



AVALIAÇÃO DE FERRAMENTAS DE BUSCA DE ANTERIORIDADES PATENTÁRIAS BASEADAS EM **INTELIGÊNCIA ARTIFICIAL**

**MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL**

Instituto Nacional da Propriedade Industrial – INPI

Presidente: Júlio Cesar Castelo Branco Reis Moreira

**Diretoria de Patentes, Programas de Computador e Topografia de Circuitos Integrados –
DIRPA**

Alexandre Dantas Rodrigues

Coordenação Geral de Estudos, Projetos e Disseminação da Informação Tecnológica - CEPIT

Alexandre Ciancio

Coordenação do Projeto

Alexandre Ciancio

Equipe de Examinadores

Adailson da Silva Santos

Bernardo Nepomuceno Pinto Mosquera

Frederico Massaud Conde

Juliana D'Avila Ferreira Viana Martha

Laura Costa Moreira Botti

Luis Carlos Oliveira da Silva

Nilda de Moraes Silva

Zaira Bruna Hoffmam

Sumário

| | |
|--|--------------------------------------|
| Resumo Executivo | 3 |
| 1. Introdução..... | 4 |
| 2. Metodologia de Análise Quantitativa dos Resultados | 5 |
| 2.1 Escolha dos Documentos | 6 |
| 2.2 Tradução dos Documentos..... | 6 |
| 2.3 Procedimento de Classificação de Documentos..... | 7 |
| 2.4 Procedimento de Análise | 7 |
| 2.5 Análise dos Resultados Quantitativos..... | 8 |
| 2.5.1 Resultado Geral..... | 8 |
| 2.5.2 Resultados por Grande Área Tecnológica | 9 |
| 3. Análise Qualitativa dos Resultados..... | 11 |
| 3.1 Resultados Qualitativos..... | 12 |
| 3.1.2 Percepção Geral dos Resultados da Busca por IA..... | 12 |
| 3.1.2 Impacto de Variáveis Inerentes ao Pedido..... | 13 |
| 3.1.3 Impacto da Ferramenta de Busca por IA no Trabalho do Examinador..... | 13 |
| 3.1.4 Desafios e Comentários Gerais..... | 14 |
| 4. Conclusões..... | 14 |
| Apêndice A: Procedimento para Busca e Classificação de Anterioridades..... | 18 |
| Apêndice B: Questionário de Avaliação da Percepção do Examinador | 18 |
| Apêndice C: Dados Completos do Questionário de Avaliação..... | Erro! Indicador não definido. |

Resumo Executivo

Este projeto investiga e avalia o impacto potencial da utilização de ferramentas de busca por Inteligência Artificial (IA) no trabalho do examinador de patentes do Instituto Nacional da Propriedade Industrial (INPI), considerando tanto a possível melhoria na eficiência dos resultados quanto o ganho de tempo no processo de busca por documento de anterioridades. A avaliação abordou tanto o desenvolvimento de uma metodologia de avaliação quantitativa como a realização de pesquisa para consideração de aspectos qualitativos por parte dos examinadores.

Na análise quantitativa foram utilizadas duas ferramentas para examinar 20 pedidos de patente, divididos em grupos com e sem busca prévia. Os resultados revelaram que, quando a IA foi utilizada com intervenção mínima dos examinadores, as buscas foram superiores ou equivalente aos resultados da busca manual em mais de 50% dos casos. Entretanto, foi observada grande variação de desempenho ao considerar-se áreas tecnológicas isoladamente.

A análise qualitativa, por meio de pesquisas de opinião junto aos examinadores, proporcionou insights valiosos. Embora a percepção inicial dos examinadores tenha mostrado uma discordância com os resultados quantitativos, destacando a busca por IA como menos eficaz, a intervenção manual durante o refinamento das buscas foi reconhecida como um componente crucial. A pesquisa indicou que a complexidade dos pedidos teve impacto mínimo na relevância dos resultados da busca por IA. Os examinadores também reconheceram o esforço inicial exigido pelas ferramentas de IA, mas expressaram a crença no potencial de contribuição significativa a longo prazo para a qualidade do exame de patentes.

Este projeto não apenas avaliou o desempenho técnico das ferramentas de IA no contexto específico do INPI, mas também lançou luz sobre as percepções, desafios e oportunidades associadas à implementação dessas tecnologias. Os resultados preliminares sugerem a importância de uma abordagem equilibrada que incorpore a experiência humana na busca por anterioridades patentárias, mesmo diante do potencial transformador da IA no processo de exame de patentes.

1. Introdução¹

O panorama contemporâneo é testemunha de uma ascensão extraordinária nas tecnologias de Inteligência Artificial (IA). Esta tecnologia tem sido notável em diversos setores, transformando fundamentalmente a maneira como desafios complexos são abordados. Como destaca o Relatório Anual 2022-2023² do Fórum Econômico Mundial, "o crescimento da IA generativa irá transformar significativamente os modelos de negócio e padrões sociais". Desde algoritmos de aprendizado profundo até sistemas de processamento de linguagem natural, as capacidades da IA têm se expandido rapidamente, proporcionando soluções inovadoras em campos tão diversos quanto medicina, finanças e, notavelmente, no setor de propriedade intelectual. O papel da IA na automação de tarefas complexas e na análise de grandes conjuntos de dados é cada vez mais reconhecido como um catalisador para a eficiência e inovação em diversas áreas profissionais.

A busca de anterioridades desempenha um papel crucial no exame de pedidos de patente, sendo um dos pilares para a concessão de uma nova patente ou indeferimento de um pedido. Identificar documentos de anterioridade relevantes é um desafio intrínseco ao processo, dada a vastidão e diversidade de informações disponíveis. A eficácia dessa busca é crucial para assegurar que invenções não sejam replicantes do que já está presente no estado da técnica. No entanto, o volume crescente de dados e a complexidade das tecnologias modernas tornam essa tarefa cada vez mais desafiadora e demorada para os examinadores de patentes. Dados da WIPO³ indicam que o número de depósitos de patente anuais alcançou 3.7 milhões de documentos em 2022 e o total já ultrapassa 150 milhões de documentos de patente no mundo.

Neste contexto, soluções de IA apresentam-se como ferramentas promissoras para otimizar o processo de busca por anterioridades. Algoritmos de *machine learning* podem analisar grandes conjuntos de dados de maneira eficiente, identificando padrões e correlações que poderiam passar despercebidos em abordagens tradicionais. A capacidade da IA em processar linguagem natural também pode facilitar a compreensão e categorização de documentos, reduzindo a carga cognitiva sobre os examinadores. A automação de tarefas repetitivas e a geração de insights mais rápidos podem potencialmente acelerar o processo de exame, permitindo uma análise mais abrangente e detalhada das anterioridades relevantes.

Já há algum tempo ferramentas de busca em bases patentárias oferecem funcionalidades baseadas em inteligência artificial como forma de simplificar o processo de recuperação de anterioridades. Este projeto visa investigar e avaliar o impacto potencial destas ferramentas de IA no trabalho do examinador de patentes, considerando tanto a possível melhoria na eficiência dos resultados quanto o ganho de tempo no processo de busca por anterioridades. Ao explorar essas contribuições, buscamos compreender como a integração de soluções de IA pode otimizar o exame de pedidos de

¹ Esta introdução é uma versão revisada, adaptada e complementada de texto gerado automaticamente pelo ChatGPT 3.5.

²Annual Report 2022-2023 – World Economic Forum (https://www3.weforum.org/docs/WEF_Annual_Report_2022-23.pdf, acessado em 20/11/2023)

³WIPO, "World Intellectual Property Indicators 2023" - <https://tind.wipo.int/record/48541>, acessado em 20/11/2023

patente, mantendo a qualidade e precisão necessárias para o processo de decisão. Considerando que tal assunto ainda não foi endereçado no INPI, este projeto possui caráter exploratório, e visa adicionalmente fomentar a discussão e estimular reflexões sobre o uso destas soluções no processo de exame.

Este projeto está alinhado com o Objetivo 1 do Planejamento Estratégico 2023-2026⁴ do INPI: “Otimizar Qualidade e Agilidade na Concessão e Registro de Direitos de Propriedade Industrial Alcançando Padrões de Desempenho de Referência Internacional” e inserido no Plano de Ação 2023⁵ do Instituto como parte do Projeto 1.13: “Busca com IA para Marcas, DI e Patentes”, que tem por objetivo “Implantar e consolidar a utilização de soluções de Inteligência Artificial aplicadas a atividades de buscas em exame de pedidos de PI.”

