

MODELOS LLM E I.A. AGÊNTICA EM ODONTOLOGIA: APLICAÇÕES, EVIDÊNCIAS, RISCOS E DESAFIOS

Edson Laranjeiras* com colaboração de SMA**

RESUMO

Os modelos de linguagem de grande escala, conhecidos pela sigla LLM, e os sistemas de inteligência artificial agêntica também denominado Sistema Multi-Agente SMA ou Multi-Agent Systems MAS - vêm ampliando as possibilidades de automação, análise de informações e apoio à tomada de decisão em saúde. Enquanto as LLMs produzem e interpretam linguagem com base em padrões estatísticos, a inteligência artificial agêntica acrescenta mecanismos de planejamento, uso de ferramentas, memória, execução de tarefas e reavaliação. Este paper discute os fundamentos dessas tecnologias, suas aplicações potenciais na saúde e, especialmente, na odontologia, além de suas limitações éticas, técnicas e clínicas. A literatura demonstra desempenho relevante de modelos de linguagem em tarefas médicas padronizadas, mas ainda não comprova segurança ou efetividade suficiente para decisões autônomas em pacientes reais. Na odontologia, os usos mais promissores concentram-se na documentação clínica, educação em saúde, recuperação de evidências, gestão e apoio à triagem. Entretanto, persistem riscos relacionados a alucinações, vieses, privacidade, segurança cibernética, baixa explicabilidade e ausência de validação prospectiva. Conclui-se que LLMs e sistemas agênticos devem ser implementados de forma assistiva, supervisionada, auditável e delimitada, sem substituir o julgamento clínico do cirurgião-dentista.

Palavras-chave: Modelos de linguagem de grande escala; inteligência artificial agêntica; saúde digital; odontologia; segurança do paciente.

1 - INTRODUÇÃO

A Inteligência Artificial (I.A.) ou Sistema Cognitivo Artificial tem assumido papel crescente na produção, organização e interpretação de informações em saúde. Entre as tecnologias mais recentes destacam-se os modelos de linguagem de grande escala ou Large Language Models (LLMs), treinados em grandes volumes de dados textuais para prever e gerar sequências linguísticas usando a probabilidades matemáticas. Esses modelos são capazes de resumir documentos, responder perguntas, traduzir conteúdos, estruturar informações e produzir textos tecnicamente plausíveis.

Embora apresentem elevada capacidade linguística, as LLMs não devem ser confundidas com sistemas dotados de compreensão clínica equivalente à humana eles utilizam da probabilidade para gerar a próxima palavra de um texto por exemplo baseado no contexto que a LLM está incluída. Seu funcionamento baseia-se na identificação de padrões estatísticos, o que pode resultar em respostas incorretas, incompletas ou descontextualizadas se. Nesse sentido, a aplicação em saúde sempre exigirá validação, supervisão profissional da área de saúde e mecanismos de controle.

A Inteligência Artificial Agêntica vem propor uma ampliação do uso convencional dos grandes modelos de linguagem - LLMs. Em vez de apenas responder a uma solicitação, que é amplamente usado, geralmente um agente pode decompor objetivos, planejar etapas, consultar bases externas, utilizar ferramentas, executar ações, armazenar informações e reavaliar os resultados obtidos. Essa arquitetura pode favorecer a integração de tarefas complexas, mas também amplia os riscos de utilizá-la, pois erros produzidos em uma etapa podem ser propagados para as etapas seguintes.

* CD, Analista de Sistemas e Pós-grad. em Agentes e Sistemas Inteligentes-UFG. ** Sistema Multi-Agentes

Este trabalho tem como objetivo discutir o estado atual das LLMs e da IA agêntica em saúde e odontologia, destacando aplicações potenciais, evidências científicas, limitações e requisitos para uma implementação segura e eticamente responsável. (Figura 01).



Figura 01 – LLMs e Inteligência Artificial Agêntica.

2 – DESENVOLVIMENTO

2.1 Modelos de linguagem de grande escala e inteligência artificial agêntica

Os LLMs são modelos computacionais capazes de processar e gerar linguagem natural a partir de padrões aprendidos em extensos conjuntos de dados. Seu emprego inclui elaboração textual, classificação, sumarização, extração de informações e interação conversacional. Os chamados modelos fundacionais podem ser adaptados para diferentes tarefas e domínios, inclusive para o processamento de dados clínicos, imagens e registros estruturados (BOMMASANI et al., 2021).

