

DeClassIFY: Uma ferramenta para auxílio à classificação de artigos científicos*

Alessander Osorio¹, Vitor Alano de Ataidés¹,
Paulo Roberto Ferreira Jr.¹, Gerson Geraldo H. Cavaleiro¹

¹Programa de Pós-Graduação em Computação
Universidade Federal de Pelotas
Pelotas – RS – Brazil

alessander.osorio@inf.ufpel.edu.br

Abstract. *Este artigo apresenta DeClassIFY, um software desenvolvido como uma extensão a navegadores de Internet para classificação de artigos científicos. O objetivo deste software é auxiliar a identificação do tipo de pesquisa em Computação documentado em um relato científico. Para a categorização dos tipos de pesquisa, o software é aplicado sobre um arquivo em formato PDF aberto no navegador, a saída indica a classificação sugerida a partir da análise realizada sobre o conteúdo. A taxonomia para classificação representa um compilado de estudos documentados pela bibliografia da área. O software foi concebido para auxiliar na formação de jovens pesquisadores e instrumentalizar disciplinas que abordam metodologia de pesquisa. O protótipo encontra-se disponível nas lojas oficiais de quatro navegadores.*

Resumo. *This article presents DeClassIFY, a software developed as a browser extension for classifying scientific articles. The aim of this software is to assist in identifying the type of research in Computing documented in a scientific report. For the categorization of research types, the software is applied to a PDF file opened in the browser; the output indicates the suggested classification based on the analysis performed on the content. The taxonomy for classification represents a compilation of studies documented in the area's bibliography. The software was designed to assist in the training of young researchers and to provide tools for disciplines that address research methodology. The prototype is available in the official stores of four browsers.*

1. Introdução

A formação de recursos humanos em nível do ensino superior usualmente inclui apresentar e desenvolver habilidades em pesquisa em disciplinas que abordem metodologias científicas. No contexto da Computação é recomendado pela Sociedade Brasileira de Computação (SBC) que este conteúdo faça parte dos currículos na área, na perspectiva de formar profissionais em Computação habilitados a pensar a ciência criticamente e estimulados a buscar soluções aos seus problemas pela investigação científica ([Mello and Mello 2022, Zorzo et al. 2017]), apoiados em valores éticos ([Bispo Jr et al. 2021, Araujo et al. 2019]).

*O presente trabalho foi realizado com apoio do Programa Nacional de Cooperação Acadêmica da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES/Brasil.

No contexto da Computação, é comum o uso de ferramentas que auxiliem na execução das usuais etapas de uma metodologia científica. Dentre estas ferramentas se apresentam aquelas oferecendo recursos para catalogação bibliográfica, como Zotero [Zot] e Mendeley [Men], úteis nas etapas iniciais da pesquisa, quando da coleta das informações iniciais para o projeto. No outro extremo, para etapas relacionadas à avaliação de resultados, pacotes como STATA [STA] e SPSS [SPS] oferecem recursos para análises estatísticas e Taguette [Tag], NVivo [NVi] e RQDA [RQD] para análises qualitativas de informações textuais, como artigos e entrevistas. No que diz respeito ao auxílio na condução das etapas da pesquisa, com aferição da produtividade e do atingimento de metas, ferramentas para gerenciamento de tarefas são usualmente adotadas. Dentre essas ferramentas, Trello [Tre] e KanbanFlow [Kan], tem inserção no meio acadêmico.

Embora a quantidade de ferramentas disponíveis para apoiar a implementação de etapas do método científico seja grande, poucas são as opções de ferramentas que ofereçam suporte à execução destas etapas. Como exemplo de ferramentas que apoiam a execução de etapas do método científico encontra-se Pasifal [Par] e Protocols.io [Pro]. Parsifal é uma ferramenta que oferece suporte à revisões sistemáticas da literatura segundo o protocolo especificado em [Kitchenham and Charters 2007a]. Já Protocols.io, embora não seja implementação de uma metodologia específica, oferece um bom suporte para registrar coletas de resultados, relacionando-os com os parâmetros de entrada, como configurações de software, parâmetros de experimentos etc.

