup>2</sup>, RDA<sup>3</sup>, WDS<sup>4</sup>, and GO FAIR<sup>5</sup>.

A **Rede VODAN** objetiva o fomento da coleta e tratamento de registros médicos de pacientes infectados com COVID-19 ao redor do mundo e foi criada devido a necessidade de estabelecer infraestrutura para suportar a captura e uso de dados de machine learning e inteligência artificial a fim de descobrir padrões significativos em surtos epidêmicos, tanto na presente situação da epidemia da COVID quanto para futuras epidemias.

Após a coleta dos registros médicos, os dados são tratados, anotados com ontologias e publicados em Fair Data Points, seguindo os princípios FAIR, a saber que os metadados sejam Findable (Encontráveis), Accessible (Acessíveis), Interoperable (Interoperáveis) e Reusable (Reutilizáveis) em novas pesquisas, para monitoramento e outros propósitos sob condições bem definidas. O fluxo de Fairificação, a publicação de FAIR Data Points, vai desde a coleta e tratamento desses dados, até a triplicação e anotação em vocabulários, sendo possível adotar abordagens variadas de implementação e escolha de ferramentas dentro de cada uma das partes.

Nesse sentido, foi criado o braço brasileiro da Rede VODAN, a **VODAN Brazil Implementation Network**, a **REDE VODAN Brasil**, estabelecida como uma parceria da Fundação Oswaldo Cruz (Fiocruz), a Universidade Federal do Rio de Janeiro (UFRJ), a Universidade Federal do Estado do Rio de Janeiro (UNIRIO) e o Hospital Municipal São José Municipal.

**Palavras-chave:** Rede VODAN. VODAN Brazil Implementation Network. Virus Outbreak Data Network(VODAN) Implementation Network.



## ABSTRACT

This paper is born as one of the works that emerged from the **Virus Outbreak Data Network(VODAN) Implementation Network<sup>1</sup>**, the **VODAN Network**, a joint effort of CODATA<sup>2</sup>, RDA<sup>3</sup>, WDS<sup>4</sup>, and GO FAIR<sup>5</sup>.

The **VODAN Network** aims to foster the collection and processing of medical records of COVID-19 infected patients around the world and was created due to the need to establish an infrastructure to support the capture and use of machine learning and artificial intelligence data to find and artificial intelligence in order to discover meaningful patterns in epidemic outbreaks, both in the present situation of the epidemic of COVID and for future epidemics.

After collecting medical records, the data are processed, annotated with ontologies and published at Fair Data Points, following FAIR principles, namely that the metadata be Findable, Accessible, Interoperable, and Reusable. Reusable in new research, for monitoring and other purposes under well-defined well-defined conditions. The flow of Fairification, the FAIR Data Points publication, goes from the collection and processing of these data, to triplication and annotation into vocabularies, and it is possible to adopt varied approaches to implementation and choice of tools within each part.

In this sense, the Brazilian branch of the VODAN Network was created, the **VODAN Brazil Implementation Network, REDE VODAN Brasil**, established as a partnership of the Oswaldo Cruz Foundation (Fiocruz), the Federal University of Rio de Janeiro (UFRJ), the Universidade Federal do Estado do Rio de Janeiro (UNIRIO) and the Hospital Municipal São José Municipal.

**Keywords:** VODAN Network. VODAN Brazil Implementation Network. Virus Outbreak Data Network(VODAN) Implementation Network.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Tabelas do módulo de Gestão de usuários . . . . .	20
Figura 2 – Tabelas do módulo de Formulário de Pesquisa Clínica . . . . .	21
Figura 3 – Tabelas do módulo de Formulário de Pesquisa Clínica . . . . .	22
Figura 4 – Ciclo de vida da pesquisa . . . . .	24
Figura 5 – Diagrama de casos de uso do problema . . . . .	26
Figura 6 – Foco nas classes relacionadas a criação e publicação de pesquisas. . . .	27
Figura 7 – Relacionamentos entre as generalizações dos casos P1 e P2. . . . .	28
Figura 8 – Relacionamentos entre os casos de P03 a P10 com os casos P01 e P02. .	29
Figura 9 – Adaptações propostas na base para o ciclo de vida da pesquisa . . . . .	31
Figura 10 – Criação de uma pesquisa Derivada . . . . .	34
Figura 11 – Pesquisa WHO Full criada a partir da WHO Rapid . . . . .	35
Figura 12 – Tela de detalhamento da Pesquisa WHO RAPID . . . . .	35
Figura 13 – Tela de detalhamento do WHO Full . . . . .	36
Figura 14 – Clique no botão de editar formulário de módulo de pesquisa publicada	37
Figura 15 – Tela de edição de Formulário de módulo publicado . . . . .	37
Figura 16 – O relacionamento da tabela tb_QuestionGroup . . . . .	39

## LISTA DE CÓDIGOS

## LISTA DE TABELAS

## LISTA DE QUADROS

## LISTA DE ABREVIATURAS E SIGLAS

Fig.	Area of the $i^{th}$ component
456	Isto é um número
123	Isto é outro número
Bibliot.	Biblioteconomia
Inform.	Informática
ABNT	Associação Brasileira de Normas Técnicas
I <sup>2</sup> C	Inter-Integrated Circuit
SRAM	Static Random-Access Memory
EEPROM	Electrically Erasable Programmable Read-Only Memory
LED	Light-Emitting Diode
MLP	Modulação por Largura de Pulso
PWM	Pulse-Width Modulation
PID	Proportional–Integral–Derivative
RAM	Random-Access Memory
API	Application Programming Interface
GPL	GNU General Public License
GNU	GNU's Not Unix
iid	Independente e identicamente distribuídas

## LISTA DE SÍMBOLOS

$\Gamma$	Letra grega Gama
$\Lambda$	Lambda
$\zeta$	Letra grega minúscula zeta
$\in$	Pertence
$\$$	subcampo

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>15</b>
1.1	MOTIVAÇÃO . . . . .	15
1.2	OBJETIVO . . . . .	16
1.3	ESTRUTURA . . . . .	17
<b>2</b>	<b>CONTEXTUALIZAÇÃO . . . . .</b>	<b>18</b>
2.1	A BASE DE DADOS ORIGINAL . . . . .	19
<b>3</b>	<b>A APLICAÇÃO CRF4FAIR . . . . .</b>	<b>24</b>
3.1	O CICLO DE VIDA DA PESQUISA . . . . .	24
3.2	DIAGRAMA DE CASO DE USO . . . . .	25
3.3	PROPOSTA DE EVOLUÇÃO . . . . .	31
3.4	STORED PROCEDURES . . . . .	32
<b>4</b>	<b>IMPLEMENTAÇÃO . . . . .</b>	<b>34</b>
<b>5</b>	<b>CONCLUSÃO . . . . .</b>	<b>38</b>
<b>6</b>	<b>TRABALHOS FUTUROS . . . . .</b>	<b>40</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>41</b>
	<b>GLOSSÁRIO . . . . .</b>	<b>42</b>
	<b>APÊNDICE A – APÊNDICE A . . . . .</b>	<b>44</b>
	<b>APÊNDICE B – APÊNDICE B. . . . .</b>	<b>45</b>
	<b>ANEXO A – ANEXO A . . . . .</b>	<b>47</b>
	<b>ANEXO B – ANEXO B . . . . .</b>	<b>48</b>



# 1 INTRODUÇÃO

Com a pandemia do Coronavírus SARS-CoV 2, que começou em dezembro de 2019 na cidade de Wuhan, criou uma avalanche de dashboards. Tal necessidade trouxe à tona a importância da atuação em agilizar mecanismos que permitam coletar, reunir e explorar dados. Diversos grupos internacionais se uniram neste propósito, em especial, os participantes da Research Data Alliance (RDA)<sup>1</sup>, da World Data Systems (WDS)<sup>2</sup>, da rede GO FAIR<sup>3</sup> e do Committee on Data (CODATA)<sup>4</sup> lançaram um apelo por ações chamado Data Together (Data Together, 2020)<sup>5</sup>.

Uma dessas iniciativas foi o projeto *Virus Outbreak Data Network* (VODAN) que surgiu em 2020 como uma rede de implementação ligada à iniciativa GO FAIR, com a necessidade de trabalhar dados epidemiológicos adotando representações que possibilitasse o uso de abordagens de Inteligência Artificial para descobrir padrões significativos em surtos epidêmicos. Tal esforço se faz por meio da criação de uma rede federada de dados usando os princípios FAIR - *Findable, Accessible, Interoperable, Reusable* - (WILKINSON, 2017) a fim de permitir sua disponibilização e reuso em novas pesquisas, respeitando as regras de privacidade exigidas. (Rodrigues, Pinto, 2021)

Os princípios FAIR propõem uma estruturação dos dados para que a computação possa trabalhar e fazer inferências sobre eles, sobretudo para conter metadados. Além disso, propõe que os dados sejam coletados, analisados, sejam definidos em um modelo semântico para que sejam descritas suas relações com outros dados de (HEALTH, BIZER 2011).

Em outras palavras, FAIR é um movimento no sentido de ampliar o compartilhamento e interoperabilidade de dados de pesquisa (GO FAIR), calçado em um conjunto de princípios gerais, que orientam para boas práticas e para a construção de uma infraestrutura de apoio para publicação Web.

