Relatório final de atividades
Projeto: SAVITS: Sistema de Avaliação e Validação de Impactos em Tecnologias Sociais

Coordenador do Projeto:	**Washington Luís Ribeiro de Carvalho Segundo**		
Nome do(a) pesquisador(a):	**Tiago Alves de Oliveira**		
Vigência da Bolsa de Pesquisa:	**20/03/2026**	**Até:**	**20/09/2026**
Encerramento em:	**20/09/2026**		
Ref. Fundação:	**32725**	**Ref. Ibict:**	**008/2026-81**

Descrição das Atividades da Bolsa de Pesquisa:

1-Planejamento, modelagem e documentação dos fluxos de coleta, tratamento, organização e validação dos dados do projeto SAVITS, com foco em qualidade, rastreabilidade e adequação metodológica; meta 1. 2- Desenvolvimento, implementação, otimização e documentação da infraestrutura analítica e dos pipelines de ingestão e processamento de dados com uso da stack Elastic (Elasticsearch, Kibana e Logstash); meta 2. 3-Desenvolvimento, evolução e manutenção de módulos de Business Intelligence e dashboards interativos para monitoramento de indicadores e apoio à tomada de decisão; meta 2. 4- Apoio à estruturação de procedimentos de governança, segurança, anonimização, organização e interoperabilidade dos dados e metadados do projeto; Meta 3.

1 INTRODUÇÃO

O presente relatório descreve as atividades desenvolvidas no período de 20 de março a 20 de setembro de 2026, tomando como referência estrita o plano de atividades registrado no formulário do bolsista: prova de conceito e coleta de dados dos grupos de pesquisa; manutenção de dashboards interativos com o Kibana; estudo aprimorado da infraestrutura da stack ELK (Elasticsearch, Kibana e Logstash); e produção de relatórios. As ações foram organizadas de modo a manter correspondência direta entre o que estava previsto, os procedimentos técnicos executados e os resultados obtidos no período.

As atividades tiveram caráter aplicado e envolveram coleta, organização e validação de dados, manutenção de recursos de visualização e acompanhamento, análise de infraestrutura de busca e indexação e documentação técnica. A stack ELK foi tratada tanto sob a perspectiva de uso analítico, por meio do Kibana, quanto sob a perspectiva de infraestrutura, considerando indexação, ingestão, segurança, operação, observabilidade e manutenção dos componentes Elasticsearch, Kibana e Logstash.

Em paralelo às atividades diretamente vinculadas ao plano, também foram executadas tarefas correlacionadas em ambientes técnicos do INPI e do BrCris. Essas ações são apresentadas neste relatório como ampliação de escopo, sem substituir ou alterar as atividades previstas na planilha. A correlação decorre do uso de tecnologias, procedimentos e problemas operacionais comuns, especialmente no que se refere à stack ELK, pipelines de ingestão, qualidade e organização de dados, segurança, snapshots, configuração de índices e produção de evidências técnicas.

2 MÉTODO

A execução foi estruturada em quatro frentes correspondentes às atividades previstas no formulário do bolsista. Em cada frente foram adotados procedimentos de análise, implementação, validação e documentação, priorizando rastreabilidade das alterações, reprodutibilidade dos testes e registro das decisões técnicas.

2.1 Prova de conceito e coleta de dados dos grupos de pesquisa

A atividade de prova de conceito e coleta de dados foi conduzida mediante identificação das fontes disponíveis, análise da estrutura dos dados, definição de recortes de teste e verificação da consistência dos registros coletados. O procedimento incluiu conferência de campos, padronização de representações, avaliação de dados ausentes ou inconsistentes e organização dos registros de forma adequada às etapas posteriores de consulta, análise e visualização.

Sempre que aplicável, a coleta foi tratada de forma incremental e controlada, evitando processamento desnecessário de todo o conjunto de dados. Essa abordagem permitiu testar fluxos de extração e transformação com volumes reduzidos antes da execução em escala maior, diminuindo risco operacional e facilitando a identificação de falhas de configuração ou de modelagem.

2.2 Manutenção de dashboards interativos com o Kibana

A manutenção dos dashboards foi realizada a partir da verificação das fontes indexadas, dos data views utilizados, dos filtros e agregações aplicados e da coerência das visualizações com os dados disponíveis. O trabalho contemplou ajustes necessários à continuidade do uso dos painéis, revisão de consultas, organização de indicadores e

validação do comportamento das visualizações após alterações na origem ou na estrutura dos dados.

Também foram considerados aspectos de desempenho e legibilidade dos painéis, buscando evitar consultas excessivamente custosas e preservar a utilidade das visualizações como instrumento de acompanhamento. As intervenções foram realizadas de forma compatível com a infraestrutura de indexação existente e com as permissões de acesso definidas para os ambientes.

2.3 Estudo aprimorado da infraestrutura da stack ELK

O estudo da infraestrutura ELK combinou análise arquitetural e verificação prática dos componentes Elasticsearch, Kibana e Logstash. Foram observados elementos relacionados à composição e aos papéis dos nós, utilização de memória e heap da JVM, distribuição de shards e segmentos, rotinas de ingestão, mappings e templates, mecanismos de segurança, certificados TLS, autenticação e autorização, além de procedimentos de backup e recuperação por snapshots.

No componente de ingestão, foram revisados fluxos e configurações com atenção à paginação, processamento em lote, variáveis de ambiente, persistência de estado e comportamento em falhas. No Elasticsearch, a análise considerou consistência dos índices, configuração de mappings, organização de shards e réplicas, avaliação de segmentos e impacto do dimensionamento de recursos. No Kibana, além dos dashboards, foram considerados aspectos de acesso, integração com o cluster e continuidade operacional.

2.4 Produção de relatórios

A produção de relatórios ocorreu de forma contínua, reunindo registros das atividades executadas, procedimentos operacionais, decisões de configuração, resultados de testes e pontos de atenção identificados durante a manutenção dos ambientes. A documentação foi utilizada como mecanismo de rastreabilidade e como apoio à continuidade do trabalho, reduzindo dependência de conhecimento exclusivamente tácito.

3 RESULTADOS

Os resultados são apresentados de acordo com as quatro atividades previstas no formulário, de modo a manter a correspondência entre planejamento e execução. Não são atribuídos resultados quantitativos quando não havia, no material de acompanhamento, uma medição consolidada que permitisse apresentar valores verificáveis.

Quadro 1 - Correspondência entre o plano de atividades e a execução relatada

Atividade prevista na planilha	**Execução relatada**	**Seção no relatório**
Prova de conceito e coleta de dados dos grupos de pesquisa	Coleta, validação, preparação de dados e testes controlados de fluxos.	2.1 e 3.1
Manutenção de dashboards interativos com o Kibana	Revisão de fontes, consultas, filtros, agregações e continuidade das visualizações.	2.2 e 3.2
Estudo aprimorado da infraestrutura da stack ELK	Análise de Elasticsearch, Kibana e Logstash, incluindo recursos, shards, segmentos, ingestão, segurança e snapshots.	2.3 e 3.3
Produção de relatórios	Registro de procedimentos, decisões técnicas, testes, diagnósticos e recomendações.	2.4 e 3.3

Fonte: autoria própria (2026).