Nesta lacuna de ferramentas, é apresentada a proposta de DeClassIFY. Esta ferramenta propõe a apresentar uma classificação do tipo de pesquisa reportado em um artigo científico. O resultado a ser atingido é auxiliar pesquisadores, em particular pesquisadores em formação, a identificar o tipo de pesquisa relatada no artigo. Em adição, esta ferramenta pode ser utilizada como um complemento aos recursos explorados em disciplinas que abordem metodologia e/ou escrita e leitura científica.

DeClassIFY foi concebido como um recurso do tipo extensão, ou simplesmente *plugin*, para navegadores Internet. Ao ser acionado esse *plugin* sobre uma página contendo um arquivo em formato PDF¹, o usuário tem acesso a visualização da classificação encontrada. A taxonomia para tal classificação encontra-se documentada em [Osorio 2024], tendo sido composta a partir de um apanhado de propostas de classificações documentadas em trabalhos anteriores.

O restante deste artigo está assim organizado. A Seção 2 resume as conclusões de uma revisão sistemática da literatura realizada visando identificar classificações das pesquisas em Computação. A Seção 3 sumariza a taxonomia adotada. O desenvolvimento da extensão e o protótipo desenvolvido encontram-se apresentados na Seção 4 e estudos de caso são apresentados na Seção 5. A Seção 7 apresenta as considerações finais.

2. Pesquisa em Computação na Literatura

Para identificar possíveis classificações de pesquisa, foi realizada uma revisão sistemática da literatura (RSL), segundo os protocolos preconizados por [Kitchenham and Charters 2007a] e [Osorio 2020]. Esta RSL encontra-se documentada em [Osorio 2024] e foi conduzida visando responder às seguintes questões:

¹PDF – Portable Document Format (formato portátil de documento).

Questão 1 - Existe alguma divisão explicativa para a pesquisa científica em Computação, considerando tipos de estudos, objetos de pesquisa e métodos avaliativos aplicáveis à área, independentemente da denominação utilizada, seja esta classificação, divisão ou taxonomia?

Questão 2 - Quais são os aspectos, características e/ou conceitos a serem considerados para se montar uma classificação ou divisão da pesquisa científica em Computação visando sua melhor compreensão?

Questão 3 - Qual a inserção da Computação dentro da Ciência e onde esta acontece dentro de uma análise cartesiana multi eixo da Computação face às divisões da Ciência? Sendo de um lado a Computação como fonte geradora de Ciência, de outro ela como ferramenta para tal.

Questão 4 - Quais as divisões de estudo atuais da Computação como área de conhecimento, as quais devem ser consideradas na montagem de um panorama abrangente da área?

O final de todo o processo da RSL resultou em 21 artigos selecionados, os quais foram considerados para responder às questões de pesquisa como segue e orientar a composição da taxonomia proposta. As respostas encontradas são:

Questão 1. A RSL identificou que sim, existem diferentes trabalhos que se utilizam de classificações para pesquisa em Computação. No entanto, não foi verificada uma uniformidade em relação às classificações utilizadas nem aos conceitos empregados para discriminar as classificações. O indicativo desta observação é a existência de uma lacuna na área para propor uma classificação que convirja as ideias apresentadas pela literatura.

Questão 2. Embora os trabalhos selecionados não possuam uma identificação uniforme para as categorias das pesquisas propostas, é possível identificar a existência de propostas de classificação de pesquisas em Computação. No entanto, a ausência de uma consistência de termos identifica a necessidade de se extrair destes trabalhos uma nomenclatura ampla capaz de contemplar todas as ideias apresentadas. A partir destes trabalhos, foi produzida uma representação homogênea das classificações propostas (Seção 3).

Questão 3. A percepção que todo cientista da Computação possui de que a Computação ao mesmo tempo representa uma ciência por si e que também permeia outras ciências se mostrou evidente a partir do estudo dos trabalhos coletados na RSL ([Ghezzi 2020, Tedre 2007, Tedre 2011, Hassani 2017]). A Computação, como Ciência, se apresenta ao mesmo tempo fonte geradora de conhecimento, ferramenta para geração de conhecimento e propriamente o objeto de estudo. Estes acabam definindo os três eixos sobre os quais pesquisa envolvendo a Computação ocorrem: a Computação como fim, como meio e como o próprio estudo da ciência em Computação. No entanto, não foi encontrada uma classificação ou divisão que contemple esses eixos de estudo. A Computação como fim, é aquela que desenvolve a técnica computacional e a valida aplicando o método científico. Como meio, é aquela que está inserida em outras áreas de pesquisa como ferramenta, permitindo a estas, com suas pesquisas apresentarem seus resultados. Como próprio objeto de estudo ou metaciência, é aquela que faz da ciência da Computação seu objeto de estudo, desenvolvendo e aprimorando o método científico.