2. Metodologia de Análise Quantitativa dos Resultados

A abordagem utilizada para este projeto definiu inicialmente uma metodologia de avaliação quantitativa que possa ser aplicada a qualquer ferramenta de busca de anterioridades de patente que utilize algoritmos de inteligência artificial. A proposta é que esta metodologia possa ser usada independentemente da solução de busca, visando padronizar a avaliação e permitir a comparação de resultados, dentro do possível. Neste sentido, as seguintes premissas foram assumidas:

1. Uma solução que utilize algoritmos de inteligência artificial deve exigir pouco ou nenhum esforço do examinador para a entrada de dados que permita a realização da busca;
2. A busca deve se basear em um documento completo de patente (chamado documento de referência) para o qual deseja-se encontrar as anterioridades mais próximas, ou em partes de seu conteúdo (tais como título, resumo, relatório descritivo e reivindicações);
3. Tais ferramentas devem retornar uma lista de documentos de patente, ranqueada por relevância ou proximidade ao documento de referência.

Nesta etapa do projeto foram avaliadas duas ferramentas que apresentam funcionalidades de busca com inteligência artificial:

- **Derwent Innovation (Smart Search):** os examinadores de patente já possuem acesso a esta ferramenta através de contrato celebrado entre o INPI e a Clarivate Analytics LLC. O *Derwent Innovation* apresenta a funcionalidade *Smart Search*⁶, baseada em busca semântica e avaliação de similaridade.
- **Orbit Intelligence:** A ferramenta *Orbit Intelligence* da Questel também oferece a funcionalidade de busca por inteligência artificial. O acesso à ferramenta foi obtido

⁴ http://intranet.inpi.gov.br/institucional/setores/cgpe/Plano_Estrategico_INPI_20232026.pdf

⁵ https://www.gov.br/inpi/pt-br/governanca/planejamento-estrategico/27-10-2023_pa2023_rev2-vf.pdf

⁶ https://support.clarivate.com/Patents/s/article/Derwent-Innovation-Smart-Search-Basics?language=en_US, acessado em 20/11/2023

através de um período de *trial* após contato com a empresa, que ofereceu uma sessão de treinamento na ferramenta automática de busca do Orbit (realizado no dia 2/5/2023).

Oito examinadores foram indicados para participar do desenvolvimento da metodologia e avaliação das ferramentas, sendo dois de cada Coordenação-Geral de Patentes. Tendo em vista o aspecto exploratório deste projeto, a metodologia foi discutida e adaptada conforme reuniões de alinhamento.

2.1 Escolha dos Documentos

Cada examinador realizou a busca nas ferramentas para 20 pedidos de patente depositados no Brasil. Os pedidos foram escolhidos a partir daqueles recentemente examinados (e que, portanto, já continham relatório de busca disponível), de forma que os examinadores tenham conhecimento e memória sobre seu conteúdo tecnológico e matéria reivindicada, visando permitir a adequada análise dos documentos de anterioridade. Os vinte pedidos foram divididos em dois grupos:

- **Grupo 1: pedidos ISA ou com despacho 6.22.** foram selecionados 10 pedidos que não possuíam busca prévia realizada por outros escritórios. Esta restrição visa eliminar a preocupação de que as ferramentas sendo avaliadas estejam de fato realizando a busca e não apenas replicando dados de relatórios publicados por escritórios.
- **Grupo 2: pedidos com despacho 6.21 ou 6.23.** foram selecionados 10 pedidos que já possuíam busca prévia realizada por outros escritórios.

Devido à limitação para a entrada de dados de fórmulas estruturais nas ferramentas para a área de química e fármacos **foram selecionados apenas documentos com fórmulas Markush menores que possuíam poucos exemplos.**

2.2 Tradução dos Documentos

Uma vez que nem toda ferramenta com base em IA possui tradução automática de idiomas distintos ao inglês, os dados de entrada para os documentos de referência precisaram ser traduzidos pelos examinadores. Entre as possibilidades para tradução do conteúdo, tais como *Google Translate* e *ChatGPT*, optou-se por utilizar o *WIPO Translate*⁷. Conforme a WIPO, o *WIPO Translate* é uma ferramenta avançada que utiliza técnicas de inteligência artificial para tradução de documentos técnicos, apresentando terminologia específica e vocabulário especializado para os diferentes campos técnicos disponíveis.

⁷ <https://www.wipo.int/wipo-translate/en/>

2.3 Procedimento de Classificação de Documentos

O procedimento de busca e classificação dos documentos recuperados consistia, em essência, em três passos: (i) entrada de dados na ferramenta, (ii) seleção dos 10 documentos com maior similaridade ao documento de referência, e (iii) classificação de cada um destes documentos mais similares em um dos tipos abaixo:

- **Documento X:** documento de particular relevância; a invenção reivindicada não pode ser considerada nova ou inventiva quando o documento é considerado isoladamente;
- **Documento Y:** documento de particular relevância; a invenção reivindicada não pode ser considerada dotada de atividade inventiva quando o documento é combinado com um outro documento ou mais de um.
- **Documento A:** documento que define o estado geral da técnica, mas não é considerado de particular relevância.
- **Documento NR:** documento não relevante, que descreve matéria distinta à contida no pedido de referência.

Cabe esclarecer que os relatórios de busca do INPI apresentam as categorias N e I para indicar documentos do estado da técnica que podem ser utilizados isoladamente para contestar novidade (documento N) ou atividade inventiva (documento I) do documento de referência. Neste relatório estas duas categorias foram unidas na categoria “X”, indicando documentos com conteúdo de relevância integral, em oposição a documentos Y, que apresentam relevância parcial. O procedimento detalhado utilizado na metodologia proposta pode ser visto no *Apêndice A: Procedimento para Busca e Classificação de Anterioridades*.

2.4 Procedimento de Análise

Após a realização das buscas automáticas para todos os 20 pedidos de referência nas duas ferramentas sendo testadas, cada resultado da busca por IA (com a classificação dos dez documentos mais próximos) foi comparada com o resultado da busca manual realizada pelo examinador durante o processo de exame do pedido. Considerando que um documento do tipo X é mais relevante que um documento Y, que um documento Y é mais relevante que um documento A e que um documento A é mais relevante que um documento NR, escolheu-se, para cada documento de referência e cada busca, o tipo do documento mais relevante para representar a relevância daquela busca (por exemplo, uma busca que retornou 1 documento Y, 6 documentos A e 3 documentos NR teve sua relevância classificada como “Y”). O mesmo foi feito para representar a busca manual, conforme relatório de busca já disponível para aquele pedido. Anterioridades não patentárias citadas no relatório de buscas não foram consideradas para fins de comparação de resultados uma vez que as ferramentas de busca utilizadas se restringiam a documentos patentários. Assim, nos raros casos em que a busca manual apresentava

somente documentos não-patentários, ainda que do tipo X, a relevância desta busca foi classificada como “NR”.

Concluída esta classificação, cada busca por IA foi categorizada como pertencendo a um dos três tipos a seguir:

1. **Busca por IA Inferior à Busca Manual:** a busca por IA retornou somente documentos menos relevantes que a busca manual;
2. **Busca por IA Equivalente à Busca Manual:** o documento mais relevante da busca por IA teve classificação de relevância equivalente ao melhor documento da busca manual;
3. **Busca por IA Superior à Busca Manual:** a busca por IA retornou pelo menos um documento mais relevante que o documento mais relevante encontrado pela busca manual.

2.5 Análise dos Resultados Quantitativos

Após a coleta, limpeza e tratamento dos dados, procedeu-se à análise quantitativa dos resultados obtidos. É importante lembrar que, nesta etapa de avaliação, foram considerados os resultados da busca usando IA **com intervenção mínima dos examinadores**. A análise das percepções dos examinadores em relação ao uso destas ferramentas em combinação com sua experiência e possibilidade de ajuste será objeto da Sessão 3.

2.5.1 Resultado Geral

O resultado global para as buscas realizadas pode ser visto na Figura 1. Percebe-se que, quando usadas com pouca ou nenhuma intervenção humana, os resultados apresentados raramente excedem a busca manual. Este resultado confirma as expectativas até das próprias empresas sobre o atual estágio de maturidade destas funcionalidades e sua eficácia quando utilizadas sem assistência. Ainda assim, é digno de nota que em mais de 50% das buscas realizadas o resultado com pouca intervenção humana é pelo menos equivalente ao da busca manual.

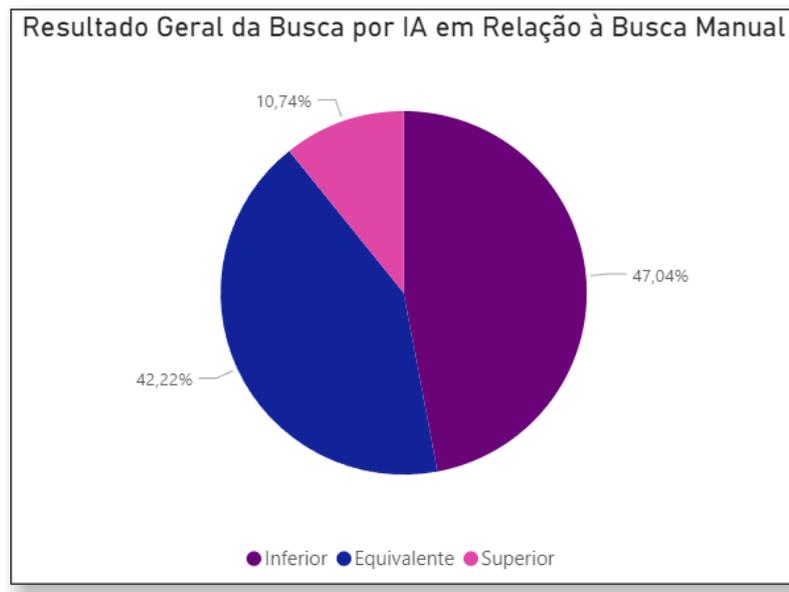


Figura 1: Resultado geral da busca por IA em comparação à busca manual.