3.1 Prova de conceito e coleta de dados dos grupos de pesquisa

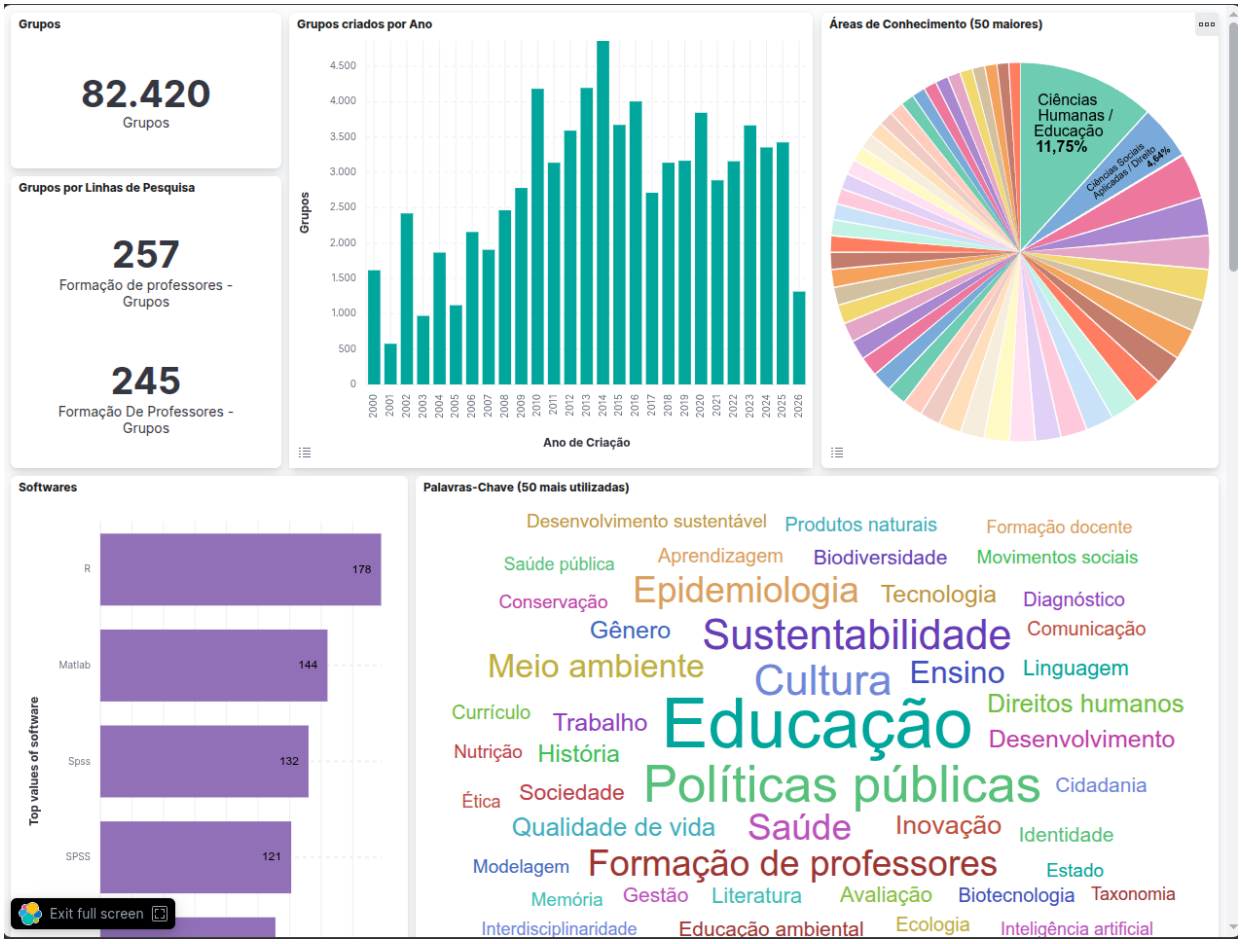
A atividade resultou na consolidação de procedimentos de coleta e validação aplicáveis aos dados dos grupos de pesquisa, com organização de etapas de extração, verificação estrutural e preparação dos registros para consumo analítico. A adoção de testes controlados permitiu avaliar o comportamento dos fluxos antes de sua utilização em volumes maiores e registrar ajustes necessários para aumentar previsibilidade e reprodutibilidade.

Como resultado técnico, foram fortalecidos os procedimentos de conferência de dados e de preparação de fontes para indexação e visualização. Esse trabalho também forneceu subsídios para a manutenção dos dashboards e para o estudo da infraestrutura ELK, uma vez que problemas de modelagem, tipagem ou consistência na origem impactam diretamente a qualidade das consultas e visualizações.

3.2 Manutenção de dashboards interativos com o Kibana

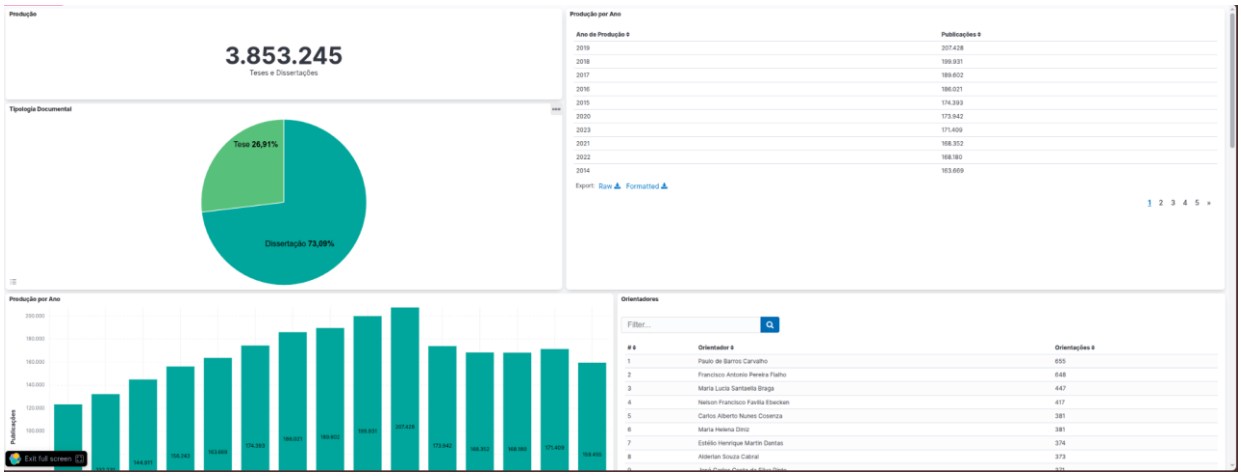
A manutenção dos dashboards preservou a disponibilidade e a coerência das visualizações utilizadas para acompanhamento dos dados. Foram revisados elementos de consulta, filtros, agregações e fontes indexadas, com ajustes associados às alterações nos dados e à infraestrutura. A atividade também contribuiu para identificar pontos em que a qualidade do dado ou a configuração do índice interferia no resultado apresentado pelo Kibana. Evidências do uso e da manutenção dos painéis analíticos estão apresentadas na Figura 1.

Figura 1 - Dashboard interativo no Kibana para análise de grupos de pesquisa.



