Gestão de Identidade de Agentes de IA no Contexto Federado

Maria Amanda de Freitas Morais[1], Luiza Kuze Gomes[2], Estefânia Angelico Pianoski Arata[3], Fiterlinge Martins de Sousa[3], Reinaldo Cézar de Morais Gomes[3]

[1]Universidade Federal do Rio Grande do Norte (UFRN)

[2]Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina (IFSC)

[3]Rede Nacional de Ensino e Pesquisa (RNP)

luiza.k12@aluno.ifsc.edu.br, amanda.morais.110@ufrn.edu.br,

{estefania.arata, fiterlinge.sousa, reinaldo.gomes}@rnp.br

Resumo. *A adoção crescente de agentes de inteligência artificial autônomos em ambientes de pesquisa científica introduz desafios inéditos para os sistemas de gerenciamento de identidade e acesso federados. O protocolo OAuth 2.0, foi projetados para humanos e não contempla cenários em que um agente age em nome de um usuário delegante e acessa recursos protegidos. Este trabalho propõe e avalia uma arquitetura de delegação federada para agentes de IA baseada no Token Exchange do OAuth 2.0 (RFC 8693), permitindo que pesquisadores deleguem suas permissões a agentes de IA com escopo reduzido.*

1. Introdução

Agentes de IA são capazes de executar ações de forma autônoma, utilizando LLMs para controlar seu comportamento. O uso crescente dessas aplicações tem revelado limitações nos protocolos de autenticação e autorização atualmente utilizados. Protocolos como OAuth 2.0 não oferecem suporte adequado para diversos desafios trazidos por esse novo paradigma. Entre eles, destacam-se o uso de credenciais estáticas (como chaves de API), com permissões amplas, por agentes com fluxos de ações imprevisíveis, bem como a representação exclusiva da identidade do usuário, sem considerar a identidade do próprio agente [Huang and Hughes 2025], gerando o risco de personificação.

Este trabalho propõe uma solução para integrar agentes de IA à federação de identidades utilizando o Keycloak 26.7.0 por meio de sua implementação do protocolo Token Exchange (RFC 8693). Esse protocolo tem como objetivo adicionar semântica de delegação aos tokens emitidos, permitindo a delegação de identidade com escopo reduzido e mantendo a separação entre a identidade do usuário e a do agente, de forma a evitar a personificação [Jones et al. 2020].

Durante a pesquisa foram identificadas limitações dessa implementação na versão utilizada do Keycloak, bem como quais práticas recomendadas de gerenciamento de agentes de IA são possíveis alcançar com ela. A arquitetura montada para o experimento permite realizar login com as credenciais da instituição de origem, escolher quais permissões serão delegadas, para qual agente e qual ferramenta será utilizada por ele.

2. Trabalhos Relacionados

Trabalhos relacionados sobre segurança e gestão de identidade em agentes de IA são apresentados na Tabela 1, em comparação com a solução proposta.

Tabela 1. Comparação com trabalhos relacionados sobre identidade e delegação para agentes de IA.

Referência	**Tipo**	**Foco principal**	**Lacuna em relação a este trabalho**
[South et al. 2025a]	Artigo	Delegação autenticada e rastreabilidade de agentes de IA	Sem implementação prática em identidade federada
[Aydeger et al. 2026]	Artigo	Identidades descentralizadas e credenciais verificáveis	Modelo descentralizado, não SAML/OIDC tradicional
[Fett et al. 2023]	Especificação	*Sender-constraining* de tokens via prova de posse de chave	Complementar; não resolve delegação por si só
[Google Cloud 2026]	Plataforma	Aplicação do Token Exchange para *workload*-to-cloud	Escopo restrito à nuvem do Google; sem federação acadêmica
[SPIFFE Project 2024]	Especificação	Identidade criptográfica de *workload* (SVIDs)	Foco em infraestrutura, não em usuário federado
Este trabalho	Artigo	Token Exchange via Keycloak em identidade federada (CAFe)	—

3. Delegação de Identidade para Agentes de IA

3.1. Agentes de IA

Agentes de Inteligência Artificial são sistemas que usam IA, funcionando como seu ”cérebro”, para raciocinar, planejar e agir com autonomia, integrando-se a ferramentas e serviços externos por meio de chamadas de API direcionadas por instruções de uma LLM [Huang and Hughes 2025]. Contudo, essa utilização de aplicações por agentes de IA expõe os serviços a ameaças de segurança que os protocolos de autenticação e autorização existentes não anteciparam, como personificação e uso indevido de credenciais delegadas.

