



Comunicação

DA ELABORAÇÃO DE UMA TAXONOMIA VOLTADA ÀS HUMANIDADES DIGITAIS NO CENÁRIO BRASILEIRO: a TaDiRAH e sua proposta multilingue

Márcia T. Cavalcanti¹, Tainá Regly^{2,3}, Ricardo M Pimenta², Elizabeth Maria Freire de Jesus³ e Renan M Castro⁴

¹ *Universidade Santa Úrsula (USU), Brasil*

² *Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Brasil*

³ *Universidade Federal do Rio de Janeiro (UFRJ), Brasil*

⁴ *Fundação Getulio Vargas (FGV), Brasil*

RESUMO: Esta comunicação trata da iniciativa de elaborar, no âmbito de um acordo de cooperação internacional, uma versão em português da taxonomia TaDiRAH, dedicada às atividades e pesquisas em Humanidades que se utilizassem de recursos digitais. Tendo em perspectiva parâmetros de acesso aberto, no tocante ao conteúdo informacional que a compõe, e no que tange à própria estrutura computacional onde ela é construída, nesse caso a partir do uso da plataforma TemaTres, realizou-se a versão em português a partir da versão argentina. O processo de versão foi conduzido integralmente de maneira remota pelos integrantes do laboratório, por ter sido realizado durante a pandemia de COVID-19. Uma vez pronta, e online, a TaDiRAH em português abre possibilidades para seu emprego em demais infraestruturas informacionais capazes de organizar e operar uma melhor recuperação da informação para pesquisas no âmbito das Humanidades. Sabe-se que o caminho é longo no sentido que o próprio campo precisa desenvolver suas próprias competências, mas igualmente se destaca a iniciativa realizada como um passo fundamental para a interoperabilidade da informação científica voltada a um campo em pleno florescimento no Brasil: as Humanidades Digitais.

Palavras-chave: Acesso aberto, Ciência da Informação, Humanidades Digitais, Taxonomia Multilingue.

THE ESTABLISHMENT OF A DIGITAL HUMANITIES RELATED TAXONOMY IN THE BRAZILIAN CONTEXT: the TaDiRAH and its multilingual proposal

ABSTRACT: This communication deals with the initiative to develop, within the scope of an international cooperation agreement, a Portuguese version of the TaDiRAH taxonomy, dedicated to activities and research in the Humanities that make use of digital resources. Bearing in mind open access parameters, regarding the informational content that composes it, and with regard to the very computational structure where it is built, in this case using the TemaTres platform, the Portuguese version was made using the Argentine version. The version process was entirely carried out remotely by the laboratory members, as it was carried out during the COVID-19 pandemic. Once ready, and online, the TaDiRAH in Portuguese opens up possibilities for its use in other informational infrastructures capable of organizing and operating better information retrieval for research in the Humanities. It is known that the road is long in the sense that the field itself needs to develop its own skills, but the initiative carried out as a fundamental step towards the interoperability of scientific information aimed at a field in full bloom in Brazil: the Digital Humanities.

Keywords: Open Access, Information Science, Digital Humanities, Multilingual Taxonomy.

Correspondência para: (correspondence to:) marciacavalcanti@gmail.com

INTRODUÇÃO

O movimento pela Ciência Aberta emerge em meio a uma tensão entre as demandas por novas formas de produção colaborativa, interativa e compartilhada da informação, do conhecimento e da cultura e as barreiras impostas por uma lógica mercadológica, que vai além do mercado editorial e que se coloca praticamente proibitiva para a realização de tais demandas, sobre determinados mecanismos de captura, armazenamento e disseminação do conhecimento. Nesse escopo, ciência, tecnologia e inovação figuram como os potenciais elementos capazes de responder tais demandas e enfrentar desafios que são de ordem ambiental, social, tecnopolítica, educacional, sanitária e econômica. Mas somente servirão ao seu propósito caso haja formas de reduzir as respectivas desigualdades que se interligam patologicamente aquelas mesmas demandas e desafios. A Ciência Aberta engloba múltiplas dimensões de abertura, tal como observado na Taxonomia da Ciência Aberta apresentada pelo Facilitate *Open Science Training for European Research*¹, onde é possível perceber o quão amplas são as possibilidades de estratégias e ações em diferentes escopos e níveis que confluem para uma efetiva democratização do conhecimento que é coletiva e socialmente produzido.