Questão 4. A área de Computação é abrangente e possui várias divisões de estudo. Dentre elas, destacam-se as classificações do European Research Council (ERC) e do Association for Computing Machinery (ACM). O ERC apresenta uma estrutura que inclui a

Tabela 1. Taxonomia da pesquisa científica em Computação – Visão sumarizada.

Categoria	Descritores	Citações
Natureza	2	3
Propósito/Paradigma	4	1
Método	16	9
Metodologia	7	2
Prova científica	3	5
Prova científica secundária	2	1
Ambiente	4	4
Natureza dos Dados	4	6
Estratégia de Validação / Coleta de Dados	35	6
Validação do Resultado	6	1

Computação nas Ciências da Física e Engenharias, enquanto a ACM oferece os Special Interest Groups (SIGs) para representar áreas de pesquisa e os Computing Classification System (CCS) para categorização precisa de publicações. No contexto brasileiro, a Sociedade Brasileira de Computação (SBC) também tem sua própria divisão de interesse na área. Para um panorama abrangente da Computação, a classificação CCS da ACM é mais indicada, embora as demais também forneçam informações importantes sobre os nichos de pesquisa na área. ([Ghezzi 2020], [Holz et al. 2006], [Ramesh et al. 2004]).

3. A Taxonomia Adotada

Conforme identificado nas respostas obtidas às questões de pesquisa trabalhadas na RSL, buscou-se conceber uma nomenclatura uniformizada para as categorias de pesquisa realizadas na área da Computação. A partir dos trabalhos estudados, foi produzida tal representação homogênea das classificações propostas, qual encontra-se apresentada na Tabela 1. A taxonomia possui dois níveis, Categoria e Descritor que identifica uma especificidade da pesquisa nas diferentes categorias. É usual, ainda, que relatos científicos descrevam pesquisas multifacetadas, fazendo com que sejam classificados por mais de um par Categoria/Descritor. Por questões tando de espaço disponível no texto e de objetivo do trabalho – que é o de apresentar a ferramenta desenvolvida – somente o nível mais alto, o nível de Categoria é descrito. A coluna Descritor indica quantos descritores são compreendidos pelas Categorias individualmente.

A coluna Citações da Tabela 1 oferece dados adicionais à resposta apresentada à Questão de Pesquisa 2. Nesta coluna são identificados em quantos artigos, dos 21 selecionados pela RSL, em que a referida Categoria é encontrada. Deve ficar claro que os termos utilizados na taxonomia, tanto para Categorias como Descritores, não se encontram nominalmente presentes em todos trabalhos coletados na RSL, havendo divergências em nomenclaturas. A uniformização apresentada, no entanto, é ampla o suficiente para contemplar todas as definições entrada nos trabalhos retirados da literatura selecionada.

A categoria *Natureza* refere-se às características fundamentais do conhecimento produzido, como sua originalidade, aplicabilidade e inovação em diversas áreas de estudo. Esta categoria engloba a criação de algo novo ou o desenvolvimento de abordagens inovadoras para questões existentes na Computação, influenciando diretamente a relevân-

cia e impacto das pesquisas. Compreender a natureza do conhecimento gerado é crucial para avaliar seu potencial de contribuição para o avanço da ciência e tecnologia na área da Computação.

Propósito/Paradigma define os objetivos subjacentes ao conhecimento produzido na área da Computação, determinando se o foco é na busca pela verdade, compreensão dos fenômenos estudados ou crítica das abordagens existentes. Essa categoria influencia diretamente a direção da pesquisa e a maneira como os resultados são interpretados e aplicados na prática, contribuindo para o desenvolvimento de novas teorias e tecnologias na Computação. Entender o propósito ou paradigma de uma pesquisa é fundamental para situá-la dentro do contexto científico e tecnológico e compreender seu potencial impacto.