## 1.1 MOTIVAÇÃO

Apesar de que muitas informações estatísticas foram disponibilizadas na forma de painéis desde o início da Pandemia, dando uma ideia geral do fenômeno e sua evolução, para estudos mais aprofundados e análise de características dos infectados e do seu contexto,

---

<sup>1</sup> <https://www.rd-alliance.org/>

<sup>2</sup> <https://www.worlddatasystem.org/>

<sup>3</sup> <https://www.go-fair.org/>

<sup>4</sup> <https://codata.org/>

<sup>5</sup> <https://www.go-fair.org/go-fair-initiative/data-together/>

outro nível de detalhe é requerido. Além disso, em boa parte deles não há acesso aos dados que deram origem, numa granularidade mínima, o que limita o uso dessas informações apenas à visualização.

Somado a isso, existe uma exigência de velocidade na busca de padrões nestes surtos epidêmicos, para além de dashboards muito objetivas e pouco interativas e que respondem não muito mais do que se possa pensar de resultados, inferir como as pandemias afetam outros aspectos da vida humana para além somente da área de saúde e obter informações estratégicas valiosas para o enfrentamento de cenários futuros a partir de uma visão bem mais ampla desses impactos.

Visando prover uma infraestrutura escalável, genérica e distribuída para atender um processo de coleta intensivo de alta heterogeneidade, disponibilizando-os em plataformas que ofereçam dados e metadados interoperáveis e processáveis por agentes de software a REDE Vodan desenvolveu um braço, A REDE VODAN Brasil, ou projeto VODAN BR<sup>6</sup>, uma parceria da Fundação Oswaldo Cruz (Fiocruz)<sup>7</sup>, a Universidade Federal do Rio de Janeiro (UFRJ)<sup>8</sup> e a Universidade Federal do Estado do Rio de Janeiro (UNIRIO)<sup>9</sup>. O desenvolvimento da infraestrutura está sob a responsabilidade da equipe do grupo GRECO (Grupo de Engenharia do Conhecimento)<sup>10</sup> da UFRJ, e tem como parceiros o Hospital Federal Gaffrée Guinle<sup>11</sup>, do Rio de Janeiro, e o Hospital Municipal São José, de Duque de Caxias.

## 1.2 OBJETIVO

Este trabalho, que faz parte de um dos muitos trabalhos financiados pelo projeto VODAN BR, propõe e implementa o módulo de Gestão de Formulário de Pesquisas Clínicas. Inicialmente realizou-se uma análise da base de dados e planejamento do processo de desenvolvimento e uma segunda etapa a implementação das alterações na base de dados e a implementação das funcionalidades do módulo de pesquisa.

O escopo limita-se ao contexto de formulários de clínica médica, mas sua realização considerou flexibilizar tanto o uso da aplicação quanto da base de dados para atender a um domínio genérico a posteriori. Este trabalho limita-se às pesquisas clínicas e o contexto da área de saúde.

---

<sup>6</sup> <https://vodanbr.github.io/>

<sup>7</sup> <https://portal.fiocruz.br/>

<sup>8</sup> <https://ufrj.br/>

<sup>9</sup> <http://www.unirio.br/>

<sup>10</sup> <http://www.greco.ppgi.ufrj.br/>

<sup>11</sup> [http://www.unirio.br/hugg\\_geral](http://www.unirio.br/hugg_geral)

### 1.3 ESTRUTURA

**Capítulo 1** contém introdução, motivação e objetivo sobre o aplicativo e estrutura de dados que será abordada no trabalho.

**Capítulo 2** apresenta a contextualização e reiteração o escopo geral do projeto do qual o esse trabalho está inserido.

**Capítulo 3** aborda o processo de desenvolvimento da aplicação.

**Capítulo 4** apresenta a implementação e o foco principal do desenvolvimento da aplicação.

**Capítulo 5** contém as conclusões, dificuldades encontradas e trabalhos futuros relacionados ao projeto.

**Capítulo 6** trabalhos futuros a partir da aplicação desenvolvida.

## 2 CONTEXTUALIZAÇÃO

Uma maneira viável de obtenção de dados de dados em uma granularidade menor seria obter os dados dos pacientes diretamente. Isso constitui um desafio em termos de acesso e permitir proteção de dados dos indivíduos. Uma solução possível seria a de captar estes dados por meio de pesquisas nos hospitais, havendo uma participação dos pacientes da área de saúde.

Essa obtenção em si constitui um desafio. Obter informações em uma granularidade implica em proteger os dados dos pacientes e manter a anonimidade e outro com relação às unidades hospitalares de softwares para prontuários eletrônicos sem maior estruturação para a entrada de dados, com muitos campos de texto livre.

Outro desafio à vista seria a questão de estruturas destes dados numa forma comum, criando um dicionário de dados que fizesse sentido para a área de saúde que permitisse a interligação com outras bases. Um caminho viável seria utilizar um modelo conhecido. A OMS<sup>1</sup> desenvolveu para a pesquisa clínica o modelo CRF (FPC, do original em inglês, *Clinical Research Form*), a pesquisa clínica Who Rapid COVID. Nesse template uma pesquisa clínica é realizada em três momentos distintos de atendimento do paciente. Na entrada no hospital ao primeiro atendimento, no meio do tratamento ou ao final, quando da alta médica, de óbito ou uma transferência de unidade.

Em virtude dos hospitais participantes adotarem prontuários eletrônicos (PE), optou-se pela captura das informações a partir desses PE, com o aplicação facilitando o complemento de informações. O complemento se tornou necessário por haver informações que estão registradas apenas nos Prontuários físicos ou que apresentam dificuldade na extração dos PE.

A fim de solucionar este problema desenvolveu-se um aplicação: inicialmente denominado VODAN BR, agora CRF4Fair. A primeira versão nasceu a partir de um trabalho final de um dos alunos na disciplina (XXX, ministrada pela professora Maria Luíza Machado Campos, Ademario XXX) do curso de Ciência da Computação.

Desenvolveu-se um sistema para registro de pesquisas a partir de PEs, usando o framework PHP Laravel, no backend, e a biblioteca Javascript React, no frontend.

---

<sup>1</sup> <https://www.who.int/>

O primeiro e principal objetivo da aplicação era o de fornecer uma entrada de dados (interface), aos profissionais de saúde, para o registro das informações solicitadas a fim de permitir a execução de pesquisas com os pacientes com COVID-19. Como tecnologias, o framework PHP Laravel, no backend, e a biblioteca Javascript React, no frontend.

O módulo de Formulários de Pesquisas Clínicas foi criado no intuito de executar o primeiro objetivo é apenas a execução de pesquisas com o template CRF foi implementada. Planejou-se também a implementação de outros módulos nas versões seguintes, o módulo de Gestão de Apoio a Pesquisa Clínica, o módulo de Gestão de Ontologias e o módulo de Gestão de Metadados.

O intuito do módulo de Apoio a Pesquisa Clínica é o dar uma vida útil para as pesquisas clínicas. Através de um sistema de versionamento e ao permitir a edição e criação de novas pesquisas.

O módulo de gestão de Ontologias e o módulo de gestão de metadados serão para que uma vez que os dados de pesquisa estiverem na base permita-se a inclusão e atualização de artefatos semânticos, associando-os a pesquisas clínicas cadastradas e o último para o gerenciamento de metadados referentes a datasets em RDF (*Resource Description Framework*), gerados com os dados das pesquisas.

## 2.1 A BASE DE DADOS ORIGINAL

A base de apoio utilizada no CRF4Fair é um trabalho de docentes da UFRJ. Ela vem sendo evoluída para atender a necessidade de prover uma estrutura relacional com dados de pacientes do COVID que possam prover pesquisas científicas permitindo sua interligação com outras aplicações. Assim sendo ela não foi desenvolvida para o projeto presente mas o projeto tem um papel de poder evoluí-la com mudanças que irão prover a implementação do versionamento de pesquisa.

Para manter a flexibilidade de uso, o princípio de construção da base de dados é não estar amarrada no contexto de pesquisa clínica, nem nas regras de negócio do template CRF, o que confere a capacidade de se aplicar a outros contextos, não somente de pesquisa clínica mas de pesquisa, em geral.

O SGBD(Sistema de banco de dados) é o MYSQL e o mecanismo de armazenamento MariaDB, posteriormente o InnoDB. Há cerca de 40 quarenta tabelas divididas entre os quatro módulos. A seguir apresenta-se cada um desses grupos. Por simplificação, estão apreciados apenas as tabelas dos módulos de interesse. Dessa forma, as tabelas dos mó-

dulos referentes a ontologias estão omitidas.

