RELATÓRIO DE CUMPRIMENTO DA META 2

Mapeamento de processos do fluxo informacional do modelo

ibict
Instituto Brasileiro de Informação em Ciência e Tecnologia

PRESIDÊNCIA DA REPÚBLICA

Luiz Inácio Lula da Silva
Presidente da República

Geraldo José Rodrigues Alckmin Filho
Vice-Presidente da República

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Luciana Santos
Ministra da Ciência, Tecnologia e Inovação

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA

Tiago Emmanuel Nunes Braga
Diretor

Reginaldo de Araújo Silva
Coordenação de Administração - COADM

Gustavo Saldanha
Coordenação de Ensino e Pesquisa em Informação para a Ciência e Tecnologia - COEPI

José Luis dos Santos Nascimento
Coordenação de Planejamento, Acompanhamento e Avaliação - COPAV

Marcel Garcia de Souza
Coordenação-Geral de Informação Tecnológica e Informação para a Sociedade - CGIT

Bianca Amaro de Melo
Coordenação-Geral de Informação Científica e Técnica - CGIC

Alexandre Faria de Oliveira
Coordenação-Geral de Tecnologias de Informação e Informática - CGTI

Milton Shintaku
Coordenação de Tecnologias para Informação - COTEC

RELATÓRIO DE CUMPRIMENTO DA META 2

Mapeamento de processos do fluxo informacional do modelo

ibict
Instituto Brasileiro de Informação em Ciência e Tecnologia

Coordenação de Tecnologias para Informação (COTEC)

Brasília

2023

© Instituto Brasileiro de Informação em Ciência e Tecnologia – Ibict 2023

EQUIPE TÉCNICA

Diretor do Instituto Brasileiro de Informação em Ciência e Tecnologia
Tiago Emmanuel Nunes Braga

Coordenador-Geral de Tecnologias de Informação e Informática – CGTI
Alexandre Faria de Oliveira

Coordenador de Tecnologias para Informação - COTEC
Milton Shintaku

Coordenador do Projeto
Diego José Macêdo

Pesquisadores
Diego José Macêdo
Caio Saraiva Coneglian
Felipe da Rocha Ferreira
Fernanda Maciel Rufino
Janinne Barcelos de Morais Silva
Larissa Moreno Silva
Monica Lima de Paiva
Marcela Virginia Cavalcanti de Albuquerque
Thays Andrade Costa

Normalização
Fernanda Maciel Rufino

Revisão
Flávia Furlan Granato
Rafael Teixeira de Souza

Diagramação e projeto gráfico
Nuielle Cristine de Medeiros da Silva

Este Relatório de Técnico é um produto do Projeto: Proposição de modelo de referência para construção de observatórios no âmbito governamental.

Ref. IBICT 01302.000166/2022-16 - Processo SEI
Ref. FUNDEP 29756

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto Brasileiro de Informação em Ciência e Tecnologia ou do Ministério da Ciência, Tecnologia e Inovação.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

Sumário

1. INTRODUÇÃO 6

2. OBJETIVOS 7
2.1 Objetivo Geral 7
2.2 Objetivos Específicos 7

3. RESULTADOS 8
3.1 Definição de métodos para coleta e seleção de dados 8
3.1.1 Coleta de Dados Estruturados 8
3.1.2 Coleta de Dados em Ambientes Web 9
3.1.3 Coleta de Dados em Bases de Dados 9
3.1.4 Ferramentas para Coleta de Dados 10
3.1.4.1 Técnica 1: Web Scraping 10
3.1.4.2 Técnica 2: APIs (Interfaces de Programação de Aplicações) 11
3.1.4.3 Técnica 3: Coleta de Dados com RPA (Automação Robótica de Processos) 11
3.1.4.4 Técnica 4: Coleta de Dados em Banco de Dados Relacionais e NoSQL 11
3.2 Indicação de técnicas para tratamento e armazenamento 12
3.2.1 Armazenamento dos Dados 12
3.2.1.1 Ferramentas para Armazenamento dos Dados 12
3.2.1.2 Bancos de Dados Relacionais 13
3.2.1.2 Bancos de Dados NoSQL 13
3.2.1.3 Data Warehouses 13
3.2.2 Normalização de Dados 13
3.2.2.1 Ferramentas para Normalização de Dados 13
3.2.3 Limpeza de Dados 14
3.2.3.1 Ferramentas para Limpeza de Dados 14
3.2.4 Transformação e Enriquecimento de Dados 15
3.2.4.1 Ferramentas para Transformação e Enriquecimento de Dados 15
3.3 Definição de estratégias para disseminação e uso das informações 16
3.3.1 Ferramentas para Visualização de Dados 16
3.4 Ferramentas para Dados Georreferenciais 17
3.5 Metodologias e Abordagens 17
3.6 Mapeamento de processos e geração do modelo para o fluxo 17
3.6.1 Produto e Serviço 1: Apresentação de indicadores 18
3.6.2 Produto e Serviço 2: Participação social 20
3.6.3 Produto e Serviço 3: Publicação de Dados e Informações 24
3.6.4 Produto e Serviço 4: Equipes e Parceiros 29
3.6.5 Produto e Serviço 5: Disponibilização de Links Úteis 31
3.6.6 Produto e Serviço 6: Pesquisa e Educação 32
3.6.7 Produto e Serviço 7: Notícias e Agenda 35

4. CONSIDERAÇÕES FINAIS 39

REFERÊNCIAS 40

1. INTRODUÇÃO

Frequentemente, a sociedade solicita aos governos melhorias na prestação dos serviços públicos, especialmente na implementação de políticas públicas, o que tem impactado diretamente na vida dos cidadãos. Nesse sentido, o governo está constantemente aprimorando os processos e serviços para melhor atender às pessoas. É essencial, portanto, estabelecer mecanismos de monitoramento dos serviços e das políticas públicas, a fim de garantir a execução de ações e otimizar os investimentos governamentais.

Uma das opções para monitorar políticas públicas e serviços governamentais é a criação de observatórios, que são espaços nos quais o governo pode atuar em diferentes níveis, monitorando temas gerais ou especializados e possibilitando a análise de várias fontes de informação, além de realizar projeções para o futuro e fornecer subsídios para a formulação de políticas públicas por meio da avaliação de seus impactos. Os observatórios também contribuem para a geração de conhecimento a longo prazo, facilitando o acompanhamento e monitoramento da evolução das políticas públicas e permitindo a antecipação de ideias e tendências por meio da construção de indicadores confiáveis. Portanto, os observatórios podem auxiliar no monitoramento da dinâmica dos serviços governamentais, identificando oportunidades e fragilidades, auxiliando na análise de informações e sendo um ambiente estratégico para que os gestores públicos tomem decisões mais embasadas.

No entanto, até o momento, existem dificuldades em compreender as funções dos observatórios e como eles podem ser estruturados para atender especialmente às demandas do ambiente digital, onde há enorme quantidade de informações disponíveis, as quais nem sempre se encontram organizadas de maneira apropriada para coleta e interpretação, a fim de gerar indicadores que auxiliem na tomada de decisões.

Portanto, o presente projeto de pesquisa tem o intuito de abordar essas lacunas e propor um modelo de referência para a construção de observatórios no âmbito governamental, com uma prova de conceito aplicada ao Governo do Distrito Federal (GDF). Esse modelo irá auxiliar na estruturação de observatórios em órgãos governamentais e no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Para alcançar os resultados do projeto de pesquisa, este relatório apresenta o cumprimento da segunda meta do projeto, na qual são apresentados os estudos e avanços referentes ao mapeamento de processos do fluxo informacional do modelo. A meta foi elaborada conforme os avanços na definição do contexto da pesquisa e os requisitos para o modelo de referência definidos previamente na primeira meta do projeto.

2. OBJETIVOS

2.1 Objetivo Geral

Mapeamento de processos do fluxo informacional do modelo.

2.2 Objetivos Específicos

a. análise do fluxo informacional;

b. definição de processos para o fluxo informacional.

3. RESULTADOS

Para cumprimento da primeira meta do projeto foram executadas diversas atividades junto com a equipe para cumprir os objetivos específicos. Nas seções a seguir são apresentados os resultados.

3.1 Definição de métodos para coleta e seleção de dados

No âmbito do observatórios, há a necessidade da reflexão dos processos e técnicas necessárias para a realização da coleta e da seleção dos dados que serão obtidos, utilizados, analisados e expostos. Ademais, verifica-se que existe um grande desafio em tal processo, devido às diversas tipologias e objetivos dos documentos que os observatórios utilizam.

Nesse sentido, a discussão e a apresentação dos métodos e ferramentas para coleta e seleção de dados devem considerar, primeiramente, a tipologia documental, e, em um segundo momento, o que será feito com os dados e as informações obtidas.

Partindo de tais premissas, apresenta-se a seguir a relação de tipologia documental identificada dos observatórios, para que, a partir dela, sejam verificados os métodos para coleta.

Quadro 1 – Tipologia Documental encontrada nos observatórios brasileiros analisados

Tipos	Características
Boletins	Documento em formato de texto (PDF)
Panorama/Estatística/Indicadores/Painéis/ Gráficos/*Dashboards* (números)	Planilhas e Dados Estruturados (Comma-Separated Values (CSV), XLS)
Relatórios/Resumos/Anuário	Documento em formato de texto (PDF)
Pesquisas/Projetos	Documento em formato de texto (PDF)
Publicações	Documento em formato de texto (PDF)
Dossiê/Legislação	Documento em formato de texto (PDF)
Artigos	Documento em formato de texto (PDF)
Notícias	Documento em formato de texto (PDF)
Mapa/Atlas	Documento em formato de texto (PDF) ou Imagens (PNG, JPG)

Fonte: Dados da pesquisa (2023).

O Quadro 1 apresenta, além dos tipos documentais, uma descrição dos documentos, destacando os possíveis formatos e características deles. A seguir, são detalhadas as técnicas e reveladas as ferramentas necessárias para a coleta de dados de diferentes formatos e estruturas.

3.1.1 Coleta de Dados Estruturados

Os dados são classificados em três tipos principais: estruturados, semiestruturados e não estruturados (EBERENDU; 2016). Nesse contexto, apresenta-se o processo de coleta de dados estruturados. Eberendu (2016, p. 48, tradução nossa) aponta que os "Os dados estruturados se referem aos dados que possuem formato definido e comprimento, fácil de armazenar e analisar com alto grau de organização".

No que diz respeito a esse tema, há técnicas específicas que possibilitam a coleta de dados estruturados, as quais possuem papel essencial nos observatórios, fornecendo informações relevantes e atualizadas para análises e tomada de decisões. Tais dados concentram especialmente planilhas, dados estruturados no formato CSV e TXT. O Quadro 2 apresenta, de forma sumarizada, o processo de coleta estruturada.

Quadro 2 – Forma de Coleta de Dados Estruturados

Tipo	Formato	Forma de Coleta
Planilhas	CSV, XLS, XLSX, TXT	Obtenção de forma manual ou por meio de técnicas de RPA (*Robotic Process Automation*) para coleta automatizada

Fonte: Dados da pesquisa (2023).

Destaca-se, no Quadro 2, que os dados podem ser coletados manualmente ou por meio de técnicas automatizadas, com o apoio de RPA, que permite sua atualização de forma controlada e sistematizada. O detalhamento de RPA encontra-se na subseção 3.1.4, com a explicação das técnicas de coleta.

3.1.2 Coleta de Dados em Ambientes Web

Complementarmente, no contexto de observatórios, é essencial a coleta de dados em ambientes Web. Tal abordagem é amplamente utilizada para obter informações de fontes on-line envolvendo a extração sistemática de dados de páginas web, incluindo informações como textos, imagens, tabelas e links. A seguir, no Quadro 3, destaca-se o processo de coleta em ambientes Web.

Quadro 3 – Forma de Coleta de Dados em Ambientes Web

Tipo	Formato	Forma de Coleta
Dados estruturados em páginas Web	HTML	Uso de ferramentas de Web *Scraping*, como Beautiful Soup[1], Selenium[2] e Scrapy[3]
Dados estruturados em ambientes Web	*JavaScript Object Notation* (JSON) ou eXtensible Markup Language (XML)	Uso de APIs (*Application Programming Interface*) para a coleta de dados.

Fonte: Dados da pesquisa (2023).

Como apresentado no Quadro 3, existem diversas ferramentas e técnicas para realizar a coleta de dados em ambientes web, como o Web *Scraping*, que envolve o uso de bibliotecas e *frameworks* como *Beautiful Soup, Selenium* e *Scrapy* para extrair informações de páginas *web* de forma automatizada. Além disso, há a opção de coleta dos dados por meio de API, que é feita a partir da solicitação de algum serviço. Na seção 3.1.4 são detalhadas as abordagens utilizadas em Web *Scraping* e API.

3.1.3 Coleta de Dados em Bases de Dados

A coleta de dados em bases envolve a extração de informações de sistemas de armazenamento estruturados, como bancos relacionais ou bancos NoSQL. Nessa abordagem, as consultas e os comandos

1 Disponível em: https://www.crummy.com/software/BeautifulSoup/bs4/doc/
2 Disponível em: https://www.selenium.dev/
3 Disponível em: https://scrapy.org/

são utilizados para recuperar os dados desejados. As linguagens de consulta mais comuns para coleta de dados em bases são *Structured Query Language* (SQL) para bancos relacionais e consultas específicas para bancos NoSQL, como o MongoDB. Essa abordagem é ideal quando os dados estão armazenados em bases e é necessário realizar consultas complexas para obter as informações desejadas.

Quadro 4 – Forma de Coleta de Dados em Bases de Dados

Tipo	Formato	Forma de Coleta
Dados contidos em ambientes de banco relacionais	SQL	Uso de ferramentas de apoio aos Sistemas de Gerenciamento de Banco de Dados (SGBDs), como MYSQL *Workbench*, que permite a consulta utilizando SQL.
Dados contidos em ambientes de banco não relacionais	JSON ou outros	Uso de ferramentas gerenciadoras que permite a consulta aos dados.

Fonte: Dados da pesquisa (2023).

No Quadro 4, destacam-se duas formas principais para coleta, de acordo com o tipo de Banco de Dados. Tais informações são detalhadas na seção 3.1.4.

3.1.4 Ferramentas para Coleta de Dados

A partir dos aspectos apresentados, foram destacadas quatro técnicas principais para realização da coleta de dados: *Web Scraping*, APIs, RPA e Coleta em Banco de Dados.

3.1.4.1 Técnica 1: Web Scraping

A técnica de Web *Scraping* é amplamente utilizada para extração de conteúdos disponíveis em páginas HTML, consistindo na extração automatizada de dados de páginas web. Essa técnica permite extrair informações estruturadas, como tabelas, listas, textos e imagens, diretamente dos elementos HTML das páginas.

Essa técnica é definida como "[...] uma ótima técnica de extração de dados não estruturados dos sites e transformação desses dados em dados estruturados que podem ser armazenados e analisados em um banco de dados" (SIRISURIYA, 2015, p. 135, tradução nossa).

Existem várias bibliotecas e *frameworks* disponíveis que facilitam a implementação do Web *Scraping*, oferecendo recursos avançados para navegar em páginas web, localizar elementos específicos e extrair os dados desejados. Algumas das bibliotecas populares para Web *Scraping* incluem:

a. *Beautiful Soup*: Biblioteca *Python* que permite analisar e extrair dados de documentos HTML e XML. Ela facilita a navegação na estrutura da página, busca de elementos por meio de seletores e extração dos dados desejados.

b. *Selenium*: Ferramenta amplamente utilizada para automação de testes em navegadores web. Ademais, o *Selenium* também pode ser empregado para o Web *Scraping*, pois permite interagir com páginas web de forma programática, preenchendo formulários, clicando em botões e coletando os dados resultantes.

c. *Scrapy: Framework Python* dedicado ao *Web Scraping*, que fornece uma estrutura completa para a construção de *spiders* (robôs de coleta de dados) altamente personalizáveis. O *Scrapy* permite definir as regras de navegação, extração e persistência dos dados de maneira flexível e escalável.

