LEARNING ANALYTICS COMUNICACIONAL: MARCADORES DISCURSIVOS PARA IDENTIFICAÇÃO DE SOFT SKILLS EM EAD CORPORATIVO

Ana Paula Alves Bleck Duque[1], Gisélia Alves de Souza[1],
Sérgio Roberto Montoro[1], Walter Teixeira Lima Junior[2]
[1] Fatec Pindamonhangaba - Centro Paula Souza –
ana.duque@cps.sp.gov.br; sergio.montoro@cps.sp.gov.br; giselia.souza@cps.sp.gov.br
[2] Universidade Federal de São Paulo, UNIFESP – SJC, walter.lima@unifesp.br

Natureza do trabalho: Resultado de Projeto
Área do Conhecimento: Ciências Humanas

RESUMO

O uso crescente de plataformas de *Learning Analytics* (LA) na Educação a Distância (EAD) corporativa tem priorizado indicadores comportamentais quantitativos, como frequência de acesso e conclusão de atividades, em detrimento do potencial de dados discursivos para identificar competências socioemocionais (soft skills) dos participantes. Diante dessa lacuna, este artigo apresenta resultado parcial de projeto de pesquisa em desenvolvimento na Fatec Pindamonhangaba, voltado à proposição de um modelo de *Learning Analytics* Comunicacional (LAC), que integra Análise do Discurso e analítica de aprendizagem. A revisão sistemática da literatura, conduzida conforme o protocolo PRISMA (2020), evidenciou predominância de abordagens quantitativo-comportamentais e escassa sistematização de marcadores discursivo-linguísticos como evidência de soft skills. Propõe-se, assim, um instrumento preliminar que associa cinco competências socioemocionais a marcadores discursivos fundamentados em categorias consolidadas da Análise do Discurso, primeira etapa de um *roadmap* de pesquisa que prevê validação empírica futura e piloto funcional do LAC.

Palavras-chave: *Learning Analytics. Soft Skills*. Análise do Discurso. Educação Corporativa. Marcadores Discursivos.

1 INTRODUÇÃO

A crescente adoção de plataformas de Educação a Distância (EAD) corporativa tem impulsionado o uso de Learning Analytics (LA) como recurso estratégico para processos formativos (Castellanos-Reyes; Koehler; Richardson, 2023). A revisão sistemática conduzida neste projeto evidencia, contudo, que grande parte das aplicações de LA ainda se concentra em métricas de comportamento, como acesso e conclusão de atividades, limitando a compreensão de dimensões qualitativas da aprendizagem.

Essa limitação torna-se particularmente relevante diante da crescente demanda do mercado de trabalho por profissionais que aliem competências técnicas a habilidades como comunicação, colaboração, liderança e inteligência emocional, coletivamente denominadas soft skills (World Economic Forum, 2023; Mittal et al., 2022). Estudos recentes já apontam a necessidade de ampliar o escopo do LA para incorporar a dimensão comunicacional das interações educacionais, sob uma perspectiva socioconstrutivista atenta à interdependência e aos relacionamentos sociais entre os participantes (Castellanos-Reyes; Koehler; Richardson, 2023). Essa dimensão comunicacional é capaz de capturar evidências linguístico-discursivas que os indicadores comportamentais tradicionais não alcançam (Chopra Et Al., 2023; Pérez Barrera et al., 2026).

Este artigo apresenta resultado parcial da Etapa I (Eixo 1) de projeto de pesquisa em Regime de Jornada Integral (RJI) em desenvolvimento na Fatec Pindamonhangaba (Centro Paula Souza), cujo objetivo geral é desenvolver e validar uma plataforma de aprendizagem profissional integrada a ferramentas de Learning Analytics Comunicacional (LAC) para identificação de soft skills. Como etapa preliminar do projeto, conduziu-se revisão sistemática da literatura conforme o protocolo PRISMA 2020, que evidenciou predominância de abordagens quantitativo-comportamentais e escassa sistematização de marcadores discursivo-linguísticos como evidência de competências socioemocionais.

A partir dessa lacuna, o presente artigo tem como objetivo propor um instrumento preliminar de marcadores discursivos para identificação de soft skills em ambientes corporativos de EAD, articulando os achados da revisão sistemática a categorias consolidadas da Análise do Discurso. O instrumento constitui a primeira etapa de um roadmap de pesquisa que prevê, nas fases subsequentes, sua validação empírica e o desenvolvimento de um piloto funcional do LAC.

