

GRUPO DE ESTUDO DE OPERAÇÃO DE SISTEMAS ELÉTRICOS - GOP

MOTOR DE BUSCA PARA SALAS DE CONTROLE - O “GOOGLE” DA OPERAÇÃO

**DIEGO JOSÉ FRAGA(1);FILIPE VENTURA MUGGIATI(1);FELIPE TREVISAN(1);BRUNO
AUGUSTO BASTIANI(1)
ITAIPU (1)**

RESUMO

Este trabalho apresenta uma ferramenta de busca, estilo “Google”, desenvolvida pela área de Operação da Usina Hidrelétrica de Itaipu para pesquisas não estruturadas em documentos de diversos sistemas corporativos utilizados cotidianamente por equipes de operação em tempo real e de pré/pós-operação. Atualmente a ferramenta conta com mais de um milhão de itens indexados, realizando buscas em tempos da ordem de milissegundos. A solução foi desenvolvida totalmente com recursos internos, utilizando tecnologias/ferramentas “open source”, sem custos de licenciamento e necessidade de aquisição de hardware dedicado. Destaca-se que a iniciativa pode ser facilmente replicada para uso nas demais áreas da organização.

PALAVRAS-CHAVE

Buscador, *Big Data*, Pesquisa não-estruturada, *elasticsearch*, Indústria 4.0

1. INTRODUÇÃO

A sistematização de processos através de aplicações computacionais têm contribuído, de forma significativa, para o aumento no volume de informações armazenadas e utilizadas nos centros de operação do sistema elétrico. Em contrapartida, a grande quantidade de dados e sua distribuição em sistemas distintos dificulta o acesso e a busca dessas informações, tanto para as equipes de operação em tempo real, quanto para as equipes de pré e pós operação, normatização e estudos elétricos.

Considerando apenas as principais instruções operativas, Itaipu possui aproximadamente 300 documentos normativos vigentes, o que equivale a mais de 2000 páginas contendo regras, as quais operadores e despachantes devem levar em conta durante suas atividades nos centros de controle. Há que se contar ainda as inúmeras ordens de serviço abertas e intervenções em andamento, num conjunto de mais de 20.000 ativos da instalação, que afetam diretamente as decisões tomadas pelas equipes de tempo real.

Como essas informações encontram-se armazenadas em diversos sistemas distintos, apenas o tempo necessário para se fazer login e abrir telas específicas de busca em cada uma das ferramentas pode comprometer o tempo de resposta durante uma situação crítica. Além disso, cabe destacar que diversos sistemas não contam com interfaces de busca práticas e eficientes, sendo necessário, muitas vezes, o preenchimento de diversos campos de um formulário para que a busca seja realizada.

O acesso a informações armazenadas em diferentes sistemas ocorre em diversas situações cotidianas da Operação. Exemplos práticos são situações de restabelecimento após perturbações na usina ou sistema interligado, nas quais a dificuldade de acesso e busca por informações essenciais para o processo decisório resultam em um aumento no tempo de resposta, crítico para equipes de operação em tempo real. Diante deste cenário, surgiu a necessidade de uma solução que auxiliasse, de forma eficiente, a busca por informações nos diversos sistemas, centralizando-as em uma única ferramenta.

2. DESENVOLVIMENTO

2.1 Breve histórico do projeto

Em maio de 2019, a equipe de Sistemas de Suporte da Operação (GSS) participou de um evento na área de TI onde foi apresentada uma ferramenta livre (Open Source) para indexação de documentos chamada ElasticSearch. Na época, a equipe do GSS já conhecia o problema das informações descentralizadas, mas ainda não dispunha de uma solução tecnológica para realizar essa integração dos sistemas e as buscas não estruturadas nos diversos tipos de documentos.

Na mesma época, um grupo de estudos estava sendo organizado no Departamento de Operação de Sistema (OPS.DT) para avaliar possíveis tecnologias da chamada “Indústria 4.0” que poderiam ser implementadas para aumentar a produtividade e eficiência do Departamento. Esta tecnologia foi apresentada ao grupo de estudos que, após avaliação, entendeu que se tratava de uma boa oportunidade de melhoria nos processos.