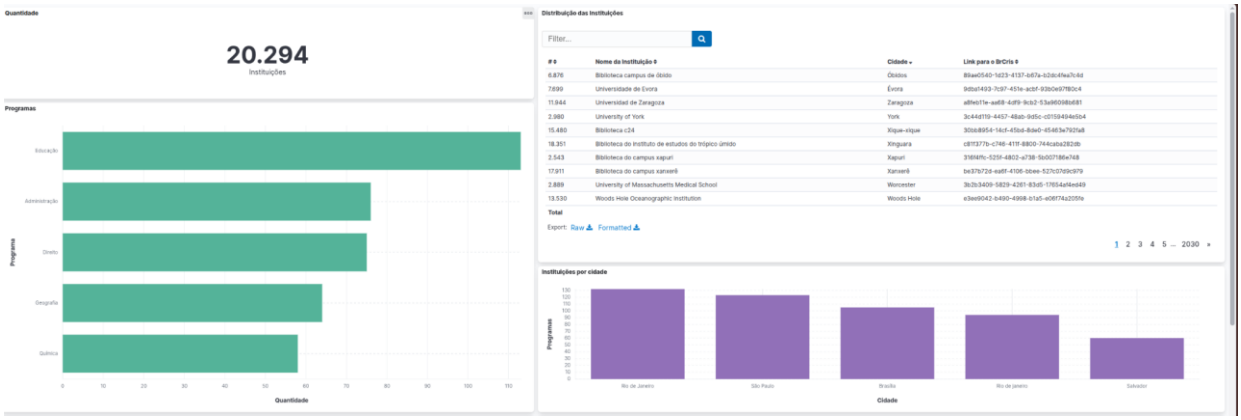
Fonte: autoria própria (2026).

Figura 5 - Dashboard com indicadores e visualizações sobre produção de teses e dissertações.



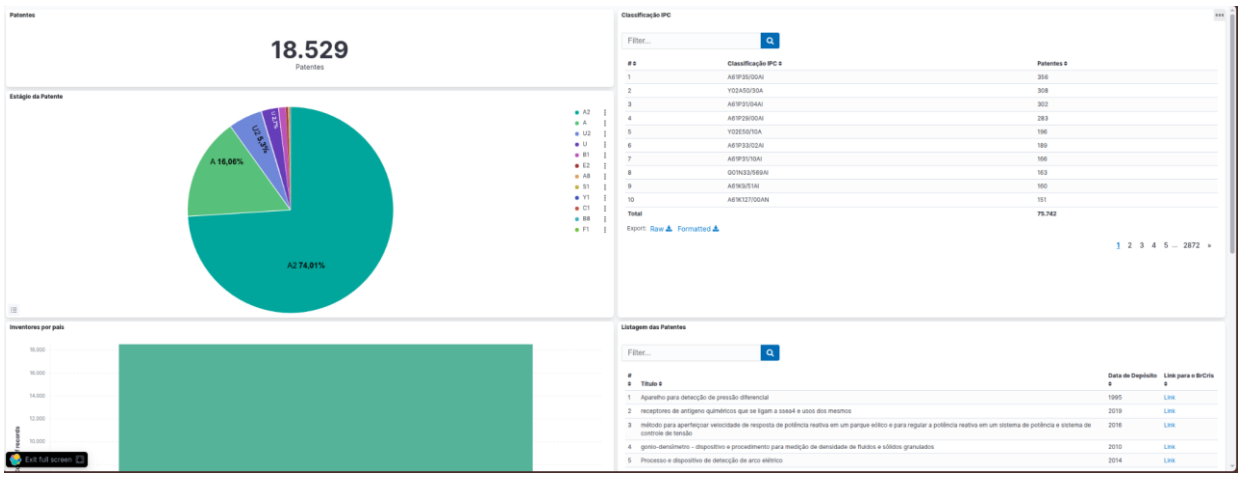
Fonte: autoria própria (2026).

Figura 6 - Dashboard com indicadores e distribuição de instituições registradas na base analisada.



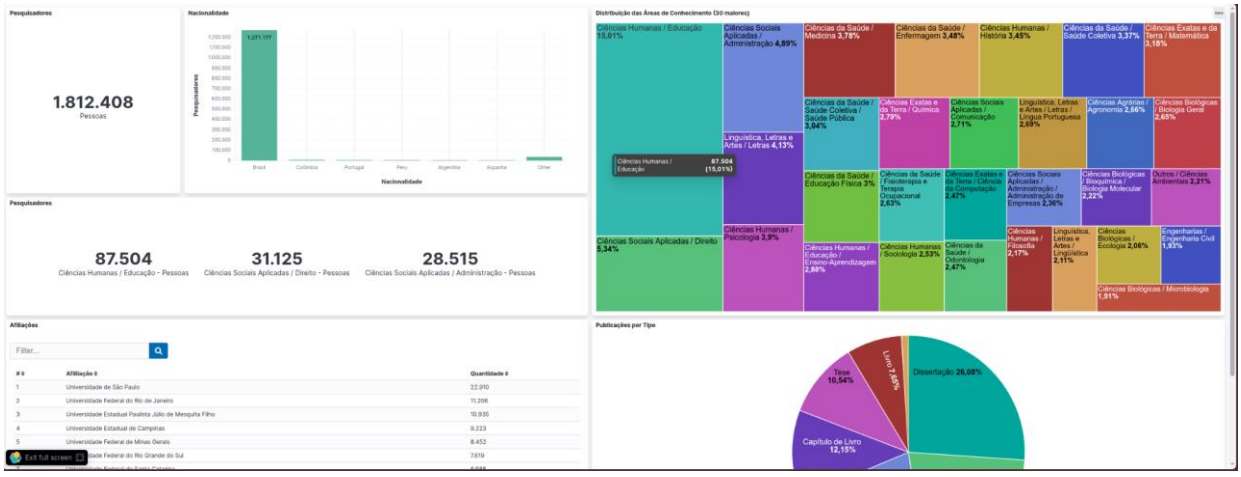
Fonte: autoria própria (2026).

Figura 7 - Dashboard com indicadores e visualizações sobre patentes, estágios e classificações associadas.



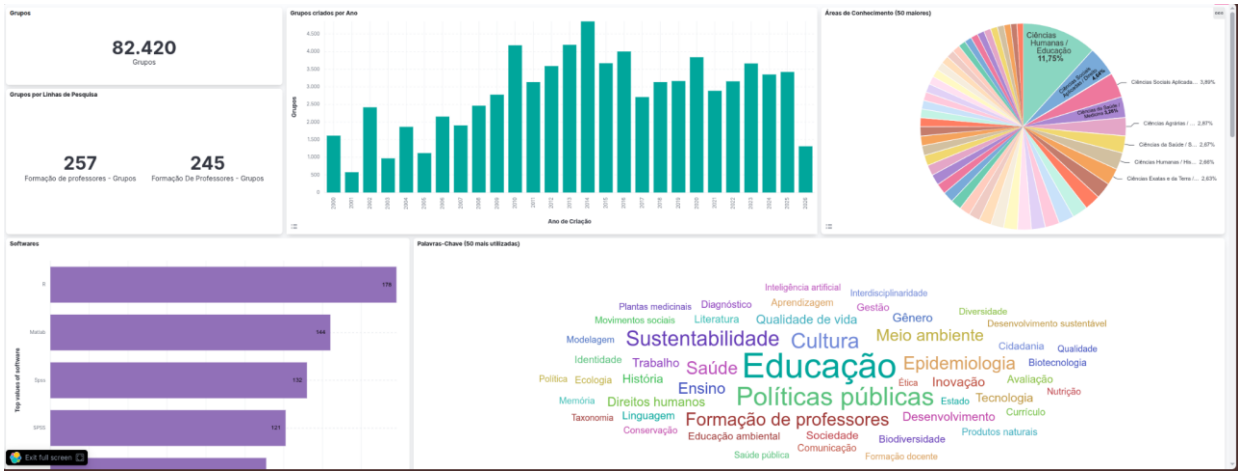
Fonte: autoria própria (2026).