3.1.4.2 Técnica 2: APIs (Interfaces de Programação de Aplicações)

Outro modo de coleta de dados em ambientes Web é por meio de APIs. As APIs são conjuntos de regras e protocolos que permitem a comunicação entre diferentes sistemas de software. Assim, diversas organizações e empresas disponibilizam APIs para permitir a coleta de dados de forma estruturada e programática.

A definição a seguir apresenta o conceito de API com mais detalhes:

> API significa Application Programming Interface (Interface de Programação de Aplicação). No contexto de APIs, a palavra Aplicação refere-se a qualquer software com uma função distinta. A interface pode ser pensada como um contrato de serviço entre duas aplicações. Esse contrato define como as duas se comunicam usando solicitações e respostas. A documentação de suas respectivas APIs contém informações sobre como os desenvolvedores devem estruturar essas solicitações e respostas (AWS, 2023, on-line).

As APIs são uma maneira eficiente e confiável de obter dados de fontes externas, pois fornecem *endpoints* (pontos de acesso) que podem ser acessados com solicitações HTTP. Esses *endpoints* retornam os dados solicitados em formato padronizado, como JSON, XML, entre outros formatos estabelecidos.

3.1.4.3 Técnica 3: Coleta de Dados com RPA (Automação Robótica de Processos)

Os observatórios podem utilizar diversos dados disponíveis publicamente em formatos de planilha, o que é uma forma de automatizar o processo de coleta. Assim, a coleta em ambientes governamentais ou de transparência pode ser facilitada com o uso da técnica de RPA. O RPA é uma abordagem que utiliza softwares robóticos para automatizar tarefas repetitivas e baseadas em regras, incluindo a extração de dados de planilhas e sistemas governamentais.

Uma das definições mais conhecidas de RPA aponta que:

> RPA é um termo abrangente para ferramentas que operam na interface do usuário de outros sistemas de computador da maneira que um ser humano faria. O RPA visa substituir pessoas por automação feita de maneira 'de fora para dentro (AALST; BICHLER; HEINZL, 2018, p. 260).

Uma das vantagens do RPA é a capacidade de interagir com interfaces de usuário, como planilhas eletrônicas, navegadores *web* e aplicativos *desktop*, de forma semelhante a um usuário humano. Isso permite que os robôs executem ações, como abrir planilhas, localizar informações específicas, copiar e colar dados, e até mesmo preencher formulários automaticamente.

No contexto da coleta de dados em ambientes governamentais ou de transparência, o RPA pode ser aplicado para automatizar o processo de busca e extração de dados de planilhas disponibilizadas pelos órgãos públicos. Os robôs de RPA podem ser configurados para acessar regularmente as fontes de dados, identificar atualizações ou novas informações, e realizar a extração dos dados relevantes.

3.1.4.4 Técnica 4: Coleta de Dados em Banco de Dados Relacionais e NoSQL

A coleta de dados em bancos pode apoiar sua obtenção para os observatórios. No entanto, há dois tipos principais de banco de dados: relacionais e NoSQL.

Os relacionais são amplamente utilizados para armazenar e gerenciar dados estruturados em tabelas, seguindo um modelo predefinido. Para realizar a coleta de dados nesse tipo de banco, é necessário utilizar consultas SQL para recuperar as informações desejadas.

A principal forma de consultar banco de dados é por meio de consultas SQL personalizadas. É possível escrever consultas SQL específicas para extrair dados de tabelas e relacionamentos. Essas

consultas podem ser executadas diretamente no banco de dados ou por meio de ferramentas de acesso, como interfaces de linha de comando ou interfaces gráficas.

Já no âmbito dos bancos de dados NoSQL, estes são projetados para lidar com volumes massivos, estruturas flexíveis e requisitos de escalabilidade. Eles oferecem uma variedade de modelos, como documentos, chave-valor, colunas amplas e grafos. A coleta em bancos de dados NoSQL requer abordagens específicas para cada modelo.

Para realizar buscas e coletar dados em bancos NoSQL, verifica-se que cada modelo NoSQL tem suas próprias consultas e métodos de acesso. Por exemplo, em bancos de dados de documentos, podem-se usar consultas baseadas em JSON para recuperar documentos específicos ou executar operações de agregação. Em bancos de dados de chave-valor, é possível obtê-los diretamente por meio das chaves.

3.2 Indicação de técnicas para tratamento e armazenamento

As técnicas e ferramentas recomendadas para o tratamento e armazenamento de informações no contexto da modelagem de um sistema para observatório são essenciais para sua compreensão no contexto atual, que se vincula fortemente à utilização de dados para a tomada de decisões nos mais diversos cenários. Complementarmente, destaca-se que o processo de tratamento e armazenamento se vincula diretamente à qualidade da informação, pois esta é de extrema importância para garantir a confiabilidade e a eficiência das operações do observatório. Portanto, é essencial adotar abordagens adequadas que permitam o correto processamento, organização e armazenamento dos dados coletados.

No que tange às técnicas de tratamento de informações, identifica-se que há diversas etapas e elementos que devem ser considerados, como é apresentado a seguir.

3.2.1 Armazenamento dos Dados

O primeiro olhar a ser lançado após a coleta de dados direciona-se à preocupação com o processo de armazenamento. Nesse contexto, é crucial considerar técnicas adequadas para garantir a eficiência e a segurança das informações considerando o contexto de um sistema para observatório. Uma abordagem recomendada é a implementação de técnicas de particionamento de dados, as quais permitem dividir grandes conjuntos de dados em partes menores, facilitando o acesso e a recuperação de informações específicas. Além disso, a replicação de dados pode ser empregada para aumentar a disponibilidade e a tolerância a falhas do sistema, garantindo que os dados estejam acessíveis mesmo em caso de falhas em componentes de hardware ou conexões de rede.

O uso de técnicas de compressão de dados também é relevante, pois reduz o espaço de armazenamento necessário, minimizando os custos associados e otimizando o desempenho na recuperação de informações. A escolha das técnicas de armazenamento deve ser feita de acordo com os requisitos e características específicas do observatório, levando em consideração a escala dos dados, as necessidades de desempenho e a segurança das informações.

3.2.1.1 Ferramentas para Armazenamento dos Dados

Há diferentes abordagens e formas para realizar o processo de armazenamento dos dados. A seguir, apresentam-se três propostas.

3.2.1.2 Bancos de Dados Relacionais

Os bancos de dados relacionais são amplamente utilizados para o armazenamento de informações estruturadas, oferecendo recursos de consulta e integridade referencial. Recomenda-se a adoção de SGBDs populares, como MySQL, *PostgreSQL* ou *Oracle*, que possuem robustez e escalabilidade para lidar com grande volume de dados.

No contexto dos observatórios, alguns tipos de dados são adequados para serem armazenados em Bancos Relacionais, em especial aqueles que são obtidos a partir de sistemas de informação transacionais. No entanto, uma proposta para conjuntos de dados menos estruturados, que muitas vezes são utilizados em observatórios, é a utilização de Bancos de Dados NoSQL, como apresentado a seguir.

3.2.1.2 Bancos de Dados NoSQL

Para o armazenamento de informações não estruturadas ou semiestruturadas, os bancos de dados NoSQL são uma opção viável. Esses sistemas permitem a flexibilidade no armazenamento de diferentes tipos de dados, como documentos, gráficos e dados em formato de chave-valor. Exemplos de bancos de dados NoSQL incluem MongoDB[4] , Apache Cassandra[5] e Redis[6].

3.2.1.3 Data Warehouses

Complementando as soluções de Bancos de Dados Relacionais e NoSQL, é essencial que exista um processo de consolidação e junção desses dados para que possa ser realizada análises e outros processos. Logo, os *Data Warehouses* são soluções que permitem a consolidação e o armazenamento de grandes volumes de dados de diversas fontes em um único local. Tal abordagem facilita a análise e a geração de relatórios, fornecendo um ambiente centralizado para consulta e processamento dos dados.

Data Warehouse são definidos como "Um sistema [que] armazena dados históricos integrados e preparados para serem analisados por OLAP [OnLine Analytical Processing] e outras ferramentas" (MARTINEZ *et al.*, 2008, p. 940, tradução nossa).

3.2.2 Normalização de Dados

A primeira técnica apresentada é a normalização de dados, que trata de um método essencial para garantir a consistência e a uniformidade dos dados armazenados nos mais diversos contextos. Por meio dela, é possível evitar redundâncias e inconsistências, facilitando a integração e análise dos dados. Recomenda-se a utilização de padrões de normalização estabelecidos, como o modelo de normalização de banco de dados relacional.

Complementarmente, a normalização de dados se define como o processo "[...] onde os dados são dimensionados para uniformidade. A normalização de dados é necessária para estudar as melhores características dos dados" (SREE; BINDU, 2018, p. 209, tradução nossa).

3.2.2.1 Ferramentas para Normalização de Dados

Há diversas ferramentas com objetivos distintos para realizar o processo de normalização dos dados. Além da própria normalização utilizando softwares de gerenciamento de banco de dados relacional, que permitem a definição de esquemas de banco de dados, incluindo a criação de tabelas, relacionamentos e restrições de integridade referencial, há outras ferramentas que propiciam a normalização de dados.

4 Disponível em: https://www.mongodb.com/
5 Disponível em: https://cassandra.apache.org/_/index.html
6 Disponível em: https://redis.io/

Fora as ferramentas específicas de gerenciamento e modelagem de dados, também existem bibliotecas e *frameworks* de programação que oferecem funcionalidades para a normalização. Essas ferramentas permitem a implementação de algoritmos e rotinas personalizadas para realizar a normalização dos dados, levando em consideração as regras de negócio específicas do observatório.

A adoção das ferramentas adequadas para a normalização de dados contribuirá para a garantia da consistência e da integridade dos dados no sistema. Ao suprimir redundâncias e inconsistências, as ferramentas de normalização proporcionam uma base sólida para a análise e a tomada de decisões com base nos dados coletados, resultando em informações confiáveis e relevantes para as atividades do observatório.

A seguir apresentam-se as ferramentas que podem ser utilizadas para a normalização de dados:

a. *OpenRefine*[7]: é uma ferramenta de código aberto que permite limpar e transformar dados de forma interativa. Ela oferece recursos avançados para detecção e correção de erros, padronização de formatos, remoção de duplicatas e enriquecimento dos dados.

b. *Alteryx Designer Cloud*[8]: essa ferramenta é voltada para a preparação de dados e possui uma interface intuitiva para limpar, transformar e estruturar dados de forma visual. Ela oferece recursos de sugestão automática, detecção de padrões e visualizações interativas para facilitar o processo de normalização.

c. *Microsoft SQL Server Integration Services* (SSIS)[9]: é uma ferramenta de integração de dados da Microsoft que permite criar fluxos de trabalho para Extração, Transformação e Carga (ETL) de dados. Ela oferece recursos para mapeamento, limpeza e normalização de dados durante o processo de integração.

3.2.3 Limpeza de Dados

Junto ao processo de normalização, é essencial a realização da limpeza de dados, pois se trata de um processo necessário para identificar e corrigir erros, omissões e inconsistências das coletas. Isso inclui a remoção de valores nulos, a padronização de formatos e a detecção de *outliers*[10].

A limpeza de dados é definida como:

> [...] uma abordagem inicial em que os conjuntos de dados são limpos para identificar quaisquer dados ausentes, remover os dados ruidosos e preparar os dados para análise. A limpeza de dados é necessária para resolver o problema de qualidade de dados. O problema de qualidade de dados é onde a análise pode dar errado em dados confusos (SREE; BINDU, 2018, p. 209, tradução nossa).

Existem diversas ferramentas e algoritmos que estão disponíveis para auxiliar no processo de limpeza, como técnicas de análise estatística e algoritmos de detecção de anomalias, conforme apresentado a seguir.

3.2.3.1 Ferramentas para Limpeza de Dados

O processo de limpeza de dados exige ferramentas capazes de apoiar a exclusão de adequação da base. Os softwares de limpeza e transformação de dados oferecem recursos avançados para identificar

7 Disponível em: https://openrefine.org/
8 Disponível em: https://www.trifacta.com/
9 Disponível em: https://learn.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver16
10 *Outlier* é um dado ou um conjunto de dados que se distancia radicalmente dos demais que compõem o conjunto analisado.

e corrigir erros, inconsistências, valores ausentes e duplicatas nos conjuntos. Eles permitem a aplicação de regras e algoritmos para padronização, normalização e validação, garantindo sua conformidade com os requisitos estabelecidos.

Um software que permite a realização de uma série de atividades para realizar a limpeza dos dados é o *OpenRefine*. O *OpenRefine* permite que sejam realizados filtros, exclusão dos dados, preenchimento de células em branco, além de permitir ao final o processo de exportação.

Outra ferramenta útil é um software de deduplicação de dados. Essas ferramentas permitem identificar registros duplicados em um conjunto de dados, seja por meio de comparação de campos-chave ou por algoritmos de similaridade. A deduplicação é crucial para garantir a precisão e a integridade dos dados, evitando a duplicação de informações e a distorção dos resultados das análises.

Destaca-se que o *OpenRefine* se mostra como uma ferramenta capaz de realizar o processo de deduplicação dos dados de forma efetiva.

Vale ressaltar que, no âmbito dos observatórios, o processo de limpeza deverá ser realizado de forma constante, com um ciclo contínuo de monitoramento e melhoria da qualidade dos dados. As ferramentas selecionadas devem permitir a automatização e a repetibilidade dessas tarefas, a fim de reduzir o tempo e os esforços necessários para manter a integridade dos dados.

No geral, o uso do software como o *OpenRefine* permite a garantia de dados confiáveis e consistentes, a redução de erros e a melhoria da precisão das análises e dos processos de tomada de decisão. Investir em ferramentas de limpeza de dados é essencial para obter insights valiosos e embasar as atividades do observatório em informações confiáveis e de alta qualidade.

Destaca-se que há outros softwares que permitem a realização dos processos de limpeza, como o *Microsoft Power BI*[11] e o Tableau, que permitem o tratamento e a limpeza dos dados no momento da importação. No entanto, o *OpenRefine* se mostra como uma ferramenta mais completa, em virtude do seu caráter mais específico, com foco na etapa anterior à análise.

3.2.4 Transformação e Enriquecimento de Dados

Após o processo de normalização e limpeza dos dados, a próxima etapa trata da sua transformação e do seu enriquecimento, buscando aplicar técnicas que permitam a conversão de dados brutos em formatos mais adequados e a incorporação de informações adicionais para enriquecer o seu conjunto. Isso pode incluir a agregação, a aplicação de regras de negócio e a incorporação de dados provenientes de outras fontes confiáveis.

3.2.4.1 Ferramentas para Transformação e Enriquecimento de Dados

Além das ferramentas de Normalização e Limpeza, aponta-se a necessidade de ferramentas adequadas para a transformação e o enriquecimento de dados. Essas ferramentas desempenham um papel fundamental na preparação e no aprimoramento dos dados coletados, permitindo que sejam convertidos em formatos mais adequados e enriquecidos com informações adicionais.

Uma das ferramentas amplamente utilizadas nesse contexto é um software de ETL. Elas possibilitam a extração dos dados de diversas fontes, como bancos, arquivos CSV ou APIs, e a aplicação de transformações para limpeza, normalização e padronização. Além disso, o ETL permite a carga dos dados transformados em um novo destino, como um banco ou um *Data Warehouse*.

11 Disponível em: https://powerbi.microsoft.com/pt-br/

Outro software diz respeito à integração de dados. Essas ferramentas facilitam a integração de dados provenientes de diferentes fontes, permitindo a harmonização e a unificação de formatos, a reconciliação de registros duplicados e a resolução de conflitos. Com a integração, é possível obter uma visão holística e unificada dos dados, promovendo uma compreensão mais completa e precisa.