Na América Latina, iniciativas em prol da Ciência Aberta têm apontado o lugar de destaque deste continente, uma vez que sustentabilidade e inclusão, fundamentais no contexto conceitual e político dessa ciência, são parâmetros quase que compulsórios no cenário da pesquisa científica em um continente que lida com o monopólio anglófono da comunicação e editoração científica mundial. Monopólio esse que sistematicamente invalida iniciativas representadas por revistas científicas abertas, e dita (direta ou indiretamente) a performance e a política referente à comunicação científica.

Com efeito, a contínua busca por meios capazes de tornar mais visível a produção científica "marginal", não anglófona por exemplo, é pauta presente para uma comunidade de práticas que frequentemente precisa encontrar meios para se infiltrar no *mainstream* das narrativas que buscam representar esse ou aquele campo científico. Tal estratégia busca fortalecer o caráter inclusivo e precisa ser sobretudo colaborativa e concretizante no sentido da viabilização de redes e sistemas que fortaleçam tais iniciativas entre seus atores.

Como a informação produzida é cada vez mais abrangente em escopo, a aplicação de vocabulários controlados também segue crescendo em necessidade para estruturar o conhecimento nos diferentes domínios do saber, sobretudo aqueles emergentes, como é o caso das Humanidades Digitais (HD), em que sua larga abrangência interdisciplinar de objetos de estudos e agentes de atuação acabou por lhe dar uma natureza transitória e transdisciplinar.

Este trabalho descreve uma primeira ação concreta da parceria estabelecida entre o Laboratório de Humanidades Digitales del Centro Argentino de Información Científica y Tecnológica del CONICET (HD CAYCIT LAB) e o Laboratório em Rede de Humanidades Digitais (Larhud)², que foi o desenvolvimento da versão em português da *Taxonomy of Digital Research Activities in the Humanities* (TaDiRAH). Originalmente, esta iniciativa se deu com o apoio da *Digital Research Infrastructure for the Arts and Humanities* europeia (DARIAH-EU) e teve como fim a organização e a categorização de conteúdos informacionais no âmbito das HD por meio de um vocabulário comum capaz de facilitar a descrição dos respectivos conteúdos e sua recuperação. Essa última versão em português passou a se chamar, portanto, Taxonomia sobre Atividades Digitais de Pesquisa nas Humanidades - TaDiRAH.

Em acordo com Domenico Fiormonte e Paolo Sordi (2019), ratificamos que realizar pesquisas na área das Humanidades que se utilizam de abordagens computacionais para a produção do conhecimento não é tarefa fácil quando produzidas no âmbito dos países do Sul Global. A insuficiência de recursos materiais, com destaque aos softwares e hardwares, se contrasta mais ainda quando o desenvolvimento de literacias computacionais por pesquisadores e estudantes das muitas áreas que compõem as Humanidades apresentam-se cotidianamente aquém daquelas ora promovidas e desenvolvidas em um mundo científico anglófono e desenvolvido. Some-se a isso a crônica escassez de formas de financiamento às pesquisas na grande área das Humanidades e já é notório que *grosso modo* estejamos relegados à condição de reprodutores das formas, técnicas e estéticas provenientes do Norte.

Em meio a isso, iniciativas que busquem elaborar acesso à informação e auxílio à organização do conhecimento por meio de recursos e/ou ferramentas abertas e, sobretudo, idealizadas, desenhadas e criadas no Sul Global sempre que possível, são verdadeiras “trincheiras” contra um crescente “mercado” simbólico no qual o *web* oligopólio da informação e do conhecimento, comumente representado pelas *Big Techs* que compõem o GAFAM (Google, Apple, Facebook, Amazon e Microsoft), se instaura e redefine pautas, valores e toda uma geopolítica do conhecimento (Riande & Fiormonte, 2022).