Método, como categoria de pesquisa, refere-se ao conjunto sistemático e ordenado de procedimentos usados para produzir conhecimento, garantindo consistência, reprodutibilidade e confiabilidade dos resultados. Esta categoria envolve técnicas e ferramentas específicas para coleta, análise e interpretação de dados, fornecendo base sólida para o desenvolvimento de novos conhecimentos em Computação. A escolha e aplicação apropriadas dos métodos são essenciais para garantir a validade e robustez dos resultados, avançando o conhecimento científico em Computação.

Metodologia se refere à abordagem adotada para conduzir a pesquisa, incluindo estratégias específicas de coleta, análise e interpretação de dados. Esta categoria é crucial para definir o escopo e os objetivos da pesquisa, além de selecionar técnicas e ferramentas adequadas na área da Computação. Uma metodologia bem elaborada é fundamental para garantir a eficácia e eficiência da pesquisa, orientando os pesquisadores em todas as etapas do processo de investigação.

Prova científica determina a validade dos resultados obtidos por meio de evidências diretas, como experimentos controlados, testes empíricos ou análises estatísticas rigorosas. Esta categoria é fundamental para validar as hipóteses e conclusões da pesquisa, proporcionando uma base sólida para o avanço do conhecimento na área da Computação. A aplicação de métodos adequados de prova científica é essencial para garantir a credibilidade e a confiabilidade dos resultados obtidos e para contribuir significativamente para o desenvolvimento da ciência e tecnologia na Computação.

Prova científica secundária valida resultados por meio de evidências indiretas, como revisões de literatura, análises de dados existentes ou comparações com estudos anteriores. Esta categoria complementa a prova científica direta, fornecendo informações adicionais e corroborando achados da pesquisa por diferentes abordagens e fontes de evidência em Computação. O uso de métodos robustos de prova científica secundária é crucial para fortalecer argumentação e fundamentação teórica, ampliando relevância e contribuição para avanço do conhecimento em Computação.

A categoria *Ambiente* refere-se ao contexto onde os dados são coletados e experimentos realizados, incluindo a infraestrutura tecnológica disponível, as condições de execução dos estudos e os recursos computacionais utilizados. Esta categoria influencia a qualidade e confiabilidade dos dados e resultados obtidos na pesquisa, afetando a interpretação e generalização dos achados na área da Computação. Uma compreensão clara do ambiente de pesquisa é fundamental para contextualizar os resultados, identificar possíveis vieses e garantir a validade e robustez dos resultados na investigação computacional.

Natureza dos Dados descreve características intrínsecas dos dados na pesquisa computacional, como estrutura, formato, volume, temporalidade e origem, influenciando métodos de análise e interpretação. Esta categoria é crucial para garantir qualidade, confiabilidade e relevância dos dados, fornecendo base sólida para percepções e soluções na Computação. Compreensão abrangente da natureza dos dados orienta seleção e aplicação de técnicas analíticas e ferramentas computacionais, contribuindo para avanço do conhecimento.

Estratégia de Validação / Coleta de Dados é a maior em termos de número de descritores e mais complexa encontrada na RSL. Essa complexidade decorre do fato de que essa estratégia determina a abordagem da pesquisa científica em relação a outras categorias mencionadas anteriormente abrangendo amplamente quase todas as etapas, especialmente a coleta e validação dos dados.

Validação do Resultado refere-se ao aspecto do objeto de estudo ao qual o conhecimento gerado se refere, incluindo a criação do objeto de estudo, seu uso, experiência de uso e custos.

Os Descritores das pesquisas de cada Categoria não se encontram explicitadas neste artigo por questões de espaço. Cada Descritor consiste em um conjunto de termos e expressões, identificados como *regras*, usuais em artigos relacionados o respectivo par Categorias/Descritor. A validação da taxonomia utilizou um procedimento algoritmo semi-supervisionado k-NN para identificar se o contexto de uso de cada regra se encontra em conformidade com a taxonomia proposta [Osorio 2024]. Este mesmo conjunto de regras foi utilizado para, na extensão desenvolvida, classificar os relatos científicos.

4. A Extensão Desenvolvida

O software apresentado neste artigo foi concebido para auxiliar o processo de classificação de relatos científicos (artigos) e oferecer acesso e publicidade à taxonomia desenvolvida. O formato de publicação deste software na forma de uma extensão a navegadores de Internet foi escolhido para facilitar sua utilização, além de ser bastante prático para ser utilizado durante as etapas de coleta de material estudo para desenvolvimento de novos projetos. Para maior fluidez no restante do texto desta seção, vamos nos referir à extensão proposta como *plugin* e navegadores de Internet como *browser*.