Adicionalmente, o enriquecimento de dados pode ser realizado por meio de ferramentas externas. Elas permitem a incorporação de informações adicionais aos dados existentes, provenientes de fontes confiáveis, como dados geográficos, dados demográficos ou informações atualizadas de terceiros. Esse enriquecimento de dados pode aumentar a qualidade, a relevância e o valor dos dados, fornecendo uma visão mais completa e enriquecida do contexto em que estão inseridos.

Alguns dos softwares de transformação foram citados no contexto da Normalização e Limpeza. São capazes também de apoiar o processo de enriquecimento, conforme a relação a seguir:

a. *OpenRefine*: é uma ferramenta de código aberto para limpeza e transformação de dados. Além de oferecer recursos para isso, o *OpenRefine* também permite a transformação de dados por meio da execução de operações em massa, como separação de colunas, remoção de espaços em branco, conversão de formatos e muito mais.

b. *Talend Data Integration*[12]: plataforma de integração que inclui recursos avançados de transformação de dados. Ela permite a criação de fluxos de trabalho para transformar dados de várias fontes, combinando, filtrando, agregando e aplicando regras de negócio. A plataforma suporta diferentes formatos de dados e oferece uma interface intuitiva para a criação e gerenciamento das transformações.

c. *RapidMiner*[13]: é uma plataforma de análise de dados que inclui recursos de transformação e enriquecimento. Ela permite a criação de fluxos de trabalho para a preparação, incluindo operações de transformação, filtragem, agregação e enriquecimento com dados externos. O *RapidMiner* suporta diferentes fontes e oferece ampla gama de algoritmos e técnicas de análise.

d. Bibliotecas Pandas[14] e NumPy[15] (Linguagem *Python*): tais bibliotecas oferecem recursos para transformação e manipulação de dados. Elas fornecem funções e métodos para realizar operações de transformação, limpeza, agregação e enriquecimento de dados de maneira eficiente.

3.3 Definição de estratégias para disseminação e uso das informações

No contexto dos observatórios, é essencial que, após todo o processo de coleta e transformação dos dados, tenha-se efetividade no âmbito da disseminação das informações. Assim, é necessário garantir que elas sejam acessíveis, compreensíveis e utilizáveis pelos usuários.

3.3.1 Ferramentas para Visualização de Dados

Uma estratégia para disseminação e uso das informações diz respeito às ferramentas de visualização de dados que desempenham um papel crucial na apresentação e compreensão. Assim, permite-se que os usuários extraiam insights valiosos dos dados e os comuniquem de maneira clara. São apresentadas, a seguir, algumas ferramentas para visualização de dados:

12 Disponível em: https://www.talend.com/products/integrate-data/
13 Disponível em: https://rapidminer.com/
14 Disponível em: https://pandas.pydata.org/
15 Disponível em: https://numpy.org/

a. Tableau: plataforma de visualização de dados que permite criar painéis interativos, gráficos e *dashboards*. O Tableau oferece recursos avançados de análise visual e suporta diferentes fontes de dados.

b. *Power BI*: ferramenta da Microsoft para análise e visualização de dados. O *Power BI* permite criar painéis interativos e relatórios personalizados, além de fornecer recursos avançados de análise e integração com outras ferramentas do ecossistema Microsoft.

c. D3.js[16]: biblioteca *JavaScript* utilizada para visualização de dados personalizada e interativa. O D3.js fornece um conjunto abrangente de ferramentas para manipular e visualizar dados, permitindo a criação de gráficos altamente personalizados.

3.4 Ferramentas para Dados Georreferenciais

Já no contexto de dados georreferenciais, que são informações relacionadas à localização geográfica de eventos, objetos ou fenômenos, aponta-se a utilização de ferramentas adequadas para análise e visualização desses dados em um observatório. Apresentam-se algumas ferramentas que podem ser utilizadas para dados georreferenciais:

a. Sistemas de Informação Geográfica (SIG): são plataformas que permitem a captura, armazenamento, análise e visualização de dados georreferenciais. Alguns exemplos de SIG: ArcGIS, QGIS e MapInfo.

b. Visão[17]: desenvolvido pelo Ibict, o Visão é um sistema para análise e visualização de dados georreferenciais, que oferece recursos para manipulação de dados espaciais e criação de mapas interativos.

3.5 Metodologias e Abordagens

Além das ferramentas supracitadas, apontam-se metodologias e abordagens adequadas para a disseminação e uso das informações. Dentre elas, destacam-se:

a. *Data Storytelling*: uma abordagem que combina dados e narrativa para comunicar informações de forma atrativa. Essa técnica envolve a criação de histórias que ajudam a transmitir mensagens claras e impactantes a partir dos dados.

b. *Design* Centrado no Usuário: uma abordagem que coloca os usuários no centro do processo de *design*, considerando suas necessidades, habilidades e preferências.

3.6 Mapeamento de processos e geração do modelo para o fluxo

A realização do mapeamento dos processos aconteceu a partir da definição dos produtos e serviços, buscando definir uma visão macro do funcionamento do produto/serviço enquanto um processo, partindo, posteriormente, para a definição dos processos de forma detalhada.

Nesse sentido, apresenta-se, a seguir, o mapeamento dos processos realizados.

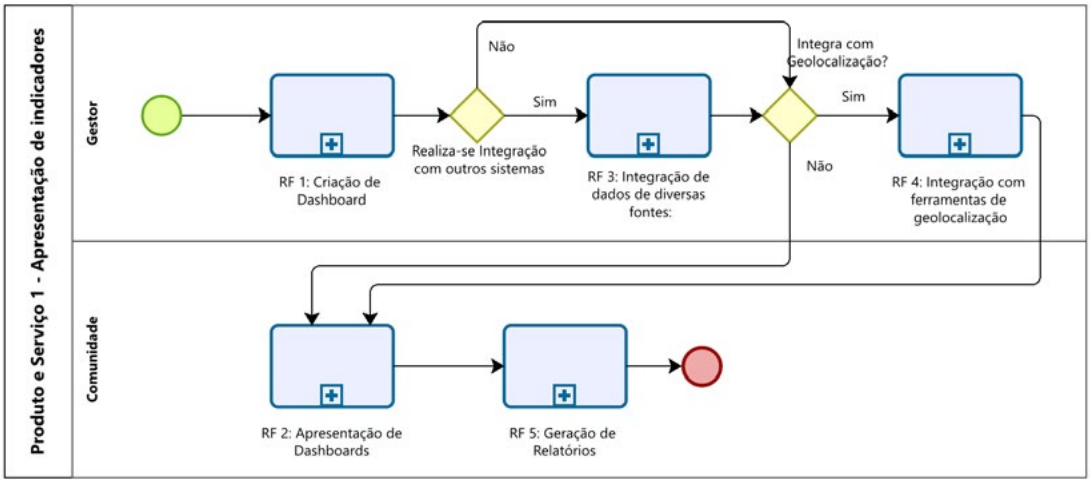
16 Disponível em: https://d3js.org/
17 Disponível em: https://visao.ibict.br

3.6.1 Produto e Serviço 1: Apresentação de indicadores

Produto e serviço 1 - Apresentação de indicadores e produção de sistema de indicadores: o sistema deve permitir que o usuário gestor crie seus próprios *dashboards* e gráficos, com os indicadores que deseja acompanhar. Ele também deve permitir que o *dashboard* criado pelo usuário gestor seja capaz de integrar dados de diversas fontes, como bancos, planilhas e APIs. Também deve ser possível que, na criação do *dashboard*, o usuário gestor realize a integração com ferramentas de geolocalização, visando à exploração desse aspecto a partir do Visão e Google Maps, por exemplo. A partir desses requisitos, o usuário comunidade poderá acessar os gráficos e *dashboards* gerados pelos gestores do observatório, como também se torna possível ao usuário comunidade gerar relatórios com base nos dados e informações apresentados nos *dashboards* e gráficos:

- Usuários: gestor e comunidade.

Figura 1 - Produto e serviço 1



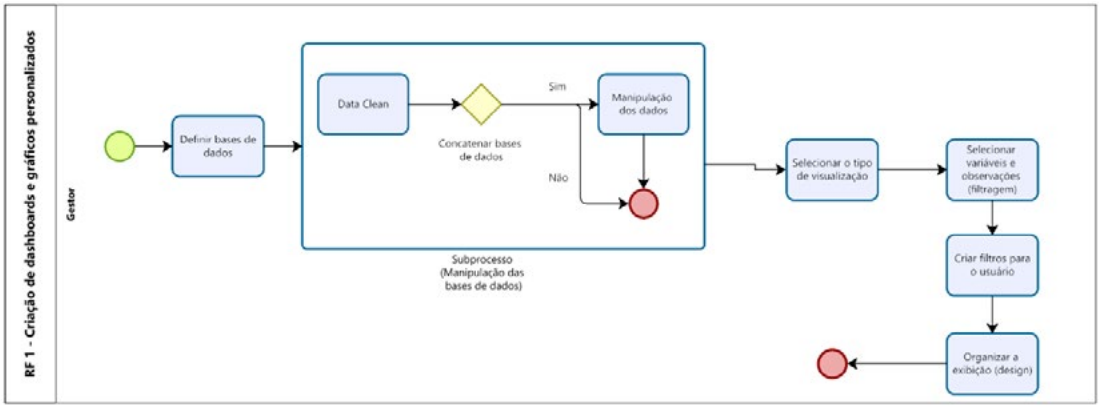
Fonte: Elaborado pelos autores (2023).

A seguir, os processos referentes ao Produto e Serviço correspondente serão apresentados de forma detalhada.

RF 1 - Criação de *dashboards* e gráficos personalizados: o sistema deve permitir que o usuário crie seus próprios *dashboards* e gráficos, selecionando os indicadores e métricas que deseja acompanhar. O usuário pode escolher o *layout*, as cores e a frequência de atualização dos *dashboards*.

- Usuário: gestor.

Figura 2 - Requisito Funcional 1

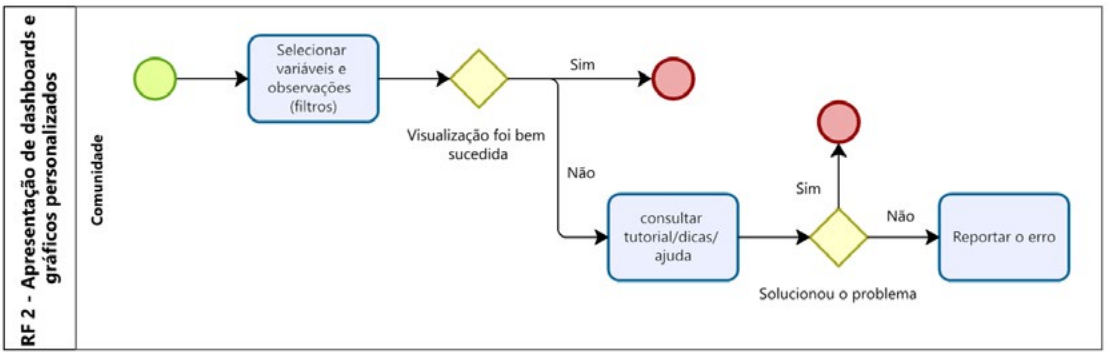


Fonte: Elaborado pelos autores (2023).

RF 2 - Apresentação de *dashboards* e gráficos personalizados: o usuário comunidade pode acessar os gráficos e *dashboards* gerados pelos gestores do observatório.

- Usuário: comunidade.

Figura 3 - Requisito Funcional 2

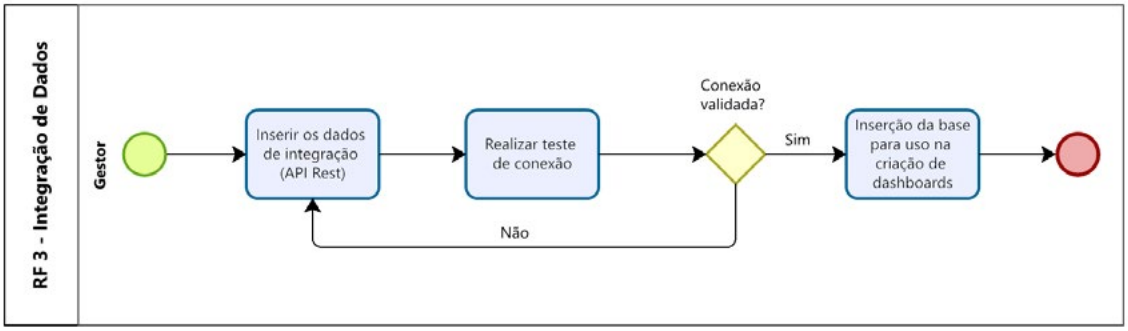


Fonte: Elaborado pelos autores (2023).

RF 3 - Integração de dados de diversas fontes: o sistema deve ser capaz de integrar dados de diversas fontes, como bancos de dados, planilhas e APIs. Isso permitirá que o usuário tenha acesso a informações completas e atualizadas em tempo real:

- Usuário: gestor.

Figura 4 - Requisito Funcional 3

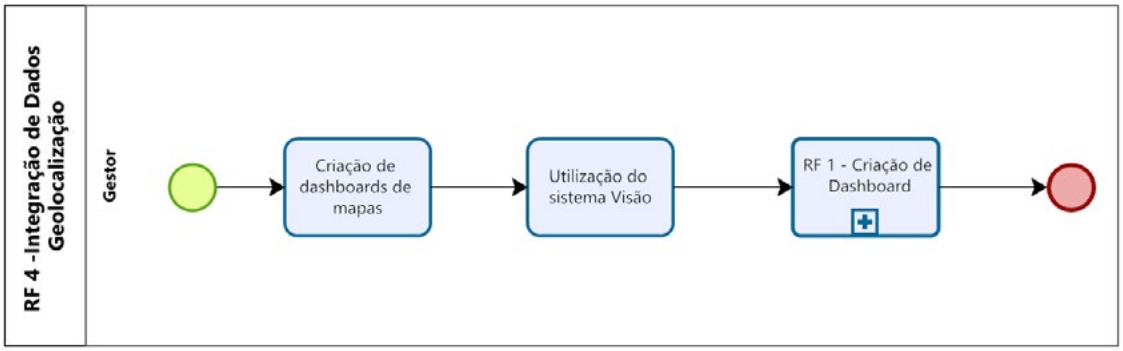


Fonte: Elaborado pelos autores (2023).

RF 4: integração com ferramentas de geolocalização: o sistema apresentará integração com sistemas de geolocalização, como Visão e Google Maps, visando à construção de *dashboards* que exploram aspectos de geolocalização.

- Usuário: gestor.

Figura 5 - Requisito Funcional 4

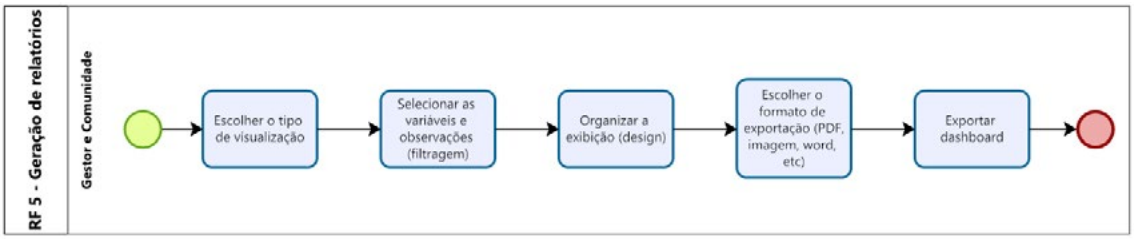


Fonte: Elaborado pelos autores (2023).

