

CAPÍTULO 11

TECNOLOGIAS DIGITAIS, INTELIGÊNCIA ARTIFICIAL E DOCUMENTAÇÃO TERMINOLÓGICA EM LIBRAS

 <https://doi.org/10.68044/a4x8m861>

Jaqueline Boldo ¹

*Doutora em Linguística, pela Universidade Federal de Santa Catarina.
Docente de Libras (LSB/CCE) /UFSC.*

Leonardo Adonis de Almeida ²

*Mestre em Estudos de Tradução pela Universidade Federal de Santa Catarina e
Profssor Substituto de Libras na Universidade Federal de Santa Catarina.*

Paulo Vinicius Trevisan ³

Mestrando em Educação pela Universidade Estadual de Maringá (UEM).

Alberto Freiburger Bernardinelli ⁴

Doutorando em Educação pela Universidade Estadual de Maringá (UEM).

Valdicley Pereira Campos ⁵

Graduando em Licenciatura em Letras Libras/Português pela Universidade Federal do Amapá (UNIFAP). Graduado em Pedagogia pela Universidade Norte do Paraná (UNOPAR), em 2016. Especialista em Libras e Educação para Surdos pela Universidade Pitágoras Unopar.

Tainara Melo Santos ⁶

Mestranda em Estudos Linguísticos pela Universidade Estadual de Feira de Santana (UEFS). Especialista em Educação Bilingue para Surdos pela Universidade Federal do Cariri (UFCA) e graduada em Letras-Libras/Português pela Universidade Estácio de Sá.

RESUMO: Este artigo discute as relações entre tecnologias digitais, inteligência artificial e documentação terminológica em Língua Brasileira de Sinais (Libras), considerando os processos de formação, registro, organização, circulação e validação de sinais-termo. Trata-se de pesquisa bibliográfica e documental, de abordagem qualitativa, fundamentada em estudos sobre terminologia, lexicografia, glossários especializados, corpora linguísticos, visão computacional, reconhecimento automático e inteligência artificial. A análise evidencia que recursos digitais ampliam as possibilidades de registro audiovisual, organização de bases terminológicas, recuperação de informações e tratamento de grandes conjuntos de dados linguísticos. Entretanto, a utilização da inteligência artificial exige validação humana, critérios linguísticos e respeito à variação regional e sociocultural da Libras. Defende-se que sistemas tecnológicos devem atuar como instrumentos de apoio, e não como autoridades autônomas de legitimação linguística. Conclui-se que corpora multimodais, plataformas interoperáveis e processos participativos de validação podem fortalecer a documentação terminológica, desde que assegurem o protagonismo das comunidades surdas.

Palavras-chave: Libras; documentação terminológica; inteligência artificial; tecnologias digitais; sinais-termo.

INTRODUÇÃO

As transformações tecnológicas ocorridas nas últimas décadas modificaram profundamente os modos de produzir, registrar, armazenar, analisar e compartilhar conhecimentos linguísticos. No caso da Língua Brasileira de Sinais (Libras), essas mudanças assumem significado particular, uma vez que se trata de uma língua de modalidade visuoespacial cuja materialidade linguística envolve configurações de mãos, movimentos, pontos de articulação, orientações, expressões não manuais, relações espaciais e outros recursos que não podem ser adequadamente documentados apenas por meio da palavra escrita. A emergência e a popularização da gravação digital, dos bancos de vídeos, das plataformas on-line, dos dispositivos

móveis, da computação em nuvem, da visão computacional e, mais recentemente, da inteligência artificial ampliaram significativamente as possibilidades de constituição de acervos linguísticos e terminológicos em Libras.

A documentação lexical das línguas de sinais não nasceu com as tecnologias digitais. A história das produções lexicográficas demonstra a existência de diferentes tentativas de registrar sinais por meio de desenhos, gravuras, fotografias, descrições verbais e outros recursos gráficos. Sofiato e Reily (2014) recordam que, no Brasil, a *Iconographia dos signaes dos surdos-mudos*, publicada em 1875, constitui uma das primeiras referências da dicionarização da língua de sinais no país. Essa trajetória evidencia que a necessidade de registro acompanha a própria história das comunidades surdas e das práticas educacionais destinadas a elas. A diferença contemporânea está na possibilidade de representar o movimento, a espacialidade e a simultaneidade de forma audiovisual, tornando os registros progressivamente mais próximos da materialidade efetiva da língua.

Sofiato e Reily (2014) demonstram que a história da dicionarização da Libras esteve associada tanto à necessidade de descrição da língua quanto às demandas de ensino. As autoras registram que os dicionários foram se consolidando como materiais de consulta, ensino e aprendizagem, ainda que apresentassem limitações decorrentes das formas de representação visual disponíveis em

diferentes momentos históricos. O desenvolvimento tecnológico, portanto, não inaugura a documentação lexical da Libras, mas transforma as condições materiais sob as quais essa documentação pode ser realizada.

Essa transformação pode ser observada no deslocamento progressivo dos materiais impressos para recursos baseados em CD-ROM, DVD, páginas eletrônicas, glossários digitais, plataformas de vídeo e bancos de dados on-line. Carvalho e Sofiato (2020) identificam essa trajetória ao analisarem materiais lexicográficos e pedagógicos destinados à educação de surdos, ressaltando que diferentes suportes foram incorporados ao longo do tempo. Os autores mostram que o advento de recursos digitais permitiu maior valorização da imagem em movimento, embora também assinalem a necessidade de aperfeiçoamento das produções on-line e de sua articulação com dados provenientes de corpora linguísticos.

O cenário contemporâneo introduz uma mudança qualitativamente distinta. Não se trata apenas de utilizar tecnologias digitais para disponibilizar vídeos, mas também de empregar recursos computacionais para organizar, recuperar, comparar, anotar e analisar grandes quantidades de dados linguísticos. A documentação terminológica passa, assim, a dialogar com bancos de dados, linguística de corpus, processamento computacional, reconhecimento de padrões, aprendizado de máquina, modelos de inteligência artificial

e visão computacional. Esse movimento estabelece uma zona de contato entre Linguística, Terminologia, Lexicografia, Educação de Surdos, Ciência da Computação e Estudos da Tradução.

A necessidade dessa aproximação decorre, entre outros aspectos, da expansão da Libras para ambientes acadêmicos, científicos e profissionais. À medida que estudantes surdos passaram a frequentar diferentes níveis educacionais e campos especializados do conhecimento, tornou-se cada vez mais evidente a demanda por recursos terminológicos que permitissem expressar conceitos de Biologia, Química, Psicologia, Radiologia, Informática, Programação e inúmeras outras áreas. Albres (2023/2024) demonstra que o desenvolvimento terminológico da Libras está diretamente associado à possibilidade de sua utilização como língua de instrução e de aprendizagem nos diferentes campos do conhecimento.

A terminologia especializada, nesse sentido, não deve ser compreendida como simples tradução lexical de palavras da língua portuguesa para sinais. O sinal-termo se articula a conceitos especializados, aos usos realizados pelas comunidades discursivas, às propriedades linguísticas da Libras e às práticas socioculturais das comunidades surdas. A documentação desses sinais requer, consequentemente, procedimentos que ultrapassem a criação pontual de equivalências. É necessário registrar definições, variantes, contextos de utilização, fontes, áreas de especialidade, informações

linguísticas, responsáveis pela validação e, quando possível, ocorrências autênticas em corpus.