A equipe do GSS iniciou então os trabalhos para desenvolver a integração dos diversos sistemas e formatos de documentos com o ElasticSearch. O plano era indexar todas as fontes de dados, utilizando diversas ferramentas e scripts e desenvolver uma interface Web para consultas pelos usuários finais.

Paralelamente, na divisão de Operação do Sistema (OPSO.DT), em Julho de 2019, o grande número de instruções operativas motivou o início do desenvolvimento de uma ferramenta para busca em arquivos PDF, que possibilitaria a despachantes a realização de buscas nas instruções de operação. A ferramenta desenvolvida na ocasião foi inspirada numa ferramenta de Furnas chamada “SOS” (Suporte à Operação de Sistemas Elétricos).

Os esforços dessas iniciativas deram origem ao desenvolvimento da ferramenta “Buscador da Operação”.

2.2 Arquitetura da solução

Inicialmente foi desenvolvida uma versão simplificada da ferramenta para testes preliminares, contando com poucos tipos de documentos. Essa versão foi apresentada às várias áreas da Superintendência da Operação, as quais indicaram usuários-chave para a realização de testes/validações iniciais. Nessa etapa, diversas sugestões de melhoria foram elencadas: indexação de novos tipos de documentos, mudanças na forma de ordenação de resultados, criação de filtros avançados, etc. Essa etapa foi fundamental para a qualidade da solução. As melhorias propostas foram implementadas na solução final agregando valor ao produto, tornando a ferramenta de buscas mais robusta, capaz de gerar resultados mais expressivos e melhorando a experiência dos usuários.

Desejava-se chegar a uma solução de pesquisas similar ao “Google”, capaz de indexar, buscar e disponibilizar acesso de maneira centralizada (através de poucos cliques), a informações de diferentes sistemas.

A arquitetura implantada para a solução é composta pelos seguintes módulos/subsistemas:

- **Motor de Busca:** mecanismo que faz o armazenamento dos dados indexados bem como realiza as buscas propriamente ditas. Trata-se do principal componente do sistema e utiliza a plataforma de código aberto *Elasticsearch*, a qual implementa de maneira nativa diversas funções de tratamento de dados não-estruturados, como *tokenization*, processamento de *stop-words* e pesquisas com *fuzziness*, que garantem resultados nas buscas mesmo que o usuário cometa erros de digitação ou entre com termos incompletos. O Elasticsearch permite realizar e combinar muitos tipos de buscas — estruturadas, não estruturadas, geográficas, métricas (1).
- **Indexador:** conjunto de *scripts* desenvolvidos em linguagem Python responsáveis por fazer o acesso e processamento das várias fontes de dados (arquivos PDF, tabelas em bases de dados) e por armazenar esses dados pré-processados no motor de busca em formato adequado para pesquisas posteriores.
- **Frontend:** interface web da ferramenta, desenvolvida utilizando o framework *React.js*, de modo que a interface seja intuitiva e permita a integração com demais sistemas corporativos, possibilitando o usuário de acessar documentos de outros sistemas com apenas um clique. O React permite a criação de componenetes encapsulados que gerenciam seus estados, podendo-se combinar vários componentes obtendo UIs mais complexas (2).
- **Gestão de acesso:** subsistema responsável por fazer a gestão de acesso de grupos de usuários ao buscador possibilitando o controle de permissão com granularidade por tipo de documento. Este módulo garante que o buscador faça pesquisas apenas nos documentos para os quais os usuários tenham acesso concedido.

A FIGURA 1 ilustra a arquitetura da solução destacando as principais tecnologias/ferramentas envolvidas no projeto, todas de característica “Open Source”: *Python*, *pdfToText*, *Elasticsearch*, *Docker*, *Apache/Tomcat*, *Node.js* e *React*.