Sendo assim, este texto trata de comunicar uma experiência de tradução e implementação de uma plataforma, que a cada nova iniciativa se torna mais multilingue, dedicada a um vocabulário controlado na forma de uma taxonomia para pesquisas e métodos digitais aplicados à grande área das Humanidades que, mesmo tendo sua origem na Europa, pretende ser

uma ferramenta que reflita e atenda as necessidades de classificação dos pesquisadores do Sul Global.

O projeto tem como objetivo auxiliar a pesquisa em Ciências Humanas, Sociais, Sociais Aplicadas, Linguística, Letras e Artes que têm se utilizado de recursos computacionais/digitais de forma exponencial, apresentando um vocabulário próprio para o cenário de interesse pelo digital que vem naturalmente se desenvolvendo nas últimas décadas.

A iniciativa da TaDiRAH como inovação e estrutura para as Humanidades Digitais brasileiras

A TaDiRAH é um sistema voltado para a classificação de forma hierárquica e facilitação do acesso à informação sobre as atividades de investigação digital aplicadas ao campo das Humanidades. Sendo assim, ela foi desenvolvida com o objetivo de estruturar as informações pertinentes e torná-las mais facilmente reconhecíveis, e utilizá-las em sites e projetos impulsionados pela comunidade das HD.

Iniciada em 2013, juntamente com o *Digital Research Tools* (DiRT), um diretório de ferramentas de HD, a TaDiRAH objetivava produzir uma taxonomia compartilhada que fosse capaz de abranger conteúdos diversos tais como a bibliografia de HD no Zotero e os recursos dos diretórios presentes no DiRT. De acordo com Dombrowski e Perkins (2014), o intuito do projeto consistiu em construir uma taxonomia focada no desenvolvimento de um subconjunto amplo dotado de categorias usuais e compreensíveis de modo a permitir a organização e fornecimento de informações sobre ferramentas, métodos, projetos e leituras sobre as HD.

No Brasil não há nada parecido quando nos debruçamos sobre o extenso campo das Humanidades. A observada lacuna de

atividades coordenadas onde metadados, vocabulários controlados e demais sistemas que auxiliam na organização do conhecimento por meio da elaboração e preservação de estruturas informacionais possam ser desenvolvidos de maneira aberta e visando a comunidade científica atual e futura, dedicada integralmente às Humanidades, é ainda algo bem embrionário se compararmos com iniciativas do Norte Global. Por sua vez, vale lembrar que tal lacuna tem suscitado investigações e práticas na área da Ciência da Informação que responde a parte significativa das problemáticas que envolvem a organização, representação, acesso e gestão de recursos informacionais.

O caráter colaborativo da TaDiRAH deu a ela um direcionamento voltado para a integração da comunidade e de recursos da área, o que auxiliou na resolução de problemas relacionados à descontextualização de seu conteúdo (Dombrowski & Perkins 2014; Borek *et al.*, 2016). Dessa forma, a taxonomia emerge como um instrumento de destaque, aquilo que Borek *et al.* (2014) consideram estar "apropriado à compreensão das Humanidades Digitais sendo melhor entendido através da interseção de seu objeto, metodologias e aplicativos" (Castro, 2020, p.63). Foi definida pelos seus produtores (Borek *et al.*, 2014, p.1, tradução nossa) como uma "taxonomia compartilhada da pesquisa em HD, na qual se pode identificar 'objetivos e métodos' (por exemplo, captura, análise), 'objetos' (por exemplo, dados, imagens, manuscritos), e 'técnicas' (por exemplo, análise de cluster, codificação, modelagem de tópico)" criada com a finalidade de colmatar o fosso entre os atores das HD e suas práticas relacionadas.

Foi sob as perspectivas aqui apresentadas que o Larhud, em parceria com o HD CAYCIT LAB, produziu uma versão em português (BR) da TaDiRAH que, dentre outras coisas, teve como intuito fortalecer o diálogo

Sul/Sul. Com a expertise de uma equipe no Larhud que é majoritariamente formada por pesquisadores e estudantes da Ciência da Informação, espera-se que as futuras pesquisas e produtos elaborados no âmbito nacional possam ser melhor recuperados e preservados a partir do uso de uma taxonomia comum e multilíngue.