Foi analisado o resultado das buscas considerando isoladamente pareceres do Grupo 1 (sem busca prévia) e do Grupo 2 (com busca prévia). Os resultados gerais (não mostrados neste relatório) foram equivalentes, sugerindo que as ferramentas de busca automática não se valem de relatórios de busca manual já publicados para a identificação de documentos de anterioridade relevantes.

Da mesma forma, foram analisados os resultados globais considerando cada uma das ferramentas de busca. As diferentes ferramentas apresentaram resultados globais similares. Este comportamento, entretanto, não foi observado quando a análise considera a segmentação por grande área tecnológica, como será visto a seguir.

2.5.2 Resultados por Grande Área Tecnológica

De forma a permitir uma análise mais aprofundada, os dados foram segmentados pelas grandes áreas tecnológicas representadas pelas Coordenações-Gerais de Patente (CGPATs). Os resultados segmentados, ilustrados na Figura 2, evidenciam a grande variância no desempenho das ferramentas analisadas dependendo da área tecnológica. Digno de nota é o desempenho das ferramentas na área de Elétrica e Eletrônica (CGPAT III), para a qual cerca de 78% dos resultados da busca por IA foram considerados superiores ou equivalentes, em comparação com a área de Biotecnologia, para a qual apenas cerca de 34% das buscas obtiveram estas classificações.

Resultado da Busca por IA em Relação à Busca Manual por Grande Área Tecnológica

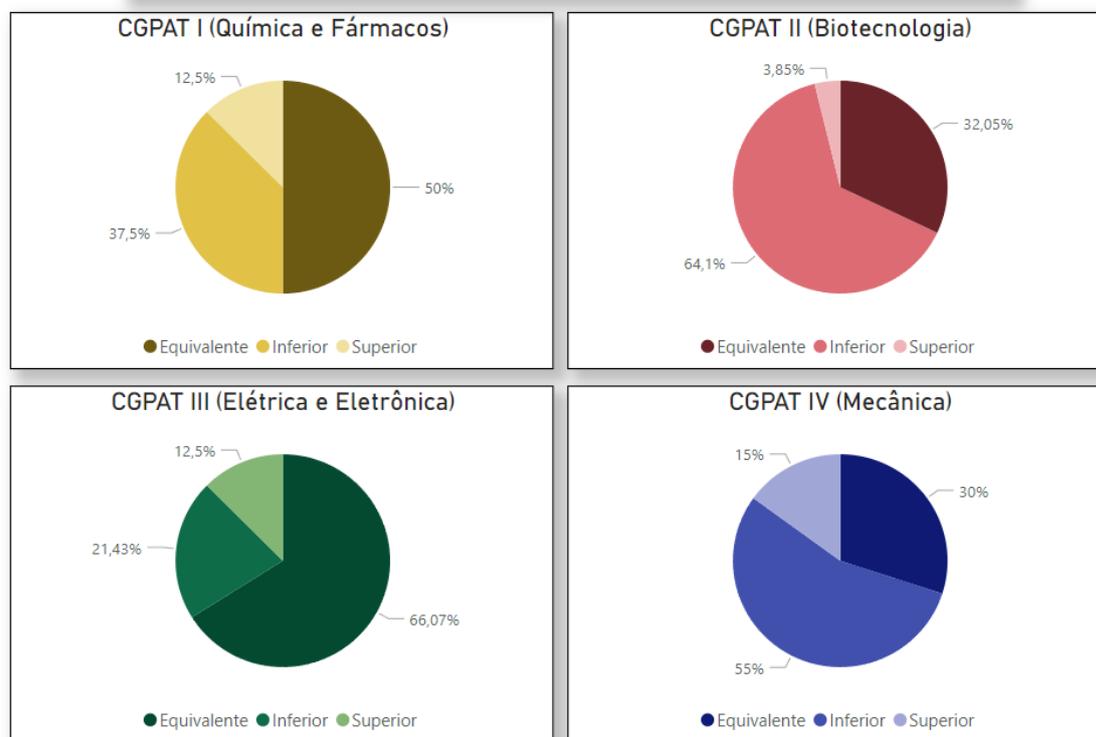


Figura 2: Resultado da Busca por IA em Relação à Busca Manual por Grande Área Tecnológica

Tal comportamento já havia sido percebido pelos examinadores e registrado nas reuniões de acompanhamento do projeto. Estas reuniões também evidenciaram o grande número de fatores que podem influenciar a recuperação de documentos relevantes por estas ferramentas. Estes aspectos serão objeto de análise na Seção 3.

É interessante notar que o comportamento semelhante das ferramentas quando analisadas globalmente não se verifica quando consideradas as particularidades das áreas tecnológicas, conforme antecipado anteriormente. A Figura 3 ilustra esta diferença de desempenho: as duas ferramentas exibem resultados com grande variação entre si em função do campo técnico em questão. Em particular, apresentam comportamentos antagônicos quando consideradas as CGPATs II e IV.

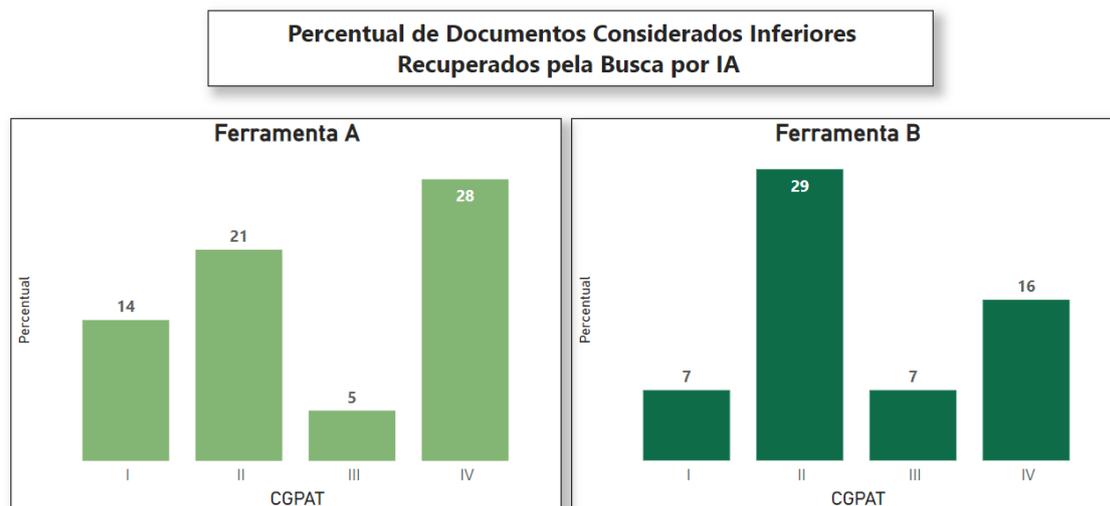


Figura 3: Percentual de Buscas realizadas por IA Considerados Inferiores à manual, por Coordenação-Geral de Patentes. As ferramentas não foram identificadas para preservar os fornecedores.

3. Análise Qualitativa dos Resultados

Durante a condução deste projeto foram realizadas três reuniões de acompanhamento e controle. Tendo em vista o caráter exploratório da avaliação, essas reuniões visavam identificar a necessidade de possíveis ajustes na metodologia e coletar impressões gerais dos examinadores. Ainda que o objetivo inicial do projeto fosse definir uma metodologia generalizável para avaliação de ferramentas de busca por IA em geral (e em boa medida isso tenha sido feito conforme apresentado na Seção 2), as discussões deixaram claro que tais ferramentas são sensíveis a um grande número de variáveis, entre elas a complexidade dos pedidos, e sua origem (pedidos nacionais ou estrangeiros).

Além disso, uma vez que um dos objetivos do projeto é também mensurar o potencial de contribuição destas ferramentas no processo de busca por anterioridades pelo examinador, uma série de fatores de interesse, baseados em percepções subjetivas, ainda deveriam ser avaliados. Com este objetivo, foi conduzida uma pesquisa de opinião junto aos examinadores que avaliou os seguintes aspectos:

- 1) Percepção da relevância dos documentos encontrados sem intervenção do examinador;
- 2) Percepção da relevância dos documentos encontrados após refinamento da busca automática pelo examinador;
- 3) Impacto da complexidade dos pedidos nos resultados obtidos pela busca por IA;
- 4) Impacto da origem dos pedidos nos resultados obtidos pela busca por IA;
- 5) Impacto da utilização de ferramentas de busca por IA no esforço de busca por anterioridades realizada pelo examinador;
- 6) Impacto da utilização de ferramentas de busca por IA na qualidade da busca por anterioridades realizada pelo examinador;

- 7) Impacto de longo prazo da utilização de ferramentas de busca por IA pelo examinador;
- 8) Principais desafios na utilização destas ferramentas.