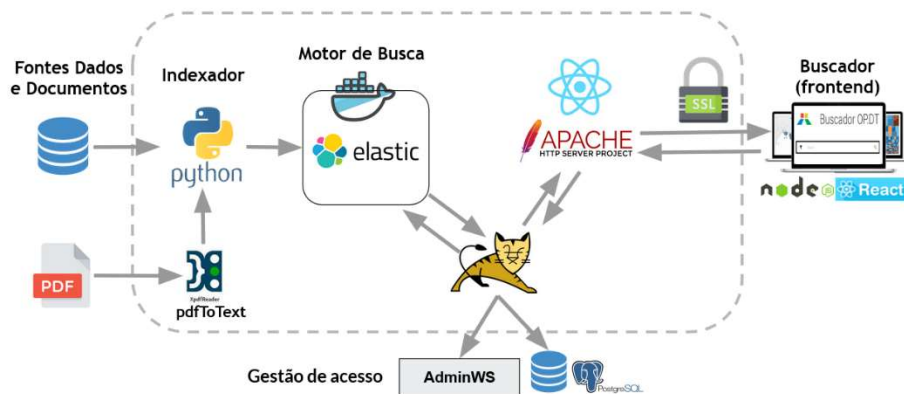


FIGURA 1 - Arquitetura de componentes da solução implantada.

3. RESULTADOS

Previamente ao desenvolvimento do Buscador da Operação, para se chegar a uma informação os Despachantes, Operadores e Engenheiros precisavam fazer a pesquisa em vários sistemas distintos. Cada sistema, em geral, possui uma tela de busca distinta, que muitas vezes não permite a realização de pesquisas não estruturadas, restringindo a busca a códigos internos ou a “textos exatos”. Além disso, em alguns casos, formulários com inúmeros campos precisavam ser preenchidos antes de se disparar uma pesquisa, cenário não aderente à realidade das equipes de operação em tempo real, que exigem a tomada de decisões críticas em poucos minutos.

Em contrapartida, o Buscador da Operação dispõe de uma interface moderna de pesquisa, similar ao “Google”, capaz de apresentar uma lista com resultados relevantes para pesquisas não estruturadas em milissegundos, considerando um volume de dados de mais de um milhão de itens indexados, e deixam o usuário a um clique de distância das principais informações relacionadas aos termos buscados.

Atualmente, o buscador conta com o seguinte volume de dados indexados: instruções operativas (744 documentos/6533 páginas), ordens de serviço (569.964), autorizações de trabalho (93.629), pedidos de desligamento (3.029), relatos de ocorrências de tempo real (574.081), e relatórios de perturbação e eventos (1.043 documentos/14.303 páginas), totalizando, aproximadamente, 1.261.701 itens indexados, ou *documents* - nomenclatura usada pelo motor de busca *Elasticsearch* (documentos textuais como instruções e relatórios foram indexados página a página).

Na sequência são apresentados exemplos de interfaces de busca de alguns sistemas corporativos (usados “antes” do buscador) juntamente com a solução de pesquisa integrada disponibilizada pelo “Buscador da Operação”.

3.1 Pesquisas por ordens de serviço

O SMA (Sistema de Manutenção Aperiódica), cuja principal finalidade é o cadastro de Solicitações de Serviço Aperiódico (SSA), não possui integrado a sua interface campos de buscas por texto livre, restringindo a consulta ao histórico de solicitações através de campos previamente conhecidos pelo usuário. Já no Buscador, além da possibilidade de buscar por esses campos, é possível realizar uma busca pela descrição da solicitação ou da descrição do serviço executado (campos textuais não estruturados). A Figura 2 apresenta uma busca realizada no SMA enquanto a Figura 3 ilustra a busca pelos mesmos documentos realizada no Buscador da Operação.