As pesquisas na área das Humanidades dotada de abordagens computacionais para a produção do conhecimento não é tarefa fácil quando realizada no âmbito dos países do Sul Global. A insuficiência de recursos materiais, com destaque aos softwares e hardwares, assim como aos financiamentos às pesquisas na grande área das Humanidades, ou a questão concernente à própria linguagem, já nos colocam em condição de desvantagem em termos materiais e de capital simbólico.

O cenário é delicado e requer atenção uma vez que produzir informação e conhecimento no cenário científico, hoje, trata de contar com recursos de ordem política, estratégica, educacional e científica. Tais facetas são, todavia, protagonistas de um *gap* ora sobre o acesso da população à internet e a serviços básicos advindos de sua mediação, ora dos meios pelos quais o conhecimento é produzido e acessado, ou ainda relacionado ao débil emprego que as Tecnologias de Informação e Comunicação (TIC) têm. O cenário é plural se olharmos para a totalidade daquilo que se propõe como Sul Global, mas detém ainda similaridades muito claras quanto às dificuldades.

Com isso, iniciativas que busquem elaborar acesso à informação e auxílio à organização do conhecimento por meio de recursos/ferramentas abertas e, sobretudo, idealizadas, desenhadas e criadas no Sul Global, são também indícios de um "sul epistemológico". A TaDiRAH é, assim, uma forma de registro realizado a partir das demandas que a comunidade Latino-americana tem em face da necessária

organização da informação e do conhecimento sobre o que se produz, e como se produz, nas HD. Do ponto de vista global, o Sul se apresenta buscando ser protagonista de nossas próprias infraestruturas informacionais e computacionais, que são cada vez mais fundamentais para a pesquisa e para a comunicação científica.

Contribuição à comunidade multilíngue das Humanidades Digitais: o processo de tradução da TaDiRAH

A elaboração de vocabulários multilíngues possui como influência fatores relativos à tempo, recursos financeiros e humanos. De acordo com Hudon (1997), ao levar esses aspectos em consideração, a construção de um vocabulário multilíngue possui três abordagens que consistem em: tradução de um vocabulário monolíngue já existente; fusão de diversos vocabulários monolíngues; e no desenvolvimento simultâneo de várias versões linguísticas do vocabulário multilíngue.

Ainda em relação a esse tipo de vocabulário, a autora salienta que a terceira abordagem referente à construção simultânea é a mais ideal por permitir um tratamento igual para todas as línguas representadas, onde nenhuma é a origem ou fonte inicial, o que reduz a artificialidade do sistema e garante o respeito à estrutura da língua e sua cultura.

Para tratar da TaDiRAH, a equipe do Larhud optou pela primeira abordagem discutida por Hudon. Quando a parceria com o CAYCIT foi iniciada, em 2020, a taxonomia já havia sido construída no idioma original e precursor, o inglês, e já possuía uma versão traduzida para a língua espanhola. Sendo assim, optou-se por realizar a tradução a partir do espanhol para o português, sendo que, em diversas situações, houve necessidade de se recorrer também à versão em inglês a fim de se verificar e/ou confirmar a equivalência de sentido atribuído ao termo em questão na língua portuguesa.

O processo de tradução da TaDiRAH iniciou-se com uma reunião online realizada entre os participantes do Larhud, uma vez que todas as etapas do projeto foram executadas durante a pandemia de COVID-19. Nesta reunião foram discutidos quais seriam os procedimentos adotados para a realização da atividade e foi definida a primeira tarefa, que consistia em decidir a respeito da divisão quantitativa dos termos entre os integrantes da equipe.

É interessante ressaltar a diversidade presente na formação acadêmica de cada um dos participantes na tradução da TaDiRAH. O grupo foi composto por uma analista de sistemas, um bibliotecário e uma cientista social/linguista, sendo todos os três doutores em Ciência da Informação (CI). Além deles, o andamento da iniciativa contou com a participação de um historiador doutor em Memória Social e uma bibliotecária mestranda em CI.

Ao acessar a lista alfabética que compõe a taxonomia, na versão em espanhol, é possível visualizar quantos termos existem na TaDiRAH e quantos estão alocados em cada letra. Partindo dessa observação, identificamos um total de 129 termos. O primeiro passo dado, então, foi relacionar o total de termos em cada letra para dividi-los entre os tradutores da taxonomia.