3.2. Gerenciamento de Identidade

A troca da personificação pela delegação é um passo importante para a identidade de agentes de IA, explicitando que o agente age em nome do usuário e mantendo identidades distintas entre delegante e delegado [South et al. 2025b]. O OAuth 2.0, embora amplamente utilizado, não foi projetado para cenários em que o cliente é uma entidade autônoma, encadeando chamadas a subagentes em nome do usuário. O Token Exchange estende o OAuth 2.0 para suprir essa lacuna, permitindo trocar um token por outro sob

semântica de delegação, com claims que expressam a identidade do usuário e do agente autorizado, estrutura que pode ser aninhada em cadeias de delegação, possibilitando auditar em nome de quem o agente opera [Jones et al. 2020].

3.3. Arquitetura Proposta

Com base no uso do Token Exchange, este trabalho propõe uma arquitetura de delegação federada para agentes de IA (Figura 1), onde encontramos os seguintes componentes: (i) Keycloak 26.7.0, atuando como *Authorization Server* e emissor de tokens OAuth 2.0; (ii) SATOSA, como proxy de interoperabilidade, viabilizando a integração do Keycloak ao Serviço de Descoberta da Federação; (iii) um agente implementado com o framework Mastra (TypeScript/Node.js), atuando em nome de um usuário previamente autenticado na federação; (iv) um portal de delegação (Express.js), responsável pela orquestração do processo; e (v) uma API protegida (Express.js), que valida os tokens antes de conceder acesso.

O código, desenvolvido com auxílio do Claude Code e revisado pelos autores, está disponível em: https://github.com/mariamandafm/wgid-ai-agent.