A revisão sistemática realizada por Albres (2023/2024) evidencia simultaneamente o crescimento e as lacunas desse campo. A autora analisou centenas de publicações e identificou uma produção relevante de pesquisas sobre sinais-termo, glossários, dicionários e lexicologia, mas também constatou a necessidade de aprofundamento conceitual e de diversificação das áreas contempladas. A própria autora ressalta:

A produção de sinais-terminos para as áreas do conhecimentos que são ensinadas nas escolas e universidades, haja vista que a “falta de sinais” é compreendida como um dos principais determinantes negativos do desenvolvimento da aprendizagem de alunos surdos [...] provocando a necessidade de criação de materiais e o registro em fichas terminológicas [...], contribuindo com os professores, alunos e intérpretes na interação em sala de aula. (Albres, 2023/2024, p. 3).

O trecho evidencia que a discussão terminológica possui consequências educacionais imediatas. Quando conceitos especializados não dispõem de formas lexicalizadas amplamente documentadas ou quando os profissionais envolvidos no processo educacional desconhecem as variantes existentes, surgem dificuldades adicionais para a mediação dos conteúdos. Isso, entretanto, não significa que todo conceito deva necessariamente corresponder a um sinal único, fixo e nacionalmente uniforme. Línguas naturais

apresentam variação, mudança, criatividade lexical e diferentes soluções discursivas. O desafio central da documentação contemporânea consiste justamente em registrar essa diversidade sem transformar o glossário em instrumento de apagamento da variação.

Ferreira et al. (2024) identificam situação semelhante ao investigar terminologias científicas. O estudo constatou tanto a escassez de sinais para determinados conceitos quanto a existência de formas diferentes para alguns itens disponíveis. Os autores defendem maior articulação entre grupos de pesquisa e a produção de ambientes compartilhados de consulta, de maneira que materiais já desenvolvidos possam ser conhecidos por estudantes, pesquisadores, docentes e tradutores-intérpretes.

A inteligência artificial acrescenta novas possibilidades e, simultaneamente, novos problemas a essa discussão. Algoritmos de visão computacional podem auxiliar na identificação de pontos corporais, mãos e face; modelos de aprendizagem de máquina podem classificar padrões; sistemas semiautomáticos podem sugerir anotações; e modelos generativos podem auxiliar na recuperação, sistematização e explicação de informações. Contudo, esses mesmos recursos podem reproduzir erros, apagar variantes, privilegiar dados de determinados grupos regionais, confundir gestos com estruturas linguísticas e produzir informações inexistentes quando utilizados sem validação especializada.

Galasso (2026) propõe uma das aproximações mais recentes entre inteligência artificial, corpus lexicográfico e variação da Libras. O autor parte justamente do problema da ausência de corpora lexicográficos de grande escala com anotação querológica e propõe uma arquitetura que articula Linguística de Corpus, Visão Computacional e Sociolinguística Computacional. A proposta utiliza estimativa de poses, representações vetoriais, anotação semiautomática, modelos computacionais e validação humana, preservando a variação linguística como objeto de investigação em vez de simplesmente tratá-la como erro ou ruído.

Paralelamente, experiências voltadas ao reconhecimento automático demonstram o crescimento desse campo. Morais et al. ([s.d.]) utilizaram recursos de aprendizado de máquina, Python, *scikit-learn* e MediaPipe para investigar reconhecimento de Libras. No conjunto experimental descrito pelos autores, foram extraídos pontos de referência das mãos, face e corpo e comparados modelos de aprendizagem de máquina, com resultados que indicaram melhor desempenho quando informações corporais e faciais foram incorporadas à representação dos sinais. Trata-se de um resultado relevante por enfatizar que a Libras não pode ser reduzida exclusivamente à configuração das mãos.

No campo da inteligência artificial generativa, Marques e Leite (2025) investigaram respostas produzidas por ChatGPT, Gemini e

Copilot acerca da Libras e da comunidade surda. No recorte analisado, as ferramentas foram capazes de produzir informações consideradas adequadas em diversas situações, mas também apresentaram conteúdos inexistentes ou incorretos. O estudo é particularmente relevante porque a avaliação contou com professor surdo de Libras e porque seus resultados reforçam a necessidade de verificação humana antes que respostas geradas por sistemas de inteligência artificial sejam utilizadas como materiais educacionais ou referências linguísticas.

Diante desse cenário, o presente artigo tem como objetivo discutir as relações entre tecnologias digitais, inteligência artificial e documentação terminológica em Libras, com atenção aos processos de formação, registro, organização, circulação e validação de sinais-termo. Busca-se compreender de que modo glossários digitais, dicionários eletrônicos, videotecas, corpora linguísticos, sistemas de reconhecimento e ferramentas de IA podem contribuir para a consolidação de infraestruturas terminológicas da Libras sem comprometer a diversidade linguística, a participação das comunidades surdas e a precisão conceitual.

A discussão é desenvolvida a partir de pesquisa bibliográfica e documental de abordagem qualitativa, tomando como corpus estudos sobre terminologia da Libras, lexicografia, glossários especializados, documentação em vídeo, tecnologias digitais, corpora, inteligência

artificial, reconhecimento automático e visão computacional. O percurso argumentativo considera a documentação terminológica como atividade linguística, tecnológica, educacional e sociopolítica, uma vez que a decisão sobre o que registrar, como representar e quem valida determinado sinal produz efeitos sobre a circulação e a legitimação das formas linguísticas.

Defende-se que tecnologias digitais e inteligência artificial possuem grande potencial para ampliar a escala, a precisão e a acessibilidade da documentação terminológica em Libras, mas sua contribuição depende da construção de modelos que reconheçam a língua em sua complexidade. A automação deve atuar como instrumento auxiliar de pesquisadores e comunidades, e não como instância autônoma de decisão linguística. Em especial, sistemas computacionais destinados ao registro lexical precisam incorporar variação regional, informações contextuais, histórico das formas registradas, critérios transparentes de validação e protagonismo de usuários surdos, evitando que eficiência computacional seja confundida com legitimidade linguística.

REFERENCIAL TEÓRICO

A expansão de uma língua para diferentes domínios sociais produz novas demandas lexicais. Quando uma língua passa a ser utilizada de maneira crescente em escolas, universidades, eventos

científicos, ambientes profissionais, serviços públicos e produção acadêmica, torna-se necessário expressar conceitos especializados que anteriormente podiam não circular com frequência em determinadas comunidades discursivas. Esse processo não representa insuficiência linguística, mas uma manifestação ordinária da dinamicidade lexical que caracteriza as línguas naturais.

No contexto da Libras, a ampliação do acesso de pessoas surdas à educação básica, ao ensino superior e à pós-graduação produziu novas condições para a circulação de discursos científicos. Termos próprios de diferentes campos passam a integrar aulas, traduções, interpretações, conferências, materiais didáticos e produções científicas sinalizadas. Consequentemente, cresce a demanda por identificação, análise, criação, negociação, validação e documentação de sinais-termo.

Albres (2023/2024) situa esse processo no interior das transformações educacionais vivenciadas pelas comunidades surdas. A autora explica que a consolidação de perspectivas bilíngues, nas quais a Libras ocupa posição de língua de instrução, implica desenvolvimento terminológico capaz de sustentar a mediação de conhecimentos provenientes de distintas áreas. A terminologia torna-se, portanto, componente da própria acessibilidade linguística.

A revisão realizada por Albres também mostra que as pesquisas terminológicas em Libras estão associadas à identificação

de processos de construção de novos sinais-termo, à elaboração de critérios terminográficos e à produção de dicionários e glossários. Isso demonstra que o trabalho terminológico não se encerra no momento de criação de um sinal. A forma precisa ser descrita, relacionada a um conceito, organizada em uma estrutura de consulta e contextualizada em relação aos usos realizados pela comunidade.

Outro aspecto importante é a diferença entre léxico geral e terminologia especializada. Um dicionário de uso geral busca representar parte ampla do vocabulário de uma língua, enquanto uma obra terminológica se concentra em conceitos e unidades linguísticas relacionadas a um domínio específico. Psicologia, Radiologia, Biologia, Química e Programação, por exemplo, possuem sistemas conceituais próprios. Um sinal-termo destinado a representar determinado conceito especializado precisa ser compreendido no interior dessas relações.