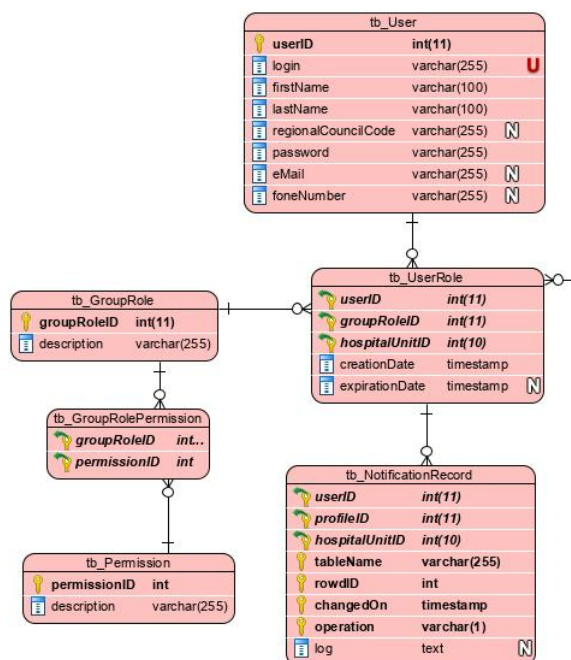


Figura 1 – Tabelas do módulo de Gestão de usuários

As tabelas em rosa representam os usuários, suas permissões e o seu histórico de operações realizadas. O usuário logado (`tb_user`) cuja permissão da categoria (`tb_UserRole`) relaciona-se com relacionamento entre as permissões e os papéis dos usuários (`tb_GroupRole`) pro meio de uma tabela (`tb_GroupRolePermission`) identificada por (`tb_Permission`).

As tabelas na cor laranja representam os registros dos pacientes coletadas pelas unidades hospitalares. As informações do paciente da pesquisa são armazenadas na tabela `tb_participant` e as da Unidade Hospitalar em `tb_HospitalUnit`.

Observe que a única informação que é registrada é o registro médico e informações pessoais, que identificam o participante, não são armazenadas. `tb_FormRecord`: Relaciona a pesquisa com o formulário de módulo. A tabela `tb_AssessmentQuestionnaire` registra a abertura de um questionário de pesquisa para um participante. O relacionamento entre cada pesquisa, os módulos e os pacientes são registrados em `tb_FormRecord`. Por último, a tabela `tb_QuestionGroupFormRecord` registra os preenchimentos de cada formulário de módulo.

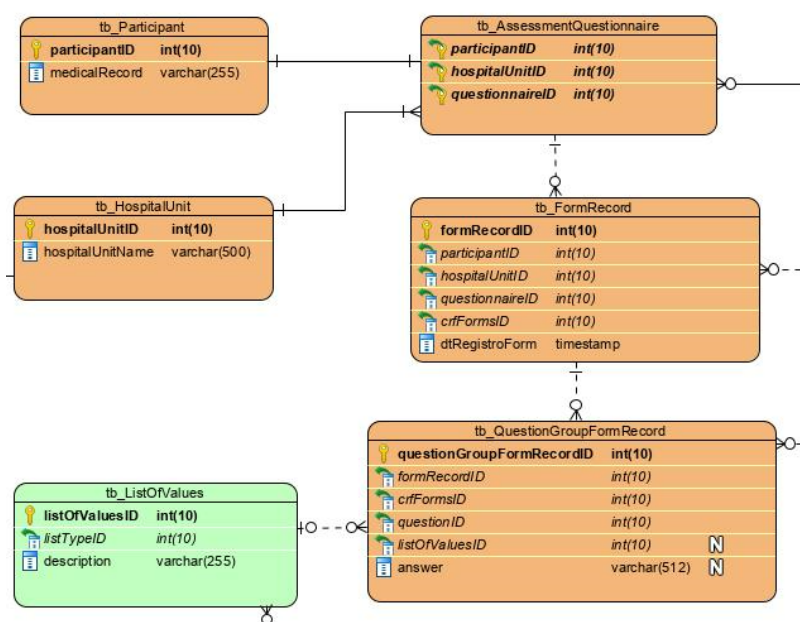


Figura 2 – Tabelas do módulo de Formulário de Pesquisa Clínica

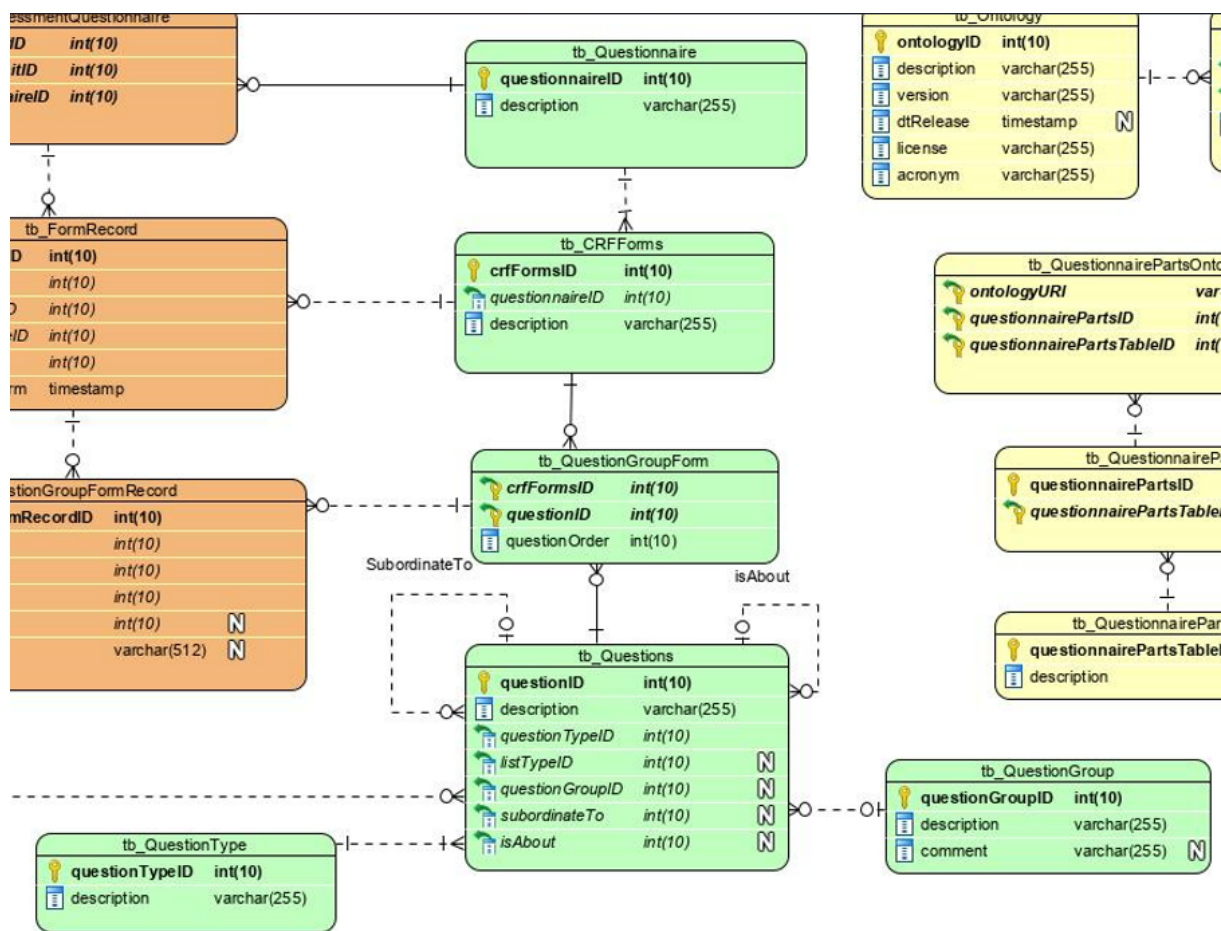


Figura 3 – Tabelas do módulo de Formulário de Pesquisa Clínica

Todavia as tabelas da parte verde correspondem a estrutura hierárquica da Pesquisa. Os formulários de módulos, as questões, agrupamentos, os tipos de questões e os relacionamentos entre eles. Um questionário de pesquisa (**tb\_Questionnaire**) tem um ou mais módulos (**tb\_CRFFForms**). Cada questionário está associado a uma questão (**tb\_Questions**) de um tipo (**tb\_QuestionType**) por uma tabela (**tb\_QuestionGroupForm**). Quando uma questão é de um tipo de lista (**tb\_ListType**) há valores resposta predefinidos (**tb\_ListOFValues**).

Além de tabelas ad hoc, há uma parte de Internacionalização para permitir a tradução de quaisquer termos para qualquer idioma. Traduções de termos (**tb\_multilanguage**) são registradas nos idiomas (**tb\_language**).

Levando em consideração a complexidade da base e em possibilitar o reuso das informações, a consultas ad hoc preferiu-se o desenvolvimento de Stored Procedures podem ser utilizadas quando o tráfego de informações na rede é grande, ou quando a aplicação tem baixo desempenho e o constatado que o consumo maior se deve à manipulação de dados entre aplicação e banco de dados.



As procedures seguem um esquema de nomenclatura de acordo com o tipo de interação com a aplicação:

- Começam com “get” quando é uma consulta SELECT.
- Começam com “post” quando é uma consulta INSERT ou de UPDATE.

Embora não necessariamente uma pesquisa clínica possa ter essas características, na primeira versão do banco de dados, implementou-se apenas uma única pesquisa clínica composta de 3 (três) formulários, a “WHO COVID-19 Rapid Version CRF”, os quais:

- Formulário de Admissão: formulário preenchido no momento em que um paciente dá entrada no hospital pela primeira vez.
- Formulário de Acompanhamento: formulário preenchido em algum momento entre a Admissão e a Alta.
- Formulário de Alta/Óbito: formulário preenchido no momento do encerramento do atendimento, seja por alta ou por óbito.