Cada um dos integrantes começou a fazer as traduções seguindo seus próprios métodos de organização: uns criaram arquivos de texto para controlar os termos já traduzidos, outros preferiram fazer a tradução diretamente na plataforma. Como era esperado, em um dado momento o processo precisou ser revisto, e todos concordaram em criar uma planilha, que foi compartilhada em nuvem própria, para que pudesse ser feito o controle das traduções e, posteriormente, o controle das revisões.

A divisão descrita anteriormente não transcorreu de forma rígida, pois a tradução

se desenrolou em uma sequência que não era contínua, mas sim com constantes retornos ao que já tinha sido traduzido para adequar aos novos formatos e para solucionar os problemas que iam surgindo.

Essa não continuidade no fluxo de tradução ocorreu também devido à relação entre os termos presentes na taxonomia. Conforme

exemplificado na figura 1, a seguir, em grande parte dos termos do vocabulário sua definição fazia menção a outros conceitos presentes na estrutura da taxonomia. Sendo assim, optou-se por seguir com a tradução de forma sequencial desses termos, apesar da divisão previamente acordada e exemplificada na planilha.



FIGURA 1 - Exemplo de termo e sua definição na TaDiRAH.

Fonte: TaDiRAH (2021)³.

Essa "quebra" no fluxo, no entanto, mostrou-se produtiva por duas razões. A primeira pelo fato de tratarmos o termo e seus termos relacionados e também, do ponto de vista do tradutor, este procedimento foi bem interessante em nível de aprendizado. Esta mudança na prática de tradução foi devidamente discutida e controlada.

Além disso, o processo de tradução passou por revisões realizadas por pelo menos dois diferentes integrantes do grupo. Algumas situações mostraram necessidade de uma terceira revisão a fim de se alcançar um consenso sobre a melhor tradução do próprio termo, como também de sua descrição.

Conforme Biscalchin e Moreira (2020, p.59), "a representação dos termos em cada língua deve basear-se nas convenções aplicadas nessa língua". Segundo os autores, a tradução deve ser realizada de forma ética e neutra sem

a imposição de valores culturais, geográficos ou religiosos. Para assegurar um vocabulário que leve em consideração o contexto sócio-histórico-cultural de seus usuários e garanta uma disseminação da informação igualitária no sistema, Hudon (1999) elenca três problemas que são tradicionalmente associados ao uso de vocabulários multilíngues. São eles:

- O alongamento de uma língua para fazê-la se ajustar na estrutura conceitual da estrangeira a ponto de se tornar irreconhecível a seus próprios falantes;
- Transferir toda estrutura conceitual de uma cultura para outra, sendo apropriada ou não;
- Traduzir literalmente termos da língua de origem para expressões sem

sentido no idioma de destino.

Michèle Hudon (1999) afirma que para evitar o acontecimento desses três problemas, o desenvolvimento da estrutura de um vocabulário multilíngue deve ser através de relações semânticas fiéis à maneira com que os falantes de uma determinada língua veem e representam o domínio.

No processo de tradução da TaDiRAH foram tomadas precauções para evitar os três problemas apontados pela autora, porém foi mantida a relação semântica entre os termos presentes na língua original. Ainda assim, o procedimento foi realizado com prudência para que a taxonomia refletisse a realidade das HD no Brasil. Dos 129 termos que deveriam ser trabalhados, *Currícula*, *Personas*, *POS-tagging* e *Recolección* ficaram sem tradução por não possuírem termos correspondentes no português e por seu uso não fazer sentido para os usuários brasileiros em sua língua materna.

Outro exemplo que vale a pena mencionar ocorreu com a tradução dos termos *Búsqueda* e *Investigación*. Esse caso demonstra de maneira clara a negociação entre os tradutores sobre quais as melhores palavras para representar cada um dos conceitos e a quais usos cada aplicação seria mais adequada. Inicialmente, ambos os termos haviam sido traduzidos como “Pesquisa”. Ao perceber o problema de homonímia, tomou-se a decisão de transformar a correspondência de *Investigación* em “Pesquisa científica”, mas após uma análise mais detalhada e um entendimento mais acertado entre os tradutores, optou-se por traduzir *Búsqueda* como “Busca” e *Investigación* como “Pesquisa”.