2 METODOLOGIA

Este projeto, situado na interseção entre tecnologia educacional e análise linguística, demanda um percurso metodológico capaz de articular rigor no mapeamento do estado da arte à proposição de contribuições originais. Adotou-se, para tanto, abordagem de revisão sistemática da literatura (RSL), conduzida conforme o protocolo PRISMA 2020 (Page et al., 2021), como etapa metodológica do projeto de pesquisa descrito na Introdução.

A busca foi realizada em seis bases de dados: *Scopus, ACM Digital Library, Web of Science (Core Collection), IEEE Xplore*, ERIC e SciELO. A consulta foi complementada por busca manual no Google Scholar, voltada à identificação de literatura cinzenta. A estratégia combinou os termos "*learning analytics", "educational data mining", "multimodal learning analytics*" e variantes, associados a termos de análise comunicacional/discursiva ("*discourse analysis", "social learning analytics*") e de competências socioemocionais ("*soft skills", "socioemotional competencies*"), unidos por operadores booleanos e adaptados à sintaxe de cada base.

Foram incluídos estudos publicados entre 2018 e 2026, em português, inglês ou espanhol, que tratassem de *Learning Analytics* associado a alguma dimensão comunicacional/discursiva e/ou identificação de soft skills em contextos de educação profissional. Excluíram-se estudos centrados exclusivamente em métricas quantitativas de acesso e desempenho, estudos de educação básica sem transposição prática para formação profissional, e registros de metadados de anais de conferência sem conteúdo de artigo individual.

A busca, realizada com apoio da plataforma *Rayyan* para gerenciamento e triagem cega, identificou 499 registros, dos quais 201 foram removidos por duplicidade. Os 298 registros únicos remanescentes foram triados por título e resumo, resultando em 217 exclusões e 81 estudos elegíveis para leitura na íntegra, etapa atualmente em curso. A partir da leitura preliminar dos estudos já processados, procedeu-se à construção do instrumento apresentado na seção de Resultados, articulando os achados da revisão a categorias consolidadas da Análise do Discurso (Halliday; Hasan, 1976; Brown; Levinson, 1987; Austin, 1962; Toulmin, 1958; Kerbrat-Orecchioni, 1980), tomadas como aporte teórico para a definição dos marcadores discursivos associados a cada soft skill.

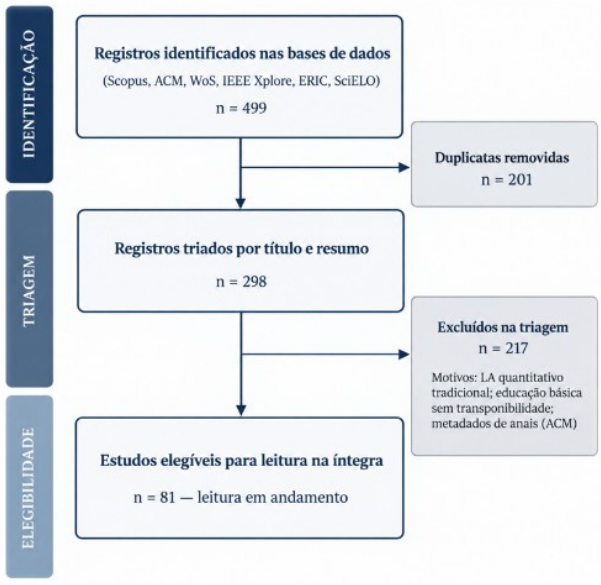
A seleção das cinco competências que compõem o instrumento não decorreu de frequência empírica na literatura revisada, mas de critério teórico-metodológico: cada soft skill foi associada a uma categoria da Análise do Discurso ou da Pragmática com correspondência conceitual direta e grau de operacionalização linguística já consolidado, como coesão referencial, dêixis, atos de fala, estrutura argumentativa e modalização epistêmica, privilegiando construtos com aparato analítico estabelecido na literatura linguística. Essa lista não é exaustiva; outras competências, como adaptabilidade, poderão ser incorporadas em etapas futuras do *roadmap*, à medida que novas categorias discursivas sejam mapeadas e testadas.

3 RESULTADOS

A aplicação do protocolo de busca e triagem descrito na Metodologia permitiu tanto mapear o tratamento dado pela literatura à interseção entre *Learning Analytics* e *soft skills* quanto fundamentar a proposição do instrumento apresentado nesta seção.

O processo de identificação e seleção dos estudos está sintetizado na Figura 1. Dos 499 registros identificados nas seis bases consultadas, 201 foram removidos por duplicidade, resultando em 298 registros triados por título e resumo. Dessa triagem, 217 registros foram excluídos e 81 estudos permanecem elegíveis para leitura na íntegra, etapa em curso.