### 3 A APLICAÇÃO CRF4FAIR

O principal objetivo do módulo de Gestão de Pesquisa é prover atributos e dar um ciclo de vida útil para as pesquisas. Contudo, antes de proposições técnicas sobre este ciclo é preciso debruçar-se sobre a sua semântica.

#### 3.1 O CICLO DE VIDA DA PESQUISA

Após o processo de análise da base de dados sugeriu-se mudanças necessárias para a implementação do módulo de Gestão de Pesquisa Clínica. Essas mudanças permitem a criação de novas pesquisas clínicas e um sistema de versionamento.

Por conseguinte, um sistema de versionamento seria um ciclo de vida útil da pesquisa clínica, onde uma pesquisa nasceria por meio de sua criação, se evoluiria por meio de pequenas edições, se reproduziria por meio da publicação de uma nova versão, que viria a substituir e em seu estágio final causar sua morte, entrando em um estado de depreciação, em desuso. De todo esse processo uma pesquisa nova nasce repetindo-se o ciclo.

Após sua morte, a pesquisa torna-se um histórico no banco de dados.

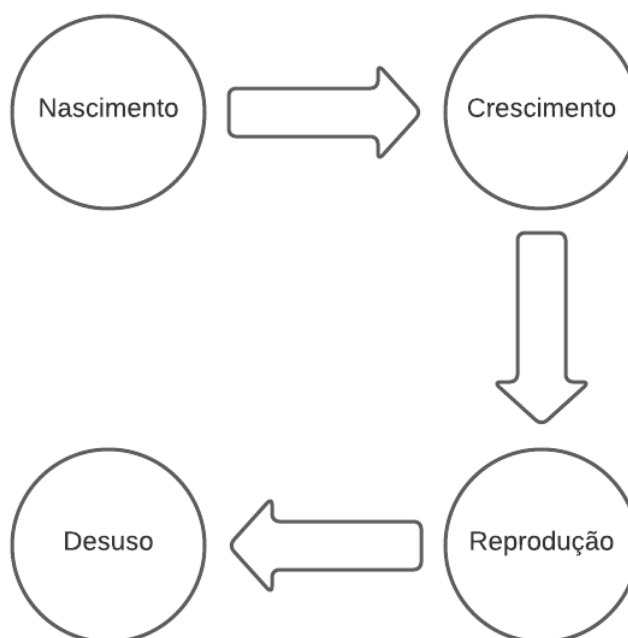


Figura 4 – Ciclo de vida da pesquisa

Todavia no nascimento de uma pesquisa poderia dar-se não por meio de outra pesquisa

que a originária mas poderia ser um evento original no qual a pesquisa seria a primeira versão da sua linhagem.

### 3.2 DIAGRAMA DE CASO DE USO

Na primeira fase do trabalho realizou-se um levantamento da documentação do projeto com um estudo das regras de negócio e uma análise do esquema da base de dados original. A partir desse levantamento, produziu-se um conjunto de diagramas de caso de usos.

Nos diagramas a seguir estão representados em verde os casos do módulo de Gestão de Usuários, em amarelo os do módulo de Formulário de Pesquisa e em azul os de Apoio a Gestão de Pesquisa.

Antes de avançar uma definição dos papéis envolvidos:

- O papel Administrador é o mais atuante no módulo de Pesquisa porque as ações desse módulo exigem afetam a forma que todos os outros papéis enxergam e utilizam o sistema.
- O papel Notificador profissional de Saúde é responsável por preencher os formulários de pesquisa clínica. Esse preenchimento é a execução de uma pesquisa clínica com dados de um determinado paciente não identificado que representa o estado clínico dele em um determinado dia.

Na figura abaixo segue um diagrama que exhibe os casos mais básicos de funcionalidade.

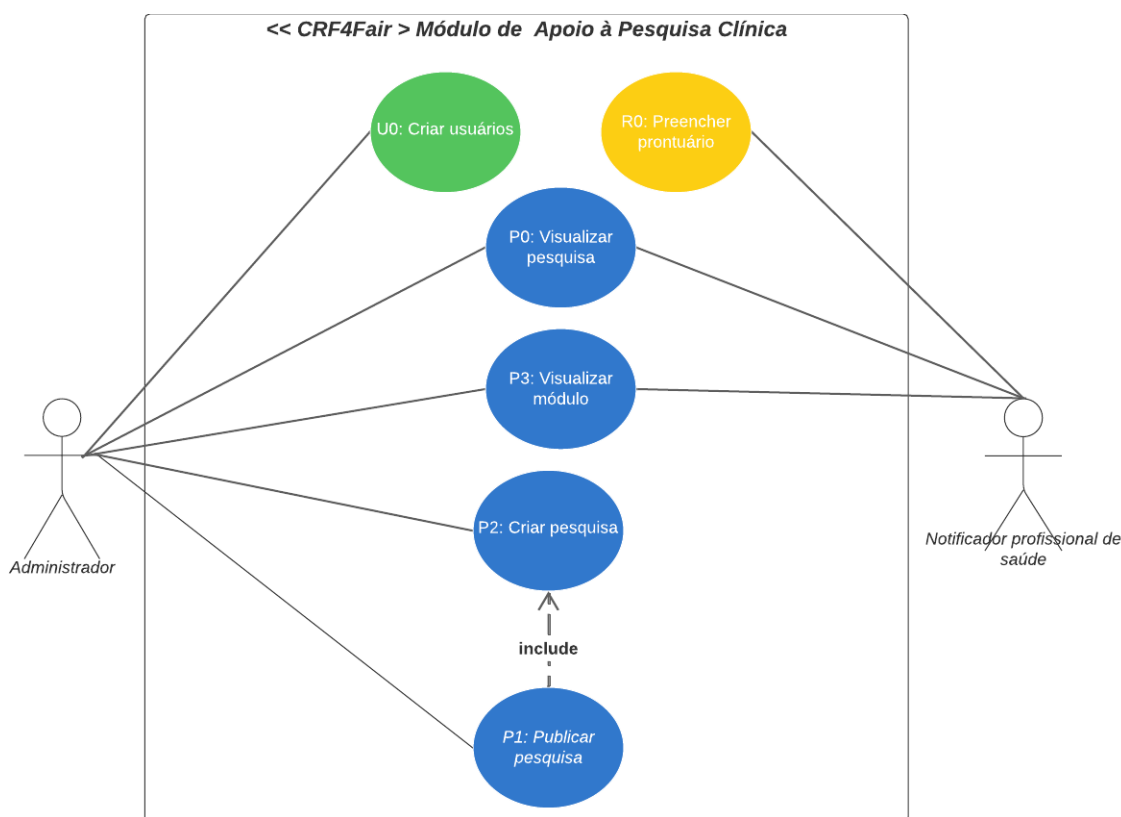


Figura 5 – Diagrama de casos de uso do problema

Observe que o papel Administrador cria usuários, cria e publica pesquisas. Esse papel tem outras opções de administração de usuários, mas por simplicidade apenas a criação está ilustrada para exemplificar que este usuário tem um papel mais alto.

Ao notificador de saúde que preenche as pesquisas apenas cabe a visualização de uma pesquisa e o preenchimento.

Publicar uma pesquisa implica que ela tenha sido criada em algum momento, daí a relação de include entre os casos P0 e P1.

Dando um zoom nos casos de uso de P1: Publicação de pesquisa e P2: criação de pesquisa. Há relações intrínsecas e fortes entre esses dois casos.

Publicar uma pesquisa acontecem há três possíveis hipóteses a se considerar:

1. Essa pesquisa publicada é uma versão nova de uma outra pesquisa.

2. A pesquisa baseou-se em outra, mas não representou um versionamento apenas tomou como ponto de partida para sua estrutura.
  3. A pesquisa é criada do zero, sem se basear em outra e não representa uma versão,
- A hipótese 1) enquadra-se no caso de uso P01. As hipóteses 2 e 3, no P02.

Os casos de uso P0 e P1 são generalizações do caso P1.

