

# 4º Congresso Latino-Americano de Casos de Open Innovation

11 A 15 | AGO

RIO DE JANEIRO

oiweek  
open innovation week  
LATAM

REALIZAÇÃO:



## UTILIZAÇÃO DE MODELOS DE LINGUAGEM DE GRANDE ESCALA PARA ANÁLISE AUTOMATIZADA DE CONFORMIDADE EM NORMAS REGULATÓRIAS

Arthur Miranda Tavares<sup>1</sup>; Edgard Henrique Coelho Tavares da Silva<sup>2</sup>; Ricardo Guidolim<sup>3</sup>;  
Ricardo Lucio Braga Reis<sup>4</sup>; Paula Abrahao<sup>5</sup>

### RESUMO

O presente trabalho tem como objetivo descrever uma solução inovadora baseada em Modelos de Linguagem de Grande Escala (LLMs) para otimizar a análise de conformidade regulatória no contexto da estrutura corporativa da B3 S.A. - Brasil, Bolsa, Balcão. O problema central reside na complexidade e na constante atualização das normativas emitidas por entidades supervisoras do mercado financeiro, que exigem interpretação precisa e revisões constantes das políticas internas. Essas normativas, frequentemente extensas e redigidas em linguagem jurídica, demandam um esforço manual significativo, resultando em altos custos operacionais, riscos de não conformidade e possíveis penalidades. A ferramenta desenvolvida emprega técnicas avançadas de processamento de linguagem natural, como embeddings e busca vetorial com a ferramenta FAISS, para automatizar a interpretação de textos jurídicos e comparar normas internas e externas com maior precisão e eficiência. Os resultados evidenciaram uma redução significativa no tempo necessário para análises, além de uma mitigação robusta dos riscos legais e a identificação de lacunas nas políticas internas da organização. A solução também demonstrou potencial de escalabilidade, podendo ser adaptada a outros setores altamente regulados. Ao automatizar tarefas repetitivas e complexas, a ferramenta libera tempo e recursos humanos para atividades estratégicas, contribuindo para a melhoria da governança corporativa e para um ambiente de negócios mais seguro e transparente. O projeto ilustra o papel transformador da inteligência artificial na modernização de processos de compatibilidade regulatória, destacando sua aplicabilidade e impacto na indústria e na academia.

**PALAVRAS-CHAVE:** Inteligência artificial; Conformidade regulatória; Processamento de linguagem natural; Modelos de linguagem de grande escala; Busca vetorial.

NOME DO ATOR	PAPEL NO ECOSISTEMA
B3 S.A. - Brasil, Bolsa, Balcão.	Corporação

<sup>1</sup> Universidade Federal da Paraíba. arthur.miranda@academico.ufpb.br

<sup>2</sup> Universidade Federal da Paraíba. edgard.tavares@academico.ufpb.br

<sup>3</sup> B3 S.A. – Brasil, Bolsa, Balcão. ricardo.guidolim@b3.com.br

<sup>4</sup> B3 S.A. – Brasil, Bolsa, Balcão. ricardo.reis@b3.com.br

<sup>5</sup> B3 S.A. – Brasil, Bolsa, Balcão. paula.abrahao@b3.com.br

**1. SITUAÇÃO PROBLEMA DE OPEN INNOVATION**

Em um ambiente altamente controlado, como o setor financeiro, a conformidade com diretrizes emitidas por órgãos reguladores é um fator crítico para garantir operações seguras, eficientes e alinhadas com as exigências legais. No Brasil, instituições como o Banco Central do Brasil (BCB) e a Comissão de Valores Mobiliários (CVM) frequentemente atualizam e publicam novas regras para o mercado, impondo desafios contínuos para empresas do setor, incluindo a B3 S.A. - Brasil, Bolsa, Balcão, que desempenha um papel central na infraestrutura do mercado financeiro do país.

Um exemplo recente que ilustra esses desafios é a resolução BCB nº 304, de 20/03/2023[1]. Com 42 páginas e 182 artigos, essa normativa demandou da B3 um esforço significativo para garantir sua conformidade. A equipe da empresa realizou diversas reuniões multidisciplinares, envolvendo áreas como negócios, tecnologia, jurídica e outras, além de dedicar mais de um mês de trabalho intensivo para analisar e implementar as mudanças necessárias. Esse processo, além de consumir tempo e recursos, destacou a complexidade inerente à interpretação e aplicação de normativas extensas e detalhadas, que muitas vezes são escritas em uma linguagem técnica e demandam uma atenção minuciosa para compará-las com políticas internas.