Essa característica explica por que a tradução direta de uma palavra em português não é suficiente. Um mesmo vocábulo pode assumir valores distintos em diferentes áreas, ao passo que conceitos técnicos podem exigir explicitação visual, classificadores, construções descritivas ou formas lexicalizadas específicas. Assim, a documentação terminológica deve começar pela compreensão conceitual e não pela busca de uma equivalência formal entre palavras e sinais.

A importância desse princípio pode ser percebida na afirmação recuperada por Albres ao discutir trabalhos anteriores. Tuxi observa que “o processo de criação dos sinais, assim como dos sinais-termo, é ainda uma área do conhecimento científico com poucas pesquisas realizadas e publicadas” (Tuxi, 2017, p. 51 *apud* Albres, 2023/2024, p. 3). Na mesma discussão, Moreira alerta que a ausência de bases conceituais pode comprometer a compreensão do conteúdo por pessoas surdas (Moreira, 2020, p. 1 *apud* Albres, 2023/2024, p. 3).

Essas observações são especialmente relevantes diante do crescimento de sistemas automáticos. Uma inteligência artificial pode localizar associações frequentes entre palavras e determinadas sequências visuais, mas frequência não equivale necessariamente a adequação terminológica. A decisão de considerar determinado sinal como correspondente a um conceito especializado requer análise das práticas discursivas, do domínio de conhecimento, das variantes e da comunidade que efetivamente emprega aquela forma.

A própria configuração do sinal-termo precisa ser considerada em sua dimensão linguística. Os dados lexicais de Libras envolvem parâmetros articulatórios e relações morfológicas e semânticas que podem participar da construção da unidade. Martins e Bidarra (2023), ao discutirem o desenvolvimento do DicELibras, analisam mecanismos e relações nos níveis morfofonético, morfossintático e semântico-lexical, demonstrando que a elaboração de uma estrutura

computacional para um dicionário de Libras precisa partir da organização linguística dos sinais.

Isso significa que a tecnologia deve ser modelada a partir da língua, e não a língua simplificada para se adequar à tecnologia. Um sistema construído apenas em torno de uma lista de palavras em português acompanhadas de vídeos corre o risco de reproduzir uma perspectiva bilíngue assimétrica, na qual a língua portuguesa organiza conceitualmente a base e a Libras aparece somente como representação complementar. Para superar essa lógica, bancos terminológicos podem incorporar mecanismos de busca diretamente relacionados aos parâmetros da Libras, às áreas conceituais, às variantes e às relações semânticas entre sinais.

A produção de sinais-termo também envolve processos coletivos de negociação. Martins e Stumpf (2016), em pesquisa sobre terminologias da Psicologia, coletaram sinais utilizados por profissionais surdos de diferentes regiões. O estudo registrou 23 autores da Psicologia e 28 sinais-termo, enfatizando justamente a importância de documentar variantes e disponibilizá-las à comunidade acadêmica surda.

A presença de variação mostra que a documentação não deve ser confundida com imposição de uma única forma. Determinados grupos podem desenvolver soluções diferentes para o mesmo conceito, e essas soluções podem circular simultaneamente. O registro

terminológico pode funcionar como instrumento para tornar visível essa diversidade, permitindo comparar formas, conhecer procedências regionais e compreender quais variantes circulam em determinados contextos.

Essa perspectiva é particularmente importante quando se discute “padronização”. Em terminologia científica, alguma estabilidade é necessária para favorecer a comunicação, especialmente em contextos educacionais e profissionais. Contudo, estabilidade comunicativa não significa obrigatoriamente eliminação das variantes. Uma plataforma terminológica pode indicar uma forma mais recorrente em determinado domínio, registrar alternativas regionais e apresentar informações acerca da validação e circulação de cada uma delas.

O desafio consiste em substituir a ideia de padronização absoluta por práticas de harmonização documentada. Nessa perspectiva, o objetivo não é declarar que apenas um sinal é correto, mas fornecer informação suficiente para que usuários compreendam as formas existentes e tomem decisões comunicativas fundamentadas. O registro histórico das alterações e das variantes torna-se, assim, tão relevante quanto a apresentação da entrada principal.

A documentação terminológica digital em Libras precisa ser compreendida a partir da tradição lexicográfica que a antecedeu. Leôncio e Zavaglia (2021) realizam um resgate histórico das

produções lexicográficas das línguas de sinais, identificando a passagem de listas e materiais elementares para manuais, glossários e dicionários mais complexos. As autoras demonstram que os diferentes formatos estão relacionados às necessidades sociais, educacionais e linguísticas de cada período.

No Brasil, Sofiato e Reily (2014) identificam a *Iconographia dos signaes dos surdos-mudos*, de Flausino da Gama, publicada em 1875, como marco histórico da produção lexicográfica de língua de sinais. O material continha centenas de verbetes ilustrados e organizados por categorias semânticas, utilizando gravuras para representar os sinais. A existência dessa obra demonstra que a relação entre visualidade, registro lexical e educação de surdos antecede amplamente a era digital.

Ao mesmo tempo, os limites da representação estática foram se tornando progressivamente evidentes. Um desenho ou fotografia pode representar uma configuração de mão ou determinada fase de um gesto, mas dificilmente registra com plena precisão aspectos temporais, mudanças de orientação, trajetórias, simultaneidade e componentes não manuais. Sofiato e Reily (2014) observaram dificuldades de estudantes ouvintes na interpretação de imagens encontradas em dicionários, mostrando que a visualidade estática não é automaticamente transparente.

As autoras acrescentam uma questão pedagógica decisiva:

Para ouvintes que se utilizam dos dicionários de línguas de sinais, os verbetes podem servir como suporte de memória [...] Entretanto, é importante lembrar que os dicionários contêm palavras isoladas e não combinadas em construções frasais específicas da língua. A tentativa de formar frases semelhantes ao português com sinais de Libras gera equívocos gramaticais, pois Libras e português são línguas de estruturas diferentes. (Sofiato; Reily, 2014, p. 112).

Essa observação continua pertinente para glossários digitais contemporâneos. A substituição do desenho pelo vídeo resolve parte do problema representacional, mas não transforma automaticamente uma coleção de sinais isolados em documentação linguística completa. Um vídeo pode mostrar adequadamente como um item lexical é articulado e ainda assim não informar como ele se comporta em construções sintáticas, como varia em diferentes contextos ou quais relações semânticas estabelece com outras unidades.

Por isso, o avanço tecnológico precisa ser acompanhado por avanço metodológico. Um dicionário eletrônico pode incorporar recursos impossíveis no material impresso: múltiplos vídeos de uma mesma entrada, câmera lenta, diferentes ângulos, exemplos contextualizados, variantes regionais, busca por parâmetros, hiperlinks entre conceitos, atualização contínua e ferramentas de pesquisa. Entretanto, a qualidade desse sistema depende da estrutura linguística e terminográfica que sustenta os dados.

Carvalho e Sofiato (2020) demonstram historicamente a passagem de materiais impressos para CD-ROM, DVD e recursos on-line. Na avaliação dos autores, os suportes digitais favoreceram a representação das línguas de sinais pelo movimento, mas as produções on-line ainda exigem aperfeiçoamento, inclusive por meio de maior fundamentação em corpora. Essa conclusão estabelece uma ponte direta entre lexicografia e tecnologias contemporâneas: digitalizar não significa simplesmente converter uma obra impressa em vídeo, mas explorar novas possibilidades de organização do conhecimento.