O questionário completo respondido pelos examinadores pode ser visto no *Apêndice B: Questionário de Avaliação da Percepção do Examinador*.

3.1 Resultados Qualitativos

Nesta Seção os resultados da pesquisa serão apresentados de forma objetiva. Os dados completos das respostas podem ser encontrados no *Apêndice C: Dados Completos do Questionário de Avaliação*.

3.1.2 Percepção Geral dos Resultados da Busca por IA

A percepção dos examinadores em relação à relevância dos resultados da busca por IA encontra-se resumida a seguir:



É interessante notar que esta percepção discorda dos dados quantitativos, que indicam um resultado equivalente ou superior para mais de 50% dos casos de busca por IA. Possíveis explicações para este efeito podem incluir a inconsistência apresentada pela ferramenta, mascarando os bons resultados, e a natural resistência à novas tecnologias. Estes fatores poderiam ser mitigados a médio e longo prazo à medida que os examinadores incorporem a prática do uso destas funcionalidades à sua rotina e os algoritmos de busca por IA tornem-se melhores.

Ainda assim, a maioria dos examinadores reconhece a melhoria dos resultados quando estas ferramentas são utilizadas não isoladamente, mas refinadas manualmente com base em sua prática e conhecimento técnico.

3.1.2 Impacto de Variáveis Inerentes ao Pedido

Foi avaliado o impacto de duas características inerentes aos pedidos de patente na relevância dos resultados obtidos por IA: a complexidade da matéria reivindicada no pedido e sua origem (nacional ou estrangeiro). Os resultados são consolidados a seguir:



Os dados indicam haver pouca influência da complexidade dos pedidos na relevância dos documentos recuperados. Ainda que esta segmentação não tenha sido avaliada, este efeito pode estar diretamente associado a áreas técnicas específicas. No caso da equivalência entre pedidos estrangeiros e nacionais, os dados também sugerem que os resultados são equivalentes na percepção dos examinadores.

3.1.3 Impacto da Ferramenta de Busca por IA no Trabalho do Examinador

Em relação ao impacto da utilização de ferramentas de IA no trabalho de busca do examinador, três aspectos foram avaliados:

- 1) Impacto geral no esforço da realização da busca;
- 2) Impacto geral na melhoria da qualidade da busca;
- 3) Impacto do uso de longo prazo da funcionalidade na qualidade do exame.

As percepções sobre redução de esforço (não ilustradas aqui) foram divididas e inconclusivas no resultado geral, indicando a necessidade de investigação mais detalhada, idealmente realizando uma segmentação dos dados por área tecnológica. Entretanto, esta possibilidade não levaria a conclusões com validade estatística para o conjunto de dados disponível, devido ao reduzido número de examinadores por coordenação.

Com relação aos impactos da utilização da busca por IA na qualidade da busca, as percepções são dicotômicas:



Ainda que o uso da busca por IA neste projeto não tenha trazido ganhos significativos de qualidade na percepção dos examinadores, fica evidente a impressão do

potencial positivo do uso destas funcionalidades a longo prazo, com 75% dos examinadores acreditando que elas trariam mais qualidade ao exame. É importante novamente realçar as diferenças entre as avaliações qualitativas aqui apresentadas e as quantitativas analisadas na Seção 2.5.

3.1.4 Desafios e Comentários Gerais

Como parte da avaliação subjetiva da experiência de busca por IA, os examinadores foram solicitados a comentar sobre sua experiência e percepção sobre os principais desafios ao usar este tipo de ferramentas. Ressaltamos a seguir alguns pontos relevantes:

- Necessidade de convencimento do corpo técnico que tais ferramentas não se propõem a substituir o examinador;
- Necessidade de entender melhor os mecanismos usados pelo algoritmo de IA visando permitir um refinamento mais adequado por parte do examinador;
- Necessidade de intervenção para o refinamento da busca automática inicial visando aumentar a relevância dos documentos recuperados;
- Necessidade de avaliar a real capacidade do algoritmo de corretamente atribuir um índice de relevância aos documentos retornados;
- Preocupação com a segurança dos dados e necessidade de automação de sistemas internos da DIRPA para alimentar o algoritmo de treinamento com o texto já traduzido para o inglês;
- Necessidade de encontrar melhores caminhos para estimular a utilização da IA como ferramenta mais familiar, através de treinamentos contínuos, aliada a amplas bases de dados, interfaces amigáveis, facilidades de downloads e registros de pesquisas feitas, dentre outras funcionalidades.

4. Conclusões

A busca de anterioridades, fundamental no processo de exame de patentes, enfrenta desafios intrínsecos dada a vastidão e complexidade do estado da técnica. O aumento exponencial no número de depósitos de patentes destaca a necessidade urgente de abordagens inovadoras para otimizar esse processo. Nesse contexto, as soluções de IA surgem como ferramentas promissoras para aprimorar a eficiência da busca por anterioridades. Algoritmos de *machine learning*, capazes de analisar extensos conjuntos de dados de maneira eficiente, e a habilidade da IA em processar linguagem natural prometem oferecer vantagens significativas. A automação de tarefas repetitivas e a geração de insights rápidos têm o potencial de acelerar o exame, permitindo uma análise mais abrangente das anterioridades relevantes.

O presente projeto teve como objetivo investigar o impacto dessas ferramentas de IA no trabalho do examinador de patentes. Ao analisar a relevância dos resultados, embora as ferramentas de IA aqui avaliadas tenham se mostrado menos eficazes quando utilizadas

sem intervenção humana, a pesquisa de opinião destaca que, apesar de algumas divergências, os examinadores reconhecem o potencial da busca por IA quando refinada manualmente e seu efeito positivo a longo prazo. A avaliação da complexidade e origem dos pedidos indica que, embora sua complexidade tenha impacto mínimo, a sensibilidade das ferramentas pode variar entre áreas tecnológicas.

Uma análise quantitativa mais aprofundada revelou que a recuperação de documentos de anterioridade pela busca por IA apresentou grande sensibilidade a uma série de variáveis, especialmente evidenciadas na análise qualitativa. Durante a fase de análise quantitativa foi realizada uma avaliação segmentada por examinador, na tentativa de identificar uma possível influência subjetiva na classificação de documentos recuperados, devido à discrepância entre as percepções dos examinadores e os resultados quantitativos. Preliminarmente, ainda que tenham sido observadas grandes variações de desempenho da busca por IA em função do examinador, esta variação parece acompanhar a área técnica a qual o examinador pertence. Desta forma, e tendo em vista que o número de examinadores participantes do projeto foi pequeno, considera-se que seria prematuro associar o impacto de fatores subjetivos à classificação das anterioridades encontradas pela busca automática.

Ainda que o projeto tenha permitido avançar expressivamente nas discussões acerca do uso destas tecnologias no trabalho diário do exame técnico de pedidos de patente, muito ainda pode ser feito. Uma investigação mais conclusiva e que permita isolar os fatores de impacto na relevância dos documentos recuperados aqui apontados e outros não conhecidos exigiria um número mais representativo de examinadores e situações de busca. Não obstante, a metodologia de avaliação quantitativa aqui proposta parece fornecer um arcabouço sólido para avaliação do potencial de contribuição de ferramentas de busca por IA na recuperação de documentos de anterioridade patentários.

Este projeto não apenas lança luz sobre a implementação dessas soluções no INPI, mas também desencadeia reflexões cruciais sobre o futuro do exame de patentes no contexto da evolução da IA. As novas tecnologias baseadas em inteligência artificial alavancadas pela capacidade de treinamento a partir de quantidades massivas de dados já fazem parte da nova realidade tecnológica. Entretanto, a compreensão de suas limitações e aplicações e do seu uso ético deixam evidentes tanto seu caráter cada vez mais notável como importantes ferramentas para potencializar o trabalho do examinador, como sua incapacidade de substituir de forma irrestrita, a experiência e conhecimento técnico do profissional de busca. Resta claro a necessidade da continuidade e ampliação de análises como as aqui propostas.