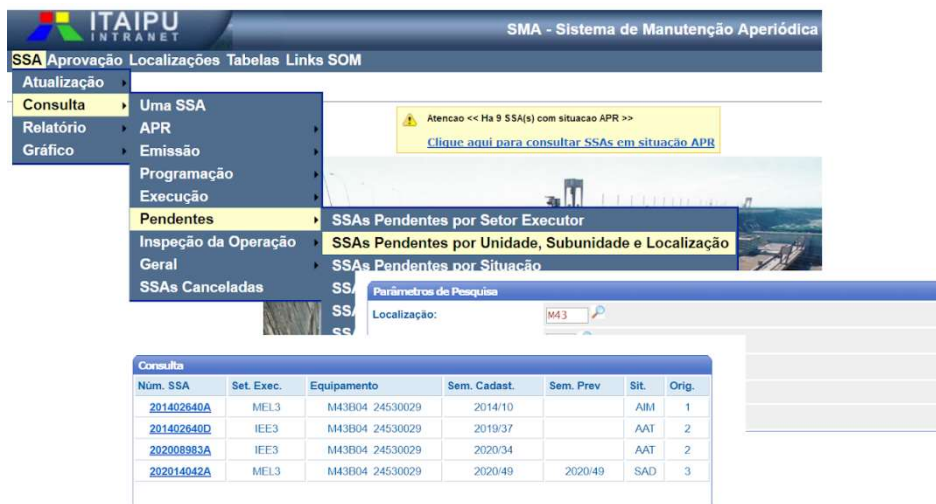


FIGURA 2 - Interface de pesquisa do SMA (Sistema de Manutenção Aperiódica).

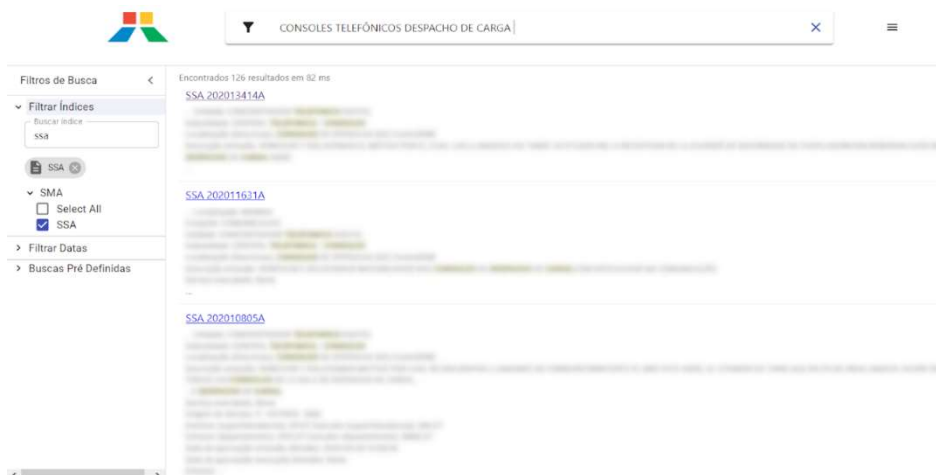


FIGURA 3 - Pesquisa por solicitações de manutenção (SSAs) realizada no Buscador da Operação.

3.2 Pesquisas por documentos normativos

O SNO, sistema que gerencia as instruções de operação, possui um sistema limitado para buscas dentro dos documentos PDF das instruções. Em muitos casos esse tipo de pesquisa acaba não retornando documentos relevantes. A FIGURA 4 apresenta a interface de pesquisa do SNO, demonstrando como é realizada a busca nas instruções de operação, enquanto a FIGURA 5 apresenta a mesma busca sendo realizada no Buscador da Operação.