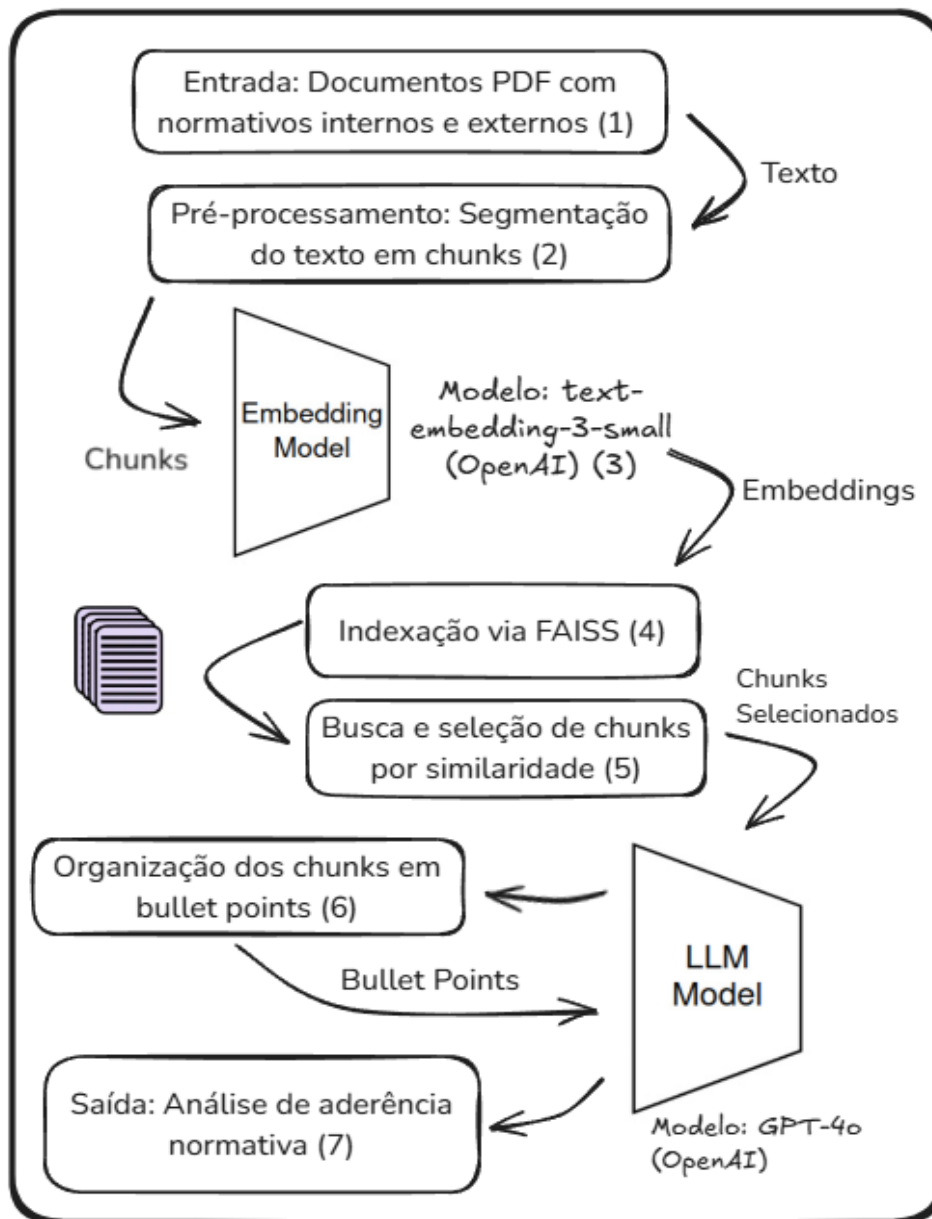
Além disso, as mudanças nas regras podem impactar diretamente a estratégia corporativa, os processos operacionais e a mitigação de riscos regulatórios. O não cumprimento de tais diretrizes pode resultar em penalidades severas, perda de credibilidade e implicações financeiras significativas.