Uma nota relevante em relação ao classificador é que é apto a tratar apenas textos de artigos na língua inglesa. Esse fato deve-se que as regras construídas para criação da taxonomia ter utilizado como fonte artigos publicados também em inglês. Consequentemente, também, toda a informação produzida pelo *plugin* é apresentada nesta língua, em particular os termos designando os nomes das Categorias e Descritores.

4.1. O código

O *plugin* foi desenvolvido em JavaScript e possui a estrutura básica para uma extensão de navegador, partindo do diretório raiz (DeClassIFY). Para manipulação dos arquivos em formato PDF foi utilizada a biblioteca Mozilla PDF.js [PDF]. O código principal se distribui em dois diretórios: `scripts`, no qual encontra-se implementada a lógica para interação com o usuário e classificação dos artigos, e `data`, contendo estruturas de dados que descrevem as regras para identificação dos Descritores – um arquivo de regras de todos os Descritores de cada Categoria. Essa estrutura de dados é representada

por um objeto representando uma determinada Categoria encapsulando um objeto para cada Descritor que a compõe contendo strings descrevendo as regras a ele aplicadas. Tal estrutura de dados permite flexibilidade em atualizações nas regras de identificação dos Descritores, bem como permite adaptação do *plugin* para outras aplicações.

O *plugin*, se incorporado ao navegador, estará disponível sempre que uma página acessada estiver exibindo um arquivo PDF. Ao ser acionado pelo usuário, a execução inicia obtendo a URL acessada pela respectiva aba para então obter o endereço do arquivo PDF correspondente. O primeiro procedimento é realizar um processo de normalização do arquivo a ser classificado pelos seguintes passos: (i) transformação do arquivo PDF em texto; (ii) conversão do texto para letras minúsculas; (iii) remoção de acentos; (iv) remoção de todo o carácter não alfabético (símbolos e dígitos); (v) garantir que todas as palavras encontram-se separadas por um, e apenas um, espaço em branco. Esta normalização é realizada para garantir que o processo de busca das regras dos descritores.

O processo de classificação pontua o número de ocorrências de cada conjunto de regras de cada Descritor da seguinte forma: caso a regra descrita tenha sido encontrada completa, pontua-se a possibilidade deste Descritor identificar a Categoria de pesquisa do artigo adicionando 2 pontos. Para identificação parcial da regra no artigo, adiciona-se 1 ponto. Os descritores mais pontuados de cada Categoria serão os apresentados como resultado da classificação. Caso nenhum Descritor seja identificado, a saída será “None” para o Descritor da saída correspondente.

É importante observar que o desenvolvimento de extensões para navegadores deve seguir regras estritas de segurança. Em particular, neste projeto, observa-se que o *plugin* executa em espaço protegido, na nomenclatura usual denominado de *sandbox*, e tem permissões limitadas ao acesso aos recursos do próprio navegador e mesmo dos dados na área local do usuário. A principal dificuldade encontrada foi a manipulação de arquivos PDF, o que foi possível com o uso da biblioteca Mozilla PDF.js. Na extensão desenvolvida, nenhuma permissão é solicitada ao usuário, garantindo total segurança em sua utilização.

4.2. Instalando e utilizando o *plugin*

A versão atual do *plugin*, como protótipo, encontra-se disponível como um repositório de acesso público² como um extensão para o Google Chrome [Goo]. Para utilizá-la, é necessário baixar o conteúdo do repositório para um diretório local, de forma não compactada. Na aba de extensões do Chrome (`chrome://extensions`), habilitar o *modo desenvolvedor* e carregar o pacote a partir de seu diretório raiz no botão *Carregar sem compactação*. A opção de utilizar o Chrome para desenvolvimento do *plugin* considerou que este navegador possui a maior fatia de mercado³, além deste navegador possuir a mesma base de implementação que outros navegadores, como o Brave, Microsoft Edge e o Opera, nos quais o *plugin* foi testado e opera corretamente. No restante deste texto, todos os usos do *plugin* registram seu uso no Chrome; testes de compatibilidade realizados no Brave, Edge e Opera mostraram que o *plugin* se mantém operacional.