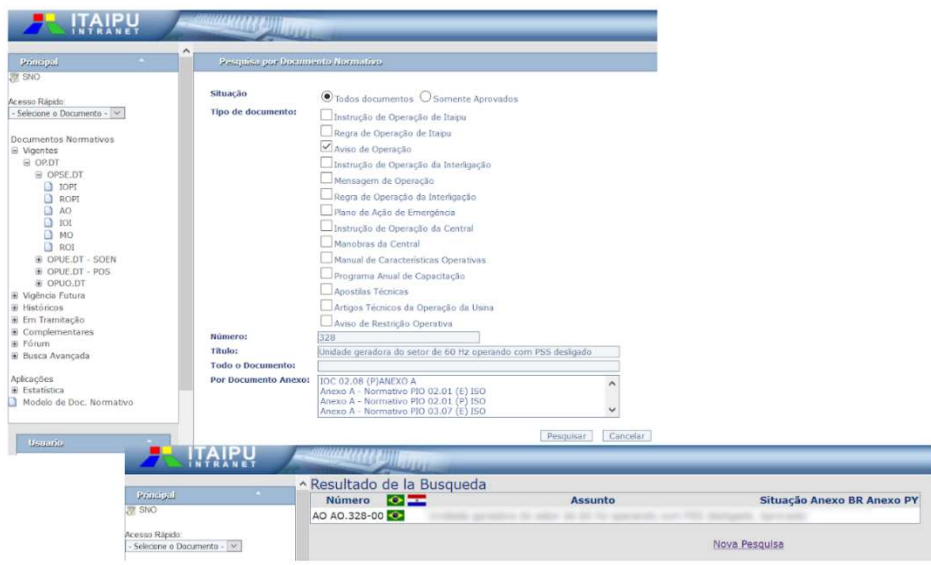


FIGURA 4 - Interface de pesquisa do SNO.

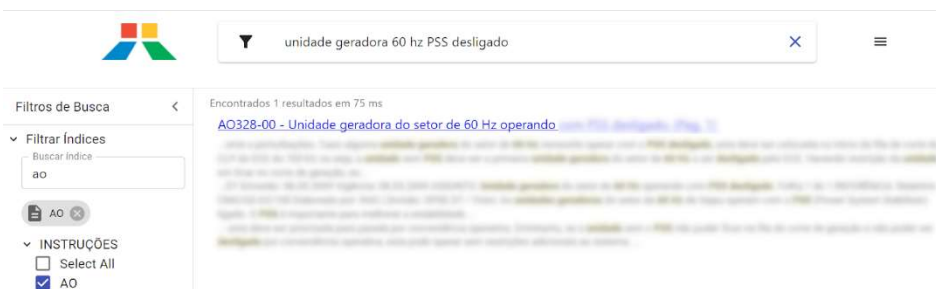


FIGURA 5 - Pesquisa por instruções de operação realizada no Buscador.

3.3 Pesquisas por intervenções em equipamentos

O SES, Sistema de Execução de Serviço, onde são tramitadas Autorizações de Trabalho (ATs) e outros documentos relacionados ao processo de intervenção em equipamentos, permite que o usuário realize consultas baseada em campos estruturados como, número da ATs, código do equipamento, datas, etc. Apesar de possuir um mecanismo de busca eficiente, o SES necessita que o usuário conheça os códigos dos equipamentos ou o número da autorização de trabalho. A FIGURA 6 apresenta a interface de pesquisa por ATs do SES.



FIGURA 6 - Interface de pesquisa do SES (Sistema de Execução de Serviço).

Em muitos casos, informações estão correlacionadas entre diferentes sistemas, por exemplo, uma Autorização de Trabalho (AT) geralmente é vinculada a uma Solicitação de Serviço Aperiódico (SSA), e ambas podem ser mencionadas no Relatório de Ocorrências da Operação ou ainda em Relatórios de Análise de pós-operação para documentação de detalhes adicionais envolvidos durante a solução de um problema. No Buscador da Operação esses vínculos são facilmente encontrados, listados e os documentos originais ficam acessíveis diretamente através de apenas um clique (FIGURA 7).



FIGURA 7 - Pesquisa de diferentes tipos de documentos realizada no Buscador.