APÊNDICE A

Procedimento para Busca e Classificação de Anterioridades

Para cada pedido de referência escolhido:

- 1) Extrair os dados de texto completo do quadro reivindicatório do pedido
 - a. A extração do texto será feita através de seleção e cópia direta a partir do PDF do quadro reivindicatório utilizado no exame.
 - b. Caso o PDF não permita a cópia direta do texto, o examinador deve informar o número do pedido para que seja feito OCR e extração do quadro reivindicatório pela CEPIT.
- 2) Traduzir o texto extraído para o inglês através da ferramenta de tradução *WIPO Translate*
 - a. Utilizar o endereço <https://patentscope.wipo.int/translate/translate.jsf>
 - b. Especificar o “*Language pair*” como sendo “*Portuguese->English*”;
 - c. Para melhores resultados na tradução, especificar o campo tecnológico mais próximo no menu “*Domain*”;
 - d. Para copiar o texto, é necessário clicar em “*Edit Translation*”
 - e. Tanto o texto completo em português quanto o traduzido pelo *WIPO Translate* devem ser informados na planilha de resultados para o pedido de referência (aba “*Texto QR*”). Caso o número de caracteres ultrapasse o limite da célula, utilizar mais de uma;
- 3) Utilizar o texto traduzido para o inglês para alimentar a ferramenta de busca automática.
 - a. No caso do *Orbit Intelligence*, utilizar a busca automática em duas etapas (Após entrar com o texto do QR traduzido, o *Orbit* retorna uma lista com anterioridades. Devem ser selecionadas entre três e cinco documentos, que serão utilizados pelo *Orbit* para gerar a lista completa de anterioridades e seus índices de similaridade.
 - b. No caso do *Derwent Innovation*, alimentar o texto traduzido diretamente na caixa do *Smart Search* (“*Get Started With Search*”, na tela inicial da plataforma) para obter os documentos de anterioridade.
- 4) Selecionar os 10 documentos considerados mais relevantes para cada ferramenta. Os números destes documentos bem como seu índice de proximidade deverão ser informados na planilha de resultados para o pedido de referência, observando que há uma aba para registrar os resultados do *Orbit Intelligence* e uma para o *Derwent Innovation*.
- 5) Classificar cada um dos 10 documentos (para cada ferramenta de busca) em relação ao documento de referência como sendo do tipo X, Y, A ou NR (não relevante), registrando esta categoria na aba de resultados da ferramenta;
- 6) Documentar brevemente qualquer observação que entender pertinente durante a avaliação de cada anterioridade, se houver;
- 7) Registrar na aba “*Busca Exame*”, os documentos mais relevantes citados no relatório de busca (tanto para pedidos para os quais a busca foi realizada pelo próprio examinador quanto para os que foi aproveitada a busca realizada por outros escritórios), e se aquela anterioridade foi citada no parecer de exame técnico (se disponível);

APÊNDICE B

Questionário de Avaliação da Percepção do Examinador

Esta pesquisa visa mensurar a percepção subjetiva dos examinadores de patente durante a utilização de funcionalidades de busca de anterioridades que utilizam inteligência artificial e seu impacto no trabalho do examinador. Como premissa, entende-se que uma ferramenta de busca que utiliza inteligência artificial deve, idealmente, exigir pouco ou nenhum esforço do examinador na definição da estratégia de busca ou no seu refinamento.

- 1) Considere a utilização da funcionalidade de busca baseada em inteligência artificial na sua forma mais direta (ou seja, que não exige grande esforço do examinador na definição da estratégia ou no seu refinamento). Na sua percepção, classifique, de forma geral, os documentos selecionados retornados pela ferramenta quanto à sua relevância para a busca de anterioridades do pedido de referência.
 - Nada relevantes
 - Pouco relevantes
 - Indiferentes
 - De alguma relevância
 - Muito relevantes.
 - Não posso avaliar.

- 2) Considere agora os casos em que a busca realizada pela ferramenta baseada em inteligência artificial foi refinada por você através da dedicação de esforço adicional, por exemplo, no ajuste da estratégia ou de parâmetros adicionais de busca. Na sua percepção os documentos de anterioridade obtidos neste caso, quanto à sua relevância:
 - Foram menos relevantes
 - Apresentaram a mesma relevância
 - Foram mais relevantes

- 3) Em sua experiência, como a relevância dos documentos de anterioridade retornados pela ferramenta varia entre pedidos com diferentes níveis de complexidade?
 - São normalmente piores que para pedidos mais simples
 - São equivalentes aos obtidos para pedidos mais simples
 - São normalmente melhores que para pedidos mais simples
 - Não tenho certeza

- 4) Em sua opinião, a relevância dos documentos de anterioridade retornados pela ferramenta é diferente para pedidos nacionais em comparação com pedidos estrangeiros?
 - Foram normalmente piores para pedidos estrangeiros
 - Foram equivalentes aos obtidos para pedidos estrangeiros
 - Foram normalmente melhores para pedidos estrangeiros
 - Não tenho experiência com pedidos nacionais/não tenho certeza

- 5) Avalie a interface da ferramenta quanto à sua facilidade de utilização
 - Interface nada amigável

- Interface medianamente amigável
- Interface amigável

Um dos propósitos deste trabalho é mensurar a possível contribuição de ferramentas de busca baseadas em IA no exame de pedidos de patente. Nas questões a seguir, avalie, de forma geral sua percepção sobre o impacto da utilização da ferramenta no seu trabalho de busca por anterioridades.

6) Em relação ao seu esforço na busca por anterioridades, como você percebe o impacto geral da ferramenta avaliada?

- reduziu consideravelmente o esforço;
- reduziu medianamente o esforço;
- não reduziu o esforço;
- aumentou o esforço;
- aumentou muito o esforço;
- Não sei avaliar

7) Em relação à melhoria da qualidade da sua busca, como você percebeu o impacto geral da ferramenta avaliada?

- Aumentou consideravelmente a qualidade da busca realizada;
- Aumentou medianamente a qualidade da busca realizada;
- Não impactou a qualidade da busca;
- Piorou a qualidade da busca realizada;
- Não sei avaliar

8) Impacto de Longo Prazo: você acredita que o uso contínuo de ferramentas de busca baseadas em IA pode afetar positivamente a qualidade do exame de patentes no longo prazo?

- Sim, de forma significativa
- Sim, de forma moderada
- Não, não afeta a qualidade
- Não tenho certeza

9) Na sua opinião, quais são os principais desafios ao usar ferramentas de busca baseadas em IA no exame de patentes? (Resposta aberta)

Este questionário visa coletar sua percepção sobre o uso de ferramentas de busca baseadas em IA no exame de patentes. Agradecemos por suas respostas detalhadas, pois elas são essenciais para nossa pesquisa.

APÊNDICE C

Avaliação de Ferramenta de Busca Baseada em IA

Resultados

Questionário 776694

| | |
|-------------------------------------|---------|
| Número de registros nesta consulta: | 8 |
| Total de registros no questionário: | 8 |
| Porcentagem do total: | 100.00% |

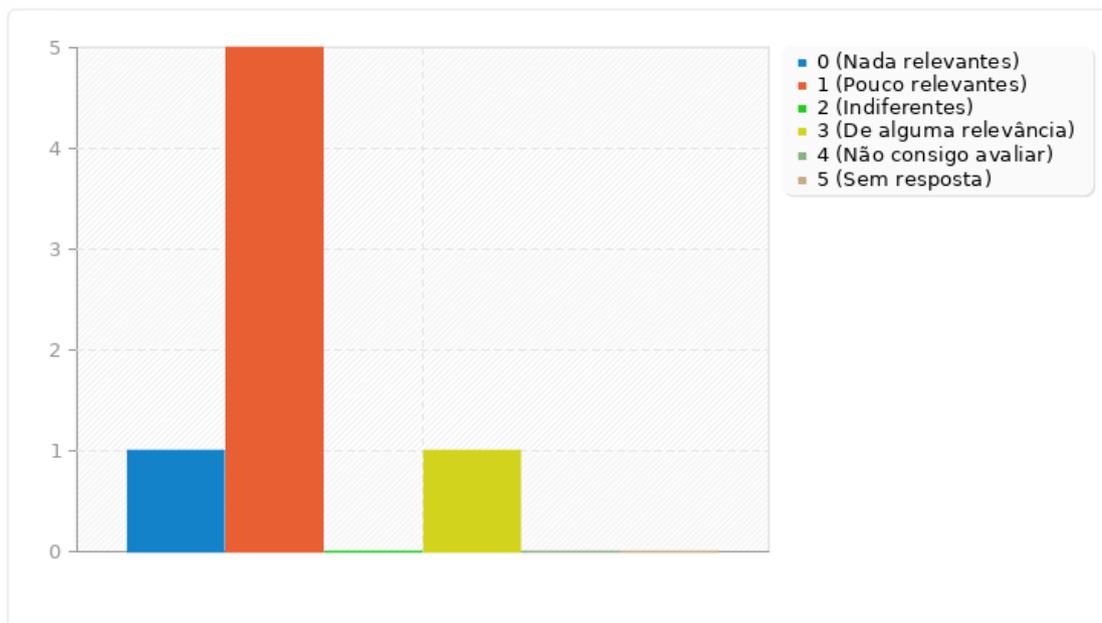
Resumo de Q01

Considere a utilização da funcionalidade de busca baseada em inteligência artificial na sua forma mais direta (ou seja, que não exige grande esforço do examinador na definição da estratégia ou no seu refinamento). Na sua percepção, classifique, de forma geral, os documentos selecionados retornados pela ferramenta quanto à sua relevância para a busca de anterioridades do pedido de referência.

| Resposta | Contagem | Porcentagem |
|-----------------------------|----------|-------------|
| Nada relevantes (AO01) | 1 | 14.29% |
| Pouco relevantes (AO02) | 5 | 71.43% |
| Indiferentes (AO03) | 0 | 0.00% |
| De alguma relevância (AO05) | 1 | 14.29% |
| Não consigo avaliar (AO04) | 0 | 0.00% |
| Sem resposta | 0 | 0.00% |

Resumo de Q01

Considere a utilização da funcionalidade de busca baseada em inteligência artificial na sua forma mais direta (ou seja, que não exige grande esforço do examinador na definição da estratégia ou no seu refinamento). Na sua percepção, classifique, de forma geral, os documentos selecionados retornados pela ferramenta quanto à sua relevância para a busca de anterioridades do pedido de referência.