Figura 8 - Dashboard com indicadores sobre pesquisadores, nacionalidade, áreas de conhecimento e afiliações.



Fonte: autoria própria (2026).

Figura 9 - Dashboard com indicadores e visualizações sobre grupos de pesquisa, áreas de conhecimento, softwares e palavras-chave.



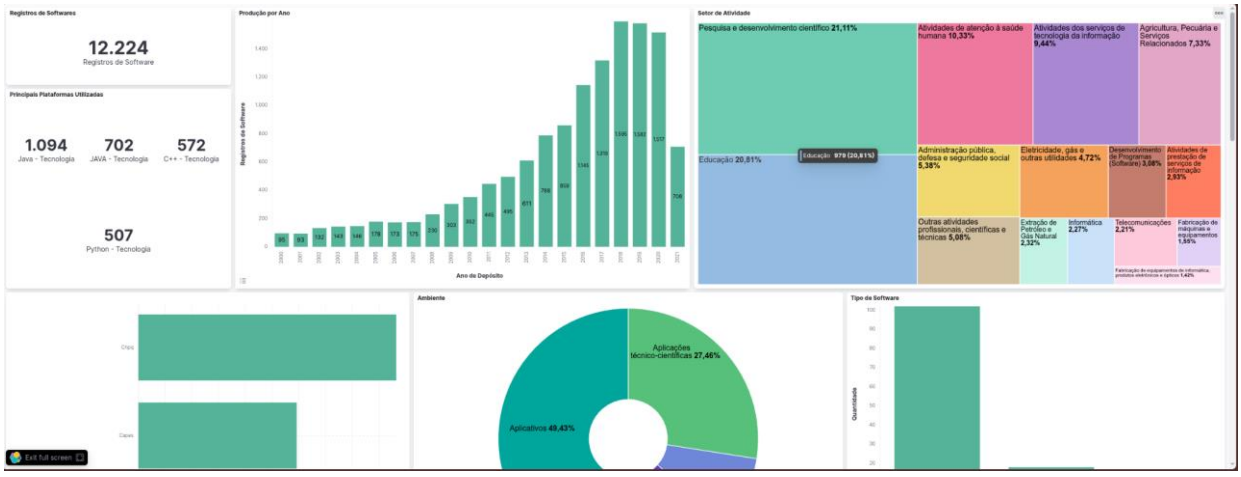
Fonte: autoria própria (2026).

Figura 10 - Dashboard com indicadores e visualizações sobre produção científica, autores e palavras-chave.



Fonte: autoria própria (2026).

Figura 11 - Dashboard com indicadores e visualizações sobre registros de software, plataformas, setores e tipos de software.



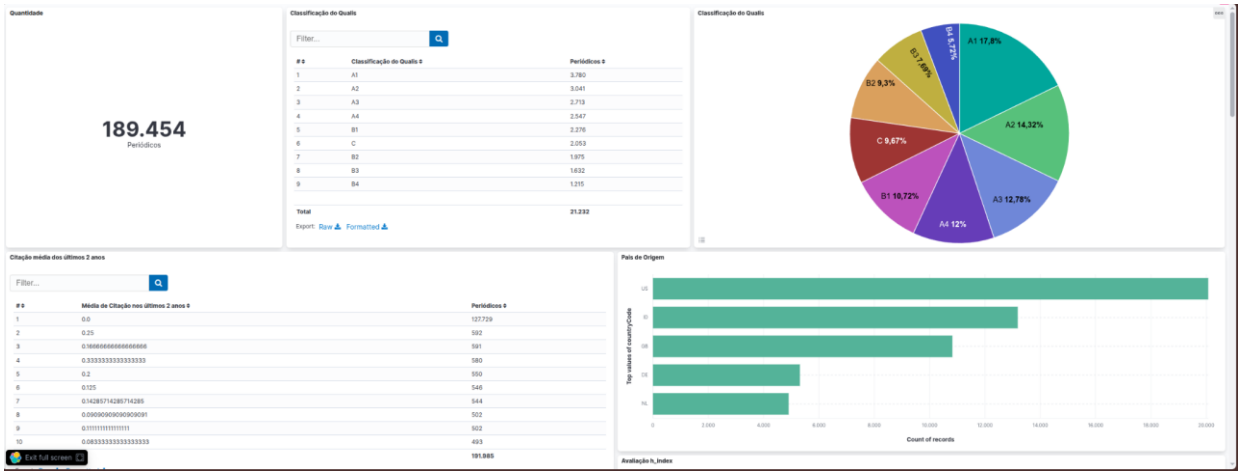
Fonte: autoria própria (2026).

Figura 12 - Dashboard com indicadores e visualizações sobre programas, áreas do conhecimento e instituições.



Fonte: autoria própria (2026).

Figura 13 - Dashboard com indicadores e visualizações sobre periódicos, classificação Qualis, citações e país de origem.



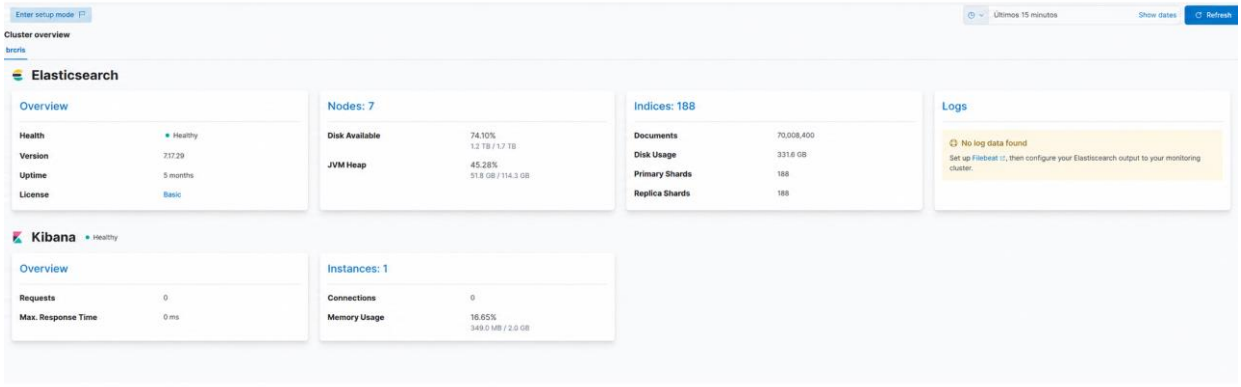
Fonte: autoria própria (2026).

3.3 Estudo aprimorado da infraestrutura ELK e produção de relatórios

O estudo da infraestrutura ampliou a compreensão operacional da stack ELK e apoiou decisões de manutenção e evolução. Entre os aspectos analisados estiveram dimensionamento de recursos, papéis de nós, comportamento do heap, organização de shards e segmentos, mappings e templates, ingestão, segurança, certificados e snapshots. A análise foi acompanhada de documentação técnica, permitindo registrar procedimentos e recomendações para uso posterior. Evidências sintéticas do acompanhamento do cluster e de seus nós são apresentadas nas Figuras A3 e A4.