4. AVALIAÇÃO DE UTILIZAÇÃO DA FERRAMENTA

Um ano após o início da operação da ferramenta, foi realizado um questionário com os usuários com o objetivo de identificar os padrões de uso bem como avaliar a satisfação com a aplicação. As seguintes perguntas foram feitas aos usuários:

- Em qual divisão trabalha?
- Com qual frequência costuma utilizar o Buscador OP.DT?
- Qual o tipo de informação normalmente pesquisa no buscador?
- Com relação aos resultados apresentados, são condizentes com os termos pesquisados?
- O item buscado normalmente é mostrado na primeira página dos resultados?
- Costuma usar as funções de filtro avançado?
- O tempo levado para execução das pesquisas se mostrou adequado às necessidades de suas atividades diárias?
- As atividades desenvolvidas por você se enquadram em qual das opções a seguir?
- A utilização da ferramenta reduziu o tempo dedicado à pesquisa de informações contidas nos documentos relacionados?
- Deixe aqui sua sugestão de melhoria para a ferramenta.

O questionário foi disponibilizado para 211 usuários, dos quais, 60 enviaram respostas à pesquisa. Na sequência são apresentadas as principais conclusões do levantamento.

Com base nos dados coletados, 35% das respostas sinalizaram o uso diário da ferramenta e 28,33% pelo menos uma vez por semana. O tempo de resposta das pesquisas se mostrou adequado em 81,67% dos casos. Além disso, 66,67% das respostas indicaram que “na maioria das vezes” o Buscador é assertivo na apresentação dos resultados logo na primeira página de resultados.

A opção de filtros avançados da ferramenta (tipos de documentos, datas, buscas pré-definidas) permite refinar as buscas para que somente os tipos de documentos marcados e/ou documentos com datas dentro do intervalo selecionado sejam apresentadas. Verificou-se que esta funcionalidade é, na maioria das vezes, utilizada por usuários com mais experiência com a ferramenta, sendo pouco utilizada ou desconhecida entre os demais usuários, inferindo que os usuários mais recorrentes tendem a filtrar suas buscas de modo a alcançar o documento desejado de forma mais eficiente.

Como desdobramento desse resultado, foram avaliadas melhorias na interface gráfica do Buscador para destacar a funcionalidade de filtragem bem como enviado e-mail informativo aos usuários destacando-se esta funcionalidade.

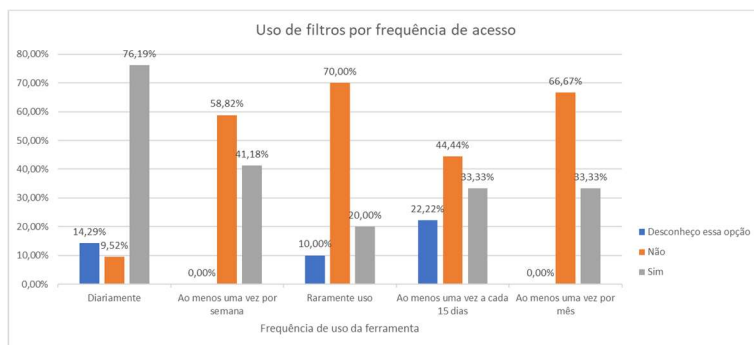


FIGURA 8 – Uso da funcionalidade de filtros avançados.

Quanto ao tipo de informação usualmente pesquisada, cada usuário poderia responder mais de uma opção. Os relatórios da operação em tempo real e instruções de operação foram os tipos de informações mais pesquisadas, conforme pode ser visto no gráfico abaixo (FIGURA 9).

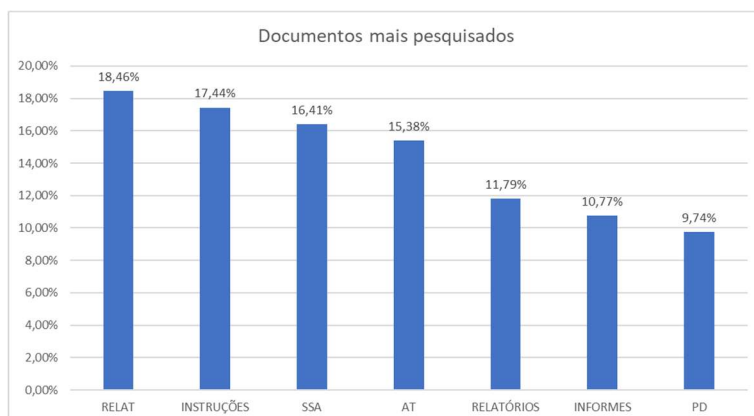


FIGURA 9 - Tipo de informação normalmente pesquisada pelos usuários.