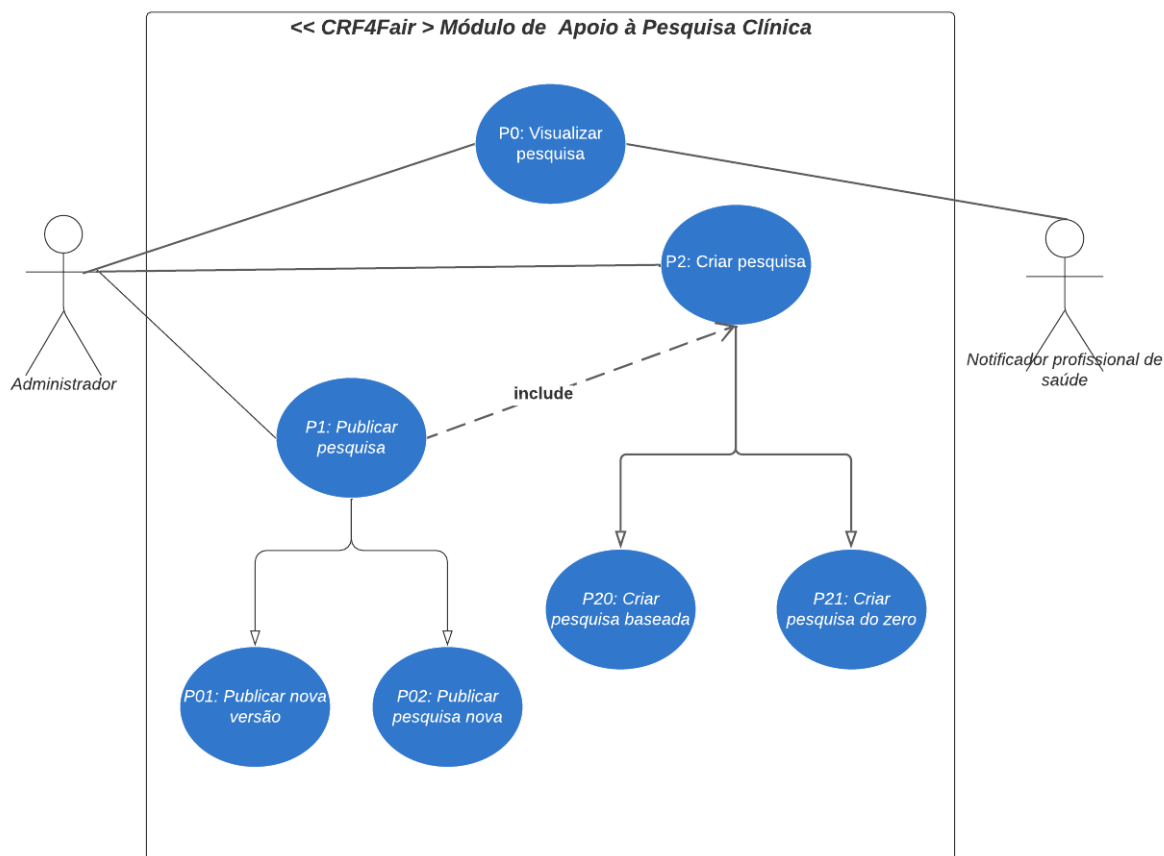


Figura 6 – Foco nas classes relacionadas a criação e publicação de pesquisas.

Os casos P20 e P21 são generalizações do caso P2.

P20, criar uma pesquisa baseada ,significa criar uma pesquisa que toma por base a estrutura copiada de outra pesquisa, como uma espécie de template. Duas hipóteses seguem neste caso:

1. A pesquisa criada era uma nova versão.
2. A pesquisa criada não era uma nova versão.

Para o caso P21 só existe um caminho possível que a pesquisa representa uma pesquisa nova mas que não representa uma nova versão (P02).

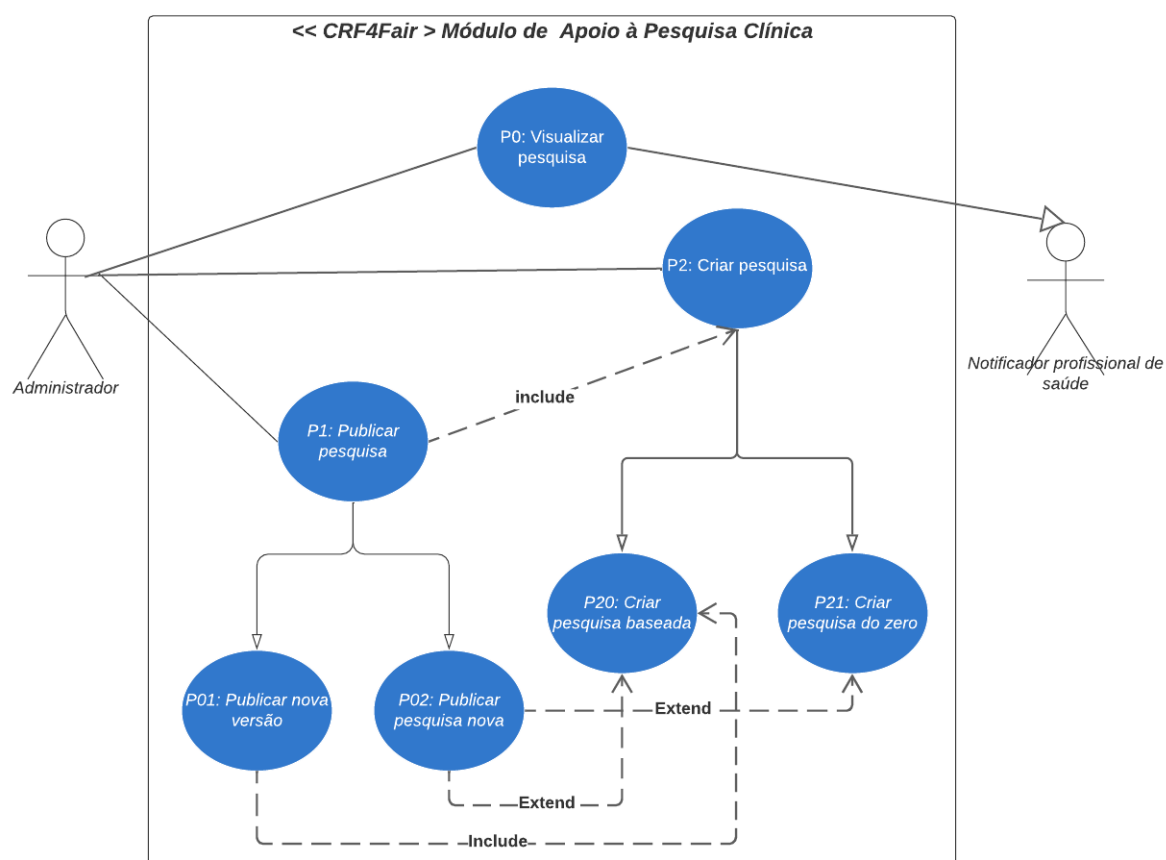


Figura 7 – Relacionamentos entre as generalizações dos casos P1 e P2.

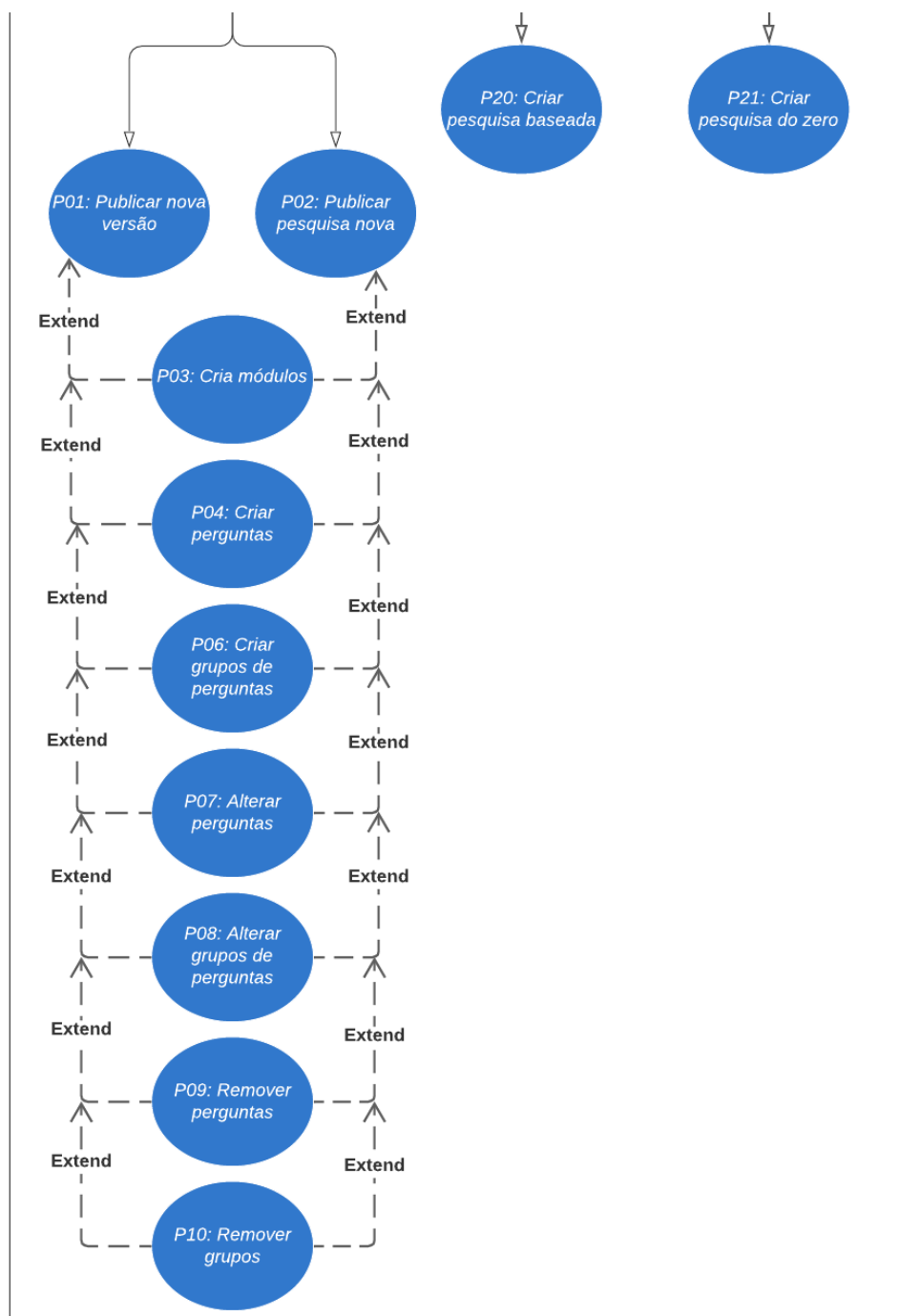


Figura 8 – Relacionamentos entre os casos de P03 a P10 com os casos P01 e P02.