Figura 1: Fluxograma PRISMA de processo de identificação, triagem e elegibilidade



Fonte: Elaborada pelos autores com auxílio da Inteligência Artificial (2026)

A análise dos motivos de exclusão revelou três padrões recorrentes: estudos de Learning Analytics quantitativo tradicional, baseados em métricas de acesso e desempenho, sem tratamento comunicacional (padrão predominante); estudos de educação básica, sem transponibilidade para formação profissional; e um artefato de exportação da ACM Digital Library, com metadados de anais indevidamente indexados como artigos, excluídos manualmente. Esses padrões constituem indício da lacuna que fundamenta este artigo: a literatura de Learning Analytics ainda trata a dimensão discursivo-comunicacional como exceção metodológica, não como eixo estruturante

Complementarmente à análise dos padrões de exclusão, procedeu-se a uma caracterização preliminar dos 23 estudos já processados, organizada em dois recortes independentes (Tabela 1): a abordagem metodológica predominante (Painel A) e a cobertura de cada soft skill do Quadro 1 (Painel B), não mutuamente exclusiva. Esse recorte corresponde a 28% dos 81 estudos elegíveis e deve ser lido como indicativo preliminar, não como resultado consolidado da revisão

Tabela 1 - Tabela 1 - Caracterização preliminar do subconjunto processado (n = 23)

Abordagem metodológica	**Nº de estudos**
Painel A — Abordagem metodológica predominante	
Processamento de Linguagem Natural (PLN) / mineração de texto	8
Análise de Redes Sociais ou Epistêmicas (SNA/ENA)	6
Abordagens multimodais (MMLA)	2
Métodos mistos ou outras técnicas quantitativo-estruturais	7
Painel B — Cobertura das soft skills do Quadro 1 (não mutuamente exclusiva)	
Colaboração	5
Empatia / Inteligência Emocional	2
Liderança	1

Abordagem metodológica	Nº de estudos
Painel A — Abordagem metodológica predominante	
Comunicação	0
Pensamento Crítico	0

Fonte: elaborado pelos autores (2026). Nota: os construtos foram identificados mesmo quando tratados sob rótulos distintos aos do instrumento (por exemplo, transatividade discursiva e modos participativos colaborativos para Colaboração).

Além dos construtos diretamente correlatos às cinco soft skills do instrumento, treze dos 23 estudos tratam de construtos adjacentes, como autoeficácia, identidade acadêmica, engajamento e presença social, sem nomeá-los no vocabulário de competências socioemocionais recorrente na literatura de empregabilidade. Essa distribuição reforça o diagnóstico central deste artigo: a literatura aborda dimensões afetivas e sociais da aprendizagem de forma consistente, mas raramente as sistematiza ou nomeia como soft skills, e duas das cinco categorias do instrumento proposto (Comunicação e Pensamento Crítico) não foram identificadas de forma explícita em nenhum dos estudos já analisados, o que reforça a lacuna que fundamenta a proposta deste artigo.

A partir desse diagnóstico e da leitura preliminar dos estudos já processados, propõe-se o instrumento apresentado no Quadro 1, que associa cinco soft skills a marcadores discursivos fundamentados em categorias consolidadas da Análise do Discurso, conforme critério de seleção detalhado na Metodologia. Cabe esclarecer que o instrumento não decorre de associações empiricamente identificadas nos estudos revisados, os quais, como discutido, não sistematizam marcadores discursivos para soft skills. Trata-se, antes, de uma construção teórica original, que articula categorias consolidadas da Análise do Discurso à tipologia de competências socioemocionais recorrente na literatura de empregabilidade (World Economic Forum, 2023; Mittal et al., 2022): uma proposta teórico-metodológica que preenche a lacuna diagnosticada na revisão sistemática, e não uma síntese descritiva dos achados desta.

Quadro 1 - Instrumento preliminar de marcadores discursivos para identificação de soft skills

Soft Skill	Marcador Discursivo	Exemplo Linguístico	Base Teórica
Comunicação	Coesão referencial e lexical; clareza na progressão temática	Uso consistente de referentes, sem ambiguidade; encadeamento lógico de ideias entre turnos	Halliday; Hasan (1976)
Colaboração	Dêixis inclusiva; estratégias de polidez positiva	Uso de "nós", "vamos", "podemos"; validação da fala do outro antes de discordar	Brown; Levinson (1987)
Liderança	Modalização deôntica; atos de fala diretivos	"É importante que...", "precisamos priorizar...", propostas de encaminhamento	Austin (1962); Bakhtin
Pensamento Crítico	Conectivos argumentativos de contraste/causalidade; marcas de ponderação	"Por outro lado...", "isso implica que...", contra-argumentação estruturada	Toulmin (1958)
Empatia/Inteligência Emocional	Modalização epistêmica; marcas de subjetividade e reconhecimento do outro	"Entendo que...", "acho que pode ser difícil...", validação emocional explícita	Kerbrat-Orecchioni (1980)

Fonte: elaborado pelos autores (2026).