Ao levar em consideração o contexto dos usuários, o vocabulário controlado multilíngue toma como prioridade a busca pelo termo mais adequado para representar um dado conceito levando em consideração a especificidade e estrutura da língua escolhida.

Caso o vocabulário apresente inconsistências relacionadas à sua hierarquia, às relações entre os termos ou até mesmo a presença indevida de um termo, a disseminação da informação de maneira igualitária será prejudicada e sua confiabilidade poderá ser posta em questão (BISCALCHIN; MOREIRA, 2020).

Assim, como exemplificado por Hudon (1997), entendemos que o arquétipo da construção de um vocabulário multilíngue seja a construção simultânea da mesma taxonomia por membros nativos das línguas que serão utilizadas. Reconhecemos a dificuldade da aplicação desse modelo na TaDiRAH, visto que a parceria entre as instituições foi firmada quando a taxonomia original já se encontrava terminada e estabelecida em sua comunidade alvo.

No processo de tradução não foi possível acrescentar termos considerados relevantes para a realidade das HD no Brasil, estabelecer novas relações entre os termos traduzidos e inserir definições aos termos que não possuíam esse tipo de conteúdo.

Um panorama dos resultados ou um horizonte que nos convida? Um balanço dos resultados parciais e propostas futuras

Uma taxonomia nunca é um produto acabado, requerendo, portanto, adaptar-se, por um lado, às mudanças no(s) domínio(s) que representa e, por outro lado, às mudanças nas necessidades e expectativas da sua comunidade alvo. Isso significa, conforme Lambe (2007, p.149), que “as taxonomias são construções fluidas, nunca permanentes”. Nessa direção, toda taxonomia deve prever manutenção e mudança. É essa condição imanente da taxonomia que é marca das contínuas ações de produção de versões em línguas diferentes e, certamente, de posteriores revisões de seu conteúdo.

No caso específico das Humanidades

Digitais, a constituição de um vocabulário para uma área emergente e pluridisciplinar torna mais tangível, desde o momento de sua adoção em projetos posteriores, que funções básicas comuns aos acadêmicos e às disciplinas construam uma orientação teórica a ser identificada por pares e responsável por uma estrutura conceitual inédita que atue como catalisador de estudos e norteador de análises no campo, do campo e para o campo. Esse processo contribui diretamente para que os estudos em Humanidades Digitais se viabilizem, por meio de uma estrutura informacional que auxilia em muito na organização da informação e em sua recuperação. Dessa forma, contribui-se para o incremento de uma agenda de pesquisa acadêmica em Humanidades Digitais característica do Sul Global lusófono.

Nesse sentido, a TaDiRAH se mostrou uma facilitadora para o acesso à informação e conhecimento apropriados para as atividades de investigação no campo das Humanidades Digitais, por representar uma espécie de “domínio emergente” do conhecimento. Com efeito, sendo a TaDiRAH organizada e representada para permitir “pontes entre domínios”, quaisquer ações que confluem em tal sentido, tendem a ser mais efetivas e expressivas à medida em que estes últimos disponham de meios e mecanismos de interação com os respectivos responsáveis pelo emprego, manutenção e atualização da mesma.

Tal cenário fala sobre a interoperabilidade, condição *sine qua non*, para que a taxonomia se torne na prática útil de fato. Sem o desenvolvimento de novos projetos, ferramentas, plataformas e sistemas que comportem dados e informação da Humanidades *lato sensu* e que façam uso da taxonomia desenvolvida e traduzida para o português, enquanto técnica de organização da informação e do conhecimento, interoperabilidade se torna tão somente um

horizonte a se alcançar sem nunca caminhar em sua direção.

Para que efetivamente a TaDiRAH alcance o seu principal propósito, é imprescindível que a comunidade alvo consiga localizá-la, acessá-la e efetivamente utilizá-la como um recurso informacional compreensível que permita explorar possibilidades de aplicações de ferramentas, métodos e técnicas computacionais em suas investigações. Para tal, torna-se imprescindível a circulação de informações acerca da existência da taxonomia bem como sua divulgação. Entendemos que uma das ações possíveis, nesse sentido, seria a interlocução do Larhud junto aos coordenadores de curso de graduação e pós-graduação nas áreas de Ciências Humanas, Sociais, Sociais Aplicadas, Linguística, Letras e Artes a fim de promover uma maior visibilidade da taxonomia para esses campos disciplinares que talvez não estejam tão familiarizados com o seu emprego e importância.