Partindo de uma comparação entre os casos P01 e P02 quanto às possibilidades em relação às ações possíveis nas pesquisas o caso Criar módulos (P03) não é requisito obrigatório (Extend) quando um questionário publicado é uma nova versão (P01) porque uma nova versão de um questionário pode ter os mesmos módulos de seu pai. Analogamente, para todos os casos de P04 a P09 aplica-se raciocínio análogo.

Em suma, todas as relações entre os casos de P03 a P09 são relações Extend em relação ao caso P01 porque é possível publicar uma nova versão que seja idêntica à sua pesquisa

mãe. Em relação ao caso P02, segue-se o mesmo porque apesar de uma pesquisa poder ter sido criada do zero neste caso, também há a possibilidade de ser um template (P20).

Observe que não há sentido em existir um caso de uso para a publicação de módulo porque um módulo é um componente de uma pesquisa e não o contrário, e a versão é um atributo dela.

Por fim, um resumo dos casos:

- P0: Visualizar uma pesquisa. Enxergar os status da pesquisa, a descrição e outros metadados, a lista de módulos que pertence a pesquisa.
- P1: Publicar uma pesquisa significa tornar a pesquisa visível para que os atores dos papéis que preenchem uma realização, a execução de uma pesquisa de um paciente.
- P01: Publicar uma nova versão. Publicar uma pesquisa que foi criada como versão nova de outra pesquisa e toma por base a pesquisa antiga como template. A nova versão substitui a pesquisa antiga, que é depreciada e se torna invisível para os papéis que preenchem execuções.
- P02: Publicar uma pesquisa nova. Publicar uma pesquisa que foi criada usando outra como template mas que não é uma nova versão de outra pesquisa ou totalmente do zero.
- P2: Criar uma pesquisa. Criar uma pesquisa para ser trabalhada, editada a fim de publicar.
- P20: Criar uma pesquisa baseada. Criar uma pesquisa que toma por base outra pesquisa como template mas que será uma pesquisa independente ou uma nova versão.
- P21: Criar uma pesquisa do zero. Criar uma pesquisa que não será uma nova versão de outra pesquisa.
- P03: Criar módulos.
- P04: Criar perguntas. Criar perguntas novas em questionário de módulo.
- P05: Criar grupos de perguntas. Criar grupos de perguntas, grupos em que as perguntas são agrupadas em que todas estão relacionadas por um mesmo assunto.
- P06: Alterar perguntas. Editar perguntas. Editar descrição e outros atributos de uma pergunta. Mudar ordem de perguntas. Mudar a ordem em que as perguntas aparecem em um questionário de módulo.



- P7: Alterar grupos de perguntas. Mudar a ordem em que os grupos de perguntas aparecem em um questionário de módulo.
- P08: Remover perguntas. Remover perguntas de um questionário de módulo.
- P09: Remover grupos. Remover grupos de um questionário de módulo.
- P3: Visualizar o módulo.

### 3.3 PROPOSTA DE EVOLUÇÃO

Tomando por base os diagramas de casos de uso apresentados na seção anterior propõem-se mudanças na base de dados original para resolver o problema do versionamento de pesquisa. As mudanças mantêm sobretudo a premissa de generalidade de contexto que é um princípio basilar de construção da base.

Prover atributos para as pesquisas que dariam significado às relações entre a pesquisa e a pesquisa que é a versão anterior. Assim como explicitar quando uma pesquisa utilizou outra contra template.

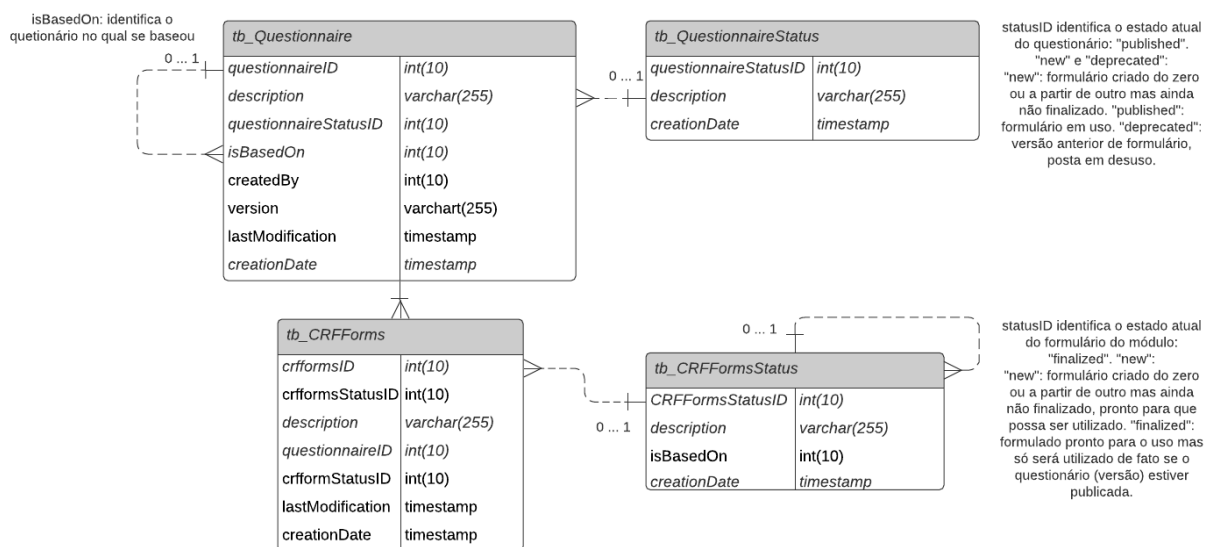


Figura 9 – Adaptações propostas na base para o ciclo de vida da pesquisa

- Alterações na tabela tb\_questionnaire:
  - questionnaireStatusID: status do questionário.
  - isBasedOn: questionário template.
  - version: o número da versão, começando a partir de 0.0.

- lastModification: a última data de modificação do questionário.
- createDate: a data de criação do questionário.
- Alterações na tabela tb\_CRFForms:
  - crfformsStatusID: status do formulário de módulo.
  - lastModification: a última data de modificação do questionário.
  - createDate: a data de criação do questionário.
- A tabela tb\_QuestionnaireStatus registro dos status possíveis para um questionário, a saber:
  - Published: publicado
  - Deprecated: deprecado, em desuso.
  - New: não publicado mas não deprecado.
- A tabela tb\_CRFFormsStatus para os possíveis status de um formulário de módulo:
  - Published: publicado
  - New: não publicado

Observe que um módulo não tem o status deprecado e nem informação de versão porque quem é versionado é o questionário. Um módulo está publicado ou não depende do status do questionário porque só é publicado quando o questionário é publicado.

### 3.4 STORED PROCEDURES

A infraestrutura de comunicação da aplicação com o banco de dados dá-se por meio de stored procedures. Para o módulo de gestão de pesquisas desenvolveu-se uma infraestrutura para atender ao conjunto de funcionalidades previstas durante o processo de análise. Por questões de padronização, a nomenclatura para o nomes das procedures segue o esquema de nomeação que se usa no banco de dados.

Criação de procedures para obter os metadados (get) de uma questionário e permitir outras manipulações:

- **getAllQuestionnaires**: procedure que obtém todos os questionários, sem nenhum parâmetro.
- **getQuestionnaire({\$id})**: obtém o status do questionário e outras informações a partir de um questionário identificado por questionnaireID da tabela tb\_Questionnaire.