O instrumento articula, para cada competência, uma categoria discursiva que possa ser identificada em registros textuais de interação (fóruns, chats, mensagens instantâneas), diferenciando-se das abordagens identificadas na triagem por operar sobre evidência linguística direta, e não sobre proxies comportamentais de acesso ou desempenho.

4 DISCUSSÃO

Os padrões de exclusão identificados na triagem, associados à leitura preliminar dos estudos elegíveis, corroboram o diagnóstico que fundamenta este artigo: a literatura sobre Learning Analytics aplicado à identificação de soft skills tende a tratar a dimensão comunicacional de forma fragmentária, sem consolidá-la em um instrumento sistemático. Ainda assim, elementos isolados dos estudos analisados dialogam com categorias específicas do Quadro 1, sugerindo que a proposta aqui apresentada não surge de um vácuo empírico, mas articula, de forma inédita, evidências dispersas na literatura.

A associação entre Liderança e atos de fala diretivos encontra ressonância em Ouyang e Dai (2022), que identificaram papéis participativos sociais (líderes, influenciadores, mediadores) como preditores do engajamento cognitivo colaborativo, sem propor instrumento linguístico-discursivo sistemático. De modo semelhante, Nasir et al. (2023) descrevem a transatividade discursiva, clarificação de metas e negociação de soluções, como marca do discurso colaborativo, sem nomeá-la como competência socioemocional. Mittal et al. (2022), por sua vez, evidenciam o papel mediador do Learning Analytics no desenvolvimento de habilidades de empregabilidade, corroborando a relevância prática do problema, ainda que sem propor os marcadores discursivos aqui apresentados

Essa constatação reforça o argumento central deste trabalho: a lacuna não reside na ausência de evidências sobre a relação entre discurso e competências socioemocionais, mas na ausência de um instrumento que sistematize tais evidências de forma replicável e aplicável a ambientes de EAD corporativo. É precisamente essa sistematização que o Quadro 1 busca oferecer, ao ancorar cada soft skill em categorias consolidadas da Análise do Discurso, e não em achados isolados e heterogêneos entre si.

Três questões metodológicas adicionais merecem esclarecimento explícito nesta etapa do trabalho. A primeira diz respeito à validade dos marcadores propostos, ou seja, à garantia de que um marcador discursivo identificado corresponda de fato à competência socioemocional associada, e não apenas a um estilo de escrita idiossincrático. Essa garantia ainda não existe, pois o instrumento, por ser preliminar, não foi submetido a nenhum teste de validade até o momento: prevê-se, na etapa de validação empírica do roadmap, verificação de confiabilidade entre avaliadores na aplicação dos marcadores a uma amostra de interações reais, além de análise de correlação entre a frequência dos marcadores e uma medida externa e independente da competência correspondente (avaliação por pares, por gestores, ou instrumentos já validados de competências socioemocionais). A segunda questão refere-se ao modo de aplicação, para o qual se prevê abordagem semiautomatizada, com codificação inicial assistida por modelos de linguagem de grande escala e validação manual por avaliadores humanos na fase de piloto, seguindo o precedente metodológico de estudos recentes que já demonstraram viabilidade técnica dessa abordagem (Chandler et Al., 2025; Chopra et Al., 2023).

A terceira questão, de caráter mais geral, é a distinção entre os resultados empíricos da revisão sistemática, apresentados na Tabela 1, e a proposição conceitual do Quadro 1: os dados da revisão evidenciam a lacuna e sua distribuição na literatura já mapeada, ao passo que o instrumento de marcadores discursivos constitui contribuição teórico-metodológica original que responde a essa lacuna, ainda não submetida a validação empírica direta. Cabe ressalvar, ainda, que os achados aqui discutidos derivam da leitura preliminar de 23 dos 81 estudos elegíveis, uma vez que a leitura na íntegra da totalidade do corpus permanece em curso: o instrumento deve, portanto, ser compreendido como resultado parcial e propositivo, o que não compromete a contribuição do artigo, mas delimita seu alcance como primeira sistematização teórica, não instrumento empiricamente validado.