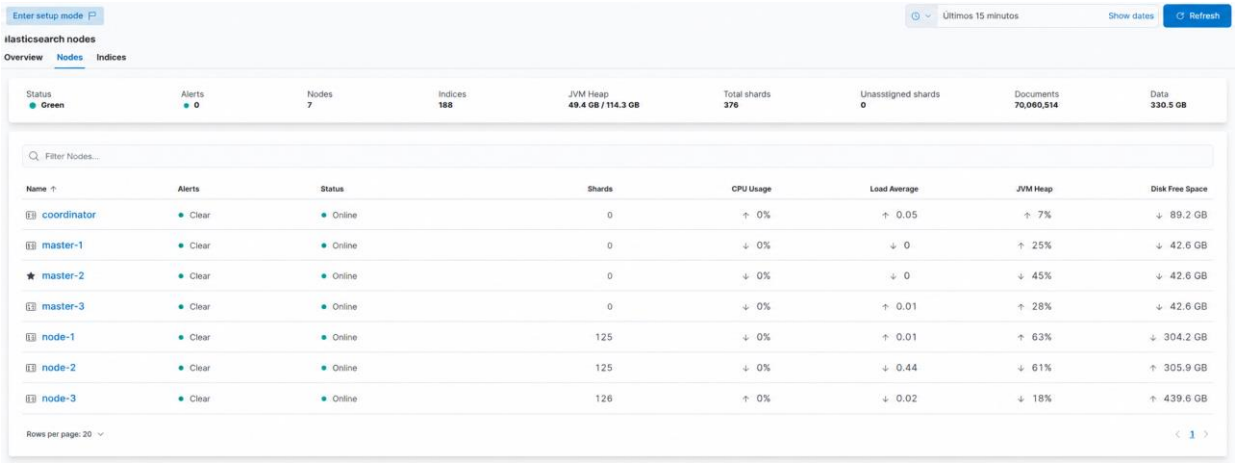
A produção de relatórios e registros técnicos consolidou as alterações realizadas e os procedimentos validados, incluindo orientações de operação, diagnóstico e manutenção. Esse resultado atende diretamente à atividade de produção de relatórios prevista no plano e contribui para a rastreabilidade das ações executadas durante a vigência.

Figura 3 - Visão geral do monitoramento do cluster Elasticsearch e Kibana no contexto do BrCris.



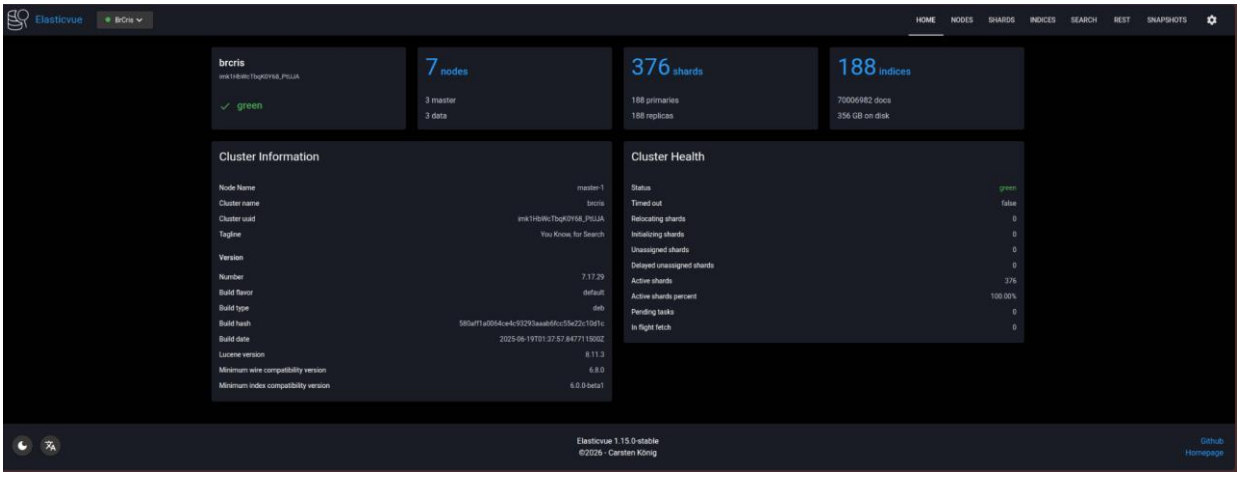
Fonte: autoria própria (2026).

Figura 4 - Acompanhamento dos nós do cluster Elasticsearch, com destaque para shards, uso de heap e espaço em disco.



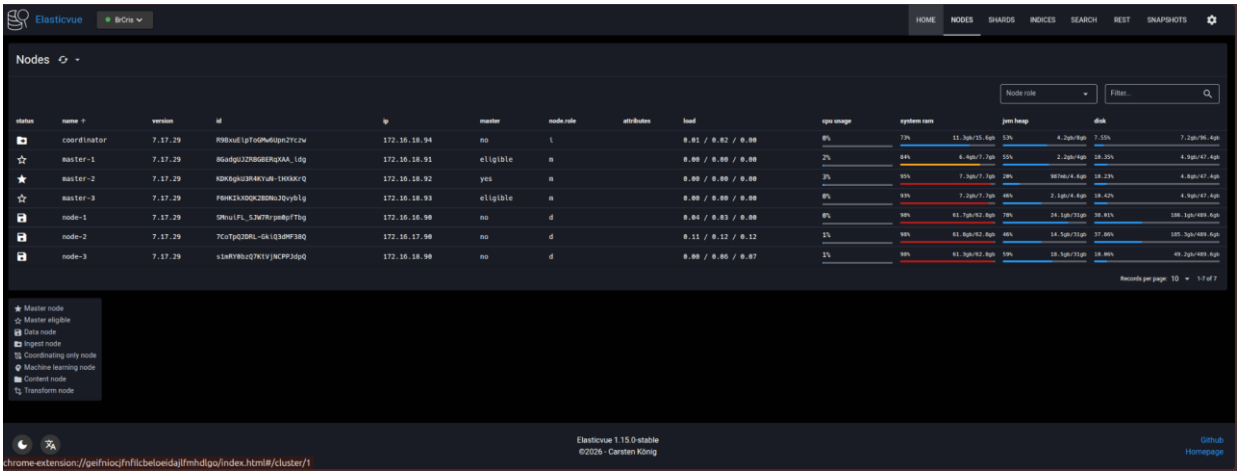
Fonte: autoria própria (2026).

Figura 14 - Visão geral do cluster Elasticsearch no Elasticvue no ambiente BrCris.