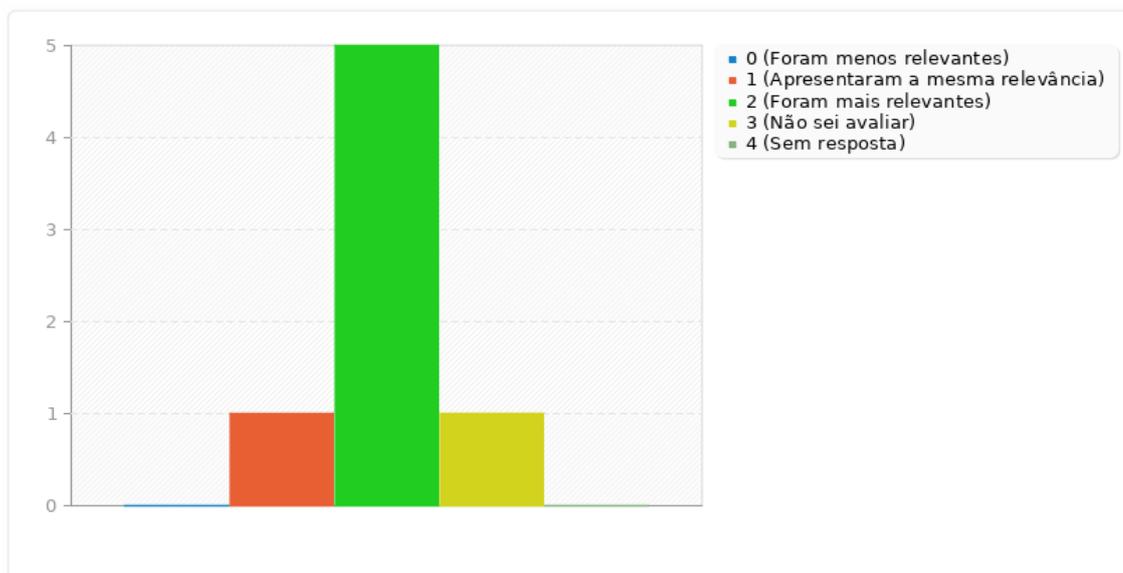
Resumo de Q02

Considere agora os casos onde a busca realizada pela ferramenta baseada em inteligência artificial foi refinada por você através da dedicação de esforço adicional, por exemplo, no ajuste da estratégia ou de parâmetros adicionais de busca. Na sua percepção os documentos de anterioridade obtidos neste caso, quanto à sua relevância:

| Resposta | Contagem | Porcentagem |
|--|----------|-------------|
| Foram menos relevantes (AO01) | 0 | 0.00% |
| Apresentaram a mesma relevância (AO02) | 1 | 14.29% |
| Foram mais relevantes (AO03) | 5 | 71.43% |
| Não sei avaliar (AO04) | 1 | 14.29% |
| Sem resposta | 0 | 0.00% |

Resumo de Q02

Considere agora os casos onde a busca realizada pela ferramenta baseada em inteligência artificial foi refinada por você através da dedicação de esforço adicional, por exemplo, no ajuste da estratégia ou de parâmetros adicionais de busca. Na sua percepção os documentos de anterioridade obtidos neste caso, quanto à sua relevância:



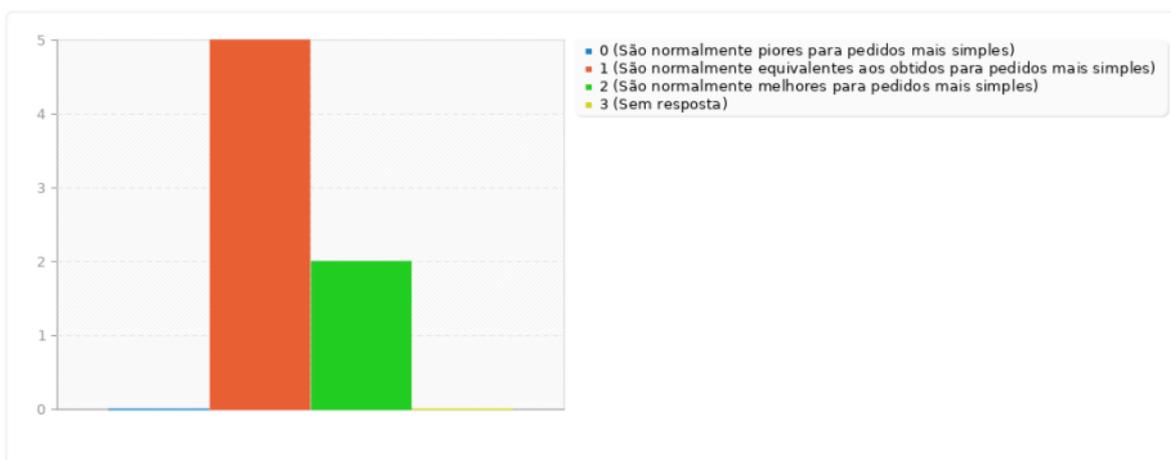
Resumo de Q03

Em sua experiência, como a relevância dos documentos de anterioridade retornados pela ferramenta varia entre pedidos com diferentes níveis de complexidade?

| Resposta | Contagem | Porcentagem |
|---|----------|-------------|
| São normalmente piores para pedidos mais simples (AO01) | 0 | 0.00% |
| São normalmente equivalentes aos obtidos para pedidos mais simples (AO02) | 5 | 71.43% |
| São normalmente melhores para pedidos mais simples (AO03) | 2 | 28.57% |
| Sem resposta | 0 | 0.00% |

Resumo de Q03

Em sua experiência, como a relevância dos documentos de anterioridade retornados pela ferramenta varia entre pedidos com diferentes níveis de complexidade?



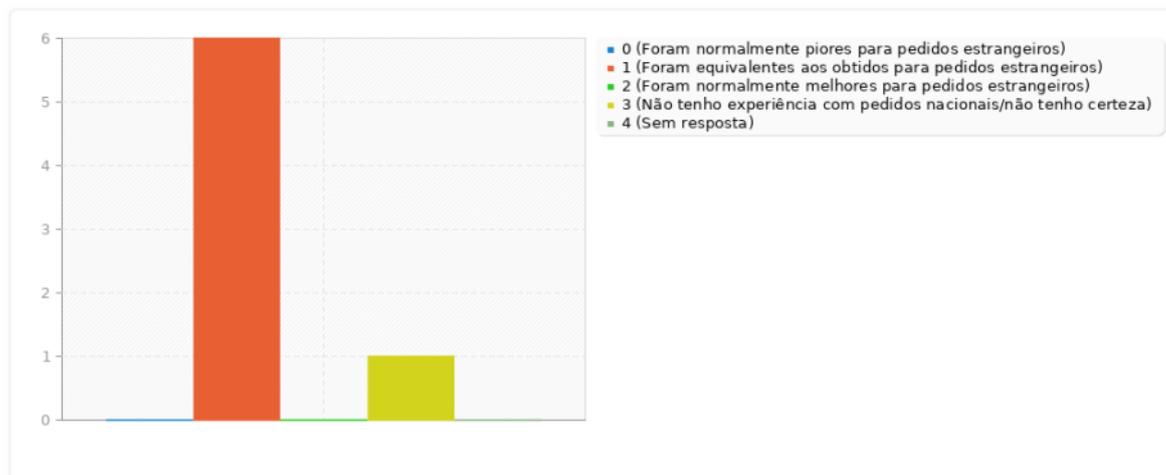
Resumo de Q04

Em sua opinião, a relevância dos documentos de anterioridade retornados pela ferramenta é diferente para pedidos nacionais em comparação com pedidos estrangeiros?

| Resposta | Contagem | Porcentagem |
|--|----------|-------------|
| Foram normalmente piores para pedidos estrangeiros (AO01) | 0 | 0.00% |
| Foram equivalentes aos obtidos para pedidos estrangeiros (AO02) | 6 | 85.71% |
| Foram normalmente melhores para pedidos estrangeiros (AO03) | 0 | 0.00% |
| Não tenho experiência com pedidos nacionais/não tenho certeza (AO04) | 1 | 14.29% |
| Sem resposta | 0 | 0.00% |

Resumo de Q04

Em sua opinião, a relevância dos documentos de anterioridade retornados pela ferramenta é diferente para pedidos nacionais em comparação com pedidos estrangeiros?