A validação empírica constitui, assim, o próximo passo do roadmap de pesquisa que orienta este projeto, a ser conduzida por meio da aplicação do instrumento a corpora reais de interação em ambientes corporativos de EAD, com vistas ao desenvolvimento de um piloto funcional do Learning Analytics Comunicacional.

5 CONCLUSÃO

Este artigo apresentou resultado parcial de projeto de pesquisa voltado ao desenvolvimento de uma plataforma de Learning Analytics Comunicacional (LAC) para identificação de soft skills em ambientes corporativos de EAD. A revisão sistemática da literatura, conduzida conforme o protocolo PRISMA 2020, evidenciou que a produção científica sobre Learning Analytics ainda trata a dimensão comunicacional das interações educacionais como exceção metodológica, concentrando-se predominantemente em métricas comportamentais quantitativas. A partir dessa lacuna, propôs-se um instrumento preliminar que articula cinco soft skills (comunicação, colaboração, liderança, pensamento crítico e empatia/inteligência emocional) a marcadores discursivos fundamentados em categorias consolidadas da Análise do Discurso.

A principal contribuição deste trabalho reside na sistematização teórica dessa articulação, ainda inédita na literatura mapeada, que oferece um ponto de partida replicável para a identificação de competências socioemocionais a partir de evidência linguística direta. Reconhece-se, contudo, que a leitura na íntegra da totalidade dos 81 estudos elegíveis permanece em curso, o que delimita o instrumento proposto como resultado propositivo e preliminar, ainda não submetido à validação empírica.

O instrumento aqui apresentado constitui, portanto, a primeira etapa de um roadmap de pesquisa que prevê, nas fases subsequentes, sua validação empírica em corpora reais de interação e o desenvolvimento de um piloto funcional do Learning Analytics Comunicacional, consolidando a contribuição deste projeto para o avanço das tecnologias educacionais aplicadas à formação profissional.

REFERÊNCIAS

AUSTIN, J. L. **How to Do Things with Words**. Oxford: Oxford University Press, 1962.

BROWN, P.; LEVINSON, S. C. **Politeness: Some Universals in Language Usage.** Cambridge: Cambridge University Press, 1987.

CHOPRA, Harshita; LIN, Yiwen; SAMADI, Mohammad Amin; CAVAZOS, Jacqueline G.; YU, Renzhe; JAQUAY, Spencer; NIXON, Nia**.** Semantic topic chains for modeling temporality of themes in online student discussion forums. In: **PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON EDUCATIONAL DATA MINING**. Bengaluru: International Educational Data Mining Society, 2023. p. 67–78. DOI: 10.5281/zenodo.8115691.

HALLIDAY, Michael Alexander Kirkwood; HASAN, Ruqaiya**. Cohesion in English**. London: Longman, 1976. 374 p.

KERBRAT-ORECCHIONI, Catheri**ne. L'énonciation: de la subjectivité dans le langage**. Paris: Armand Colin, 1980.

MITTAL, Prabhat; KAUR, Amrita; JAIN, Renu. Online learning for enhancing employability skills in higher education students: the mediating role of learning analytics. **TEM Journal**, v. 11, n. 4, p. 1469–1476, 2022. DOI: 10.18421/TEM114-06.

NASIR, Afifah Quraishah Abdul; SHAH, Norshima Zainal; AB AZIZ, Nur Syafiqah**.** Insufficiency of soft skills among graduates within the services and manufacturing industries. **International Journal of Modern Trends in Social Sciences**, v. 6, n. 25, p. 25–38, 2023. DOI: 10.35631/IJMTSS.625003.

OUYANG, Fan; DAI, Xinyu; CHEN, Si**.** Applying multimodal learning analytics to examine the immediate and delayed effects of instructor scaffoldings on small groups' collaborative programming. **International Journal of STEM Education**, v. 9, n. 1, art. 45, 2022. DOI: 10.1186/s40594-022-00361-z.

PAGE, Matthew J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, v. 372, n. 71, 2021. DOI: 10.1136/bmj.n71.

PÉREZ BARRERA, Hendy Maier et al. Predicción del compromiso estudiantil mediante analítica del aprendizaje y análisis del discurso en foros Moodle. **Revista de Ciencias Sociales**, v. 32, Número Especial 13, p. 211–232, 2026. DOI: 10.31876/rcs.v32i.45756.

TOULMIN, Stephen Edelston. **The Uses of Argument**. Cambridge: Cambridge University Press, 1958. 264 p.

WORLD ECONOMIC FORUM**. The Future of Jobs Report 2023**. Geneva: World Economic Forum, 2023. ISBN 978-2-940631-96-4.