Com relação às sugestões registradas pelos usuários no formulário de avaliação da ferramenta, podem ser destacados os seguintes pontos, os quais estão sendo tratados como melhorias futuras para o Buscador:

- Inclusão de demais documentos/artefatos ao conjunto de documentos indexados pelo Buscador (ex: atas de reunião, memorandos, procedimentos internos das áreas, documentação de projetos, etc).
- Aplicar tokenização e filtros customizados visando melhorar os resultados em ocasiões onde tem-se palavras com sentidos diferente porem de mesmo radical.
- Apresentar sugestões de busca ao usuário baseado seu histórico de utilização.

Por fim, utilizando os dados estatísticos de acesso e consultas feitas ao Buscador, é possível traçar um perfil de utilização da ferramenta. Os gráficos abaixo ilustram essa análise. Obs: os dados de consultas passaram a ser coletados a partir de junho de 2021.

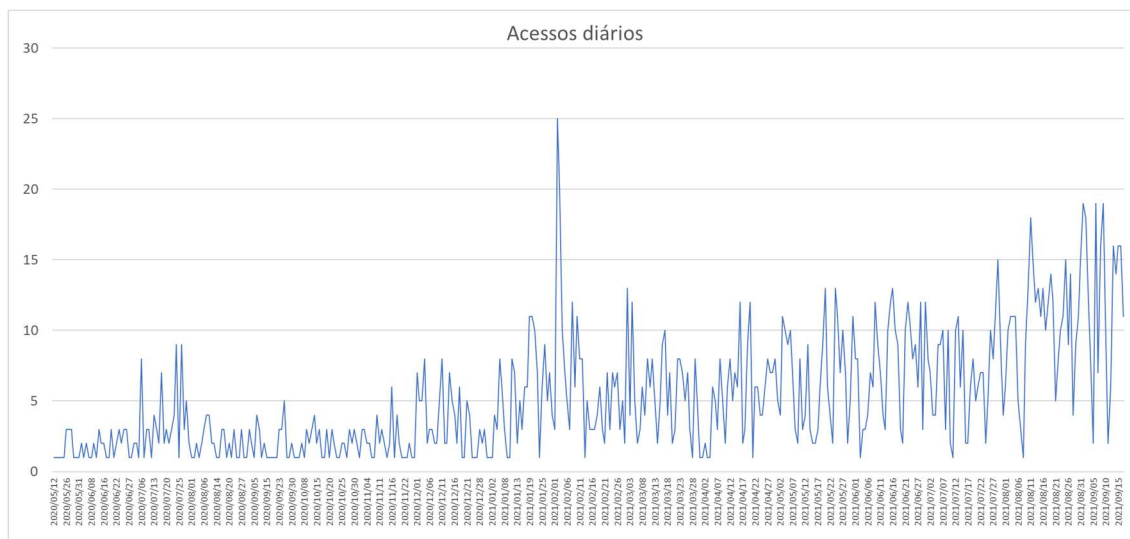


FIGURA 10 - Número de usuários conectados ao Buscador por dia.



FIGURA 11 - Número de consultas realizadas por dia.

5. CONCLUSÕES

A ferramenta desenvolvida possibilita a pesquisa/busca em diversas instruções operativas, facilitando e agilizando a busca por informações para as equipes de tempo real, principalmente em situações de perturbação e restabelecimento, onde o tempo para a tomada de decisões é fundamental.

Na pré-operação, contribui no planejamento de intervenções, facilitando acesso a informações de trabalhos em andamento permitindo análise de possíveis conflitos com trabalhos a serem aprovados. Auxilia também as equipes de pós-operação na elaboração de relatórios de análise e estudos estatísticos, fornecendo rápido acesso aos registros diários de ocorrências documentadas pelos centros de controle.