- **postQuestionnaire**: cria um novo questionário. ({questionnaireID}, {questionnaireStatus}, {description}, {isNewVersionOf}, {isBasedOn}, {createdBy}, {version}, {last Modification}, {creationDate})
- **putQuestionnaire**: altera os metadados de um questionário ({questionnaireID}, {questionnaireStatus}, {description}, {version}, {lastModification}, {questionnaireStatus})

Para os formulários de módulos:

- **getCRFForm({\$crfformsID})**: obtém um módulo (crfforms). Retorna um formulário cujo identificador é \$crfformsID. Serve para identificar se um formulário existe, e seu status, a data de modificação, etc metadados de tb\_CRFForms.
- **postCRFForms**: cria novos módulos (tb\_CRFForms) ({\$crfformsID}, {description}, {questionnaireID}, {lastModification}, {creationDate}), {\$excluded}).
- **putCRFForms**: Atualiza um formulário existente, como status e o nome (descrição). Serve também para remover um módulo usando o parâmetro \$excluded. ({\$crfformsID}, {description}, {questionnaireID}, {lastModification}, {\$excluded}).
- **putQuestions**: Alterar descrições, o agrupamento, ordem, tipos das questões, o tipo de lista se for uma pergunta do tipo de pergunta com listas de respostas predefinidas e as listas de valores associadas. Remoção. ({\$crfformsID}, {questionID}, {\$groupsID}, {questionsTypeID}, {\$listTypesID}, {\$listOfValuesID}, {\$excluded}, {questionOrder}).
- **putGroups({\$crfformsID}, {\$questionGroupID}, {\$description}, {\$comment}, {\$excluded})**: alterar e remover grupos de perguntas.
- **putQuestionsTypes({\$questionTypeID}, {\$description}, {\$excluded})**: alterar e remover os tipos de perguntas (boolean..)
- **putListTypes({\$listTypeID}, {\$description}, {\$listTypeDescription}, {\$excluded})**: alterar e remover tipos de perguntas com respostas já pré definidas.
- **putListOfValues({\$listOfValuesID}, {\$description}, {\$comment}, {\$excluded})**: alterar e remover as opções de respostas de perguntas do tipo lista já predefinidas.

## 4 IMPLEMENTAÇÃO

O foco do presente trabalho é versionar a pesquisa WHO Rapid COVID-19 CRF (WHO Rapid) a partir da criação do template WHO COVID-19 Full Version CRF (WHO Full), sua versão mais completa. Essa versão é estruturalmente idêntica a Full apenas possui mais agrupamentos de questões.

A primeira parte é a criação da pesquisa (caso P2) WHO Full como uma pesquisa derivada de WHO RAPID (caso P20) estabelecida uma relação de versionamento. Como usuário logado na aplicação devidamente autorizado para realizar gerenciamento de pesquisas uma tela de seleção de hospital é aberta. Nesta tela o usuário após selecionar um dos hospitais a opção “Gerenciar pesquisas” é aberta uma tela que lista todas as pesquisas ativas (caso P0). Nesta tela há um botão para “Adicionar Pesquisa” que abre as opções “Adicionar pesquisa do zero” (P21) e “Adicionar Pesquisa Derivada” (P20). O usuário seleciona a opção “Adicionar Pesquisa Derivada”.

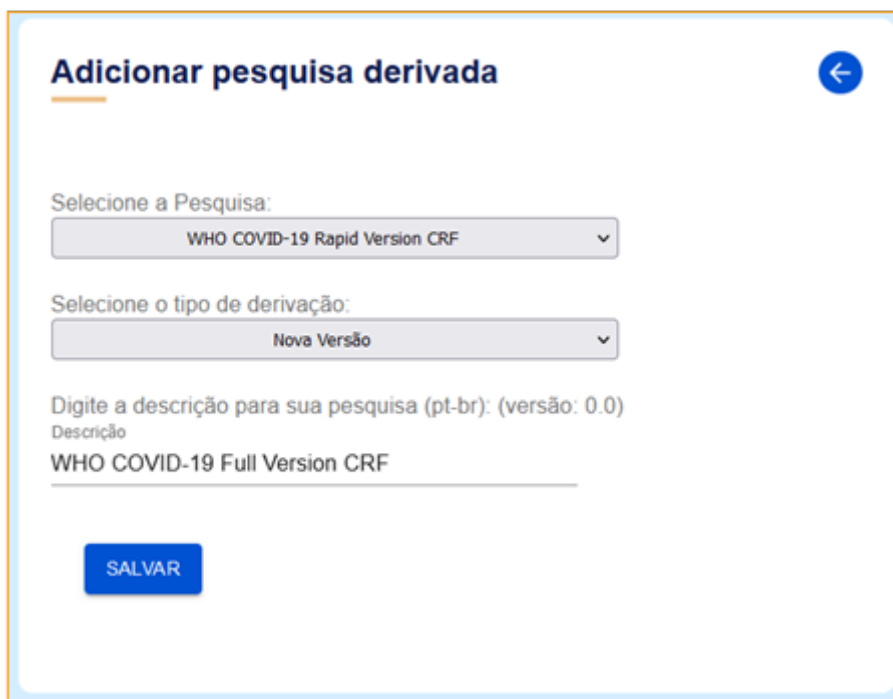


Figura 10 – Criação de uma pesquisa Derivada

Nesta tela o usuário seleciona qual a pesquisa a nova irá se basear em um dropdown e em outro seleciona o tipo de relacionamento com ela. Dentre as opções “Nova versão” e “Como template”. A primeira opção estará sendo criada uma nova versão e na segunda uma apenas uma pesquisa que se baseará em sua estrutura como template.



**Gerenciamento de Pesquisas Clínicas**

Lista de todas as pesquisas clínicas

Descrição da pesquisa

PESQUISA CLÍNICA	VERSÃO	STATUS	CRIADO EM	MODIFICADO EM	VER
WHO COVID-19 Rapid Version CRF	1.0	Publicado	17/11/2021	17/11/2021	
WHO COVID-19 Full Version CRF	0.0	Novo	26/02/2022	26/02/2022	

Figura 11 – Pesquisa WHO Full criada a partir da WHO Rapid

Na página de detalhamento da pesquisa Full consta que ela é uma versão da pesquisa que deu origem e uma lista de seus módulos é exibida (caso P0).



Pesquisa: **WHO COVID-19 Rapid Version CRF**

Versão: 1.0 Status: Publicado Data de criação: 17/11/2021 Última modificação: 17/11/2021

Descrição do módulo

MÓDULO	STATUS	CRIADO EM	MODIFICADO EM	EDITAR
Formulário de Admissão	Publicado	17/11/2021	17/11/2021	
Acompanhamento	Publicado	17/11/2021	17/11/2021	
Formulário de alta/óbito	Publicado	17/11/2021	17/11/2021	

Figura 12 – Tela de detalhamento da Pesquisa WHO RAPID

Pesquisa: **WHO COVID-19 Full Version CRF**

Versão: **0.0** Status: Novo Data de criação: 26/02/2022 Última modificação: 26/02/2022

Essa pesquisa é uma versão de: WHO RAPID COVID-19 Rapid Version.

Descrição do módulo  **BUSCAR**

**+ ADICIONAR NOVO FORMULÁRIO DE MÓDULO +**

**+ PUBLICAR PESQUISA**

**+ ADICIONAR PESQUISA DERIVADA**

MÓDULO	STATUS	CRIADO EM	MODIFICADO EM	EDITAR
--------	--------	-----------	---------------	--------

Figura 13 – Tela de detalhamento do WHO Full

Após clicar em um dos módulos a tela de edição do formulário é exibida. Como a WHO Full não está publicada, nem todas as opções de edição estão disponíveis. No caso da Rapid as opções para edição são mínimas, haja visto que a pesquisa está publicada e isso poderia produzir resultados indesejáveis, como perdas de prontuários preenchidos.

Na tela de edição de formulário de módulo de pesquisa publicada há uma visualização de um relatório de toda a estrutura do formulário. Do lado esquerdo é exibido um menu dropdown flutuante que mostra âncoras para todos os agrupamentos de questões. A edição de formulário limita a edição dos textos das perguntas. Do lado direito são exibidas todas as questões e ordenadas por seus grupos. As questões sem agrupamento são exibidas primeiro e as demais são ordenadas de acordo com a ordem no grupo. Para cada uma delas é exibido o seu tipo e um input de texto com a descrição e ao final um botão para salvar as alterações.

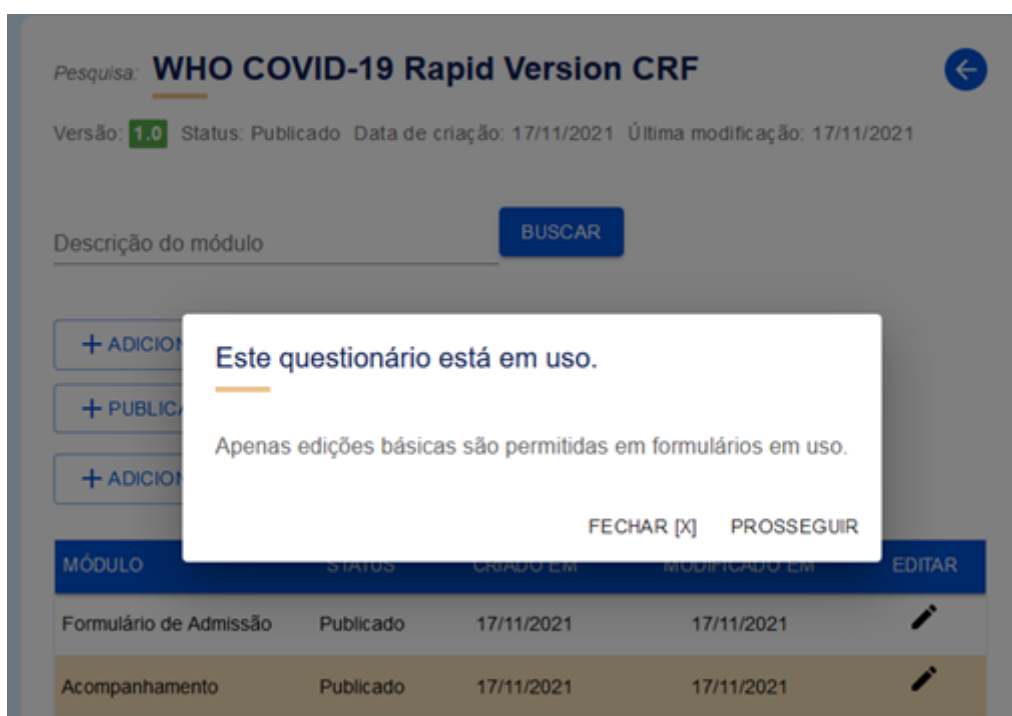


Figura 14 – Clique no botão de editar formulário de módulo de pesquisa publicada

MENU DE NAVEGAÇÃO

- Critérios Clínicos para Inclusão
- Dados demográficos
- Início da doença e sinais vitais na admissão
- Comorbidades
- Pré-admissão e medicamentos de uso contínuo
- Sinais e sintomas na hora da admissão
- Medicação
- Cuidados
- Resultados laboratoriais