Albres (2023/2024) também recupera esse percurso ao observar a incorporação da videogravação, inicialmente em obras distribuídas em CD-ROM e posteriormente em glossários hospedados on-line. A autora identifica ainda uma produção realizada por integrantes das próprias comunidades surdas em blogs, canais de YouTube e redes sociais.

A multiplicação desses espaços produz uma situação paradoxal. De um lado, democratiza a produção e a circulação lexical, permitindo que grupos acadêmicos, professores, estudantes, intérpretes e comunidades regionais disponibilizem suas soluções. De outro, fragmenta os registros em dezenas ou centenas de páginas, canais e bases não interoperáveis. Um usuário que busca determinado sinal-termo precisa frequentemente consultar diversas fontes sem

saber se elas utilizam os mesmos critérios ou se registram variantes equivalentes.

Nesse contexto, o desenvolvimento do Dicionário Eletrônico de Libras, o DicELibras, discutido por Martins e Bidarra (2023), aponta para uma etapa mais sofisticada. Os autores não se concentram apenas na apresentação audiovisual dos itens, mas na construção de um modelo on-line monolíngue orientado por macroestrutura e microestrutura lexicográficas. A proposta analisa parâmetros dos sinais, relações morfossintáticas e semântico-lexicais e critérios para organização das entradas.

A implementação computacional, portanto, deixa de ser um elemento periférico e passa a fazer parte da própria arquitetura lexicográfica. Isso permite imaginar bancos de dados nos quais cada entrada não seja apenas “palavra + vídeo”, mas um objeto complexo formado por identificador, domínio, conceito, definição em Libras, definição escrita, forma sinalizada, variantes, informações articulatórias, contexto, relações semânticas, fonte, local de coleta e status de validação.

Esse tipo de estrutura é particularmente adequado à documentação terminológica porque possibilita atualizações sem reconstruir toda a obra. Um novo sinal registrado em determinada região pode ser associado à entrada existente; uma variante pode ganhar novos exemplos; uma definição pode ser revisada; e a história

das alterações pode ser preservada. A tecnologia digital transforma, desse modo, a obra lexicográfica de produto relativamente fechado em infraestrutura documental continuamente atualizável.

Os glossários especializados constituem uma das principais formas pelas quais a terminologia acadêmica e profissional da Libras vem sendo registrada. Diferentes trabalhos do corpus analisado apresentam iniciativas em Psicologia, Radiologia, Ciências Biológicas e Programação, demonstrando que a necessidade terminológica atravessa campos muito distintos do conhecimento.

Martins e Stumpf (2016) pesquisaram sinais-termo relacionados à Psicologia a partir de usuários surdos de diferentes regiões do país. A motivação incluiu justamente a ausência de registro de variantes terminológicas, situação capaz de interferir na negociação de sentidos e conceitos entre integrantes da comunidade acadêmica surda. O resultado foi posteriormente disponibilizado em glossário para facilitar o acesso.

Esse exemplo revela que um glossário pode desempenhar mais de uma função. Ele não é somente ferramenta de consulta imediata, mas também registro de formas linguísticas utilizadas em determinado momento histórico e por determinados grupos. Ao conservar variantes, o material adquire valor para investigações futuras sobre mudança lexical, difusão regional e processos de convencionalização.

Moreira e Murta (2021), por sua vez, desenvolveram glossário de termos da Radiologia destinado ao ensino técnico. Foram produzidos 31 vídeos com termos de uso cotidiano no setor, sinalizados em Libras e acompanhados de legendas em português. A equipe responsável incluía docentes, estudante ouvinte, estudante surdo do curso técnico de Radiologia, intérprete de Libras e profissional da área de informática.

A composição dessa equipe é significativa para a discussão sobre documentação terminológica digital. A produção exige competências distintas: conhecimento da área especializada, domínio linguístico da Libras, experiência de usuários surdos, tradução/interpretação e conhecimentos técnicos de produção e disponibilização digital. Não se trata, portanto, de tarefa que deva ser delegada exclusivamente ao profissional de informática ou exclusivamente ao especialista de determinada disciplina.

A participação de usuários surdos é especialmente relevante porque a legitimidade de um sinal não pode ser determinada apenas por correspondência conceitual elaborada externamente à comunidade linguística. A solução precisa ser linguisticamente viável, compreensível e passível de uso. A documentação deve conservar informação sobre quem utiliza e quem validou a forma, permitindo ao consulente interpretar adequadamente o estatuto daquele registro.

No campo da Programação, Melo *et al.* ([s.d.]) propõem glossário de termos técnicos destinado a estudantes surdos. O trabalho destaca a complexidade terminológica da Tecnologia da Informação e propõe que cada sinal seja associado a uma definição clara do conceito, em vez de se limitar à apresentação da forma sinalizada. Os autores utilizaram computadores, recursos de internet, programas de edição e planilhas para organizar o material.

A associação entre sinal e definição conceitual representa princípio fundamental da terminografia. Um estudante precisa saber não apenas como determinado termo é sinalizado, mas o que ele significa no sistema conceitual da disciplina. Essa necessidade torna-se ainda maior em áreas como Programação, nas quais uma palavra cotidiana pode assumir significação técnica muito específica.

Os glossários digitais podem, assim, ser compreendidos como interfaces entre língua, conhecimento especializado e tecnologia. A interface destinada ao usuário é apenas uma camada. Por trás dela precisa existir uma estrutura de dados coerente, capaz de documentar as relações entre os elementos. A maturidade de um projeto terminológico pode ser avaliada justamente pela riqueza dessa estrutura subjacente.

A partir dos estudos analisados, uma entrada terminológica digital robusta deveria permitir identificar o conceito representado, registrar definição contextualizada, apresentar o sinal em vídeo,

documentar variantes, informar região e grupo de usuários, registrar fontes de coleta, apresentar exemplos, relacionar itens semanticamente próximos, indicar responsáveis pela validação e conservar a data das diferentes versões. Essa síntese decorre da necessidade recorrente, identificada nas pesquisas, de articular representação visual, precisão conceitual, variação e possibilidade de atualização.

A existência de campos estruturados também possui importância para aplicações de inteligência artificial. Quanto mais organizada e semanticamente explícita for a base, maior a possibilidade de empregar sistemas computacionais para localizar relações, sugerir correspondências e identificar ocorrências semelhantes. Bases compostas exclusivamente por vídeos nomeados de modo irregular são muito menos úteis para análise automática do que recursos com metadados linguísticos e terminológicos bem definidos.

Uma diferença central precisa ser estabelecida entre uma coleção de vídeos e um corpus linguístico. Uma coleção pode reunir sinais selecionados previamente para determinada finalidade. Um corpus procura constituir um conjunto organizado de dados linguísticos passíveis de análise, acompanhado de critérios de seleção, informações contextuais, metadados e formas de anotação.

Silva-Oliveira, Wanderley e Stumpf (2020) apresentam contribuição importante ao propor o ENEM em Libras como corpus

linguístico para estudos léxico-terminográficos. A videoprova, segundo as autoras, pode ser compreendida como documento de registro da Libras em diferentes áreas do conhecimento. O material torna possível observar termos efetivamente empregados no processo tradutório da avaliação e utilizá-los como ponto de partida para a construção de glossários digitais.

A proposta é relevante porque desloca a documentação do modelo em que os sinais são produzidos exclusivamente para preencher entradas de glossário. Em um corpus, as unidades podem ser examinadas em contextos discursivos reais ou institucionalmente situados. Isso permite observar como são combinadas, quais recursos visuais acompanham a sinalização, como conceitos são explicados e quais soluções terminológicas foram escolhidas em determinadas circunstâncias.

O corpus favorece ainda a investigação da variação. Se diferentes produções forem incorporadas ao banco, torna-se possível comparar regiões, períodos, áreas acadêmicas e perfis dos sinalizantes. Um mesmo conceito pode aparecer sob formas distintas, e essas diferenças deixam de ser tratadas imediatamente como problema para se tornarem dados de pesquisa.