A I.A. agêntica utiliza uma LLM como componente de coordenação ou raciocínio dentro de um sistema mais amplo, semelhante a um “Cérebro de um ser humano” para atingir os seus objetivos. Esse sistema pode incluir um objetivo definido, um plano de execução, ferramentas externas, memória, mecanismos de verificação e possibilidade de intervenção humana. Dessa forma, o agente não se limita a produzir uma resposta textual, mas pode, por exemplo, consultar uma diretriz, recuperar dados de um prontuário, aplicar regras clínicas e apresentar alternativas para validação profissional. Um exemplo é este próprio texto, o qual definimos junto ao Odontinho – nosso avatar para criar este paper porém, com a nossa visão e responsabilidade. Houve correções para incluí-lo do domínio da área da saúde.

Entre as abordagens relevantes está o método ReAct, que integra raciocínio e ação, permitindo que o modelo alterne entre a elaboração de etapas e a interação com ferramentas externas (YAO et al., 2023). Outra proposta é o Toolformer, voltado à utilização de ferramentas por modelos de linguagem em tarefas específicas (SCHICK et al., 2023). Essas estratégias evidenciam que o desempenho de uma LLM pode ser ampliado por meio de sistemas modulares. Entretanto, a autonomia operacional não significa confiabilidade clínica automática.

2.2 Evidências na área da saúde

A literatura recente indica que LLMs podem apresentar desempenho expressivo em tarefas de conhecimento médico. Singhal et al. (2023), ao avaliarem o modelo Med-PaLM, observaram resultados próximos ou superiores a referências estabelecidas em determinadas questões médicas. Apesar disso, os autores também identificaram limitações relacionadas à precisão, ao raciocínio e à segurança das respostas; isso reforça aquela fala inicial que a responsabilidade é do profissional.

Estudos com modelos avançados, demonstraram capacidade de responder a questões médicas padronizadas. Contudo, resultados obtidos em exames, bancos de perguntas ou casos simulados não comprovam, isoladamente, eficácia em pacientes reais. A prática clínica envolve dados incompletos, incerteza diagnóstica, comunicação interpessoal, contexto social e responsabilidade profissional.

Moor et al. (2023) destacam que os modelos fundacionais podem favorecer a criação de sistemas generalistas de I.A. médica, capazes de integrar diferentes tipos de informação. Entre as aplicações possíveis encontram-se a sumarização de prontuários, o apoio à documentação, a educação de pacientes, a tradução, a busca de evidências e a organização de dados clínicos. Thirunavukarasu et al. (2023), entretanto, ressaltam que a incorporação de LLMs em saúde deve considerar problemas de confiabilidade, privacidade, viés, responsabilidade e validação.

Neste contexto, IA Agêntica acrescenta complexidade a esse cenário. Um Agente Inteligente Artificial pode receber uma pergunta clínica, procurar recomendações em fontes selecionadas, cruzar informações, identificar contraindicações e elaborar uma resposta estruturada. Por outro lado, uma falha na recuperação de dados, na interpretação de uma diretriz ou na utilização de uma ferramenta pode comprometer todo o fluxo de trabalho e dependendo do caso pode levar a situações gravíssimas caso este agente não seja validado por um profissional.

2.3 Aplicações potenciais em odontologia

Na odontologia, o uso mais adequado de LLMs e Agentes Inteligentes, no estágio atual, é como suporte às atividades profissionais. A documentação clínica constitui uma das aplicações mais viáveis. Sistemas desse tipo podem transformar a fala em registros estruturados, organizar informações da anamnese e auxiliar na elaboração de notas clínicas. A revisão pelo cirurgião-dentista é indispensável para evitar omissões, erros de interpretação ou registros incompatíveis com o atendimento realizado.

Outra possibilidade consiste na educação e comunicação com pacientes. A tecnologia pode auxiliar na produção de explicações individualizadas sobre higiene bucal, doença periodontal, cirurgia, endodontia, prótese e cuidados pós-operatórios a chamada Literacia em Odontologia. O conteúdo, contudo, deve ser verificado quanto à adequação clínica, à linguagem e à compatibilidade com o plano de tratamento.