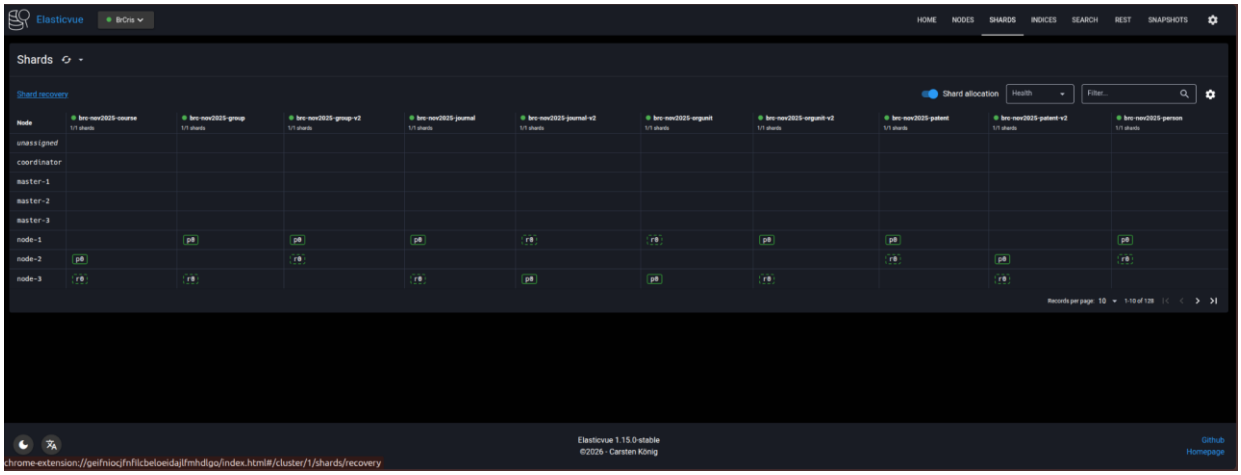
Fonte: autoria própria (2026).

Figura 15 - Painel de nós do cluster Elasticsearch no Elasticvue no ambiente BrCris.



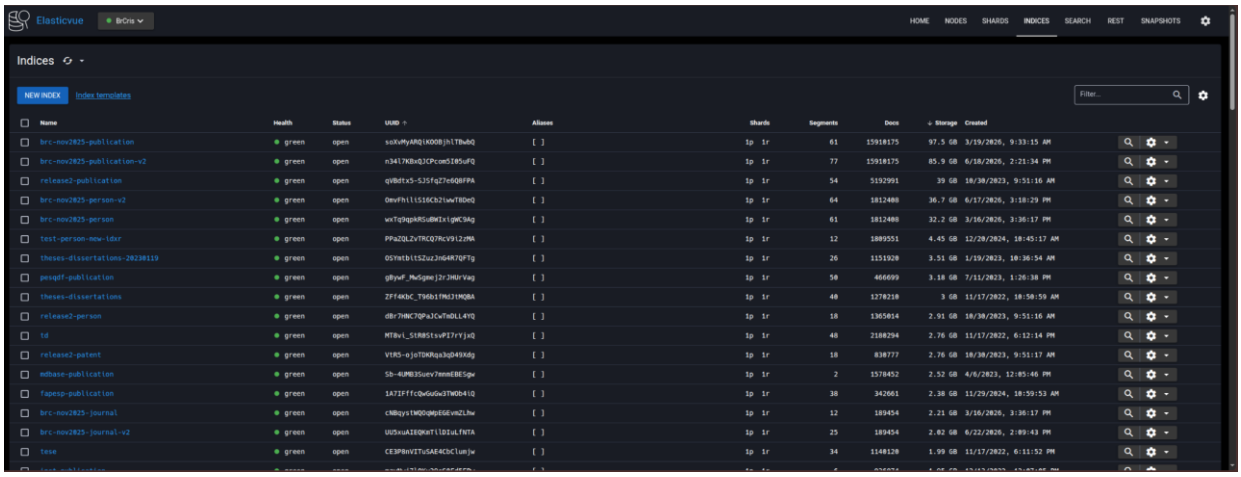
Fonte: autoria própria (2026).

Figura 16 - Visualização da distribuição de shards no Elasticvue no ambiente BrCris.



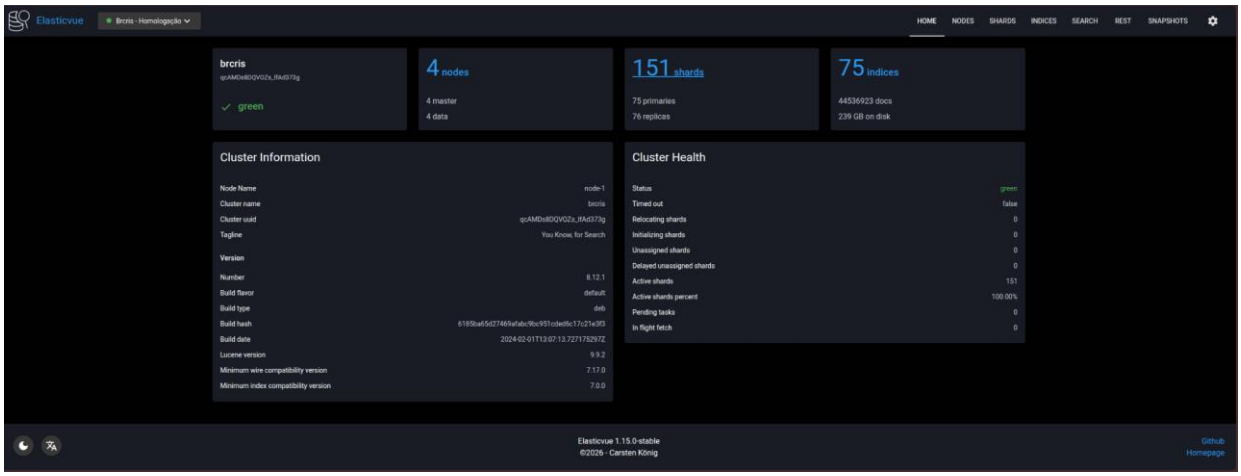
Fonte: autoria própria (2026).

Figura 17 - Listagem de índices do cluster no Elasticvue no ambiente BrCris.



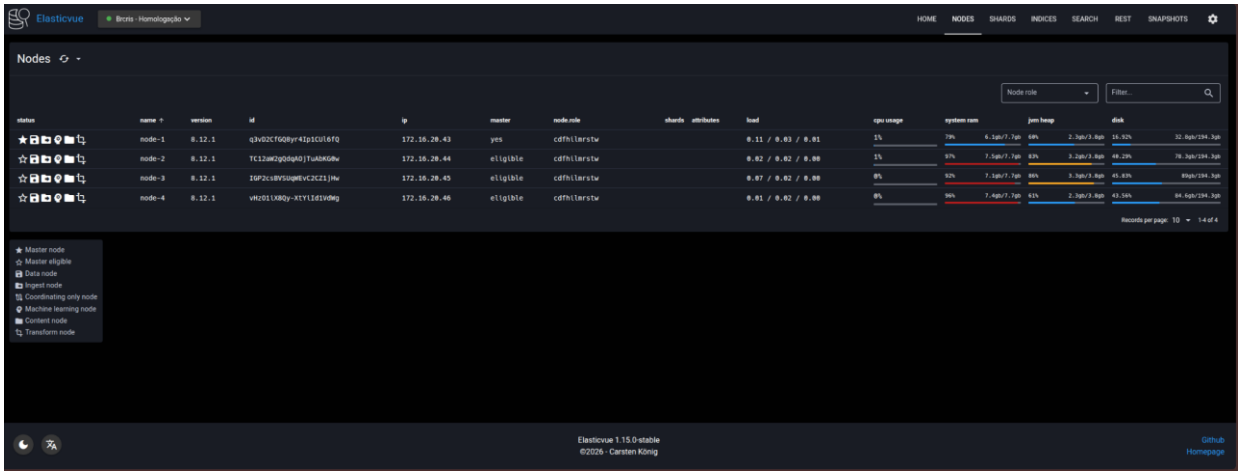
Fonte: autoria própria (2026).