O *plugin* também está disponível na loja de aplicativos do Chrome (Chrome Web

²DeClassIFY no GitHub em <https://github.com/bundz/DeClassIFY>

³Segundo StatCounter, em <https://gs.statcounter.com/browser-market-share>, Google Chrome respondia por 65,3% do mercado de navegadores em uso em Dezembro de 2023.

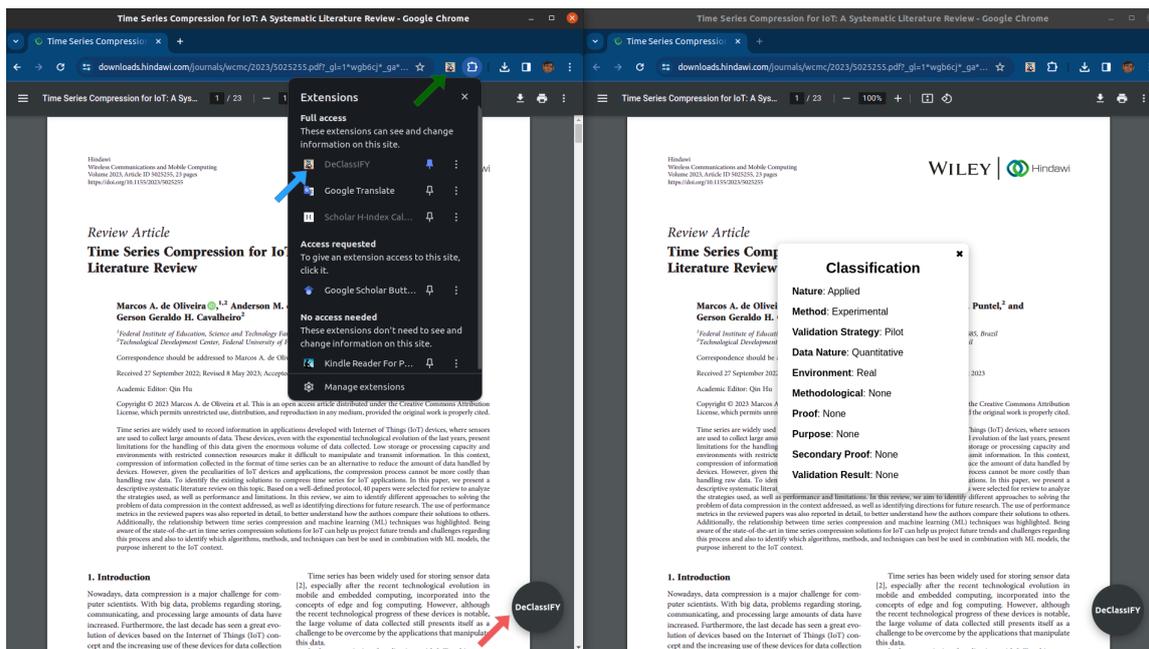


Figura 1. Interface com um PDF aberto após a instalação do *plugin*.

Store, <https://chromewebstore.google.com>), bastando procurar por DeClassIFY e instalando conforme passos indicados na própria loja.

A Figura 1 apresenta uma tela com duas janelas do Chrome abertas. A janela do lado esquerdo indica a visualização de que o *plugin* está instalado e a da direita sua utilização. Uma vez instalado o *plugin*, ele estará disponível no navegador, seu ícone será visível no *menu de plugins*, caso este menu seja aberto (seta azul na figura). O *plugin* pode ser fixado na barra de ferramentas (opcional), neste caso, seu ícone também será visível ao lado da barra de endereços (seta verde). Estando ativado e sendo aberta uma página contendo um arquivo PDF, o botão do DeClassIFY ficará disponível – no local indicado pela seta vermelha na Figura 1. O acesso ao DeClassIFY somente estará habilitado nos arquivos abertos após a instalação do *plugin*. As páginas com arquivos já abertos precisam ser recarregados para ativá-lo.

A ativação do *plugin* se dá selecionando o botão *DeClassIFY*, sendo apresentada a classificação observada. A visualização da classificação se dá em uma janela *pop-up*, como mostra o lado direito da Figura 1.

5. Estudos de Casos

A atual versão do protótipo contempla a classificação de todos os Descritores contidos na taxonomia. A versão do *plugin* na loja contém a última versão estável e a versão no repositório aquela em trabalho.