A relação entre corpus e glossário também modifica a noção de evidência terminográfica. Em vez de afirmar que determinado sinal existe simplesmente porque foi registrado por uma equipe, o sistema

pode apontar suas ocorrências documentadas, os contextos e a distribuição. Essa rastreabilidade fortalece a qualidade científica da base.

Albres (2023/2024) identifica crescimento de materiais disponíveis em plataformas eletrônicas, mas observa que muitos deles não foram objeto de estudos sistemáticos. A autora constata ainda que grande parte dos glossários documentados em pesquisas está ligada a projetos desenvolvidos em universidades, dissertações e teses.

Isso evidencia uma lacuna entre produção e documentação científica. Há sinais circulando em redes, aulas, vídeos e comunidades que podem não integrar bases terminológicas estruturadas. A ampliação de corpora permite aproximar esses dois universos, desde que coleta, consentimento, atribuição de autoria e informações contextuais sejam tratados com rigor.

O trabalho de Galasso (2026) avança justamente nessa direção ao propor um corpus lexicográfico de grande escala associado à visão computacional. O autor parte da dificuldade de construir corpora massivos por meio de anotação exclusivamente manual e discute um modelo semiautomático no qual a IA auxilia na extração de características e na sugestão de rótulos, mantendo validação humana.

Essa proposta altera substancialmente a escala possível da documentação. Anotações minuciosas de vídeos demandam elevado tempo de trabalho quando realizadas integralmente de maneira

manual. A visão computacional pode identificar pontos de referência no corpo e converter determinadas propriedades do movimento em representações numéricas. Sistemas de aprendizagem podem localizar padrões recorrentes e apresentar sugestões que posteriormente são verificadas pelos especialistas.

Entretanto, a dimensão mais relevante da proposta de Galasso (2026) talvez seja a decisão de considerar a variação como informação linguística. Em muitas aplicações de reconhecimento automático, diferenças regionais e individuais podem ser tratadas como ruído que reduz a acurácia do classificador. Para a Sociolinguística, exatamente essas diferenças constituem objeto de pesquisa. Um sistema de documentação precisa, portanto, preservar informações que um sistema convencional de classificação poderia tentar eliminar.

O problema evidencia que inteligência artificial e documentação possuem objetivos potencialmente distintos. Um reconhecedor automático pode desejar reduzir várias formas a um único rótulo para maximizar a taxa de acertos. Um corpus lexicográfico, ao contrário, precisa saber que formas diferentes ocorreram e sob quais condições. O desenho da tecnologia depende, conseqüentemente, da pergunta linguística que orienta o projeto.

A IA torna-se mais produtiva quando empregada para localizar, organizar e comparar dados, e não para decidir previamente quais diferenças são linguisticamente relevantes. Essa decisão requer

conhecimento especializado e participação dos sujeitos que utilizam a língua.

A análise das tecnologias aplicadas à Libras não deve restringir tecnologia a equipamentos ou programas. Schlindwein e Rocha (2023) propõem compreensão mais ampla ao discutir as relações entre tecnologias digitais e práticas linguísticas de produtores surdos de conteúdo. Para as autoras, tecnologia está inserida em relações sociais e influência, ao mesmo tempo em que é influenciada, pelas práticas dos usuários.

O impacto dos ambientes digitais sobre a Libras é descrito de maneira explícita:

As mudanças e inovações no campo das tecnologias digitais trouxeram novas formas de comunicação e aprendizagem, em especial para a comunidade surda brasileira. Ligações telefônicas com vídeo permitiram que a Língua Brasileira de Sinais (Libras) fosse a língua de expressão dos surdos. Canais de vídeos disponíveis em diversas plataformas tornaram possível a divulgação de expressões artísticas, informações científicas e culturais. (Schlindwein; Rocha, 2023, p. 1).

Esse cenário possui consequências diretas para a terminologia. Antes da popularização do vídeo digital, a circulação de um novo sinal dependia fortemente de interação presencial ou de materiais produzidos institucionalmente. Com plataformas de vídeo e redes sociais, sinais podem circular rapidamente entre cidades e estados, ser

discutidos publicamente e sofrer alterações em períodos relativamente curtos.

A terminologia deixa, desse modo, de ser produzida exclusivamente em projetos formais. Estudantes, professores, pesquisadores, profissionais e criadores surdos podem apresentar sinais utilizados em suas comunidades. A documentação acadêmica precisa reconhecer essa produção distribuída e desenvolver mecanismos de incorporação e descrição sem retirar das comunidades a autoria de suas criações.

Albres (2023/2024) observa justamente que integrantes das comunidades surdas criam glossários e os disponibilizam em blogs, canais de YouTube e redes sociais. Tal fenômeno amplia o repertório disponível, mas também coloca desafios relativos à descoberta, à preservação e à avaliação de qualidade.

Um vídeo postado em rede social pode desaparecer, ter endereço modificado ou não conter metadados suficientes para permitir sua recuperação anos depois. A documentação terminológica institucional pode atuar preservando referências e informações contextuais sem necessariamente retirar o conteúdo de seu ambiente original.

Esse movimento sugere a necessidade de infraestruturas que funcionem menos como dicionários centralizadores e mais como redes de documentação. Diferentes grupos poderiam registrar entradas em

uma plataforma comum ou interligar seus repositórios por padrões compartilhados. O usuário poderia visualizar as formas disponíveis e suas origens, em vez de receber uma única solução apresentada como universal.

A questão também envolve autoria e reconhecimento intelectual. Um sinal-termo pode resultar de processo coletivo desenvolvido em sala de aula, grupo de pesquisa ou comunidade profissional. Plataformas precisam conservar essas informações, evitando que sistemas automáticos apresentem sinais sem informar as pessoas ou os grupos responsáveis por sua documentação.

A inteligência artificial, quando treinada sobre vídeos disponíveis na internet, aumenta a importância desse cuidado. Um modelo pode reconhecer padrões sem explicitar de quais comunidades vieram os dados. Para a documentação linguística, entretanto, procedência é informação essencial. A eficiência algorítmica não deve apagar a história social das unidades registradas.

A aplicação da inteligência artificial às línguas de sinais reúne diferentes linhas de investigação. Entre elas encontram-se reconhecimento de sinais, tradução automática, detecção de mãos e corpo, extração de parâmetros articulatórios, segmentação de vídeos, classificação lexical, identificação de variantes e anotação semiautomática de corpora.

Uma das abordagens mais recorrentes utiliza visão computacional para converter imagens ou vídeos em conjuntos de pontos de referência. Esses pontos podem representar posições relativas das mãos, articulações corporais e regiões faciais ao longo do tempo. Em vez de analisar diretamente todos os pixels de uma imagem, determinados sistemas trabalham com séries de coordenadas que sintetizam parte da informação visual.

Morais *et al.* ([s.d.]) utilizaram MediaPipe para extrair referências corporais e treinar classificadores destinados ao reconhecimento de Libras. O experimento comparou uma abordagem baseada principalmente nas mãos com outra que incorporou informações adicionais de face e corpo. Segundo o resumo apresentado, a segunda abordagem obteve melhor desempenho no conjunto avaliado, sendo destacado o modelo Random Forest e uma medida F1 de 99,56%.

Esse percentual deve ser interpretado dentro das condições específicas do experimento e não como indicação de que o reconhecimento automático da Libras esteja resolvido. Bases controladas, conjuntos delimitados de classes e tarefas de classificação são muito diferentes da compreensão de sinalização contínua produzida espontaneamente em condições variadas.

Ainda assim, o estudo demonstra princípio relevante: sinais não podem ser reduzidos às mãos. Expressões faciais, movimentos

corporais e relações espaciais possuem papel linguístico. Sistemas que ignoram esses elementos correm o risco de perder distinções fundamentais.