Hospital municipal São José - Duque de Caxias > **Admissão**

**Módulo 1 - Admissão - Publicado [Edição]**

Questionário: WHO COVID-19 Rapid Version CRF ( 1.0 ) Publicado

Questão:

Nomee da Instalação

Tipo da questão: Text\_Question

Questão:

País

Figura 15 – Tela de edição de Formulário de módulo publicado

## 5 CONCLUSÃO

O desenvolvimento de artefatos computacionais para fornecer dados confiáveis e de fonte rastreável para pesquisas científicas é um desafio tanto do ponto de vista infra estrutura quanto do ponto de vista do acesso às informações. A Vodan Internacional, representada pela Vodan BR, por meio de sua iniciativa de fomento de projetos que disponibilizam dados para este fim, desenvolveu uma plataforma para publicação de dados e fomenta iniciativas como o desenvolvimento do presente trabalho.

Essa experiência rendeu algumas lições ao longo do caminho. A partir do processo de análise do banco de dados relacional, nota-se algumas ressalvas quanto ao esquema de nomenclatura adotado. A tabela relativa à a um prontuário clínico de um paciente em dia chama-se `tb_questionGroupFormRecord`, uma combinação de grupo de questões com o formulário, o que não faz jus ao uso. `tb_CRFForms`, denomina a tabela dos formulários de módulos. A tabela que relaciona as questões com os grupos `tb_QuestionGroupForm`, apenas ordena as questões dentro de um formulário, porém o nome transmite a ideia de ser a relação de um agrupamento com um formulário. A sigla CRF designa a versão do template do questionário da OMS que está implementada WHO COVID-19 Rapid Version CRF, uma amarração de contexto em um banco de dados que se propõe a ser genérico e atender a qualquer pesquisa.

Na parte debaixo do modelo há algumas observações a considerar, com relação a ordem das questões. A ordem de uma questão está associada ao formulário, todavia as questões são organizadas em agrupamentos temáticos, e não existe a ordem do grupo de exibição nem a ordem no grupo, que é determinada apenas à ordem que a questão é inserida no banco. Com isso perde-se a semântica da ordem lógica de apresentação estrutural dos formulários, vulnerável a um erro de inserção no banco.

Outro ponto, é os agrupamentos não terem uma ordem de exibição, não seria possível realizar alterações nas ordens dos grupos. Uma possibilidade muito viável de ser necessária em um módulo que se propõe a permitir alterações em formulários.

Uma abordagem para resolver este problema seria relacionar a tabela `tb_QuestionGroup` diretamente com a tabela `tb_QuestionGroupform` pois dessa maneira a ordem da questão também significaria a ordem no grupo.



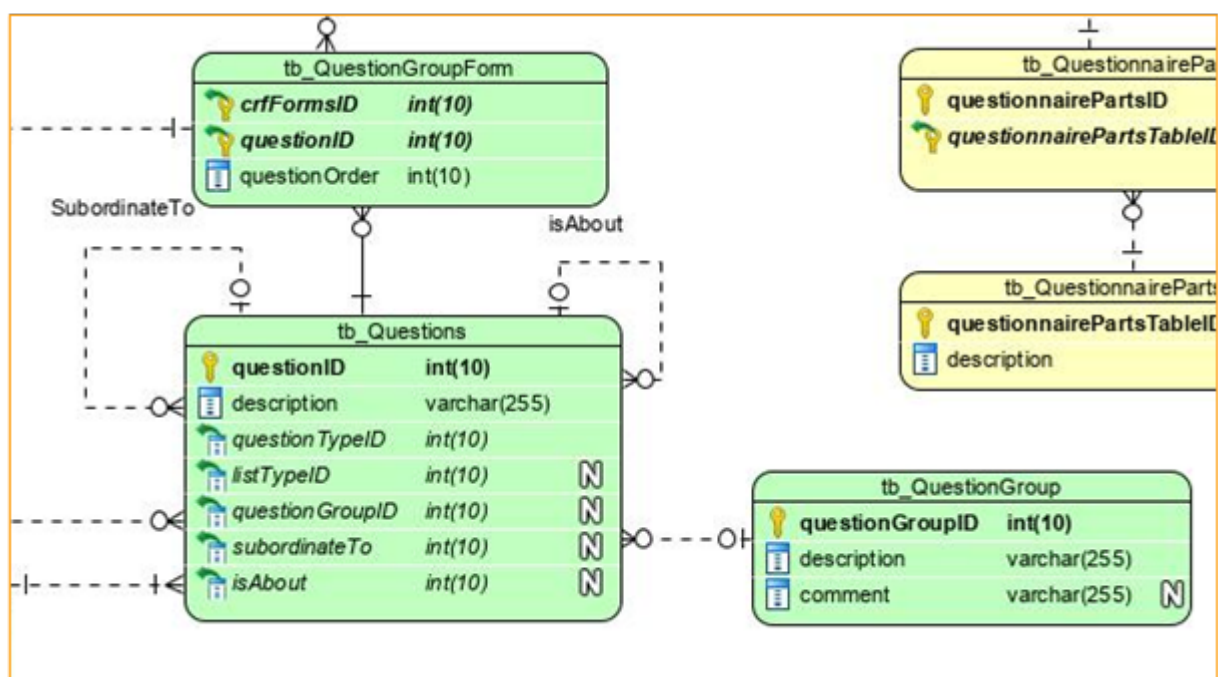


Figura 16 – O relacionamento da tabela tb\_QuestionGroup

## 6 TRABALHOS FUTUROS

Estender uma integração com os módulos de internacionalização. Embora o banco de dados esteja preparado para receber traduções dos termos, a aplicação não conta com funcionalidades em sua interface gráfica que possam contemplar ações.

O presente trabalho se propôs a criar uma interface para atender aos requisitos mínimos do versionamento de Pesquisa e há muito espaço para melhorias nas funcionalidades. O desenvolvimento focou-se nas regras de negócio de um template de pesquisa clínica da COVID específico, mas existe potencial para que o aplicativo atenda a uma pesquisa clínica qualquer ao permitir a adoção de regras de negócio para cada uma.

Aprimorar a ferramenta de edição de formulários da aplicação é um outro caminho a seguir através da agregação de plugins e extensões que possam melhorar sua usabilidade bem como permitir sua interligação com outras apis e aplicações, expandindo as possibilidades de uso-fim.

## REFERÊNCIAS

## GLOSSÁRIO

**Agrupamento ou grupo de questões** é uma pergunta que pode ter um ou mais tipos e estar agrupada com outras perguntas relativas a um mesmo assunto. *veja* Agrupamento ou grupo de questões

**Formulário** é um sinônimo para módulo e formulário de módulo. *veja* Formulário de módulo

**Formulário de Acompanhamento** formulário preenchido em algum momento entre a Admissão e a Alta. *veja* Formulário de Acompanhamento

**Formulário de Admissão** formulário preenchido no momento em que um paciente dá entrada no hospital pela primeira vez. *veja* Formulário de Admissão

**Formulário de Alta/Óbito** formulário preenchido no momento do encerramento do atendimento, seja por alta ou por óbito. *veja* Formulário de Alta/Óbito

**Formulário de módulo** ou módulo é um formulário composto de uma ou mais perguntas sobre um paciente não identificado sobre condições de saúde, entre outras, como por exemplo, região. *veja* Formulário de módulo

**Pesquisa e Questionário** são sinônimos. *veja* Questionário

**Prontuário** é um conjunto de status de saúde do paciente no momento em que ele foi atendido e este atendimento pode ter sido uma entrada de atendimento na emergência, uma entrada para uma consulta de acompanhamento ou uma alta médica. *veja* Prontuário

**Questionário** representa a pesquisa que será realizada. Ele pode ser entendido como um pacote de “um” ou “n” formulários organizados em módulos como, por exemplo, o questionário da OMS estruturado em três módulos: Admissão, acompanhamento e desfecho. Sinônimo de Pesquisa Clínica, um questionário entende-se no contexto da pesquisa clínica como uma pesquisa aplicada em um ou mais momentos da saúde de um paciente. *veja* Questionário

## APÊNDICES

**APÊNDICE A – APÊNDICE A**

**APÊNDICE B – APÊNDICE B.**

## ANEXOS



**ANEXO A – ANEXO A**

**ANEXO B – ANEXO B**