RF 5 - Geração de relatórios: o sistema deve permitir que o usuário gere relatórios com base nos dados e informações apresentados nos *dashboards* e gráficos. Isso facilitará a documentação e o compartilhamento de informações com outras pessoas:

- Usuários: comunidade e gestor.

Figura 6 - Requisito Funcional 5



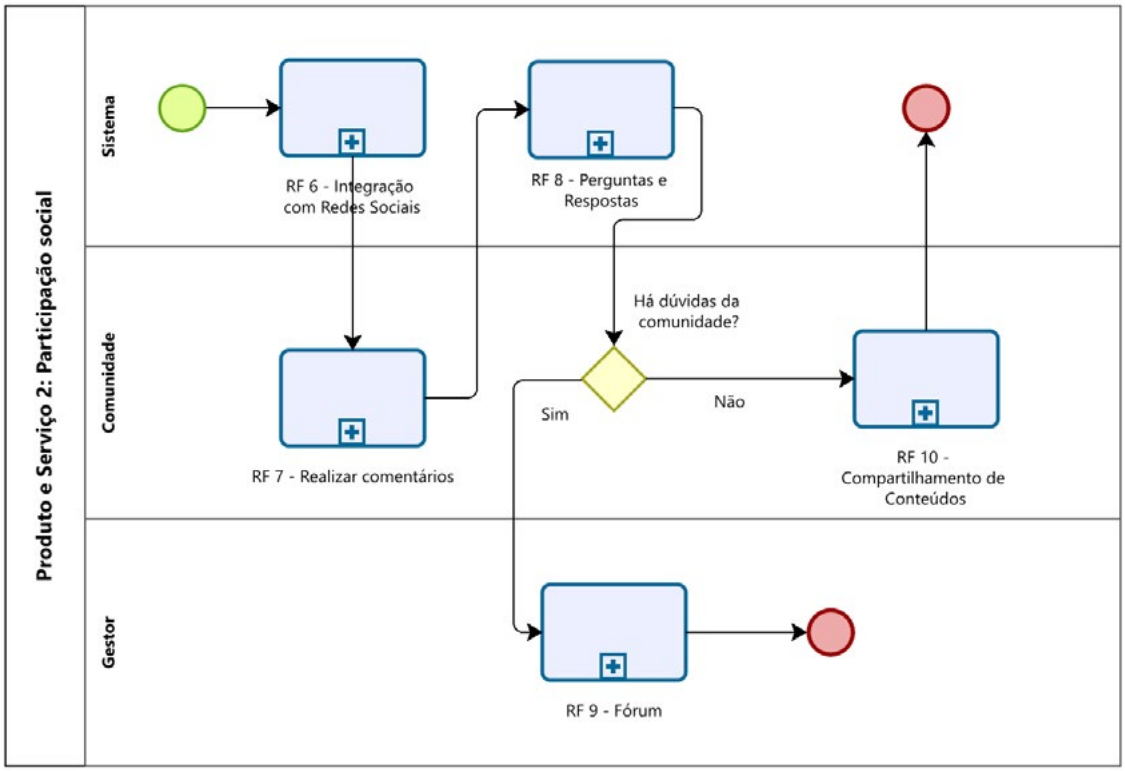
Fonte: Elaborado pelos autores (2023).

3.6.2 Produto e Serviço 2: Participação social

Produto e serviço 2 - Participação social: o sistema deve permitir a integração com diversas redes sociais. O usuário comunidade pode realizar comentários nos conteúdos disponíveis no observatório via integração com redes sociais ou por meio do próprio ambiente. O sistema disponibilizará um ambiente com perguntas e respostas. Caso haja alguma dúvida da comunidade, o sistema apresenta um fórum que possibilitará que os gestores e a comunidade do observatório possam interagir. O sistema ainda permitirá que o usuário compartilhe o conteúdo nas redes sociais, alcançando um público amplo e diversificado e promovendo o observatório nas redes sociais.

- Usuários - sistema, comunidade e gestor.

Figura 7 - Produto e serviço 2



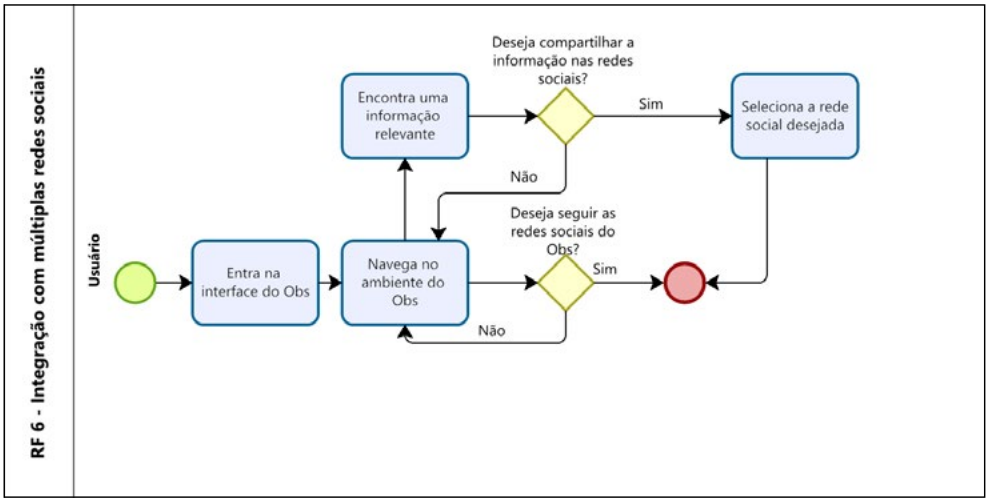
Fonte: Elaborado pelos autores (2023).

A seguir, de forma detalhada, apresentam-se os processos pertencentes ao Produto e Serviço correspondente.

RF 6 - Integração com múltiplas redes sociais: o sistema deve permitir a integração com diversas redes sociais, como Facebook, Twitter, Instagram e LinkedIn, garantindo maior compartilhamento das informações.

- Usuário: Sistema.

Figura 8 - Requisito Funcional 6

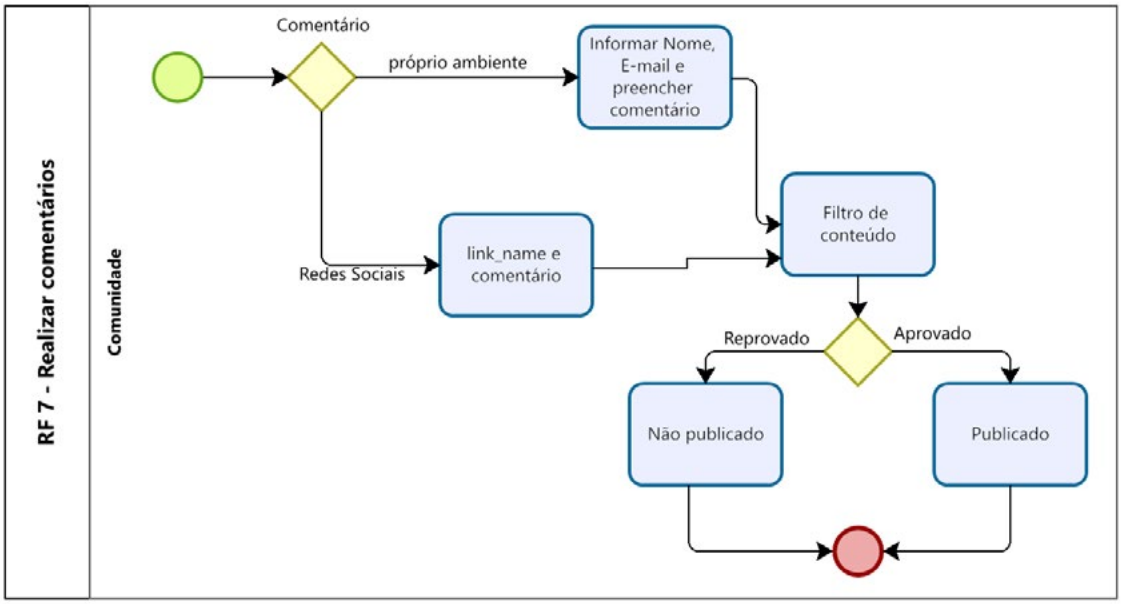


Fonte: Elaborado pelos autores (2023).

RF 7 - Realizar comentários: nos conteúdos disponíveis no observatório, o usuário pode comentar no próprio ambiente e/ou via integração com redes sociais.

- Usuário: comunidade.

Figura 9 - Requisito Funcional 7

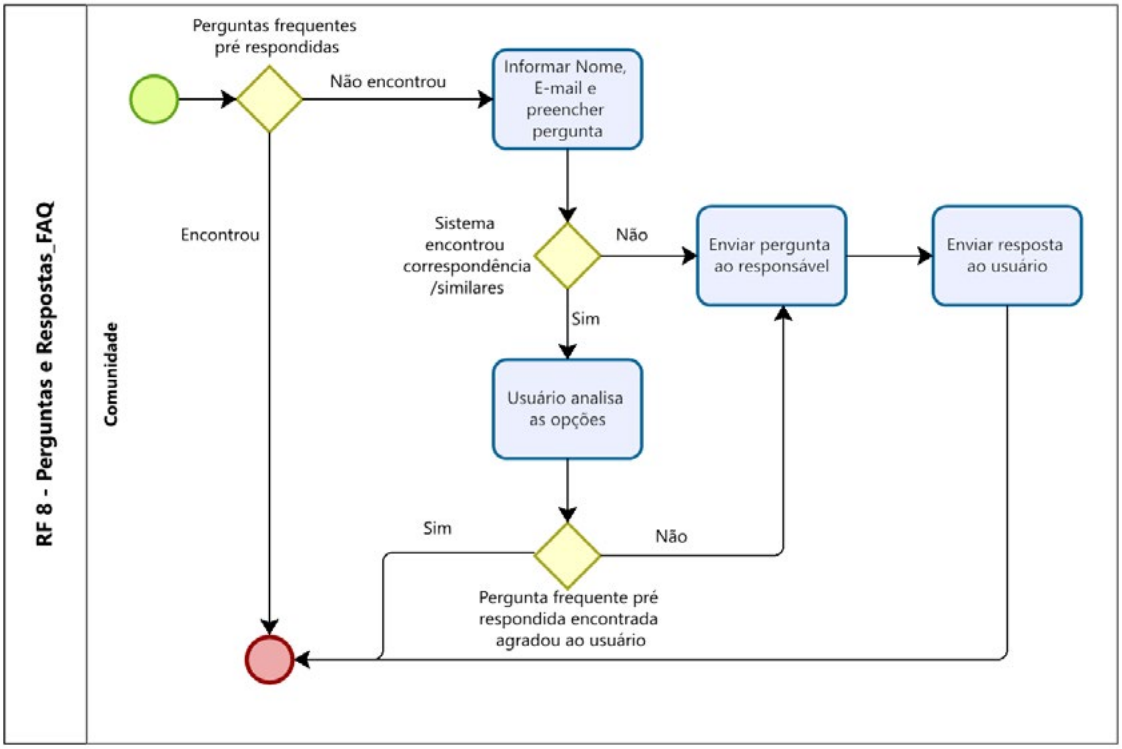


Fonte: Elaborado pelos autores (2023).

RF 8 - Perguntas e Respostas/FAQ: o sistema disponibiliza um ambiente com perguntas e respostas, FAQ.

- Usuário: comunidade.

Figura 10 - Requisito Funcional 8

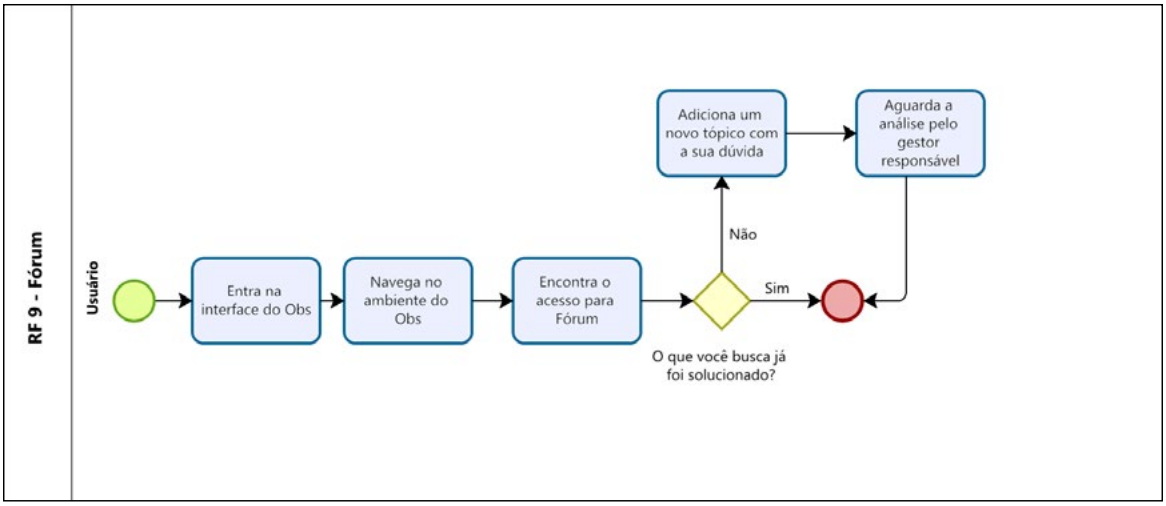


Fonte: Elaborado pelos autores (2023).

RF 9 - Fórum: o sistema apresenta um fórum que possibilita que gestores do observatório e da comunidade possam interagir.

- Usuário: sistema.

Figura 11 - Requisito Funcional 9

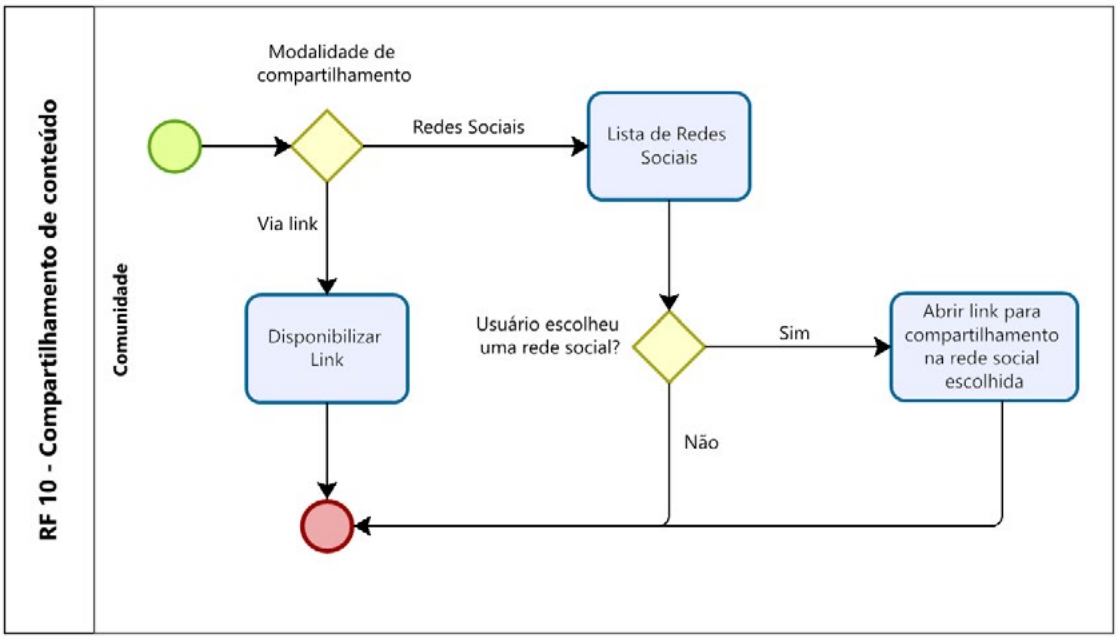


Fonte: Elaborado pelos autores (2023).