Diante desse cenário, identificou-se a necessidade de uma solução inovadora que pudesse otimizar o processo de análise de compatibilidade regulatória, reduzindo o tempo gasto pelas equipes e aumentando a precisão na verificação de aderência entre os normativos internos da B3 e os regulamentos externos. Com os avanços recentes na inteligência artificial, os Modelos de Linguagem de Grande Escala (Large Language Models – LLMs) surgiram como uma alternativa promissora para enfrentar esse desafio.

**2. INTERVENÇÃO**

A partir da evolução dos Modelos de Linguagem de Grande Escala (LLMs), novas abordagens para a análise regulatória tornaram-se viáveis, permitindo a automação de tarefas antes exclusivamente humanas. LLMs, como o BERT (Bidirectional Encoder Representations from Transformers) [2] e o GPT (Generative Pre-trained Transformer), têm demonstrado grande eficiência na compreensão contextual e na geração de textos, sendo amplamente utilizados em diversas aplicações de Processamento de Linguagem Natural (PLN) (Devlin et al., 2019) [2]. Nesse sentido, diante do desafio de manter a aderência das normas internas às regulamentações externas, mostrou-se necessária uma solução que reduzisse o tempo e aumentasse a precisão para a interpretação e comparação de grandes volumes de documentos com a utilização dessas tecnologias.

Para atender a essa necessidade, foi desenvolvida uma ferramenta baseada em LLMs que correlaciona normativas externas com as diretrizes internas da B3. Utilizando técnicas de processamento de linguagem natural e aprendizado de máquina, a solução automatiza a interpretação de textos jurídicos complexos e aprimora a identificação de aderência regulatória.



[1] Imagem: Diagrama da Arquitetura do Sistema

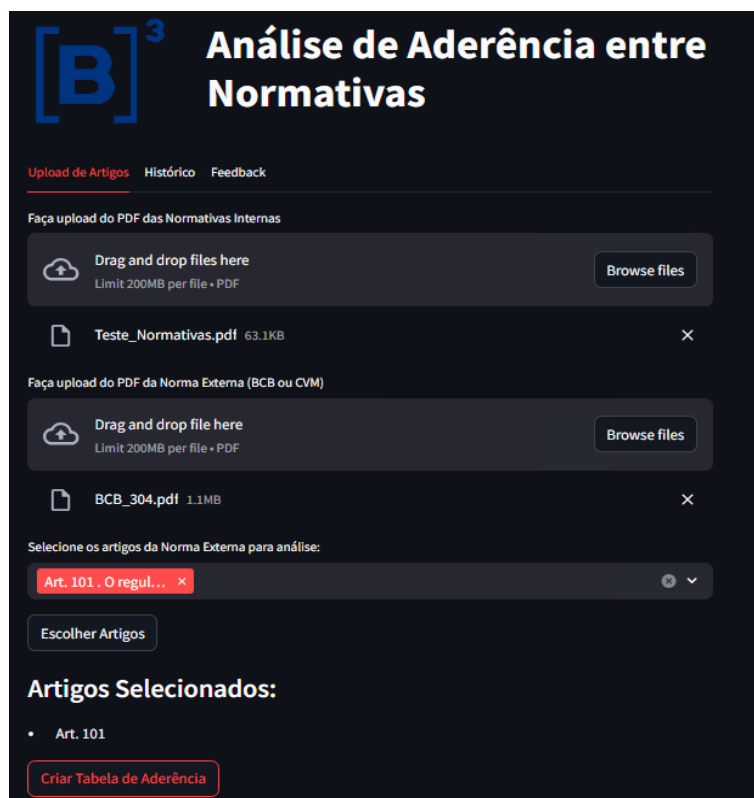
Com base na visualização da imagem 1, a arquitetura do sistema fundamenta-se na extração automatizada de textos normativos em documentos PDF (1.1), que passam por um pré-processamento onde são divididos em fragmentos menores, chamados chunks (1.2). Essa divisão considera a estrutura do texto, garantindo que cada segmento mantenha um contexto coerente. Para capturar relações semânticas entre normativas internas e externas, esses chunks são convertidos em vetores (Embeddings) por meio do modelo text-embedding-3-small da OpenAI (1.3) [5]. Embeddings são representações numéricas de conceitos, convertidas em sequências de números, cujo uso permite identificar padrões e similaridades, uma vez que embeddings numericamente similares também são semanticamente próximos [6]. Para otimizar a busca dessas semelhanças em grandes volumes de dados, a ferramenta emprega o Facebook AI Similarity Search (FAISS) (Johnson, Douze & Jégou, 2019) [3], uma biblioteca eficiente para busca vetorial em larga escala (1.4).