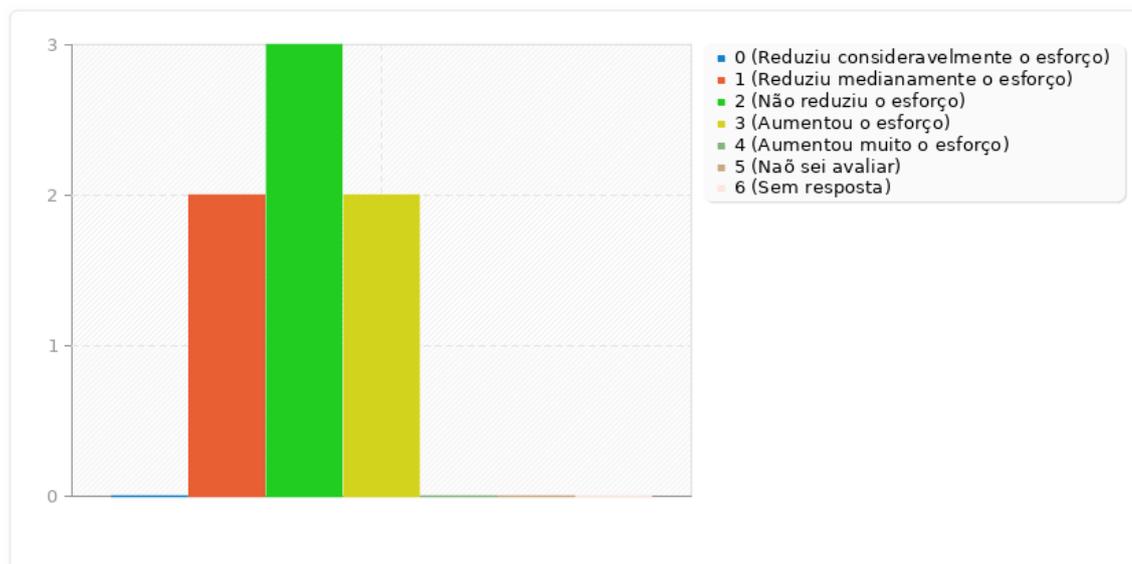
Resumo de Q06

Em relação ao seu esforço na busca por anterioridades, como você percebe o impacto geral da ferramenta avaliada?

| Resposta | Contagem | Porcentagem |
|--|----------|-------------|
| Reduziu consideravelmente o esforço (AO01) | 0 | 0.00% |
| Reduziu medianamente o esforço (AO02) | 2 | 28.57% |
| Não reduziu o esforço (AO03) | 3 | 42.86% |
| Aumentou o esforço (AO04) | 2 | 28.57% |
| Aumentou muito o esforço (AO05) | 0 | 0.00% |
| Naõ sei avaliar (AO06) | 0 | 0.00% |
| Sem resposta | 0 | 0.00% |

Resumo de Q06

Em relação ao seu esforço na busca por anterioridades, como você percebe o impacto geral da ferramenta avaliada?



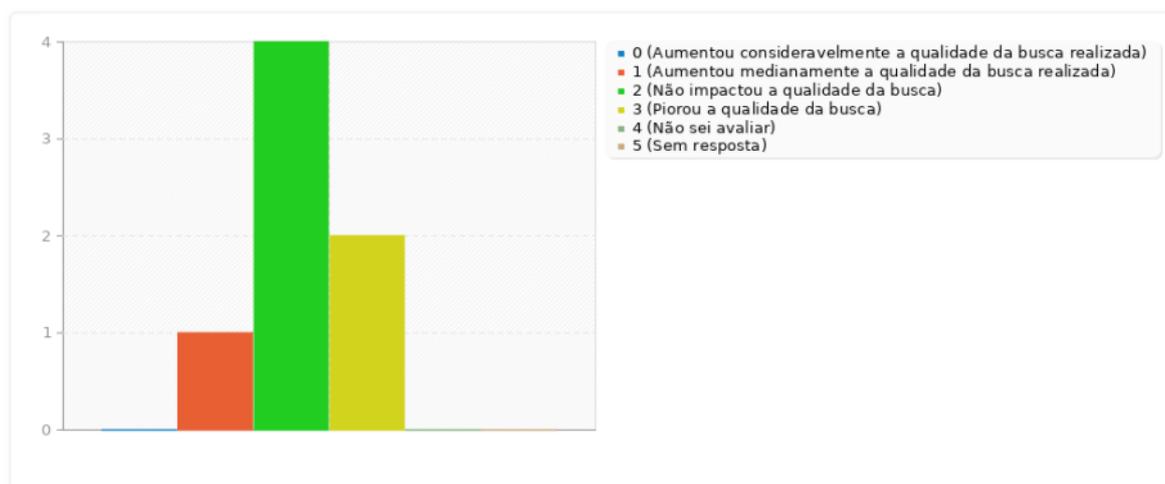
Resumo de Q07

Em relação à melhoria da qualidade da sua busca, como você percebeu o impacto geral da ferramenta avaliada?

| Resposta | Contagem | Porcentagem |
|--|----------|-------------|
| Aumentou consideravelmente a qualidade da busca realizada (AO01) | 0 | 0.00% |
| Aumentou medianamente a qualidade da busca realizada (AO02) | 1 | 14.29% |
| Não impactou a qualidade da busca (AO03) | 4 | 57.14% |
| Piorou a qualidade da busca (AO04) | 2 | 28.57% |
| Não sei avaliar (AO05) | 0 | 0.00% |
| Sem resposta | 0 | 0.00% |

Resumo de Q07

Em relação à melhoria da qualidade da sua busca, como você percebeu o impacto geral da ferramenta avaliada?



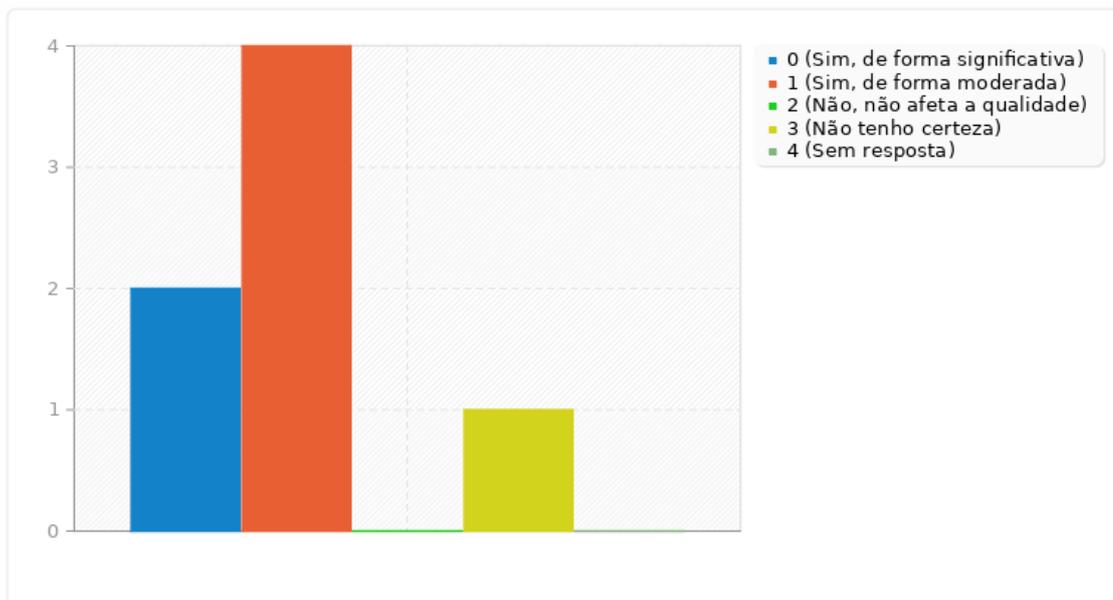
Resumo de Q08

Impacto de Longo Prazo: você acredita que o uso contínuo de ferramentas de busca baseadas em IA pode afetar positivamente a qualidade do exame de patentes no longo prazo?

| Resposta | Contagem | Porcentagem |
|------------------------------------|----------|-------------|
| Sim, de forma significativa (AO01) | 2 | 28.57% |
| Sim, de forma moderada (AO02) | 4 | 57.14% |
| Não, não afeta a qualidade (AO03) | 0 | 0.00% |
| Não tenho certeza (AO04) | 1 | 14.29% |
| Sem resposta | 0 | 0.00% |

Resumo de Q08

Impacto de Longo Prazo: você acredita que o uso contínuo de ferramentas de busca baseadas em IA pode afetar positivamente a qualidade do exame de patentes no longo prazo?