RF 10 - Compartilhamento de conteúdo: o sistema deve permitir que o usuário compartilhe conteúdo do observatório nas redes sociais, como notícias, artigos, gráficos e vídeos. Isso permitirá que o usuário alcance um público amplo e diversificado, promovendo o observatório nas redes sociais.

- Usuário: comunidade.

Figura 12 - Requisito Funcional 10



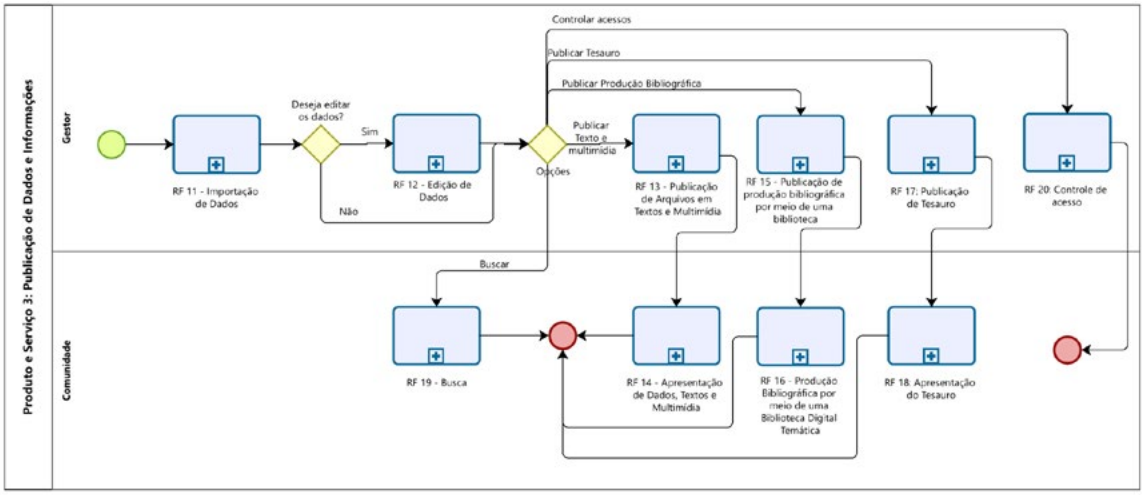
Fonte: Elaborado pelos autores (2023).

3.6.3 Produto e Serviço 3: Publicação de Dados e Informações

Produto e serviço 3 - Publicação de dados e informações: o sistema deve permitir que o usuário gestor importe dados em diferentes formatos, como CSV, Excel e XML. Isso permitirá que o usuário carregue grandes quantidades de dados no sistema de forma eficiente. Caso o usuário gestor deseje editar os dados importados, o sistema deve permitir a edição diretamente no sistema do observatório, sem a necessidade de usar ferramentas externas. Isso permitirá que o usuário faça alterações e atualizações rapidamente. O usuário gestor poderá publicar arquivos de texto e multimídia, como documentos, relatórios, boletins, fotos, vídeos, infográficos etc. O usuário comunidade pode acessar os dados, os textos e os arquivos multimídia inseridos pelos gestores do observatório. O usuário gestor pode publicar a sua produção bibliográfica, como artigos, periódicos, teses, dissertações, livros, entre outros, por meio de uma biblioteca digital temática. O usuário comunidade pode acessar a produção bibliográfica inserida na Biblioteca Digital Temática do Observatório. O usuário gestor pode publicar um tesauro que vai ficar disponível para o usuário comunidade na área do observatório. O sistema deve permitir o controle de acesso aos diferentes tipos de informações disponíveis, garantindo que apenas usuários autorizados (usuários gestores) possam visualizar, editar ou baixar determinados arquivos ou dados. A fim de que o usuário comunidade encontre informações específicas de forma rápida e eficiente, o sistema deve permitir que ele realize buscas simples e avançadas por palavras-chave, data, autor, tipo de arquivo e outras informações relevantes.

- Usuários: gestor e comunidade.

Figura 13 - Produto e serviço 3



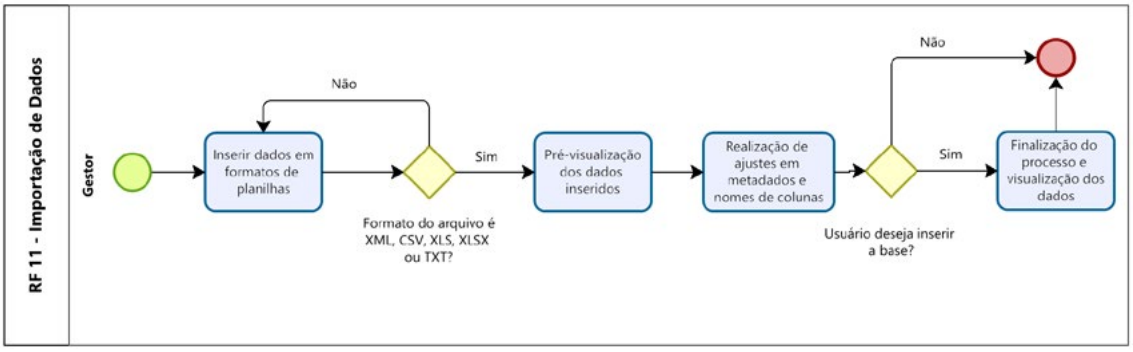
Fonte: Elaborado pelos autores (2023).

A seguir, de forma detalhada, os processos pertencentes ao Produto e Serviço correspondente são apresentados.

RF 11 - Importação de dados: o sistema deve permitir a importação de dados em diferentes formatos, como CSV, Excel e XML. Isso permitirá que o usuário comporte grandes quantidades de dados no sistema de forma eficiente.

- Usuário: gestor.

Figura 14 - Requisito Funcional 11

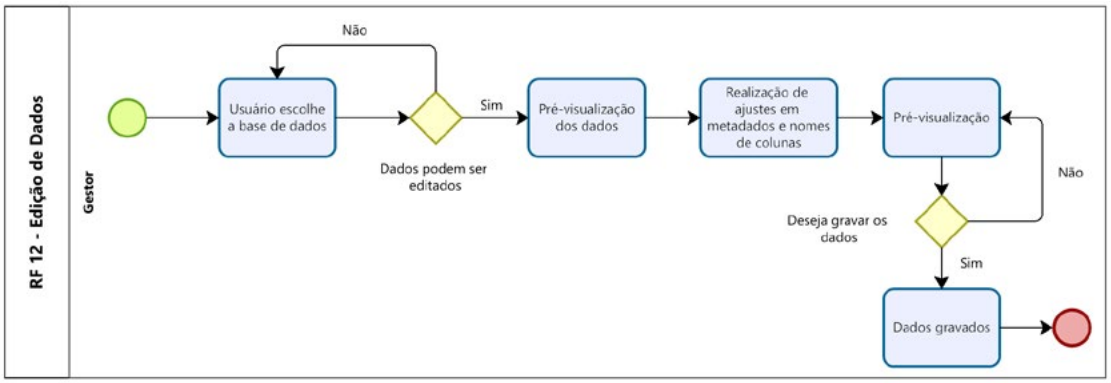


Fonte: Elaborado pelos autores (2023).

RF 12: Edição de dados: o sistema deve permitir a edição de dados diretamente no sistema, sem a necessidade de usar ferramentas externas. Isso permitirá que o usuário faça alterações e atualizações rapidamente.

- Usuário: gestor.

Figura 15 - Requisito Funcional 12

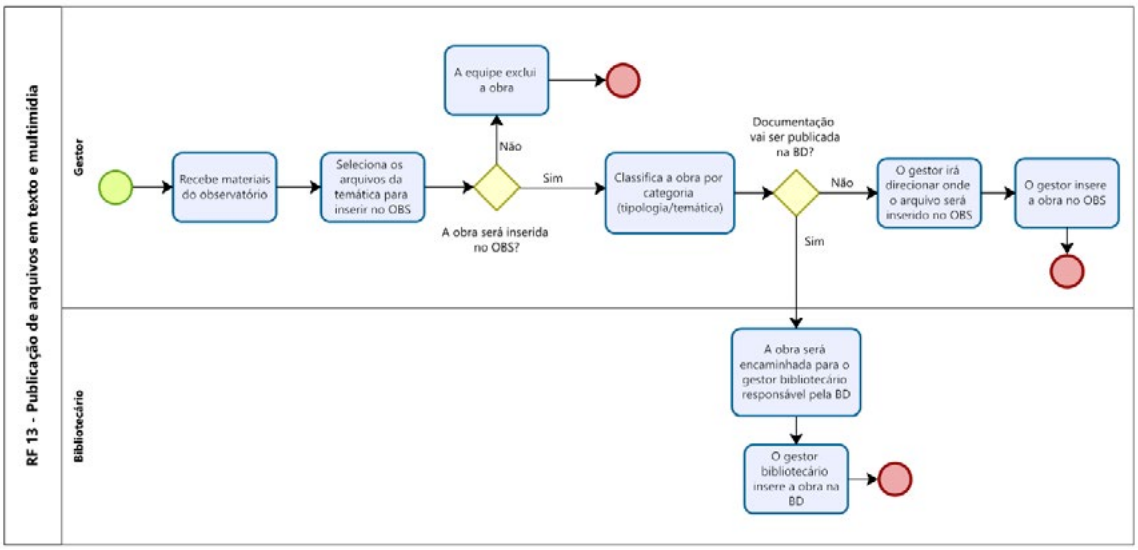


Fonte: Elaborado pelos autores (2023).

RF 13 - Publicação de Arquivos em Texto e Multimídia: publicação de arquivos de texto e multimídia no observatório, como documentos, relatórios, boletins, fotos, vídeos, infográficos, entre outros materiais em formatos como .docx, .pdf, .txt, .png, .jpg, .mp3, mp4. Esses arquivos podem ser acessados por usuários autorizados e estão organizados em categorias que facilitam a sua localização.

- Usuário: gestor.

Figura 16 - Requisito Funcional 13

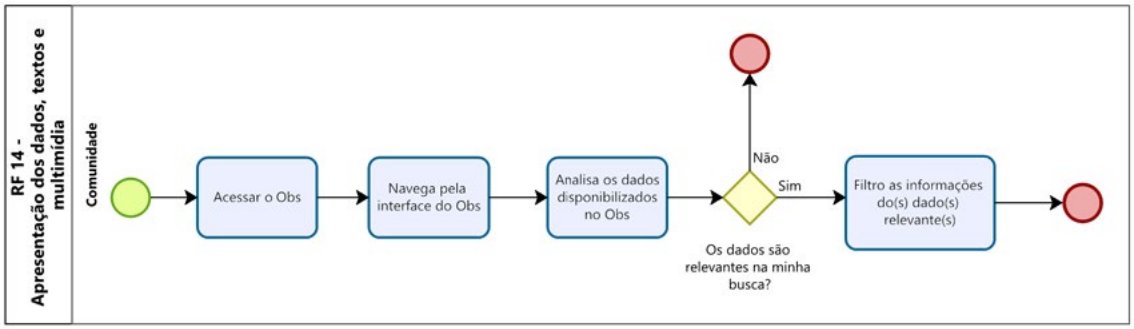


Fonte: Elaborado pelos autores (2023).

RF 14 - Apresentação dos dados, textos e multimídia: o usuário comunidade pode acessar os dados, os textos e os arquivos multimídia inseridos pelos gestores do observatório.

- Usuário: comunidade.

Figura 17 - Requisito Funcional 14

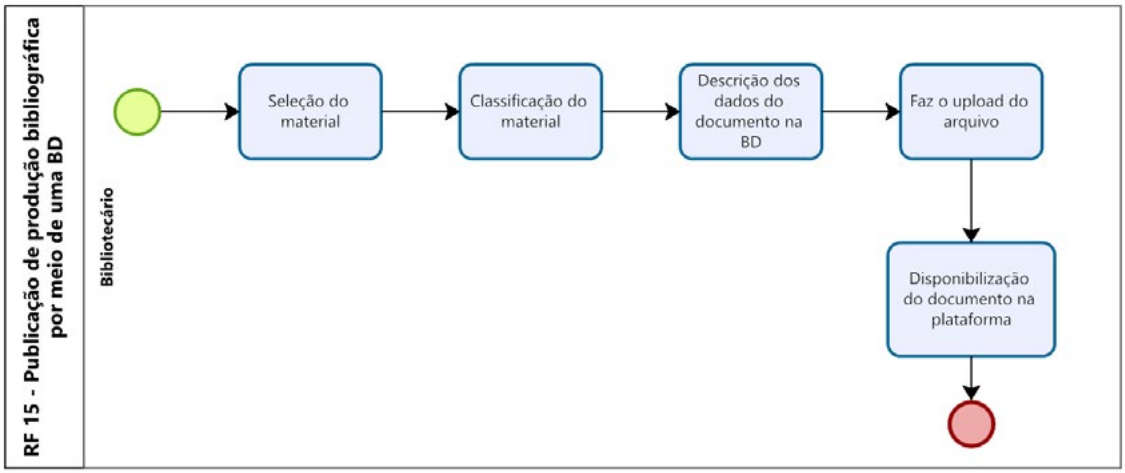


Fonte: Elaborado pelos autores (2023).

RF 15 - Publicação de produção bibliográfica por meio de uma biblioteca digital temática: publicação de artigos, periódicos, teses, dissertações, livros, documentação técnica, cartilhas, manuais, entre outros, da temática do observatório.

- Usuário: gestor.

Figura 18 - Requisito Funcional 15

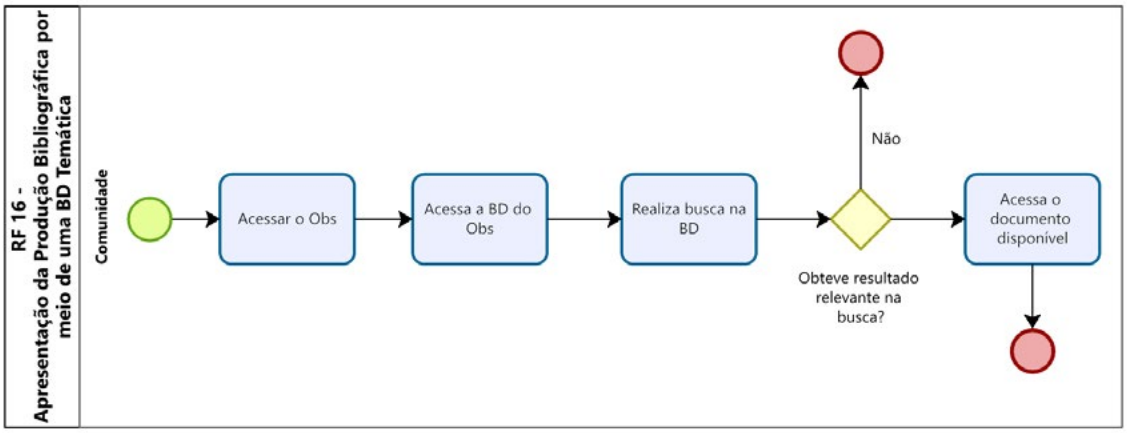


Fonte: Elaborado pelos autores (2023).

RF 16 - Apresentação da produção bibliográfica por meio de uma biblioteca digital temática: o usuário comunidade pode acessar a produção bibliográfica da temática do observatório.

- Usuário: comunidade.