Diante dessa busca de correspondências, os chunks mais similares identificados (1.5) são organizados em bullet points (1.6) pelo GPT-4o [4] via API da OpenAI, criando um fluxo lógico de informações (chain of thought). Esse encadeamento favorece a compreensão dos dados e aprimora significativamente a análise de aderência. Com essa abordagem, o modelo gera o relatório com a análise final (1.7) obtendo interpretações mais precisas, identificando compatibilidades, divergências e lacunas. Desse modo, permitindo uma visão mais clara das normativas internas e suas relações com os regulamentos externos.

O desenvolvimento dessa ferramenta envolveu pesquisas aprofundadas sobre inteligência artificial aplicada à análise regulatória. O sistema foi estruturado com base no estado da arte na recuperação de informações e modelos generativos de IA, utilizando embeddings e busca vetorial para garantir eficiência e escalabilidade. Além disso, testes iterativos foram conduzidos em colaboração com especialistas de compliance para validar os resultados e ajustar a solução às necessidades da B3, assegurando sua eficácia na otimização da conformidade legal.

### 3. RESULTADOS

Os resultados dos testes realizados demonstram que a solução baseada em LLMs provou ser eficaz na extração, análise e comparação de grandes volumes de documentos, reduzindo significativamente o tempo necessário para as equipes da B3 realizarem verificações de aderência entre a empresa e os órgãos reguladores. Em contraste com os métodos tradicionais, que demandam esforço manual intensivo e leitura minuciosa, a inteligência artificial permitiu uma análise mais ágil, precisa e capaz de identificar padrões e similaridades entre normativas com elevado grau de acurácia. Isso facilitou a tomada de decisões estratégicas no desenvolvimento e revisão de diretrizes internas.



[2] Imagem: Interface da Ferramenta



[3] Imagem: Análise de Aderência

As imagens 2 e 3 ilustram um fluxo simulado para a avaliação de normativas internas, criando um ambiente seguro para testes sem comprometer dados sensíveis. O processo inicia-se com o upload de documentos PDF contendo normativas de teste e a resolução pública BCB 304, permitindo ao usuário selecionar artigos específicos dessa regulamentação para análise detalhada. A ferramenta gera automaticamente a comparação de correspondência, organizando a correlação entre diretrizes experimentais e o artigo 102 da BCB nº 304, selecionado para compor o teste de análise de conformidade. A saída contempla aspectos como governança corporativa e gestão de riscos, segundo o estabelecido nos incisos do artigo 102, evidenciando pontos de convergência e divergências de maneira clara e objetiva.

Além de identificar a aderência entre documentos normativos, a ferramenta fornece recomendações detalhadas para aprimorar a compatibilidade regulatória, detectando lacunas e inconsistências nas diretrizes internas. Isso possibilita ajustes estratégicos que elevam a qualidade dos processos internos, fortalecem a governança corporativa e asseguram o alinhamento da B3 às exigências de compliance, reduzindo riscos institucionais e prevenindo impactos financeiros negativos.

Em síntese, os testes indicaram que a ferramenta possui capacidade de evolução contínua, incorporando melhorias baseadas em feedbacks dos usuários e nos avanços da tecnologia de LLMs. Essa adaptabilidade garante que a solução permaneça alinhada às constantes mudanças regulatórias, consolidando-se como um elemento inovador na transformação digital empresarial.

#### 4. CONTRIBUIÇÕES

Este projeto representa um avanço significativo na aplicação de inteligência artificial no setor financeiro, demonstrando como a tecnologia pode otimizar processos complexos e garantir a conformidade com as exigências legais e empresariais. Ao unir academia e indústria, buscamos não apenas resolver desafios específicos da B3, mas também contribuir para o desenvolvimento de soluções inovadoras aplicáveis em outros contextos de compliance e corporativos.