Cabe destacar que a solução é escalável e está preparada para um contexto de Big Data, permitindo a expansão do volume de dados indexado através da execução do motor de busca em modo cluster realizando-se apenas configurações na plataforma Elasticsearch, sem a necessidade de desenvolvimentos adicionais.

A solução foi implantada na Superintendência de Operação da Usina Hidrelétrica de Itaipu, disponível para colaboradores de todos os seus departamentos e divisões, contando com um grupo de aproximadamente 200 usuários atualmente. Considerando o baixo esforço requerido para integração de novos tipos de sistemas e documentos, bem como pela natureza escalável da arquitetura e componentes de software utilizados, esta mesma solução de indexação e busca possui potencial enorme para utilização nas demais áreas de Itaipu, bem como, por demais empresas do Setor Elétrico.

Com relação aos custos envolvidos, a solução foi desenvolvida totalmente com recursos internos da Itaipu utilizando tecnologias/ferramentas "open source" (sem custos de licenciamento associado) e sem necessidade de

aquisição de hardware dedicado, uma vez que utiliza a infraestrutura já disponível no datacenter da TI corporativa. Do ponto de vista de Hxh empregado, o desenvolvimento foi feito com dedicação parcial de 4 colaboradores da Superintendência de Operação durante aproximadamente um ano de trabalho.

Sobre os benefícios, como um resumo dos pontos já apresentados, destaca-se a melhora significativa no acesso a informações essenciais à rotina de Operação de Itaipu, tornando acessíveis através de um clique os principais documentos que norteiam as decisões tomadas no dia a dia por equipes de operação em tempo real e de pré/pós-operação.

A solução desenvolvida resolve o problema raiz da descentralização de informações e da falta de mecanismos eficientes de busca, gerando como principal resultado uma redução considerável no tempo despendido para acesso às informações necessárias para a tomada de decisão pelas equipes de tempo real e de pré e pós-operação.

A ferramenta indexa de maneira centralizada informações disponíveis em diversos sistemas, possui uma interface intuitiva, no estilo dos buscadores tradicionais da Internet, e um mecanismo robusto de indexação e busca em diferentes formatos de documentos. Além disso, dispõe de um mecanismo eficiente de pesquisa que permite encontrar informações dispersas em fontes de dados distintas rapidamente, mesmo que o usuário cometa erros de digitação, digite termos incompletos ou fora de ordem.

6. Referências Bibliográficas

[1] - <https://www.elastic.co/guide/index.html>

[2] - <https://pt-br.reactjs.org/>

DADOS BIOGRÁFICOS



Engenheiro eletricista pela Universidade Estadual do Oeste do Paraná - Unioeste, há dois anos tem atuado como desenvolvedor back e front-end no Grupo de Sistemas de Suporte da Operação, na Itaipu Binacional.

(2) **FELIPE VENTURA MUGGIATI**
Mestre em Computação pela Universidade Federal Fluminense (UFF) e graduado em Engenharia Elétrica pela Universidade Estadual do Oeste do Paraná (UNIOESTE). Já atuou como engenheiro no Departamento de Comercialização de Energia da Eletrobras e atualmente trabalha em Itaipu no desenvolvimento de sistemas de suporte à operação em tempo real e ferramentas de apoio às atividades de pré e pós-operação.

(3) **FELIPE TREVISAN**
Felipe Trevisan possui graduação em Engenharia Elétrica (2007) e mestrado em Sistemas de Energia (2011) pela UFSC. Atualmente trabalha em Itaipu no desenvolvimento de sistemas de suporte à operação em tempo real (Scada/EMS) e ferramentas de apoio às atividades de pré e pós-operação.

(4) **BRUNO AUGUSTO BASTIANI**
Graduado em Engenharia Elétrica pela UTFPR, Mestre em Engenharia Elétrica pela UTFPR. Já atuou como desenvolvedor Python e C. Atualmente, atua como engenheiro de operação em tempo real na Itaipu Binacional.