



PÉROLA: INTELIGÊNCIA ARTIFICIAL PARA DADOS DE SAÚDE ARQUITETURA, APLICAÇÕES EM PRODUÇÃO E PERSPECTIVAS

Cicero Isidro da Silva, Rodrigo Hiroto Moraes, Luiza Volfa, Ricardo Evandro Pereira de Oliveira.
Centro de Estudos e Pesquisas Dr. João Amorim (CEJAM), Gestão da Informação, São Paulo, Brasil

Eixo Temático: LT5 - Sistemas Digitais e Dados em Saúde

Introdução

Evoluções clínicas contêm dados extremamente valiosos que raramente são registrados em campos de dado estruturado. Elas também contêm informações pessoais de saúde altamente sensíveis, protegidas pela Lei Geral de Proteção de Dados (LGPD), que estabelece requisitos regulatórios rígidos no *Capítulo II: Do Tratamento de Dados Pessoais*.

Além disso, essas evoluções às vezes apresentam textos relacionados a violência que violam as *guardrails* (camadas de proteção) estabelecidas pelos provedores de serviços de IA. Tais sistemas possuem etapas que envolvem a revisão humana manual, o que violaria a LGPD.

Nós, da equipe de Gestão da Informação do CEJAM, desenvolvemos a *Pérola* como uma arquitetura de IA que respeita as regras definidas pela LGPD e garante governança e auditoria centralizadas.

Objetivo

Descrever a arquitetura da *Pérola*, aprovada para uso em dados de saúde, as aplicações já em produção e as perspectivas para a gestão.

Métodos

Relato de implantação, em três camadas.

Conformidade. A Microsoft aprovou o acesso *Modified Guardrails* (código 3548301, em 08/05/2026) e *Modified Abuse Monitoring* (código 3559556, em 09/05/2026). Esses acessos garantem que o fornecedor não armazena dados sensíveis e que não há revisão humana. Devido a falta dos modelos mais modernos na *Data Zone* (DZ, Zona de Dados) brasileira, a implantação é feita na DZ Europeia, com a maior parte do processamento feito na Alemanha, e a transferência internacional apoia-se em dois fundamentos independentes da LGPD: a adequação Brasil-UE (Resolução CD/ANPD nº 32, de 26/01/2026) e as cláusulas-padrão contratuais.

Arquitetura. A *Pérola* define o acesso aos modelos de IA através de um Gateway *Bifrost* centralizado. Isso garante a governança de acessos, custos, e auditoria de todas as chamadas feitas para cada modelo.

Agente. Sobre ele roda a agente *Pérola*, construída a partir do software *Pi*, em contêiner *sandbox Docker*, com diferentes perfis de permissão e memória compartilhada entre sessões de usuários em um banco SQL.

Conclusão

Com aprovação formal do fornecedor, base legal explícita e um ponto único de acesso auditado, a *Pérola* opera modelos de linguagem sobre texto clínico em produção. A mesma arquitetura abre caminho para qualificação e auditoria de bases, descoberta de dados por agente dedicado e agentes ligados ao lakehouse para outras finalidades.

Resultados

Conformidade

O Art. 33 da LGPD, permite o trânsito internacional de dados apenas “para países ou organismos internacionais que proporcionem grau de proteção de dados pessoais adequado ao previsto nesta Lei” e quando “há garantias de cumprimento dos princípios, dos direitos do titular e do regime de proteção de dados”.

A Resolução CD/ANPD nº 32, de 26/01/2026 reconhece a União Europeia “como organismo internacional com grau de proteção de dados pessoais adequado ao previsto na Lei nº 13.709, de 14 de agosto de 2018, para fins de transferência internacional de dados”, permitindo que modelos de IA com residência na DZ Europeia da Azure sejam utilizados no processamento de dados de saúde de Brasileiros.

Os acessos modificados para *Abuse Monitoring* e *Guardrails* nos garantem contratualmente que esses filtros de segurança não são aplicados aos dados sob a tutela do CEJAM. Esse acesso foi conquistado devido a posição da instituição como organização de saúde e a parceria entre a instituição e a Azure Brasil.

Arquitetura

Nossa arquitetura de IA conta com um Gateway (*Bifrost*) centralizado que faz a gestão de acessos, controle de custos e auditoria de uso. É possível assim, estabelecer limites de gasto e acompanhar em detalhes a utilização da IA, com uma auditoria histórica completa e detalhada.

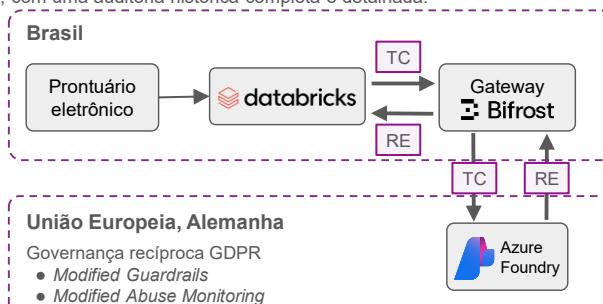


Fig 1. Caminho do texto clínico (TC) até o modelo e de volta ao indicador como resposta estruturada (RE), com a única travessia de fronteira.

Agente

Dentro dessa arquitetura, o time de informação faz uso do agente de IA *Pérola*, criada a partir do *Pi*, restrita a um *sandbox Docker*, com ferramentas e permissionamento limitados, fortalecendo a segurança contra *prompt injections* (injeções de prompt) e *data leaks*.

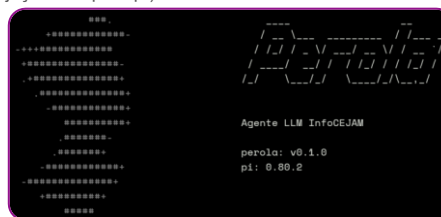


Fig 2. Captura de tela do agente de IA *Pérola*

Acesse o trabalho
na íntegra!