Figura 19 - Requisito Funcional 16

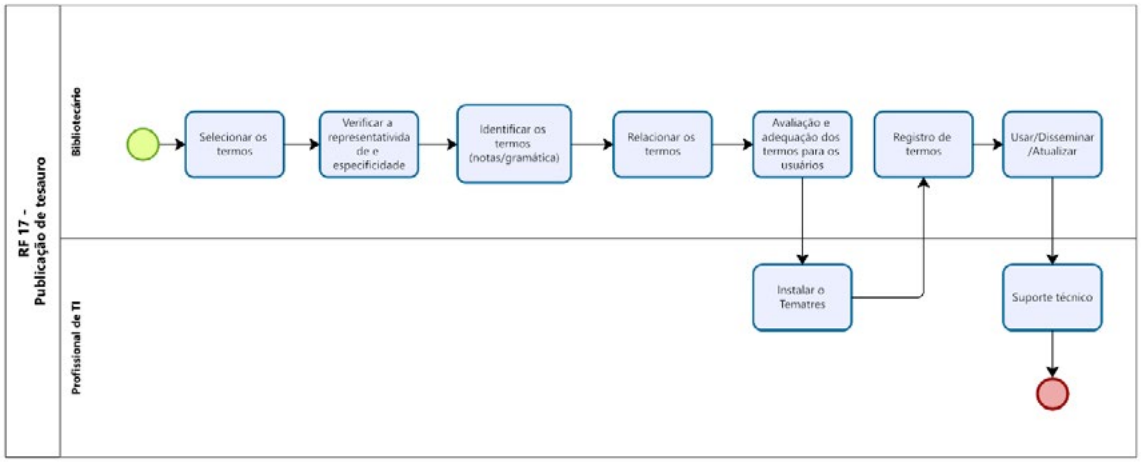


Fonte: Elaborado pelos autores (2023).

RF 17 - Publicação de Tesauro: publicação de tesauros na área do observatório.

- Usuário: gestor.

Figura 20 - Requisito Funcional 17

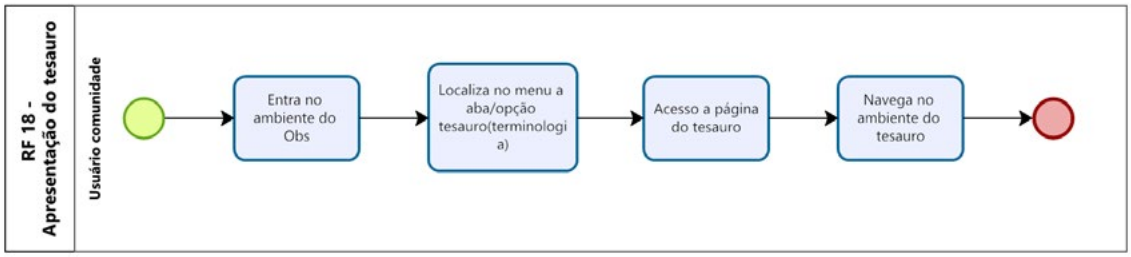


Fonte: Elaborado pelos autores (2023).

RF 18 - Apresentação do Tesauro: apresentação de tesauros para o usuário comunidade na área do observatório.

- Usuário: comunidade.

Figura 21 - Requisito Funcional 18

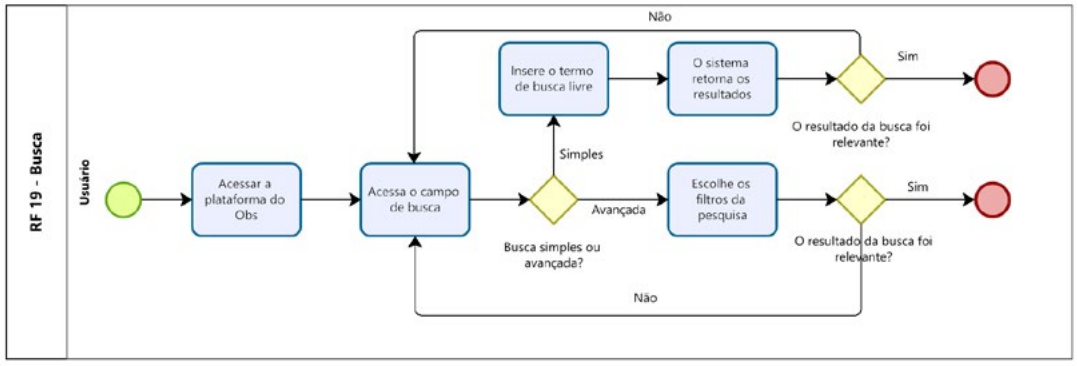


Fonte: Elaborado pelos autores (2023).

RF 19 - Busca: o sistema deve permitir que o usuário realize buscas simples e avançadas por palavras-chave, data, autor, tipo de arquivo e outras informações relevantes. Isso permitirá que o usuário encontre informações específicas de forma rápida e eficiente.

- Usuário: gestor e comunidade.

Figura 22 - Requisito Funcional 19

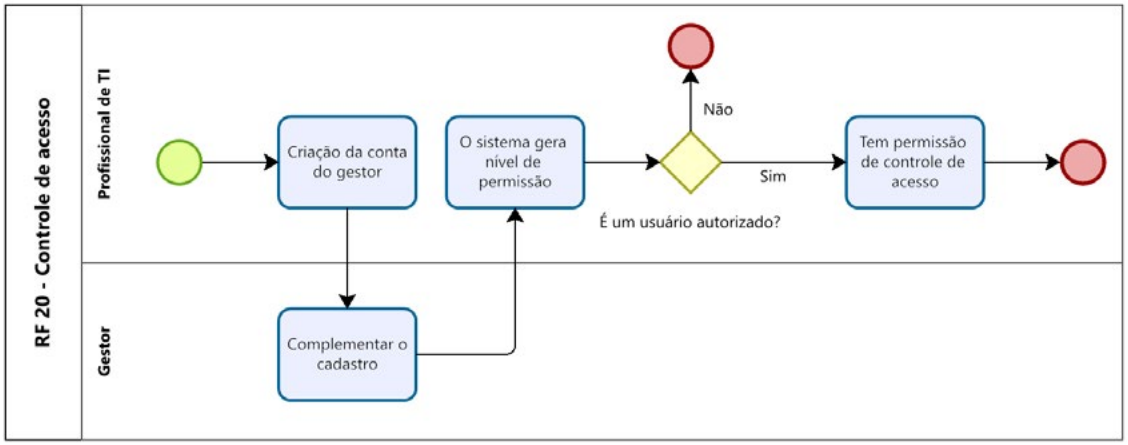


Fonte: Elaborado pelos autores (2023).

RF 20 - Controle de acesso: o sistema deve permitir o controle de acesso aos diferentes tipos de informações disponíveis, garantindo que apenas usuários autorizados possam visualizar, editar ou baixar determinados arquivos ou dados.

- Usuário: gestor.

Figura 23 - Requisito Funcional 20



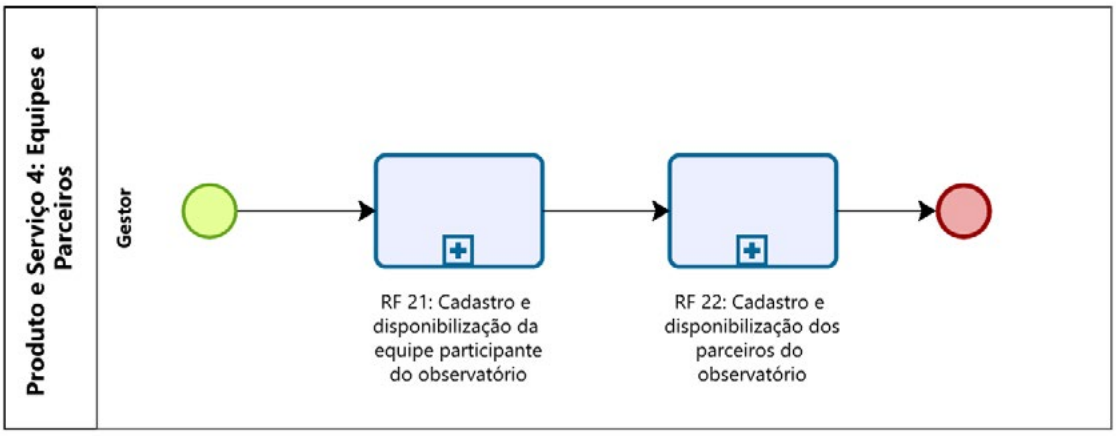
Fonte: Elaborado pelos autores (2023).

3.6.4 Produto e Serviço 4: Equipes e Parceiros

Produto e serviço 4 - Equipes e parceiros: o usuário gestor irá cadastrar e disponibilizar informações sobre a equipe que participa do observatório, como nome, cargo, instituição, e-mail, telefone, área de atuação etc. Além disso, há possibilidade de atualizar e editar essas informações, assim como definir permissões de acesso para diferentes tipos de usuários visualizarem e editarem os dados. O usuário gestor também poderá cadastrar e disponibilizar informações sobre os parceiros e instituições relacionadas ao observatório, como nome, ramo de atuação, descrição etc. Há a possibilidade de atualizar e editar essas informações, além de definir permissões de acesso para diferentes tipos de usuários visualizarem e editarem esses dados.

- Usuário: gestor.

Figura 24 - Produto e serviço 4



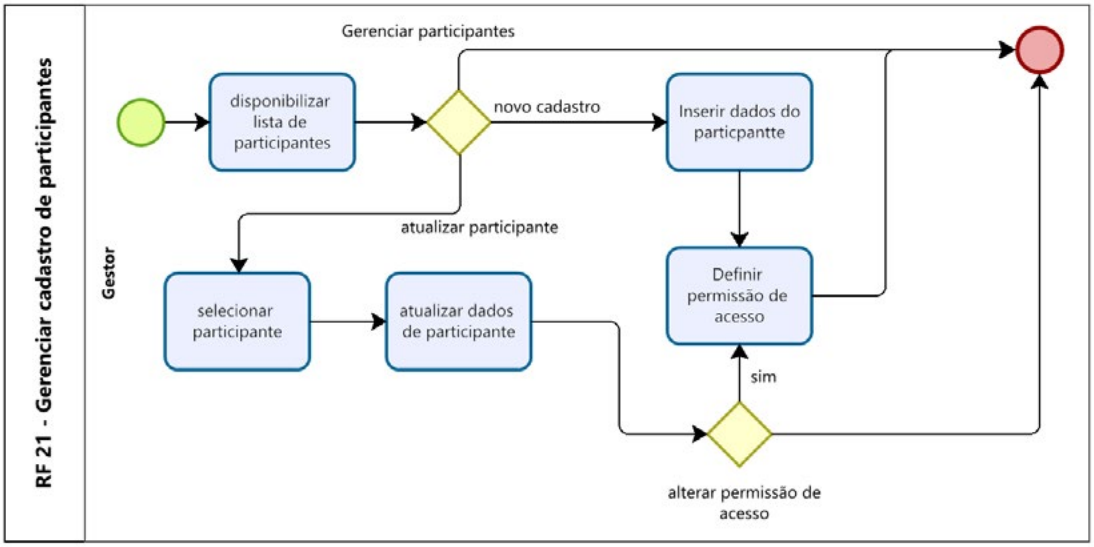
Fonte: Elaborado pelos autores (2023).

A seguir, são apresentados, de forma detalhada, os processos pertencentes ao produto e serviço correspondente.

RF 21 - Cadastro e disponibilização da equipe participante do observatório: cadastrar e disponibilizar informações sobre a equipe, como nome, cargo, instituição, e-mail, telefone, área de atuação etc. Além disso, há possibilidade de atualizar e editar essas informações, além de definir permissões de acesso para diferentes tipos de usuários visualizarem e editarem esses dados.

- Usuário: gestor.

Figura 25 - Requisito Funcional 21

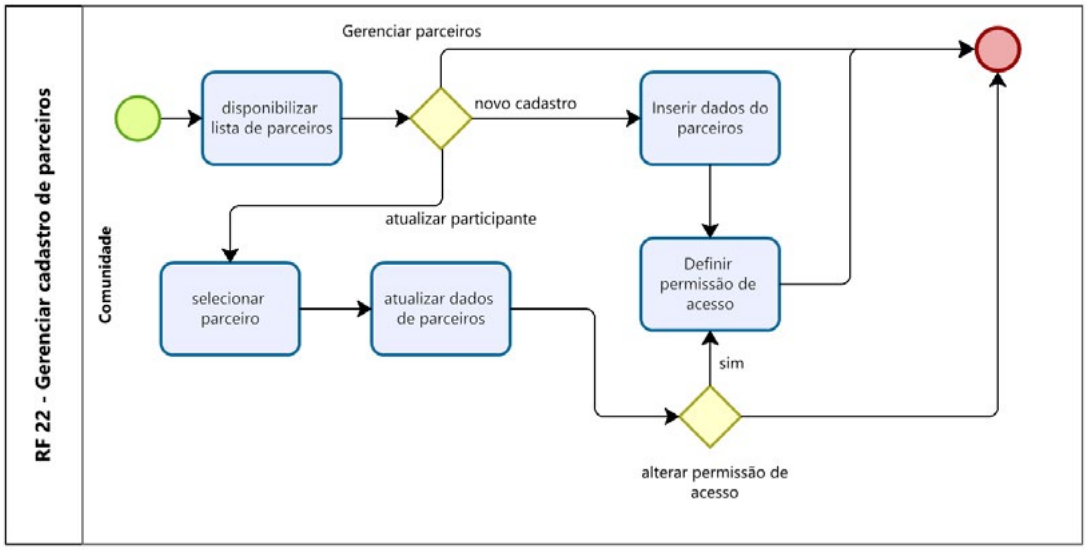


Fonte: Elaborado pelos autores (2023).

RF 22 - Cadastro e disponibilização dos parceiros do observatório: cadastrar e disponibilizar informações sobre os parceiros e instituições relacionadas, como nome, ramo de atuação, descrição etc. Há a possibilidade de atualizar e editar essas informações, além de definir permissões de acesso para diferentes tipos de usuários visualizarem e editarem esses dados.

- Usuário: comunidade.

Figura 26 - Requisito Funcional 22



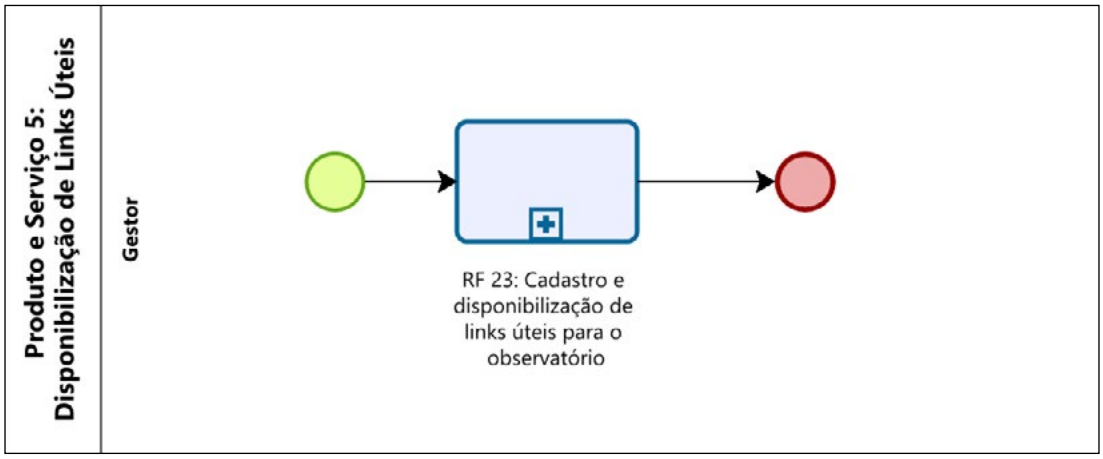
Fonte: Elaborado pelos autores (2023).

3.6.5 Produto e Serviço 5: Disponibilização de Links Úteis

Produto e serviço 5 - Disponibilização de links úteis: o usuário gestor pode cadastrar e disponibilizar uma lista de links úteis para o usuário do observatório, como sites de referência, instituições parceiras, bancos de dados, entre outros recursos relevantes para o tema. É importante que esses links sejam organizados em categorias e que haja a possibilidade de pesquisar e filtrar por palavra-chave ou categoria a fim de facilitar a navegação.