A atuação do Larhud no desenvolvimento da TaDiRAH fala mais dessa contribuição para o campo das Humanidades, provendo os meios para tal “caminhada”, que é uma ação que se propõe a atender apenas sua própria pauta. A atuação em Humanidades Digitais, no desenvolvimento de recursos de toda sorte, idealmente deve ser em direção à ciência aberta, à colaboração e ao favorecimento do multilinguismo. Entendemos que não se viabiliza muito se não tivermos meios capazes de atuar com a importância e sustentação que uma taxonomia atua.

Não é surpresa que haja uma vontade comum presente nas instituições responsáveis pelas diversas versões da TaDiRAH de promover um trabalho colaborativo no sentido de avaliar, revisar e atualizar a taxonomia, por exemplo. Um outro desafio, que guarda forte relação com o apresentado acima, diz respeito à promoção de mecanismos de interação com

a comunidade alvo. Nessa direção, o Larhud tem em seu horizonte de pesquisas futuras, a princípio, implementar sempre que possível a TaDiRAH enquanto um Sistema de Organização do Conhecimento (SOC) nas demais estruturas informacionais a serem desenvolvidas. O desenvolvimento de ferramentas e sistemas acompanhado do aproveitamento de algumas dessas no desenvolvimento de estruturas mais complexas parece ser imprescindível no âmbito do Larhud. Plataformas e ferramentas ligadas à análise quanti e quali, à curadoria digital, à preservação digital ou a formas de recuperação da informação *grosso modo* se potencializam com o emprego de um vocabulário controlado e é esse papel que a TaDiRAH intenta jogar.

CONSIDERAÇÕES FINAIS

O arquétipo da construção de um vocabulário multilíngue consiste na construção simultânea da mesma taxonomia por membros nativos das línguas que serão utilizadas. Reconhecemos a dificuldade da aplicação desse modelo na TaDiRAH, visto que a parceria entre as instituições foi firmada quando a taxonomia original já se encontrava terminada e estabelecida em sua comunidade alvo.

No processo de tradução não foi possível acrescentar termos considerados relevantes para a realidade das HD no Brasil, estabelecer novas relações entre os termos traduzidos e inserir definições aos termos que não possuíam esse tipo de conteúdo. Porém, existe uma vontade comum presente nas instituições responsáveis pelas diversas versões da TaDiRAH de promover uma atualização da taxonomia. Nesse cenário, representantes brasileiros seriam eleitos para colaborar e construir de maneira simultânea um vocabulário multilíngue que levasse em consideração a sua realidade e seguisse os preceitos ideais abordados por Hudon (1997).

A tradução da taxonomia TaDiRAH

representou minimamente um duplo desafio: o primeiro refere-se à tradução interlíngua, do espanhol para o português, sendo que em diversas situações houve a necessidade de se recorrer também à versão em inglês, a fim de verificarmos e/ou confirmarmos a equivalência de sentido atribuído ao termo em questão. Um segundo desafio diz respeito à comunicabilidade da taxonomia, considerando-se o domínio deste instrumento que abarca métodos, objetos e técnicas computacionais em sua relação com o público-alvo ou usuários a que se destina, que são pesquisadores das HD e que, potencialmente, poderão se beneficiar do suporte conceitual abarcado pelo âmbito da taxonomia. Para tal, sabemos que a capacidade de comunicação e divulgação desse instrumento entre pares, na comunidade científica em geral, é fundamental.

O desenvolvimento da taxonomia TaDiRAH e seu reconhecimento pela comunidade dedicada às HD busca gerar possibilidades de interoperabilidade entre o vocabulário controlado e as demais plataformas e artefatos advindos de novas pesquisas de fundo humanístico debruçadas sobre recursos computacionais. Formar-se-á, em tal contexto de práticas e tecnologias, uma “cadeia metodológica” em prol de atividades investigativas. Obviamente, esta “boa” construção não está desprovida de pensamento crítico. É ele, inclusive, que nos faz chegar até aqui. Comunicar a construção e reprodução de novas estruturas para o emprego nas Humanidades está atrelada ao entendimento de sua contribuição para o desenvolvimento do emprego dos recursos digitais nas pesquisas da área, bem como para sua permanente recuperação da informação.