Figura 18 - Visão geral do cluster Elasticsearch no Elasticvue no ambiente BrCris - Homologação.



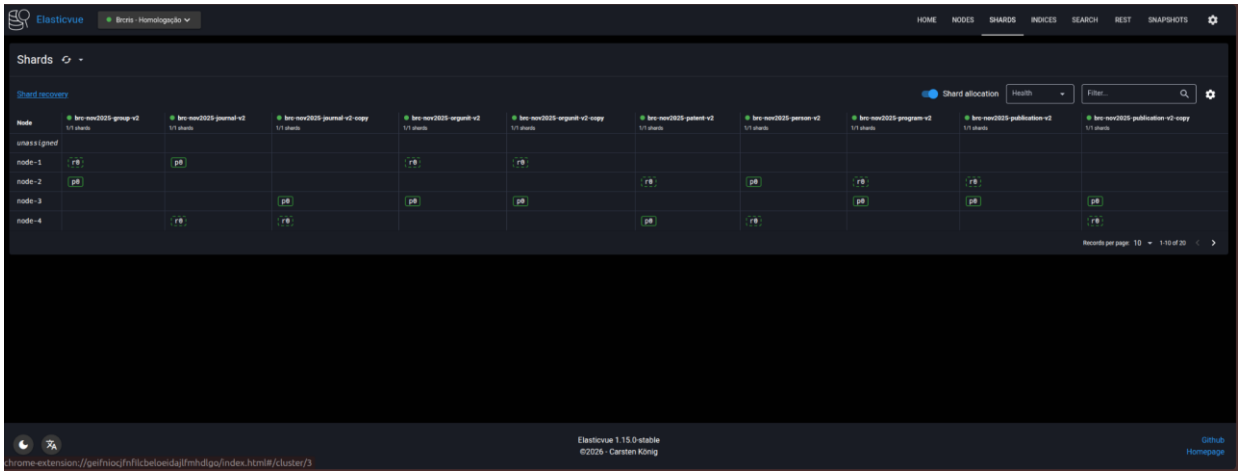
Fonte: autoria própria (2026).

Figura 19 - Painel de nós do cluster Elasticsearch no Elasticvue no ambiente BrCris - Homologação.



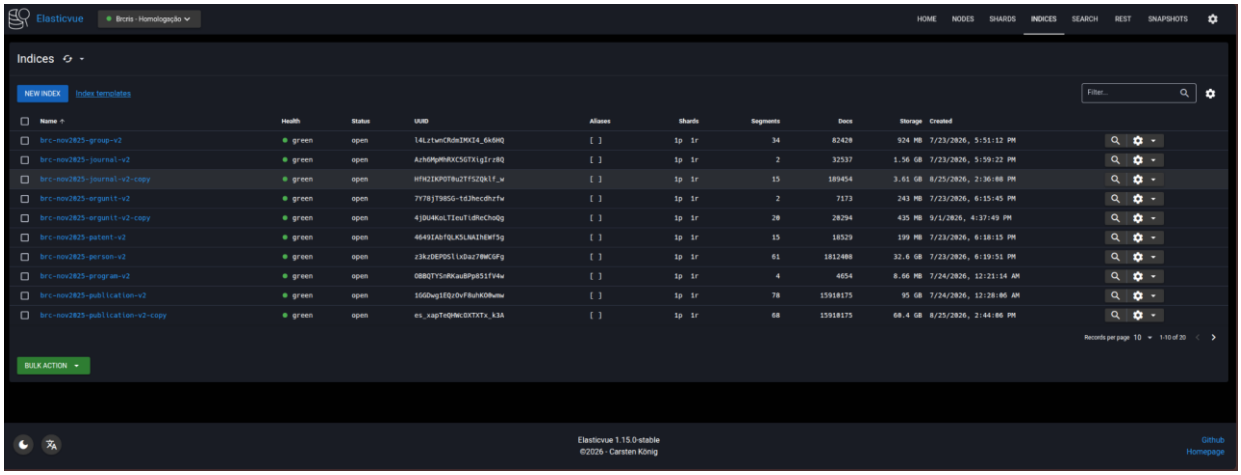
Fonte: autoria própria (2026).

Figura 20 - Visualização da distribuição de shards no Elasticvue no ambiente BrCris - Homologação.



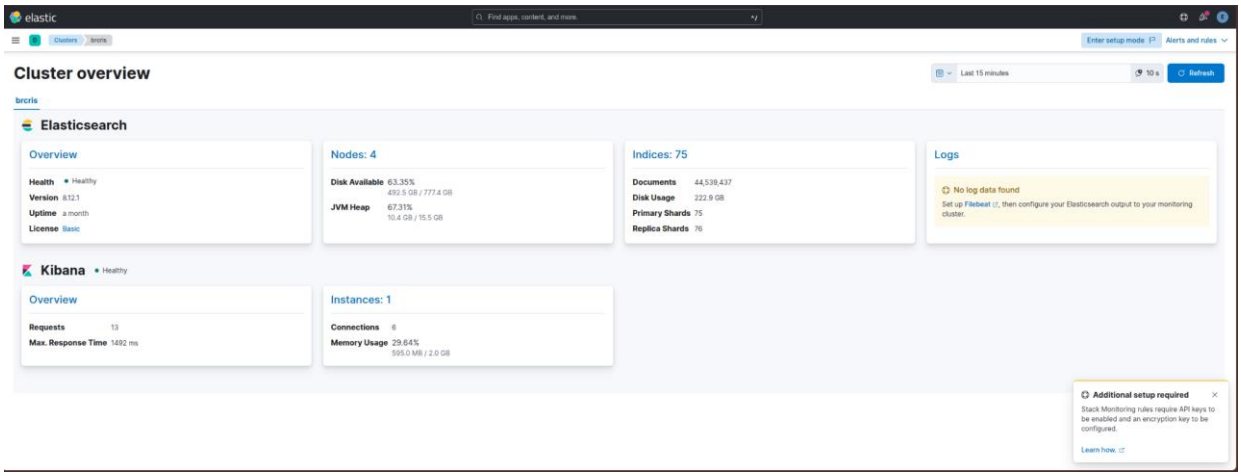
Fonte: autoria própria (2026).

Figura 21 - Listagem de índices do cluster no Elasticvue no ambiente BrCris - Homologação.