Para a documentação terminológica, técnicas semelhantes podem auxiliar na indexação de sinais. A extração de características pode permitir buscas por semelhança articulatória, agrupamento preliminar de variantes e localização de ocorrências de determinada configuração em um corpus. O objetivo deixa de ser apenas traduzir uma sequência e passa a incluir organização e investigação linguística.

É nessa direção que o trabalho de Galasso (2026) se torna particularmente significativo. Sua arquitetura propõe um fluxo bifásico: inicialmente, informações visuais são extraídas dos vídeos e convertidas em representações quantificáveis; posteriormente, modelos computacionais apresentam sugestões de anotação. A validação permanece associada a especialistas, incluindo linguistas surdos, em processo *human-in-the-loop*.

O modelo *human-in-the-loop* representa alternativa importante tanto à anotação inteiramente manual quanto à automação irrestrita. A máquina realiza operações repetitivas em grande escala e identifica padrões, enquanto especialistas avaliam os resultados, corrigem erros e fornecem novos dados ao sistema.

Na documentação de Libras, essa distribuição de tarefas é especialmente apropriada porque determinadas decisões dependem de

interpretação linguística e sociocultural. Um algoritmo pode indicar que duas sequências são visualmente semelhantes, mas não necessariamente determinar se constituem variantes do mesmo item, sinais diferentes ou realizações contextuais.

Outro ponto relevante é a possibilidade de utilização de IA para investigação sociolinguística. Quando o corpus contém dados regionais e metadados adequados, algoritmos de agrupamento podem revelar padrões que posteriormente são analisados linguisticamente. Nesse caso, a IA não determina previamente categorias, mas ajuda a localizar regularidades.

Essa abordagem contrasta com sistemas desenhados para eliminar diversidade. Se o treinamento exige que toda ocorrência seja reduzida a uma única etiqueta, variantes legítimas podem ser classificadas como erros. Um projeto terminológico deve, ao contrário, tornar possível associar múltiplas realizações a um mesmo conceito sem apagar suas diferenças.

A arquitetura tecnológica precisa refletir essa decisão. O banco de dados pode conter uma camada conceitual, à qual se ligam diversas formas sinalizadas. Cada uma delas recebe informações próprias sobre região, contexto, frequência, fonte e validação. O reconhecimento automático pode sugerir qual forma se aproxima da ocorrência analisada, sem obrigar o sistema a transformar todas as variantes em um vídeo único considerado padrão.

A popularização de sistemas de inteligência artificial generativa introduziu outro conjunto de questões. Diferentemente dos sistemas especializados em visão computacional, ferramentas conversacionais geram textos, explicações, atividades, traduções e respostas a partir de grandes modelos treinados em volumosos conjuntos de dados.

No campo educacional, esses recursos podem auxiliar professores e estudantes na busca de informações, planejamento de atividades, produção de materiais e organização de conteúdo. Porém, quando a consulta envolve línguas de sinais, terminologia especializada ou culturas surdas, problemas de precisão tornam-se particularmente relevantes.

Marques e Leite (2025) investigaram ChatGPT, Gemini e Copilot em perguntas relacionadas à Libras e à comunidade surda. No estudo, respostas foram submetidas à análise de professor surdo de Libras. Os resultados indicaram a presença de conteúdos adequados e coerentes, mas também de informações inexistentes e incorretas. O desempenho variou entre as ferramentas e entre as respostas avaliadas.

O resultado impede que sistemas generativos sejam tratados como fontes terminológicas autossuficientes. Um modelo pode produzir uma explicação com aparência convincente mesmo quando a informação não está adequadamente fundamentada. Esse fenômeno é particularmente grave quando o usuário pergunta “qual é o sinal de

determinado termo”, pois a resposta pode atribuir validade a uma forma inexistente, regional ou descontextualizada.

A documentação terminológica oferece justamente uma possível solução para parte desse problema. Em vez de solicitar que o modelo produza a informação exclusivamente a partir de seus parâmetros internos, sistemas podem ser conectados a bases terminológicas verificadas. A IA atuaria como interface de recuperação e explicação, enquanto os dados linguísticos seriam provenientes de registros curados.

Esse modelo altera a função da IA generativa. Em lugar de “inventar” uma resposta, a ferramenta identifica entradas relevantes em uma base, apresenta as variantes disponíveis, informa a fonte e direciona o usuário ao vídeo correspondente. A confiabilidade passa a depender da qualidade da documentação e da transparência da recuperação.

Para isso, os registros precisam ser estruturados de modo compatível com processamento computacional. Informações como domínio, definição, palavras relacionadas, variantes e região podem melhorar a recuperação. A existência de identificadores estáveis permite que diferentes aplicações se refiram à mesma entrada.

A IA também pode ajudar em tarefas terminográficas preliminares: identificar termos candidatos em textos, organizar documentos, sugerir agrupamentos semânticos, comparar definições e

localizar vídeos potencialmente relacionados. Contudo, a seleção final e a validação linguística continuam exigindo especialistas.

A possibilidade de geração automática de avatares ou representações sinalizadas acrescenta outra camada de complexidade. Uma sequência visual tecnicamente bem formada não é necessariamente uma construção natural de Libras. Aspectos gramaticais, discursivos, prosódicos e culturais precisam ser considerados. A documentação terminológica não deve ser confundida com tradução automática integral.

Há ainda risco de circularidade. Se conteúdos gerados por IA forem publicados sem verificação e posteriormente incorporados a novas bases de treinamento, erros podem adquirir aparência de frequência e legitimidade. Bases terminológicas curadas e claramente identificadas tornam-se, nesse contexto, instrumentos para estabelecer pontos de referência verificáveis.

Um dos temas mais recorrentes nos estudos analisados refere-se à coexistência entre necessidade de estabilidade terminológica e variação linguística. Ferreira *et al.* (2024) encontraram ausência de sinais para diversos conceitos científicos e falta de uniformidade entre formas disponíveis. Os autores defendem maior colaboração entre grupos responsáveis pela produção de glossários.

No estudo, é proposta inclusive a criação de ambiente compartilhado no qual sinais de diferentes projetos possam ser consultados. Os autores afirmam:

Acreditamos que é necessário um trabalho em conjunto dos diversos grupos que pesquisam e produzem sinais em todo o Brasil, para que o registro deles seja realizado em ambiente compartilhado para que todos tenham acesso aos sinais produzidos. Tal ação pode contribuir para evitar o retrabalho, aproveitando o que já foi produzido, e para unificar sinais-termos científicos e acadêmicos. (Ferreira *et al.*, 2024, p. 13).

A proposta de ambiente comum é altamente pertinente, sobretudo porque a dispersão atual dificulta localizar materiais já existentes. Contudo, a ideia de “unificação” precisa dialogar com estudos sociolinguísticos que compreendem variação como característica constitutiva das línguas.

A contribuição de Galasso (2026) é importante nesse ponto porque desloca a variação do lugar de obstáculo para o lugar de dado linguístico. Um corpus nacional não deveria necessariamente eliminar regionalismos, mas registrá-los e permitir análise de sua distribuição.

A distinção entre unificação documental e uniformização linguística oferece caminho conceitual produtivo. Pode-se unificar a infraestrutura de acesso sem exigir uma única forma linguística. Uma plataforma nacional pode reunir diferentes variantes em um ambiente

comum, preservando sua origem e fornecendo informações que ajudem o consulente a compreender os usos.

O que deve ser padronizado prioritariamente são os procedimentos de documentação: campos obrigatórios, qualidade audiovisual, formas de identificar conceitos, metadados, critérios de coleta, informações de autoria e modos de registrar validação. As variantes linguísticas, por outro lado, podem ser documentadas em sua diversidade.