- Usuário: gestor.

Figura 27 - Produto e serviço 5



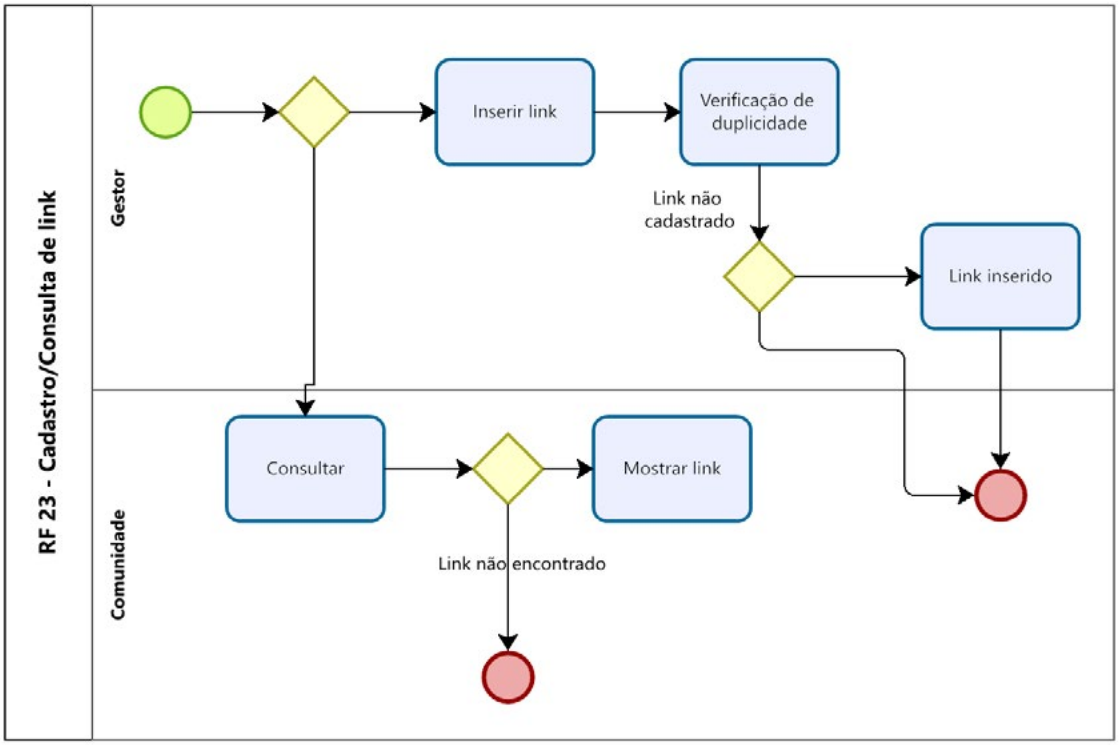
Fonte: Elaborado pelos autores (2023).

A seguir, de forma detalhada, são apresentados os processos pertencentes ao produto e serviço correspondente.

RF 23 - Cadastro e disponibilização de links úteis para o observatório: cadastrar e disponibilizar uma lista de links úteis para o usuário, como sites de referência, instituições parceiras, bancos de dados, entre outros recursos relevantes para o tema. É importante que esses links sejam organizados em categorias e que haja a possibilidade de pesquisar e filtrar por palavra-chave ou categoria a fim de facilitar a navegação do usuário.

- Usuário: comunidade e gestor.

Figura 28 - Requisito Funcional 23



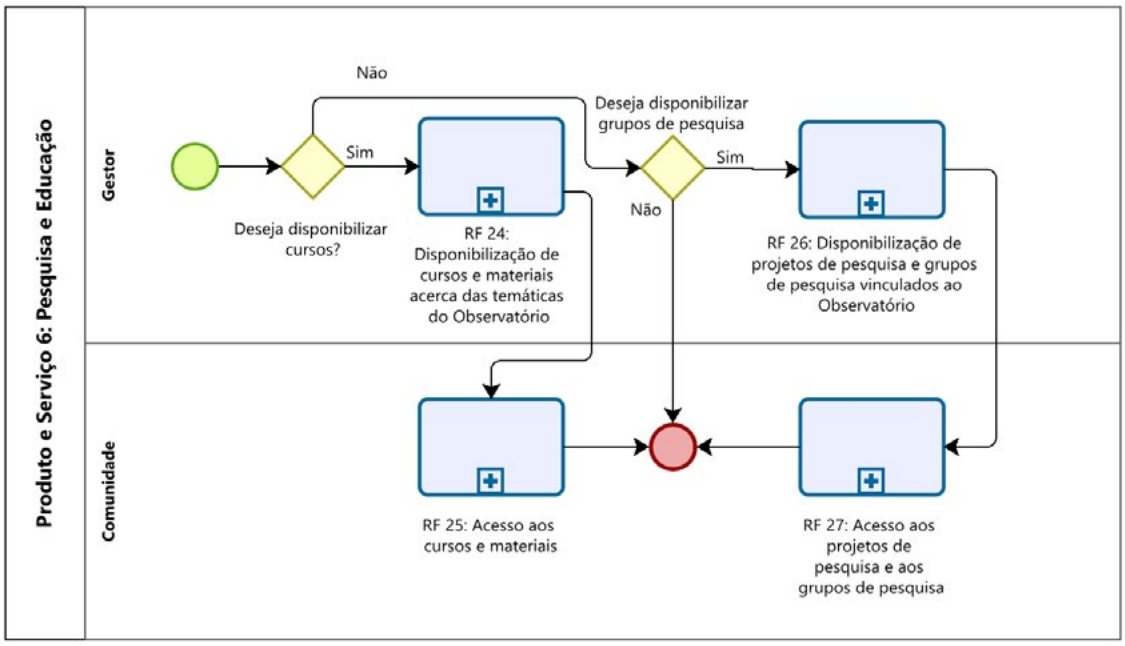
Fonte: Elaborado pelos autores (2023).

3.6.6 Produto e Serviço 6: Pesquisa e Educação

Produto e serviço 6 - Pesquisa e educação: o usuário gestor pode pretender disponibilizar cursos e materiais educativos relacionados às temáticas do observatório, com o objetivo de promover a capacitação e formação de competências em seus usuários. Caso disponibilizado pelo usuário gestor o acesso aos cursos e materiais para o usuário comunidade, ocorrerá a garantia ao acesso aos cursos e materiais educativos disponibilizados pelo observatório, permitindo que os usuários possam se capacitar e se informar sobre as temáticas abordadas. O usuário gestor também terá a opção de disponibilizar projetos de pesquisa e grupos de pesquisa que estejam vinculados ao observatório, permitindo que seus usuários possam se informar e conhecer mais sobre as pesquisas em andamento. Se o usuário gestor realizar essa disponibilização, o sistema permitirá ao usuário comunidade o acesso aos projetos de pesquisa e grupos de pesquisa vinculados, permitindo que seus usuários possam acompanhar o desenvolvimento das pesquisas e conhecer mais sobre os grupos de pesquisa atuantes nas temáticas abordadas.

- Usuário: gestor e comunidade.

Figura 29 - Produto e serviço 6



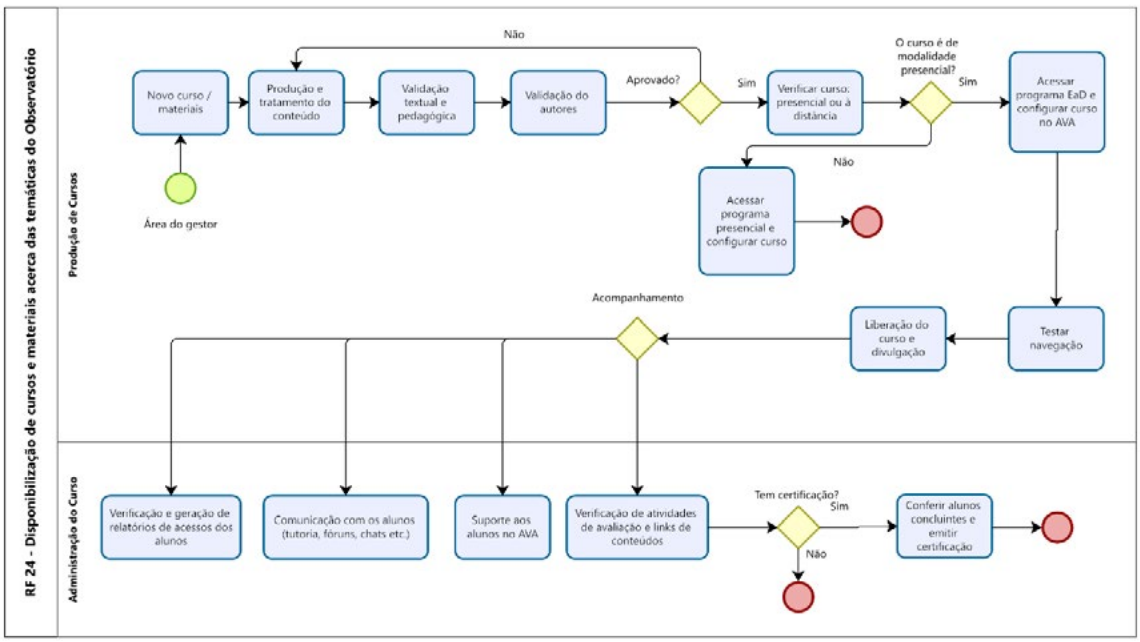
Fonte: Elaborado pelos autores (2023).

A seguir, são apresentados detalhadamente os processos pertencentes ao produto e serviço correspondente.

RF 24 - Disponibilização de cursos e materiais acerca das temáticas do observatório: disponibilizar cursos e materiais educativos relacionados às temáticas, com o objetivo de promover a capacitação e formação de competências em seus usuários.

- Usuário: gestor.

Figura 30 - Requisito Funcional 24

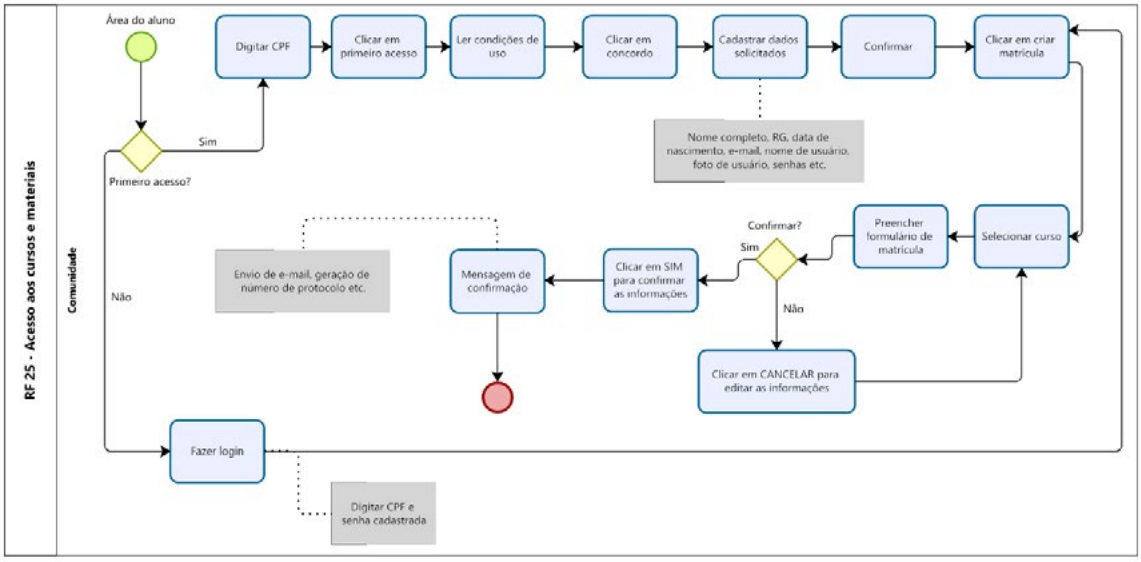


Fonte: Elaborado pelos autores (2023).

RF 25 - Acesso aos cursos e materiais: garantir o acesso aos cursos e materiais educativos disponibilizados pelo observatório, permitindo que os usuários possam se capacitar e se informar sobre as temáticas abordadas.

- Usuário: comunidade.

Figura 31 - Requisito Funcional 25

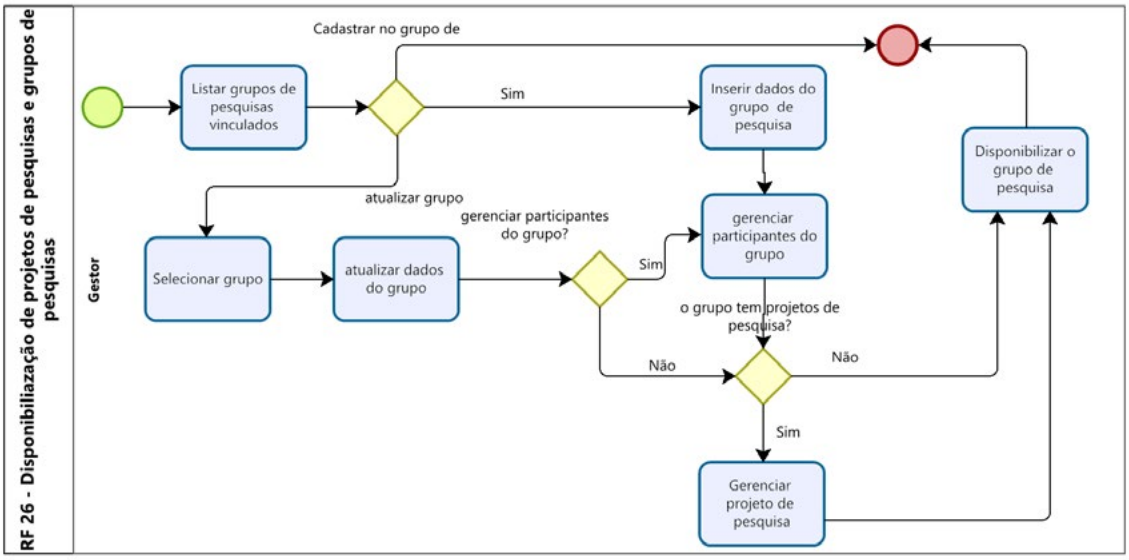


Fonte: Elaborado pelos autores (2023).

RF 26 - Disponibilização de projetos de pesquisa e grupos de pesquisa vinculados ao observatório: o sistema deve disponibilizar projetos de pesquisa e grupos de pesquisa que estejam vinculados ao observatório, permitindo que seus usuários possam se informar e conhecer mais sobre as pesquisas em andamento.

- Usuário: gestor.

Figura 32 - Requisito Funcional 26

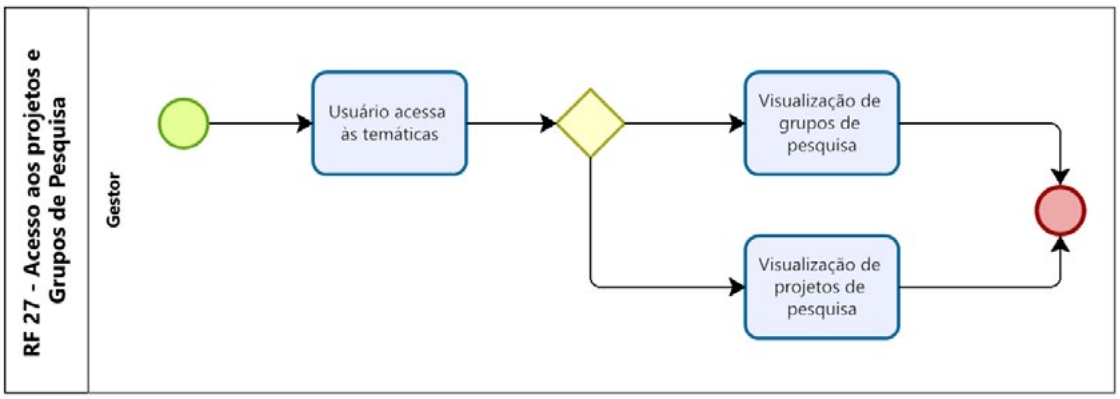


Fonte: Elaborado pelos autores (2023).