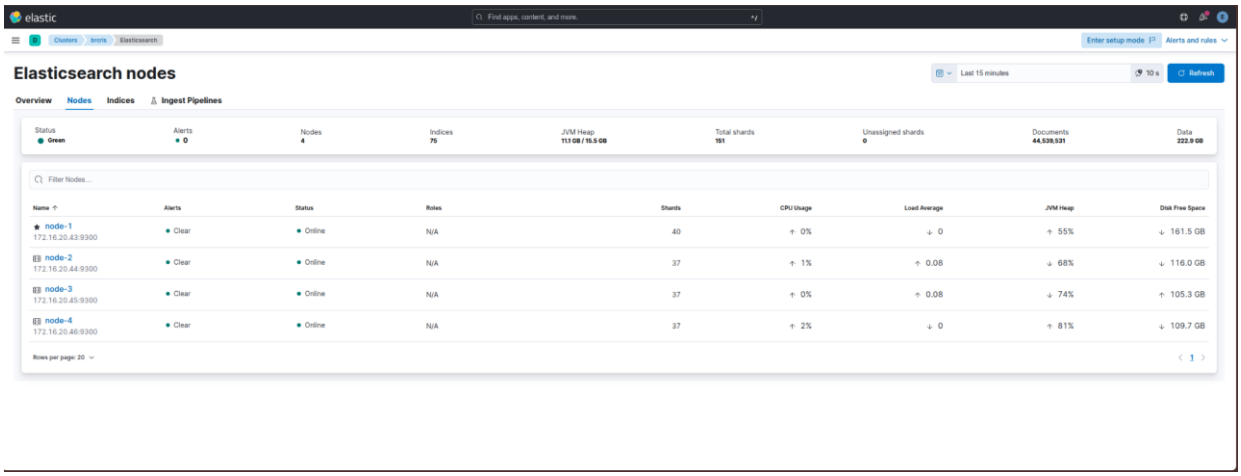
Fonte: autoria própria (2026).

Figura 22 - Visão geral do monitoramento do cluster no Stack Monitoring do Elastic.



Fonte: autoria própria (2026).

Figura 23 - Painel de nós do cluster no Stack Monitoring do Elastic.



Fonte: autoria própria (2026).

3.4 Ampliação do escopo da pesquisa - atividades correlacionadas no INPI e no BrCris

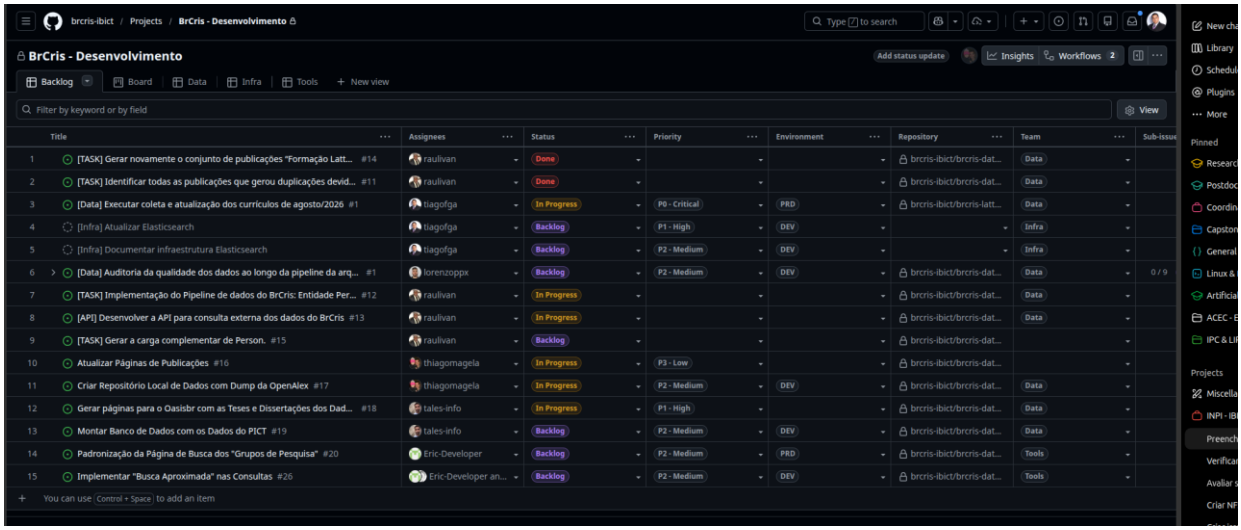
Além das atividades diretamente previstas, foram desenvolvidas tarefas correlacionadas em ambientes do INPI e do BrCris. Essas tarefas foram mantidas separadas nesta seção porque pertencem a contextos institucionais e operacionais distintos, mas apresentam aderência técnica às atividades da bolsa por envolverem os mesmos domínios de coleta e tratamento de dados, indexação, dashboards, infraestrutura ELK, segurança e documentação.

No contexto do BrCris, foram realizadas atividades de manutenção e evolução da infraestrutura ELK, com análise da topologia do cluster, papéis dos nós, dimensionamento de heap, comportamento de segmentos, ingestão e continuidade operacional dos serviços. Também foram tratados procedimentos associados a snapshots em NFS e à organização da infraestrutura necessária para recuperação e manutenção dos dados. Essas ações funcionaram como um ambiente prático adicional para validar procedimentos de administração e evolução da stack. Como apoio à governança técnica e à organização do trabalho, também foi mantida a estrutura das equipes e do acompanhamento das atividades no GitHub, conforme evidenciado na Figura 2.

No contexto do INPI, foram executadas tarefas correlacionadas de manutenção e evolução de ambientes de dados e busca, incluindo revisão de ingestão, atualização e validação de mappings e índices, tratamento de certificados TLS, criação e revisão de usuários, papéis e chaves de API, rotinas de snapshots e diagnóstico de componentes Elasticsearch e Kibana. Também foram analisados fluxos de carga e transformação de dados destinados à indexação.

A relação dessas atividades com o plano da bolsa é técnica e metodológica: os conhecimentos obtidos na operação de ambientes reais de Elasticsearch, Kibana e Logstash reforçaram procedimentos de coleta, validação, manutenção de dashboards, análise de infraestrutura e documentação. Assim, as tarefas correlacionadas complementaram o trabalho previsto sem descaracterizar as quatro atividades registradas na planilha.

Figura 2 - Organização das equipes técnicas e dos repositórios no GitHub do projeto.



Fonte: autoria própria (2026).

4 SÍNTESE DO ATENDIMENTO AO PLANO DE ATIVIDADES

As quatro atividades registradas no formulário do bolsista foram contempladas durante a vigência. A prova de conceito e a coleta de dados foram tratadas por meio de procedimentos de extração, validação e preparação; os dashboards do Kibana foram mantidos e revisados; a infraestrutura ELK foi estudada e acompanhada em seus principais aspectos operacionais; e os procedimentos, resultados e decisões técnicas foram consolidados em relatórios e documentação. As atividades do INPI e do BrCris foram registradas separadamente como tarefas correlacionadas e de ampliação de escopo.

REFERÊNCIAS

ELASTIC. Elasticsearch reference. [S. l.], 2026. Disponível em: https://www.elastic.co/docs/reference/elasticsearch. Acesso em: 4 set. 2026.

ELASTIC. Kibana. [S. l.], 2026. Disponível em: https://www.elastic.co/docs/reference/kibana. Acesso em: 4 set. 2026.

ELASTIC. Logstash. [S. l.], 2026. Disponível em: https://www.elastic.co/docs/reference/logstash. Acesso em: 4 set. 2026.

Documento assinado digitalmente
gov.br
TIAGO ALVES DE OLIVEIRA
Data: 23/09/2026 22:51:31-0300
Verifique em https://validar.iti.gov.br

Tiago Alves de Oliveira
Pesquisador

De acordo:

Documento assinado digitalmente
gov.br
WASHINGTON LUIS RIBEIRO DE CARVALHO SEG
Data: 28/09/2026 17:47:56-0300
Verifique em https://validar.iti.gov.br

Washington Luís Ribeiro de Carvalho Segundo
Coordenador do projeto