Esse modelo reduz ainda o problema do retrabalho apontado por Ferreira *et al.* (2024). Antes de criar um sinal, uma equipe poderia consultar a base e verificar formas já registradas. Se nenhuma delas fosse adequada ao contexto, um novo registro poderia ser acrescentado, acompanhado da justificativa e da comunidade de uso.

A plataforma deixaria de operar segundo uma lógica binária de “sinal correto” e “sinal incorreto”. Determinadas entradas poderiam receber informações como variante regional, forma utilizada em instituição específica, sinal em processo de validação, forma histórica ou ocorrência rara. A documentação torna-se mais informativa e menos normativa.

Esse princípio é coerente com os resultados de Martins e Stumpf (2016), cujo trabalho se justifica justamente pela necessidade de registrar variantes terminológicas na Psicologia. A existência de

variantes não impediu a elaboração do glossário; ao contrário, constituiu uma das razões para documentá-lo.

Albres (2023/2024) também chama atenção para a necessidade de maior participação de pesquisadores surdos e das comunidades surdas, não apenas na condição de informantes ou consulentes. Esse princípio deve ocupar posição central na governança de qualquer plataforma terminológica baseada em IA.

Sistemas automáticos aprendem a partir dos dados e das categorias fornecidas. Se a comunidade participa apenas ao final, validando decisões previamente estruturadas por outros grupos, parte importante de sua agência linguística já terá sido limitada. Pesquisadores surdos precisam participar da definição das categorias, das políticas de inclusão das variantes, dos critérios de qualidade e das prioridades de pesquisa.

CONSIDERAÇÕES FINAIS

A relação entre tecnologias digitais, inteligência artificial e documentação terminológica em Libras revela um campo de pesquisa em rápida transformação. A trajetória histórica das produções lexicográficas demonstra que a necessidade de registrar sinais antecede o ambiente digital, mas as tecnologias contemporâneas modificaram profundamente as possibilidades de representação, armazenamento, atualização, circulação e análise dos dados.

As obras lexicográficas impressas foram fundamentais para preservar e difundir parte do léxico da língua de sinais. Entretanto, a modalidade visuoespacial da Libras impõe limitações às representações exclusivamente estáticas. O vídeo digital possibilitou registrar movimento e expressões não manuais com maior fidelidade, enquanto os ambientes on-line tornaram possível atualizar e compartilhar conteúdos em escala anteriormente impraticável.

Os glossários especializados analisados demonstram a expansão da Libras para diferentes domínios científicos e profissionais. Psicologia, Radiologia, Ciências e Programação constituem exemplos de áreas nas quais pesquisadores identificaram necessidade de registrar e disponibilizar terminologias específicas. Esses trabalhos mostram que a documentação possui efeitos educacionais, profissionais e sociais, auxiliando estudantes, professores, profissionais surdos e tradutores-intérpretes.

Ao mesmo tempo, a pesquisa evidencia que um glossário terminológico não deve ser reduzido a uma coleção de vídeos correspondentes a palavras da língua portuguesa. A unidade de trabalho é o conceito especializado e sua realização linguística em Libras. O registro precisa conservar definição, contexto, variantes, procedência, autoria, parâmetros linguísticos e critérios de validação.

A incorporação dos corpora representa avanço metodológico importante. O ENEM em Libras demonstra como produções

audiovisuais podem funcionar como registros linguísticos e alimentar estudos léxico-terminográficos. O corpus permite deslocar o olhar do sinal produzido exclusivamente para o glossário para formas efetivamente utilizadas em contextos discursivos.

A inteligência artificial amplia significativamente esse potencial. Visão computacional, extração de pontos corporais, aprendizagem de máquina e sistemas de anotação semiautomática podem reduzir o volume de trabalho repetitivo necessário para tratar grandes quantidades de vídeos. A proposta de Galasso (2026) é particularmente representativa dessa mudança ao articular corpus lexicográfico, visão computacional, sociolinguística e validação *human-in-the-loop*.

Entretanto, a automação não elimina a necessidade de conhecimento linguístico. Os experimentos de reconhecimento demonstram que informações das mãos, corpo e face são necessárias para representar adequadamente sinais, ao passo que as pesquisas com inteligência artificial generativa evidenciam que sistemas podem produzir conteúdo plausíveis e ainda assim incorretos. A IA, portanto, deve ser utilizada como instrumento de apoio à documentação, e não como autoridade autônoma para determinar o léxico da Libras.

Os resultados também mostram que a variação linguística precisa ocupar posição central nas futuras infraestruturas terminológicas. A existência de diferentes sinais para determinado

conceito não constitui, por si mesma, um defeito a ser corrigido. Regionalismos, soluções institucionais e mudanças históricas fornecem informações importantes sobre a dinâmica da língua.

Nesse sentido, considera-se mais adequado unificar infraestruturas de documentação do que necessariamente uniformizar sinais. Uma plataforma nacional poderia reunir materiais provenientes de diferentes grupos e permitir consulta integrada, preservando as variantes e indicando seus contextos de uso. Essa organização diminuiria duplicações de esforços sem transformar a diversidade em erro.

Ferreira *et al.* (2024) sugerem precisamente a construção de ambiente compartilhado que reúna sinais produzidos por diferentes grupos. A contribuição de uma infraestrutura desse tipo seria ampliada caso incorporasse mecanismos de versionamento, rastreabilidade, metadados regionais e participação contínua das comunidades responsáveis pela produção das formas.

A inteligência artificial poderia atuar sobre essa base em múltiplas tarefas: localização de vídeos semelhantes, sugestão de anotações, agrupamento preliminar de variantes, recuperação semântica, identificação de termos candidatos e apoio à pesquisa. Sistemas generativos, por sua vez, poderiam funcionar como interfaces conversacionais para a base, desde que suas respostas

fossem fundamentadas nos registros documentados e apresentassem fontes verificáveis.

Uma arquitetura dessa natureza reduziria o risco de que o usuário receba um sinal inventado ou descontextualizado como se fosse consensual. Em vez de responder simplesmente “este é o sinal”, o sistema poderia informar que determinado conceito possui diferentes formas documentadas, apresentar regiões ou grupos associados a elas e disponibilizar os vídeos correspondentes.

A análise conduz ainda à necessidade de protagonismo surdo em todas as fases. Comunidades surdas não podem aparecer apenas como fornecedoras de imagens para conjuntos de dados. Sua participação deve abranger definição das prioridades, coleta, análise, validação, governança e atualização dos recursos. A tecnologia deve responder às necessidades da comunidade linguística, e não reorganizar a língua exclusivamente segundo conveniências computacionais.

Esse princípio é decisivo diante do crescimento da inteligência artificial. Modelos aprendem a partir de categorias previamente estabelecidas. Se essas categorias ignorarem variação e práticas reais de uso, o sistema reproduzirá o mesmo apagamento em escala ampliada. Por outro lado, se os dados forem elaborados segundo critérios linguísticos e sociolinguísticos robustos, a IA poderá auxiliar na preservação e compreensão da diversidade.

Conclui-se, portanto, que tecnologias digitais e inteligência artificial possuem potencial expressivo para contribuir com a documentação terminológica em Libras, sobretudo por possibilitarem representação audiovisual, tratamento de grandes corpora, recuperação de informações, análise de padrões e atualização contínua. Esse potencial, entretanto, somente poderá ser plenamente realizado quando inovação tecnológica estiver articulada à Terminologia, à Lexicografia, à Sociolinguística, à Educação de Surdos e ao protagonismo das próprias comunidades surdas.

O futuro da documentação terminológica em Libras não reside na substituição do trabalho humano pelo algoritmo, mas na construção de uma colaboração na qual máquinas realizem tarefas computacionais de grande escala e especialistas assumam as decisões linguísticas, conceituais e socioculturais. A inteligência artificial pode acelerar processos; não pode definir sozinha o que é legítimo na língua.