Resumo de Q09

Na sua opinião, quais são os principais desafios ao usar ferramentas de busca baseadas em IA no exame de patentes? (Resposta aberta)

| Resposta | Contagem | Porcentagem |
|--------------|----------|-------------|
| Resposta | 7 | 100.00% |
| Sem resposta | 0 | 0.00% |

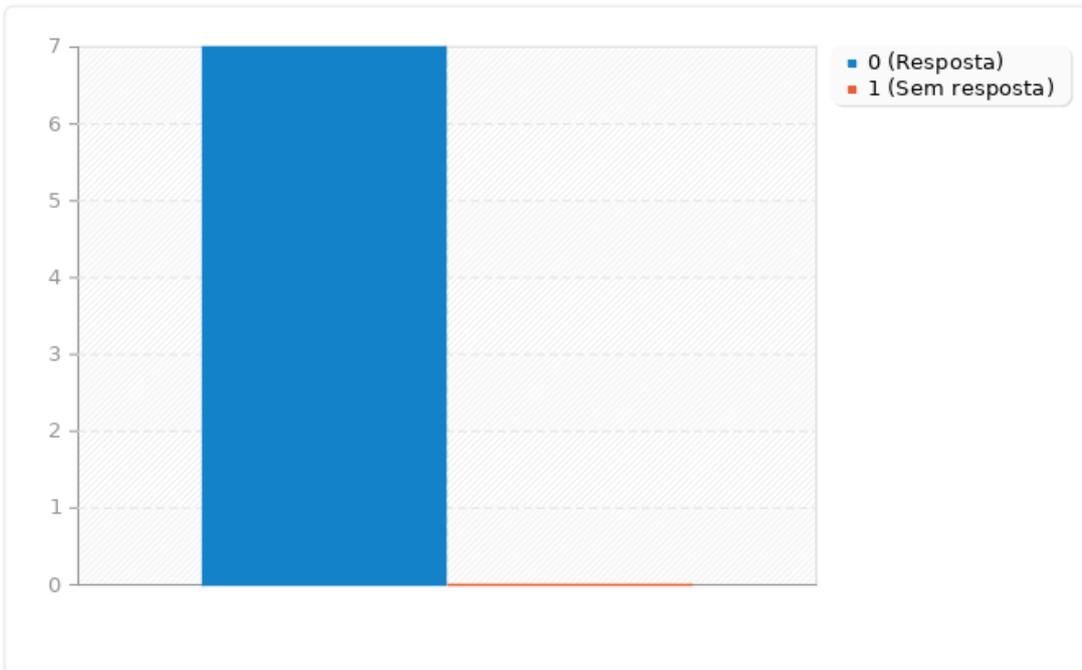
| ID | Resposta |
|----|--|
| 1 | <p>Desafio maior será convencer as pessoas que as ferramentas de busca com IA não são um substituto ao examinador nas buscas, mas uma ferramenta que pode ou não ser utilizada por eles.</p> <p>Outro desafio é saber melhor os mecanismos como as IAs trabalham para que nós possamos com os ajustes do examinador fazer escolhas mais apuradas nos filtros, por exemplo.</p> <p>Por fim, outro desafio são pedidos com quadros maiores, que acabam poluindo as buscas. Para estes casos a metodologia do teste poderia ser aprimorada, por exemplo há pedidos que pedem um processo, um produto e um aparelho e colando tudo no campo de texto de buscas até a IA tem dificuldades. Na prática, para estes casos, busco cada campo individualmente (uma busca por aparelho, outra para o processo e outra para o produto).</p> <p>Na questão que compara pedidos nacionais e internacionais, na minha opinião é difícil comparar, pois alguns dos nacionais estavam muito amplos e simples o que facilita a busca. Já os nos pedidos internacionais a ferramenta encontrou alguns documentos que são similares aos citados nos exames internacionais. Achei difícil comparar.</p> <p>Na questão do esforço a escala da pergunta foi "muito", "mediano" "nada" (...), mas não tem "pouco". Marcaria que reduz pouco o esforço. Infelizmente só tive esta percepção na hora de marcar as alternativas e não na hora que foi pedida opinião sobre o questionário.</p> |
| 2 | Implementar a avaliação para o "Get Prior Documents" presente no STN. |
| 4 | <p>Sobre as penúltimas duas perguntas:</p> <p>*Em relação ao seu esforço na busca por anterioridades, como você percebe o impacto geral da ferramenta avaliada?</p> <p>**Em relação à melhoria da qualidade da sua busca, como você percebeu o impacto geral da ferramenta avaliada?</p> <p>Respondi que aumentou o esforço e não impactou a qualidade de busca considerando o projeto executado, ou seja, sem mexer nos parâmetros. Dessa maneira, entendo que a execução automática sem mexer em parâmetros é uma grande perda de tempo, porque os documentos voltam sem relevância para a esmagadora maioria dos pedidos. Algumas áreas muito específicas tiveram resultado, mas creio ter sido mais por baixa complexidade da área do que pelo algoritmo em si.</p> <p>Uma observação, noto que a ferramenta de smart search do Derwent foi melhor porque ela extrai e permite manipular os conceitos. Então estou utilizando isso agora para minhas buscas, e de fato o Derwent ajuda em alguma extensão porque permite manipular e "limpar" os conceitos extraídos do pedido. O Orbit não aponta os conceitos que extrai, então não permite interação nossa no momento da busca. Ou seja, não há como utilizar o Orbit de maneira interativa, uma vez que a plataforma "esconde" os conceitos que extrai para busca.</p> |
| 5 | <p>No geral, o uso da ferramenta por si só não retornou resultados significantes. Em muitos casos, a ferramenta retornou documentos completamente fora do escopo dos pedidos, e atribuo isto ao fato de a mesma se basear apenas em uma seleção de palavras dentro do texto, o que pode realmente levar a documentos de áreas técnicas completamente diferentes. Além disso, dado o enorme número de resultados retornados, é possível que as ferramentas retornem algum documento relevante, mas que o mesmo não seja entendido pela IA como mais próximo ao pedido buscado e fique de fora da seleção inicial feita para refinamento. Mesmo com o refinamento por documentos mais próximos (como o feito para o Orbit) acredito que se possa perder alguma anterioridade que seria relevante. Também não</p> |

considero factível que o examinador passe por todos os documentos retornados para selecionar aqueles que, em sua opinião, seriam os mais relevantes, uma vez que ocorre de a IA entregar milhares de documentos. Um refinamento por parte do examinador, em cima dos resultados retornados, então, se faria necessário no estágio atual das ferramentas (minimamente IPC / CPC). Acredito que as ferramentas de IA se tornariam mais precisas e poderosas se de alguma forma pudesse incorporar classificações (ou outros filtros) já em sua busca, o que ainda reduziria etapas a serem realizadas pelo examinador.

- 6 Acredito que os maiores desafios são:
- 1- criação de metodologia para treinar a ferramenta de maneira específica e inteligente
 - 2- análise do volume de dados gerados
 - 3- segurança dos dados
 - 4- necessidade de automação de sistemas internos da DIRPA para alimentar a IA: tradução dos docs para ingles (por exemplo)
 - 5- criar cultura de Educação Continuada dentro da dirpa
- 7 Acredito que, no momento atual, e tendo em vista a ferramenta avaliada, o uso da IA necessita, como toda ferramenta, ser explorada para que possa ser aproveitada da maneira mais adequada, passando por treinamentos contínuos, e assim encontrar um caminho de maior aproveitamento. Até o momento, a ferramenta se mostrou interessante, apenas em poucos casos, como busca complementar, não tendo sido suficiente para chegar aos documentos mais relevantes dentro da metodologia escolhida para a avaliação. A ferramenta em questão apresenta uma interface de refinamento de busca que exige comandos específicos e linhas de textos que saem do usual das ferramentas de buscas já conhecidas, cujos filtros são mais intuitivos. A tela de resultados da ferramenta avaliada não supera a interface das ferramentas já conhecidas, onde a agilidade e facilidade de visualizar os documentos, baixar e traduzir , por exemplo, são muito importantes durante a análise da matéria. De forma geral, o desafio é encontrar os caminhos para guiar de maneira mais precisa a utilização da IA como ferramenta, através de treinamentos e desde que aliada a amplas bases de dados, interfaces amigáveis, facilidades de downloads e registros de pesquisas feitas, dentre outras funcionalidades .
- 8 O maior de todos os desafios seria o INPI se aproveitar dessa tecnologia de terceiros para promover a formação de um banco de dados nacional que possibilite um melhor acesso às descrições das invenções pelos institutos de pesquisa e universidades além, é claro, do próprio INPI no exercício de suas funções.

Resumo de Q09

Na sua opinião, quais são os principais desafios ao usar ferramentas de busca baseadas em IA no exame de patentes? (Resposta aberta)





INPI INSTITUTO
NACIONAL DA
PROPRIEDADE
INDUSTRIAL

GOVERNO FEDERAL
BRASIL
UNIÃO E RECONSTRUÇÃO



gov.br/inpi



[inpi brasil](#)



[inpi brasil](#)



[inpi brasil](#)



[inpi brasil](#)