AGRADECIMENTOS

Agradecemos o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Fundação Carlos Chagas Filho de Amparo à Pesquisa do

Estado do Rio de Janeiro (FAPERJ) para a contínua ação de pesquisa que permitiu esta e demais atividades a serem realizadas. Ao Laboratório de Humanidades Digitales (HD Lab) del Centro Argentino de Información en Científica y Tecnológica del CAICYT-CONICET pelo apoio institucional na forma de acordo de cooperação técnica, no qual se circunscreveu a atividade de desenvolvimento da versão em português do TaDiRAH. Por fim, o reconhecimento da cooperação de todos os membros do Laboratório em Rede de Humanidades Digitais envolvidos neste projeto que figuram aqui como coautores.

REFERÊNCIAS

BISCALCHIN, R., MOREIRA, W.

Construção de vocabulários multilíngues: perspectivas culturais. **Perspectivas em Ciência da Informação**, v.25, n.4. 2020. <https://periodicos.ufmg.br/index.php/pci/article/view/26929/20513>

BOREK, L., DOMBROWSKI, Q., PERKINS, J., & SCHÖCH, C. TaDiRAH: A Case Study in Pragmatic Classification. **Digital Humanities Quarterly**, v. 010n.1. 2016.

BOREK, L.; DOMBROWSKI, Q.; PERKINS, J. & SCHÖCH, C. Building Bridges to the Future of a Distributed Network: From DiRT Categories to TaDiRAH, a Methods Taxonomy for Digital Humanities. DCMi International Conference on Dublin Core and Metadata Applications. **Proceedings... Austin**, Texas, p.8-11 October 2014.

CASTRO, RM. **Análise da literatura das humanidades digitais: uma proposta bibliométrica para descrição de seu escopo e congruência conceitual**. 2020. 200 f. Tese (Doutorado em Ciência da Informação) –

Escola de Comunicação, Universidade Federal do Rio de Janeiro; Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, RJ, 2020. <https://ridi.ibict.br/handle/123456789/1096>.

DOMBROWSKI, Q. & PERKINS, J. **TaDiRAH: Building Capacity for Integrated Access**. Dh+lib, 2014. <https://acrl.ala.org/dh/2014/05/21/tadirah-building-capacity-integrated-access/>.

FIORMONTE, D., SORDI, P. Humanidades digitais do sul e GAFAM. Para uma geopolítica do conhecimento digital. **Liinc Em Revista**, v.15, n.1. 2019. <https://doi.org/10.18617/liinc.v15i1.4730>

HUDON, M. (). Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts. **Knowledge Organization**, Würzburg, v.24, n. 2, 1999. https://www.ergon-verlag.de/isko_ko/downloads/ko_24_1997_2_d.pdf.

HUDON, M. Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts. **Knowledge Organization**, Würzburg, v.24, n. 2, p. 84-91, 1997. https://www.ergon-verlag.de/isko_ko/downloads/ko_24_1997_2_d.pdf.

Lambe, P. **Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness**. England: Chandos Publishing Oxford, 2007.

RIANDE, G. R.& FIORMONTE, D. Una vez más sobre los sures de las digital humanities, **Acervo**. Rio de Janeiro, v. 35, n. 1, jan./abr 2022. <https://revista.an.gov.br/index.php/revistaacervo/article/view/1850/1711>.

NOTAS

¹ Disponível em:
<https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>. Acesso em: 15 set. 2022.

² Pertencente à Coordenadoria de Ensino e Pesquisa, Ciência e Tecnologia da Informação (COEPE) do

Instituto de Informação em Ciência e Tecnologia (IBICT).

³ TaDiRAH. Disponível em:
<https://www.vocabularyserver.com/tadirah/pt/index.php?tema=26/análise-de-conteúdo>. Acesso em: 22 set. 2022.