A consolidação de corpora nacionais multimodais, plataformas terminológicas interoperáveis e processos participativos de validação aparece, assim, como importante agenda futura. Esses recursos poderão fortalecer o ensino bilíngue, a formação científica de estudantes surdos, o trabalho de tradutores-intérpretes, a produção de materiais didáticos, a pesquisa linguística e a circulação da Libras em domínios de especialidade.

Mais do que armazenar sinais, documentar terminologias significa registrar conhecimento, história linguística e modos surdos de produzir e compartilhar conceitos. Quando orientadas por essa perspectiva, as tecnologias digitais e a inteligência artificial deixam de representar somente inovação instrumental e passam a constituir meios para ampliar o acesso, a investigação linguística e a autonomia das comunidades de usuários da Libras.

REFERÊNCIAS

ALBRES, Neiva de Aquino. Produção de sinais-termos em Libras para conceitos acadêmicos: uma revisão sistemática e suas contribuições para surdos, professores e tradutores e intérpretes educacionais. **Revista GTLex**, Uberlândia, v. 9, e0917, 2024. DOI: 10.14393/Lex-v9a2023/24-17. Disponível em: <https://seer.ufu.br/index.php/GTLex/article/view/72942>. Acesso em: 3 jul. 2026.

CARVALHO, Paulo Vaz de; SOFIATO, Cássia Geciauskas. Materiais lexicográficos e pedagógicos para a educação de surdos: revisitando a história e as produções. **Perspectiva**, Florianópolis, v. 38, n. 4, p. 1-24, 2020. DOI: 10.5007/2175-795X.2020.e65934. Disponível em: <https://periodicos.ufsc.br/index.php/perspectiva/article/view/65934>. Acesso em: 7 jul. 2026.

FERREIRA, Alessandra Teles Sirvinskas; VASCONCELOS, Iara Alves Hooper; DAWES, Tathianna Prado; BRAZ, Ruth Maria Mariani; ALVES, Gustavo Henrique Varela Saturnino; FRAGEL-MADEIRA, Lucianne. Sinais-termos científicos em Libras: uma reflexão sobre a escassez e a necessidade de padronização. **Ciência**

& Educação, Bauru, v. 30, e24007, 2024. DOI: 10.1590/1516-731320240007. Disponível em: <https://www.scielo.br/j/ciedu/a/ZpSLhkqJmNRqMvRrkS64THq/>. Acesso em: 10 jul. 2026.

GALASSO, Bruno José Betti. Corpus lexicográfico e visão computacional: uma metodologia baseada em IA para a coleta e anotação semiautomática de sinais em larga escala e a análise de variações lexicais (regionalismos) em Libras. **Domínios de Lingu@gem**, Uberlândia, v. 20, e020004, 2026. DOI: 10.14393/DLv20a2026-4. Disponível em: <https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/80463>. Acesso em: 14 jul. 2026.

LEÔNCIO, Érika Lourrane; ZAVAGLIA, Claudia. Lexicografia das línguas de sinais: resgate histórico e estudo descritivo. **Signótica**, Goiânia, v. 32, e63091, 2021. DOI: 10.5216/sig.v32.63091. Disponível em: <https://revistas.ufg.br/sig/article/view/63091>. Acesso em: 18 jul. 2026.

MARQUES, Rafael Emil Korossy; LEITE, Bruno Silva. Utilização de chatbots para o ensino de Libras: uma análise crítica dos impactos e desafios. **Revista Teias**, Rio de Janeiro, v. 26, n. 83, e85575, out./dez. 2025. DOI: 10.12957/teias.2025.85575. Disponível em: <https://www.e-publicacoes.uerj.br/revistateias/article/view/85575>. Acesso em: 22 jul. 2026.

MARTINS, Francielle Cantarelli; STUMPF, Marianne Rossi. Coleta e registro de sinais-termos psicológicos para Glossário de Libras. **Revista Leitura**, Maceió, v. 1, n. 57, p. 35-59, jan./jun. 2016. DOI: 10.28998/2317-9945.201657.35-59. Disponível em: <https://periodicos.ufal.br/revistaleitura/article/view/2654>. Acesso em: 25 jul. 2026.

MARTINS, Tânia Aparecida; BIDARRA, Jorge. Dicionário Eletrônico de Libras (DicELibras): das estruturas lexicais internas à implementação computacional. **TradTerm**, São Paulo, v. 45, p. 46-73, 2023. DOI: 10.11606/issn.2317-9511.v45p46-73. Disponível em: <https://revistas.usp.br/tradterm/article/view/220836>. Acesso em: 29 jul. 2026.

MELO, Ana Júlia Stradioto; SOARES, Irvna Maria Costa; MELO, Silvana Morita; PRATES, Jorge Marques. Glossário de termos técnicos em Língua Brasileira de Sinais (Libras) para ensino de programação a estudantes com deficiência auditiva. **Anais do Seminário de Educação Inclusiva da UEMS**, Dourados, MS, v. 1, n. 1, p. 31-34, 2026. Disponível em: <https://anaisonline.uems.br/seiUEMS/article/view/11155>. Acesso em: 31 jul. 2026.

MORAIS, Leticia Maria Gonçalves de; ALMEIDA, Welizangela Moreira; LIMA, Náthalee Cavalcanti de Almeida; REGO, Rosana Cibely Batista. Inteligência artificial para reconhecimento de língua de sinais. **Revista Eletrônica do Seminário de Iniciação Científica da UFERSA**, Mossoró, v. 31, n. 1, 2025. Disponível em: <https://periodicos.ufersa.edu.br/index.php/rsemic/article/view/14452>. Acesso em: 2 ago. 2026.

MOREIRA, Neliane Cristina; MURTA, Cláudia Rodrigues. Glossário em Libras de termos da Radiologia para o ensino técnico. **Letras & Letras**, Uberlândia, v. 37, n. 2, p. 144-156, jul./dez. 2021. DOI: 10.14393/LL63-v37n2-2021-08. Disponível em: <https://seer.ufu.br/index.php/letraseletras/article/view/57096>. Acesso em: 5 ago. 2026.

SCHLINDWEIN, Ana Flora; ROCHA, Daniele Silva. LIBRAS e tecnologia: práticas translíngues na produção de youtubers surdos. **D.E.L.T.A.: Documentação e Estudos em Linguística Teórica e Aplicada**, São Paulo, v. 39, n. 1, p. 1-27, 2023. DOI: 10.1590/1678-

460X202359803. Disponível em:
<https://www.scielo.br/j/delta/a/jDnPP9Z8bKr8SkDfpQMx6Mb/>.
Acesso em: 8 ago. 2026.

SILVA-OLIVEIRA, Gláucia Caroline; WANDERLEY, Débora Campos; STUMPF, Marianne. Enem em Libras como corpus linguístico: metodologia para produção de glossários em Libras. **Revista da Anpoll**, v. 51, n. 1, p. 106-117, 2020. DOI: 10.18309/anp.v51i1.1358. Disponível em:
<https://revistadaanpoll.emnuvens.com.br/revista/article/view/1358>.
Acesso em: 12 ago. 2026.

SOFIATO, Cássia Geciauskas; REILY, Lucia Helena. Dicionarização da língua brasileira de sinais: estudo comparativo iconográfico e lexical. **Educação e Pesquisa**, São Paulo, v. 40, n. 1, p. 109-126, jan./mar. 2014. DOI: 10.1590/S1517-97022014000100008. Disponível em:
<https://www.scielo.br/j/ep/a/vY3XRbKqCzKG6kLpQdhd3dN/>.
Acesso em: 17 ago. 2026.