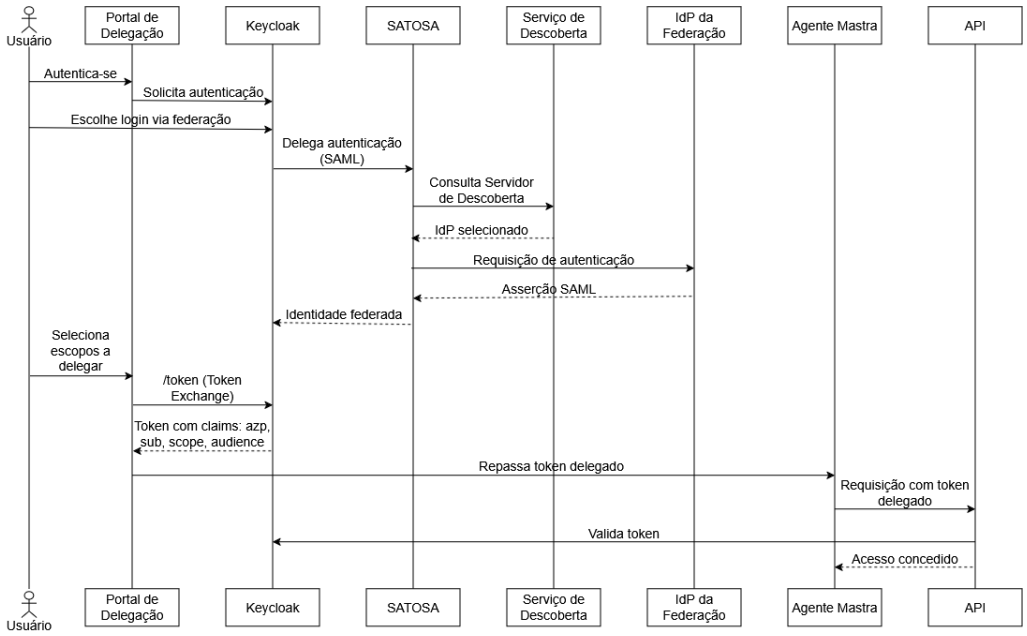

Figura 1. Fluxo da solução

3.4. Metodologia

Para validar a proposta, foi implementado o fluxo de Token Exchange do Keycloak 26.7.0 a partir da identidade federada do usuário. O usuário autentica-se no portal de delegação com a identidade mediada pelo proxy SATOSA, após escolher os escopos a serem delegados, o portal obtém junto ao Keycloak, por meio do endpoint */token*, um novo token com escopo reduzido e os seguintes claims de delegação: (i) *azp*, identificando o agente; (ii) *sub*, preservando a identidade federada do usuário; (iii) *scope*, com as permissões reduzidas do agente; e (iv) *audience*, definindo a aplicação alvo (a API). A API valida o token junto ao Keycloak, reconhece o contexto de delegação e concede acesso restrito ao escopo delegado.

4. Resultados e Discussões

O fluxo atualmente suportado pelo Keycloak 26.7.0 permite tratar aspectos relevantes da gestão de identidade para agentes de IA, como a identificação do agente responsável pela requisição, delegação explícita, redução de privilégios e auditoria das permissões delegadas e de seus respectivos destinatários. No entanto, esse fluxo apresenta uma limitação significativa: a ausência de suporte à cadeia de delegação prevista na definição do protocolo Token Exchange, ainda não implementada pelo Keycloak. Como consequência, delegações indiretas permanecem ocultas, impossibilitando um rastro auditável completo e comprometendo a rastreabilidade das ações executadas em nome do usuário. No OAuth 2.0 tradicional, a redução de privilégios é apenas parcial, pois os escopos são definidos na emissão inicial do token e reutilizados pelo agente. Da mesma forma, a auditoria também é parcial, pois, embora permita identificar o usuário e os escopos concedidos, não representa explicitamente a identidade do agente responsável pela execução das ações.

5. Conclusão

A proposta avalia a utilização do OAuth 2.0 Token Exchange por meio do Keycloak em uma federação acadêmica aos moldes da Comunidade Acadêmica Federada da RNP para permitir a delegação federada de identidades para agentes de IA. A solução desenvolvida demonstrou que é possível delegar permissões federadas a agentes de IA com escopo reduzido e separação entre a identidade do agente e a do usuário delegante. A avaliação evidenciou, contudo, uma limitação relevante na implementação do Token Exchange pelo Keycloak que impede a rastreabilidade completa da cadeia de delegação no token resultante.

Referências

Aydeger, A., Zeydan, E., Mangues-Bafalluy, J., Turk, Y., Khowaja, S. A., and Dev, K. (2026). Decentralized digital identity management for large language model agents. *IEEE Communications Standards Magazine*.

Fett, D., Campbell, B., Bradley, J., Lodderstedt, T., Jones, M., and Waite, D. (2023). OAuth 2.0 Demonstrating Proof of Possession (DPoP). RFC 9449.

Google Cloud (2026). Workload Identity Federation.

Huang, K. and Hughes, C. (2025). *Securing AI Agents: Foundations, Frameworks, and Real-World Deployment*. Springer Nature.

Jones, M. B., Nadalin, A., Campbell, B., Bradley, J., and Mortimore, C. (2020). OAuth 2.0 Token Exchange. RFC 8693.

South, T., Marro, S., Hardjono, T., Mahari, R., Whitney, C. D., Greenwood, D., Chan, A., and Pentland, A. (2025a). Authenticated delegation and authorized ai agents. *arXiv preprint arXiv:2501.09674*.

South, T., Nagabhushanaradhya, S., Dissanayaka, A., Cecchetti, S., Fletcher, G., Lu, V., Pietropaolo, A., Saxe, D. H., Lombardo, J., Shivalingaiah, A. M., et al. (2025b). Identity management for agentic ai: The new frontier of authorization, authentication, and security for an ai agent world. *arXiv preprint arXiv:2510.25819*.

SPIFFE Project (2024). SPIFFE: Secure Production Identity Framework For Everyone. `https://github.com/spiffe/spiffe`. Accessed: 2026-08-05.