Sistemas baseados em recuperação aumentada por geração, conhecidos como Retrieval-Augmented Generation - RAG, podem consultar diretrizes, revisões sistemáticas, protocolos institucionais e publicações científicas antes de elaborar uma resposta. Na odontologia, esse mecanismo pode contribuir para a localização de evidências e para a organização de alternativas terapêuticas. Ainda assim, a simples apresentação de citações não garante que a fonte tenha sido interpretada corretamente ou que seja aplicável ao caso individual.

A triagem também pode ser apoiada por agentes capazes de organizar relatos de dor, edema, trauma ou sangramento e identificar sinais que indiquem necessidade de atendimento prioritário. Essa função não deve substituir o exame clínico, os exames de imagem ou a avaliação de urgência por profissional habilitado.

Os Sistemas Cognitivos Artificiais neste momento já apresentam aplicações relevantes na odontologia convencional, especialmente na análise radiográfica e na identificação de alterações como cárie, lesões periapicais e doença periodontal. Revisão sistemática conduzida por Khanagar et al. (2021) descreveu desenvolvimento e desempenho promissores da I.A. em diferentes áreas odontológicas. Entretanto, os resultados devem ser obrigatoriamente interpretados e com cautela devido à heterogeneidade metodológica, à utilização de bases de dados limitadas e à insuficiente validação externa.

Schwendicke, Samek e Krois (2020) apontam que a IA pode ampliar a eficiência e a qualidade de determinadas atividades odontológicas, mas também apresenta desafios relacionados à interpretação dos resultados, à responsabilidade e à integração com o trabalho clínico. No caso específico das LLMs agênticas, a evidência ainda é mais limitada, não havendo base suficiente para recomendar o uso autônomo desses sistemas no diagnóstico ou no planejamento terapêutico.

2.4 Limitações, riscos e requisitos de segurança

Um dos principais riscos destes citados modelos - LLMs é a produção de “alucinações”, isto é, informações falsas ou referências inexistentes apresentadas com aparência de segurança. Em saúde, esse problema pode induzir a condutas inadequadas. O risco é ampliado em sistemas agênticos, nos quais uma informação incorreta pode ser utilizada por outras etapas automatizadas.

Também devem ser considerados os vieses presentes nos dados de treinamento. Diferenças linguísticas, culturais, étnicas e socioeconômicas podem produzir desempenhos desiguais entre grupos populacionais. Além disso, respostas não determinísticas podem variar diante de entradas semelhantes, dificultando a padronização e a auditoria.

A privacidade representa outra preocupação central. Prontuários, fotografias intraorais, radiografias e informações pessoais são dados sensíveis e devem ser processados conforme requisitos legais, institucionais e éticos aplicáveis. O uso de ferramentas externas ou serviços em nuvem pode aumentar o risco de exposição não autorizada. Estes sistemas deverão seguir as diretrizes legais da Lei Geral de Proteção de Dados – LGPD no Brasil.

Sistemas agênticos também estão sujeitos a ataques de *prompt injection*, manipulação de documentos, acionamento indevido de ferramentas e falhas de integração com sistemas odontológicos. Por isso, recomenda-se a adoção de registros de auditoria, controle de permissões, validação das fontes, limites de escopo, confirmação humana antes de ações críticas e mecanismos de interrupção.

A Organização Mundial da Saúde (WORLD HEALTH ORGANIZATION, 2024) recomenda transparência, governança, proteção de dados, supervisão humana, avaliação contínua e responsabilização no desenvolvimento e no uso de modelos multimodais de I.A. em saúde. Esses princípios são igualmente aplicáveis à implantação de LLMs e agentes em clínicas odontológicas, hospitais, universidades e serviços públicos.

2.5 Lacunas científicas e agenda de pesquisa

A evidência disponível, no momento deste texto, ainda é insuficiente para determinar o impacto das LLMs e dos Sistemas Multi-agênticos sobre desfechos clínicos odontológicos. Grande parte dos estudos permanece baseada em benchmarks, casos artificiais, avaliações retrospectivas e provas de conceito. Em um breve futuro serão necessários estudos prospectivos em ambientes reais, com comparação entre agentes, profissionais e sistemas tradicionais.