Um dos benefícios mais expressivos da solução foi a significativa economia de tempo nas análises de aderência às diretrizes. Métodos tradicionais exigem que analistas jurídicos e demais equipes realizem leituras minuciosas e manuais, demandando dias ou até semanas para verificar a conformidade de políticas internas com as exigências dos órgãos fiscalizadores. Com a adoção da ferramenta baseada em LLMs, esse processo foi reduzido para minutos ou horas, dependendo do volume de documentos analisados. A automação permitiu uma redistribuição mais eficiente dos recursos humanos, direcionando os especialistas para a interpretação dos resultados fornecidos pelo sistema.

Outro impacto relevante foi a possibilidade de escalabilidade da solução. A tecnologia desenvolvida pode ser aplicada a outras instituições financeiras e até mesmo a outros setores altamente regulados, como seguradoras e empresas de telecomunicações, que também enfrentam desafios semelhantes na interpretação e adaptação de normativas externas. Isso abre novas oportunidades de colaboração e inovação no ecossistema de regulação e compliance.

Além dos benefícios diretos para a B3, o impacto social da solução também merece destaque. Ao tornar os processos mais ágeis e precisos, a ferramenta contribui para um ambiente de negócios mais transparente e seguro, aumentando a confiança de investidores e clientes. Com esses avanços, o projeto estabelece a consolidação da inteligência artificial como um elemento-chave na inovação e modernização de processos empresariais.

Por fim, a parceria entre a B3 e a UFPB proporciona aos alunos inúmeras oportunidades de crescimento profissional, permitindo que adquiram experiências valiosas ao colaborar com uma empresa nacional com alcance global. Essa colaboração facilita a inclusão dos estudantes no mercado de trabalho, oferecendo visibilidade e reconhecimento pelo trabalho desempenhado com excelência. As iniciativas desenvolvidas conjuntamente abrem portas para futuras carreiras, destacando o talento dos alunos e fortalecendo sua inserção em um ambiente competitivo. Assim, o trabalho em conjunto da B3 e a UFPB enaltece o sistema de ensino superior no Brasil, valorizando e potencializando as habilidades dos estudantes para que se destaquem no mercado profissional.

## **AGRADECIMENTOS**

Agradecimentos especiais à Universidade Federal da Paraíba, aos laboratórios acadêmicos TRIL e LASER e a B3 S.A. - Brasil, Bolsa, Balcão, pelo suporte essencial e incentivo ao desenvolvimento deste trabalho.

## REFERÊNCIAS

- [1] **BRASIL. Banco Central do Brasil. Resolução BCB nº 304, de 20 de março de 2023.** Dispõe sobre [...]. Diário Oficial da União: seção 1, Brasília, DF, 22 mar. 2023. Disponível em: <https://www.bcb.gov.br/estabilidadefinanceira/exibenormativo?tipo=Resolu%C3%A7%C3%A3o%20BCB&numero=304>. Acesso em: 18/02/2025.
- [2] **DEVLIN, J.; CHANG, M.; LEE, K.; TOUTANOVA, K.** BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*, arXiv:1810.04805, 2019. Disponível em: <https://arxiv.org/abs/1810.04805>. Acesso em: 18/02/2025.
- [3] **JOHNSON, J.; DOUZE, M.; JÉGOU, H.** Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, v. 7, n. 3, p. 535-547, 2019. Disponível em: <https://arxiv.org/abs/1702.08734>. Acesso em: 18/02/2025
- [4] **OpenAI.** Hello GPT-4o. 2024. Disponível em <https://openai.com/index/hello-gpt-4o/>. Acesso em: 20/02/2025
- [5] **OpenAI.** New embedding models and API updates. 2024. Disponível em <https://openai.com/index/new-embedding-models-and-api-updates/>. Acesso em: 20/02/2025
- [6] **NEELAKANTAN, A; WENG, L.; POWER, B.; JANG, J.** Introducing text and code embeddings. 2022. Disponível em <https://openai.com/index/introducing-text-and-code-embeddings/>. Acesso em: 20/02/2025