RF 27 - Acesso aos projetos de pesquisa e aos grupos de pesquisa: o sistema permite o acesso aos projetos de pesquisa e grupos de pesquisa vinculados ao observatório, permitindo que seus usuários possam acompanhar o desenvolvimento das pesquisas e conhecer mais sobre os grupos de pesquisa atuantes nas temáticas abordadas.

- Usuário: comunidade.

Figura 33 - Requisito Funcional 27



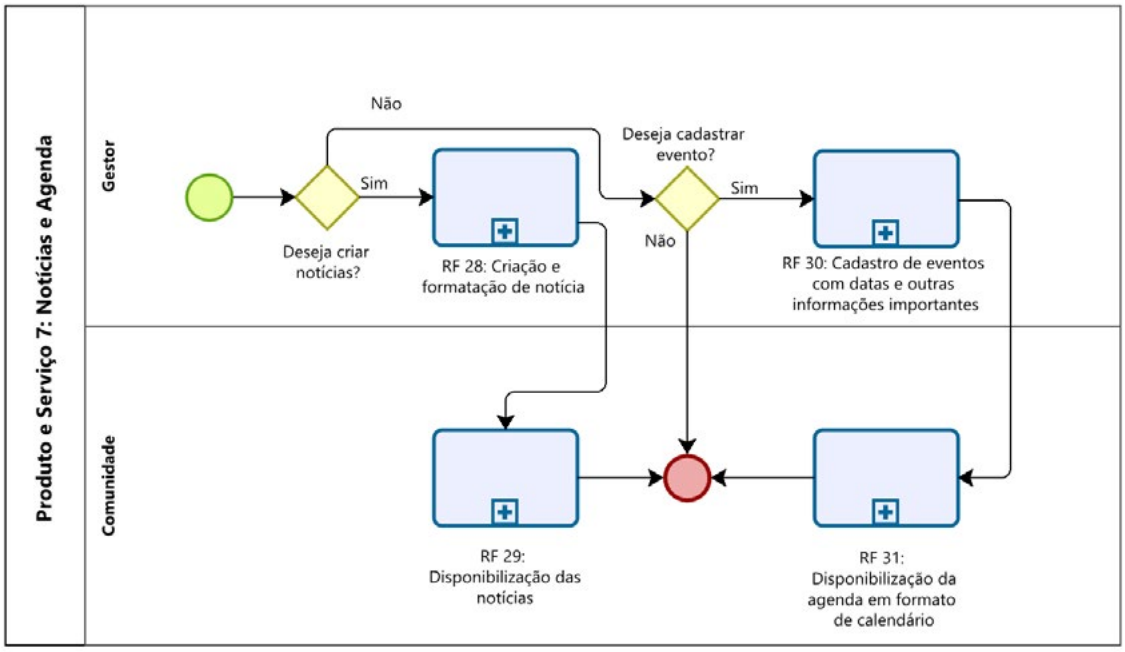
Fonte: Elaborado pelos autores (2023).

3.6.7 Produto e Serviço 7: Notícias e Agenda

Produto e serviço 7 - Notícias e agenda: caso o usuário gestor deseje criar um ambiente de notícias no observatório, o sistema deve permitir criar e editar notícias na plataforma, concedendo a inserção de conteúdos como texto, imagens e vídeos, além de formatação e estilização. O sistema deve ter um ambiente que apresenta as notícias criadas pelo usuário gestor, permitindo que os usuários acessem o conteúdo por meio da plataforma. Caso o usuário gestor deseje cadastrar um evento futuro, o sistema deve ter uma área que permita esse cadastro na plataforma, incluindo informações como data, horário, local e descrição do evento. O mesmo deve ter uma agenda de eventos em um formato de calendário na interface do usuário comunidade, facilitando a visualização e possibilitando o acesso às informações sobre cada evento cadastrado.

- Usuário: gestor e comunidade.

Figura 34 - Produto e serviço 7



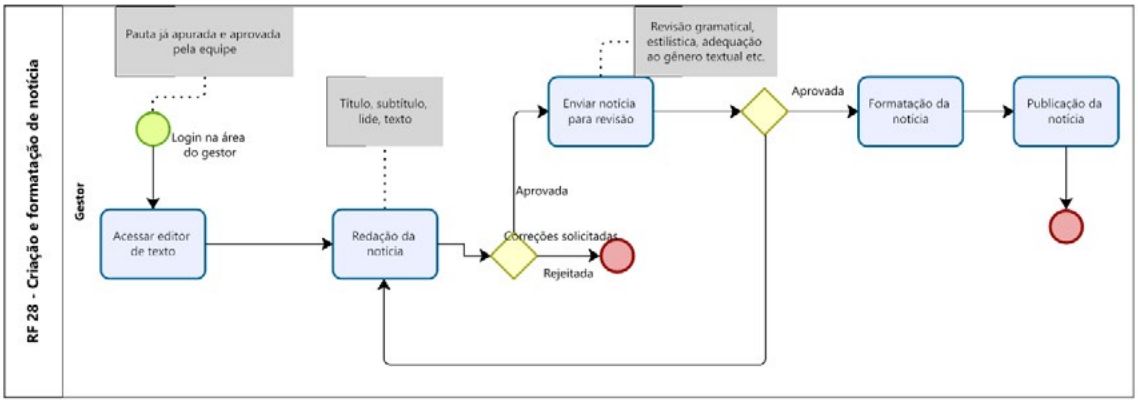
Fonte: Elaborado pelos autores (2023).

A seguir, apresentam-se os processos pertencentes ao produto e serviço correspondente.

RF 28 - Criação e formatação de notícia: o sistema deve permitir a criação e edição de notícias na plataforma do observatório, possibilitando a inserção de textos, imagens e vídeos, além de formatação e estilização do conteúdo.

- Usuário: gestor.

Figura 35 - Requisito Funcional 28

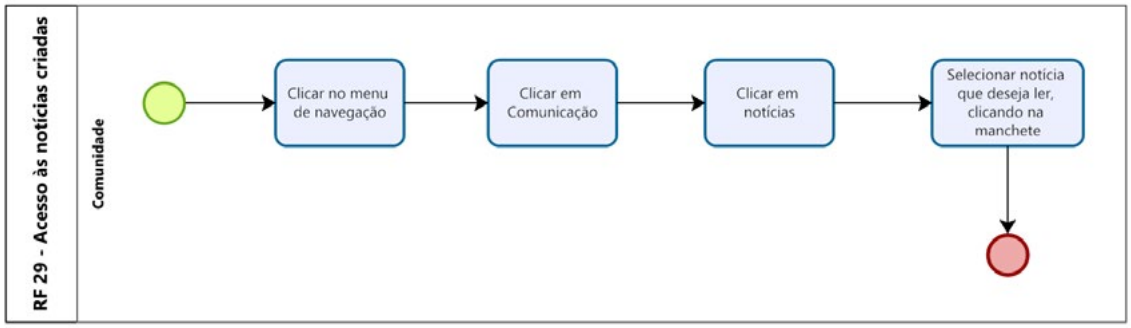


Fonte: Elaborado pelos autores (2023).

RF 29 - Disponibilização das notícias: o sistema deve dispor de um ambiente que apresente as notícias criadas, permitindo que os usuários acessem o conteúdo por meio da plataforma do observatório.

- Usuário: comunidade.

Figura 36 - Requisito Funcional 29

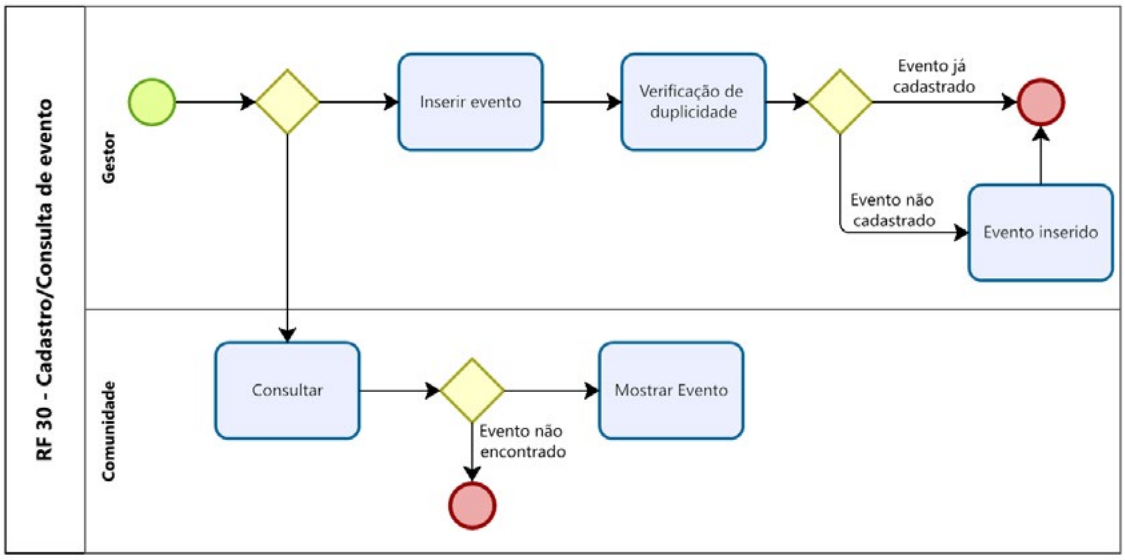


Fonte: Elaborado pelos autores (2023).

RF 30 - Cadastro de eventos com datas e outras informações importantes: o sistema deve ter uma área que permite o cadastro de eventos futuros na plataforma, incluindo informações como data, horário, local e descrição do evento.

- Usuário: gestor.

Figura 37 - Requisito Funcional 30

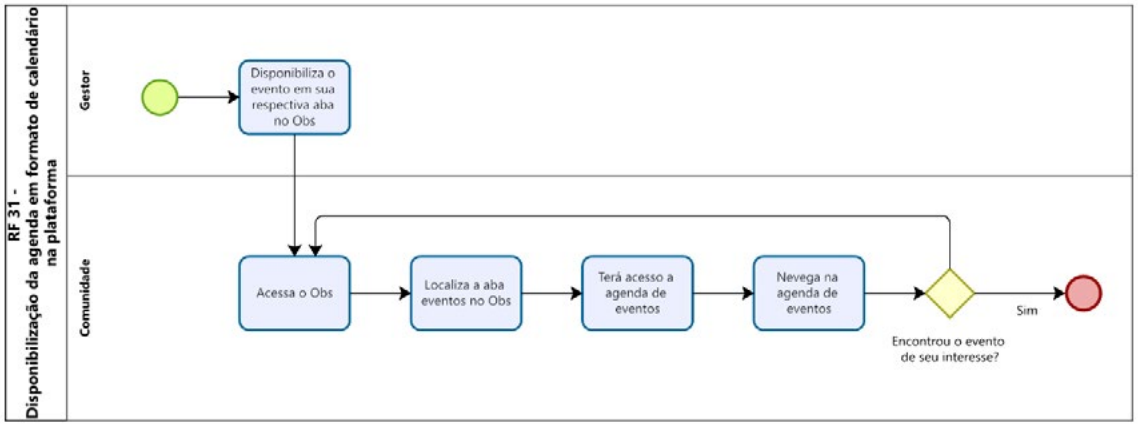


Fonte: Elaborado pelos autores (2023).

RF 31 - Disponibilização da agenda em formato de calendário na plataforma: o sistema deve ter uma agenda de eventos em forma de calendário, facilitando a visualização dos usuários e possibilitando o acesso às informações sobre cada evento cadastrado.

- Usuário: comunidade e gestor.

Figura 38 - Requisito Funcional 31



Fonte: Elaborado pelos autores (2023).

4. CONSIDERAÇÕES FINAIS

A avaliação do progresso da segunda meta do projeto, que envolveu o mapeamento de processos do fluxo informacional para o modelo, alcançou os objetivos estabelecidos para a etapa. Cada fase da meta apresenta resultados que fornecem informações para dar continuidade ao projeto. No entanto, os dados gerados poderão ser revisados à medida que a elaboração do modelo avançar, permitindo melhor entendimento e uma validação contínua até o final.

Os resultados apresentados no relatório serão subsídios para a próxima meta do projeto, na qual o mapeamento das fontes de informação, produtos/serviços e indicadores desempenham um papel integrante para o modelo. Com base nas informações coletadas até o momento, espera-se que os subsídios sejam fundamentais para a consecução dos objetivos propostos. Ao explorar e compreender melhor as fontes de informação, os produtos/serviços e os indicadores, será possível aprimorar significativamente o modelo e otimizar sua aplicação prática.

A expectativa é que, cumprida a próxima meta, sejam obtidos resultados ainda mais sólidos e relevantes. O contínuo avanço do mapeamento e a integração de informações adicionais permitirão aperfeiçoar a precisão e a eficácia do modelo, garantindo uma abordagem cada vez mais completa e abrangente.

REFERÊNCIAS

AALST, Wil M. P. Van Der; BICHLER, Martin; HEINZL, Armin. Robotic process automation. **Business & Information Systems Engineering**, [*s. l.*], v. 60, n. 4, p. 269-272, 14 May 2018. DOI: https://doi.org/10.1007/s12599-018-0542-4. Disponível em: https://aisel.aisnet.org/bise/vol60/iss4/1/. Acesso em: 25 maio 2023.

AWS. **O que é uma API?** [*s. l.*], © 2023. Disponível em: https://aws.amazon.com/pt/what-is/api/#:~:text=API%20significa%20Application%20Programming%20Interface,de%20servi%C3%A7o%20entre%20duas%20aplica%C3%A7%C3%B5e/. Acesso em: 25 maio 2023.

EBERENDU, Adanma Cecilia *et al.* Unstructured Data: an overview of the data of Big Data. **International Journal of Computer Trends and Technology**, [*s. l.*], v. 38, n. 1, p. 46-50, Aug. 2016. Disponível em: https://ijcttjournal.org/archives/ijctt-v38p109/. Acesso em: 14 mar. 2023.

MARTINEZ, Juan Manuel Perez *et al.* Integrating data warehouses with web data: a survey. **IEEE transactions on knowledge and data engineering**, Canada, v. 20, n. 7, p. 940-955, July 2008. DOI: 10.1109/TKDE.2007.190746. Disponível em: https://ieeexplore.ieee.org/document/4490177/. Acesso em: 14 mar. 2023.

SIRISURIYA, SCM de S. . A comparative study on web scraping. 2015. *In:* INTERNATIONAL RESEARCH CONFERENCE, 8., 2015, Sri Lanka. **Anais** [...]. Sri Lanka: KDU, 2015. p. 135-140. Disponível em: http://ir.kdu.ac.lk/bitstream/handle/345/1051/com-059.pdf?sequence=1&isAllowed=y/. Acesso em: 4 fev. 2023.

SREE, K. Dhana; BINDU, C. Shoba. Data analytics: Why data normalization. **International Journal of Engineering and Technology (UAE)**, [*s. l.*], v. 7, n. 4.6, p. 209-213, 2018. DOI: https://doi.org/10.14419/ijet.v7i4.6.20464. Disponível em: https://www.sciencepubco.com/index.php/ijet/article/view/20464/. Acesso em: 30 abr. 2023.

SAS - Quadra 05 - Lote 06 - Bloco H – Sobreloja
Cep: 70070-912 - Brasília / DF

Telefone: +55 61 3217 6213
E-mail: cotec@ibict.br

ibict
Instituto Brasileiro de Informação em Ciência e Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

GOVERNO FEDERAL
BRASIL
UNIÃO E RECONSTRUÇÃO