Também devem ser desenvolvidos bancos de dados odontológicos multilíngues, representativos e anonimizados, além de protocolos para avaliação de equidade, segurança, custo-efetividade, satisfação dos pacientes e tempo de atendimento. É importante medir não apenas a acurácia das respostas, mas também a frequência de erros, a capacidade de detecção de incerteza e o impacto sobre as decisões profissionais.

Outra frente relevante consiste na criação de mecanismos para auditoria automática de referências, rastreabilidade das ações do agente, monitoramento contínuo e interrupção de processos inseguros. A pesquisa deve ainda estabelecer critérios para consentimento, governança institucional, responsabilidade profissional e integração com prontuários eletrônicos.

3 CONCLUSÃO

As LLMs e a I.A. Agêntica ou Multi-Agêntica representam tecnologias promissoras para a organização, recuperação e produção de informações em saúde. As LLMs apresentam capacidade relevante em tarefas linguísticas e em determinados testes de conhecimento médico, enquanto os sistemas agênticos ampliam essa capacidade ao incorporar planejamento, uso de ferramentas, memória e execução de fluxos compostos.

Na odontologia, as aplicações mais plausíveis no curto prazo concentram-se na documentação clínica, educação de pacientes, busca de evidências, gestão, pesquisa e apoio limitado à triagem. A IA convencional já apresenta resultados promissores na análise de imagens odontológicas, mas a literatura sobre agentes autônomos permanece insuficiente.

Os riscos relacionados a alucinações, vieses, privacidade, segurança cibernética, baixa explicabilidade e automação excessiva impedem, no momento, que esses sistemas sejam considerados “substitutos” dos profissionais da saúde em casos específicos. A adoção responsável exige supervisão humana, validação clínica, transparência, proteção de dados, registro de auditoria e delimitação clara das funções automatizadas.

Conclui-se que o potencial das LLMs e da I.A. Agêntica deve ser explorado prioritariamente em modelos assistivos, supervisionados e com escopo definido. O desenvolvimento de evidências prospectivas e de mecanismos robustos de governança será essencial para que essas tecnologias contribuam efetivamente para a segurança, a qualidade e a equidade da atenção odontológica.

REFERÊNCIAS

1. BOMMASANI, Rishi et al. *On the opportunities and risks of foundation models*. Stanford: Stanford Center for Research on Foundation Models, 2021.
2. ODONTINHO, Literacia no domínio da odontologia. <https://odontinho.com.br>. 2026.
3. KHANAGAR, Santosh B. et al. Developments, application, and performance of artificial intelligence in dentistry: a systematic review. *Journal of Dental Sciences*, v. 16, n. 1, p. 508-522, 2021. DOI: 10.1016/j.jds.2020.06.015.
4. MOOR, Michael et al. Foundation models for generalist medical artificial intelligence. *Nature*, v. 616, p. 259-265, 2023. DOI: 10.1038/s41586-023-05881-4.
5. SCHICK, Timo et al. *Toolformer: language models can teach themselves to use tools*. In: *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, 2023.
6. SCHWENDICKE, Falk; SAMEK, Wojciech; KROIS, Joachim. Artificial intelligence in dentistry: chances and challenges. *Journal of Dental Research*, v. 99, n. 7, p. 769-774, 2020. DOI: 10.1177/0022034520915714.
7. SINGHAL, Karan et al. Large language models encode clinical knowledge. *Nature*, v. 620, p. 172-180, 2023. DOI: 10.1038/s41586-023-06291-2.
8. THIRUNAVUKARASU, A. J. et al. Large language models in medicine. *Nature Medicine*, v. 29, p. 1930-1940, 2023. DOI: 10.1038/s41591-023-02448-8.
9. WORLD HEALTH ORGANIZATION. *Ethics and governance of artificial intelligence for health: guidance on large multi-modal models*. Geneva: World Health Organization, 2024.
10. YAO, Shunyu et al. *ReAct: synergizing reasoning and acting in language models*. In: *INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS*, 2023.