O primeiro estudo de caso encontra-se ilustrado na Figura 1. A janela *pop-up* com a classificação identificou Descritores em cinco das Categorias da taxonomia, **Nature/Applied**, **Method/Experimental**, **Validation Strategy/Pilot**, **Data Nature/Quantitative** e **Environment/Real**. Os Descritores identificados para tais Categorias foram, respectivamente, *Applied*, *Exploratory* e *Literature Review*. Para as demais

Categorias, o *plugin* não identificou uma classificação possível, sendo informado *None*. O artigo documenta uma Revisão Sistemática de Literatura. DeCLASSIFY classificou o trabalho com um bom nível de precisão. O artigo apresenta uma Revisão Sistemática de Literatura (RSL) sobre questões de natureza aplicada, envolvendo a identificação de técnicas para compressão de séries temporais em um contexto específico. Enquanto método, o artigo foi classificado como Experimental. Este Descritor corresponde a um modelo de pesquisa que envolve a coleta de observações e formação de hipóteses. Neste sentido, a RSL realizada documenta uma série de material coletado, culminando em uma análise das oportunidades de pesquisa na área em questão. Enquanto natureza de dados e ambiente, estes foram caracterizados como Quantitativo e Real. De fato, os objetos trabalhados em artigo que documenta uma RSL consistem em artigos e estes artigos são objetos reais e manipulados em grandes quantidades. A estratégia de validação, caracterizada como Piloto deve ser observada. A RSL documentada no artigo segue um protocolo [Kitchenham and Charters 2007b] de pesquisa, no entanto, nem este protocolo de pesquisa nem algum outro um objeto existente está sendo avaliado ou testado – conforme preconiza o descritor para Piloto [Osorio 2024].

O segundo caso de estudo submeteu um artigo na língua portuguesa ao classificador. Como indicado anteriormente, o classificador é apto a tratar apenas textos na língua inglesa. Para todas as classificações possíveis, os Descritores obtidos foram *None*, indicando não ter sido possível encontrar um Descritores adequados para cada situação. Essa situação era a esperada.

O terceiro caso de estudo foi realizado sobre um artigo apresentando uma análise de desempenho de um benchmark sobre arquiteturas do tipo GPU⁴. A classificação identificada foi **Nature/Applied**, **Method/Experimental** e **Validation Strategy/Pilot**, *None* para todas as demais. O estudo documentado neste artigo documenta um conjunto de experimentos sobre variações de códigos implementando o benchmark NAS, as classificações obtidas para natureza e método são bastante apropriadas. Neste caso, a estratégia de validação ser dita como Piloto é adequada, pois existe uma implementação base servido como parâmetro de para verificar as variações de desempenho observadas pelas implementações alternativas.

Dos casos de estudo considerados, entende-se que as classificações apresentadas estão de acordo com a taxonomia proposta. A experiência de uso é bastante simples e o tempo necessário ao processo de classificação não é significativo. Ainda assim observamos que podem ocorrer situações em que a classificação encontrada é sujeita à interpretação, como no primeiro caso em que um relato científico de uma RSL foi classificado como sendo um tipo de pesquisa de Natureza Aplicada.

6. Nota Legal

O *plugin* desenvolvido apresenta uma classificação para relatos científicos publicados como PDF em língua inglesa e disponíveis na Internet como arquivos PDF. Nem todas as bases de dados oferecem acesso direto ao arquivo PDF pelos navegadores, como é o caso da ACM Digital Library e da IEEE Xplorer. Nestes casos, o arquivo PDF é um objeto apresentado para visualização. Na versão protótipo construída, o tratamento de arquivos

⁴Título do artigo: NAS Parallel Benchmarks with Python: a performance and programming effort analysis focusing on GPUs, publicado em *The Journal of Supercomputing*.

disponibilizados pela ACM Digital Library foi possível, não tendo ainda sido possível a inclusão do acesso aos artigos disponibilizados pela plataforma IEEE Xplorer. Nos estudos de caso apresentados, os artigos foram acessados diretamente da base dos editores, Wiley/Hindawi e Link Springer. Assim, é possível detectar a não operacionalidade em alguns repositórios e também observar a alteração na manipulação de seus objetos por algum provedor. Nesses casos, o *plugin* desenvolvido deverá sofrer ajustes.

Outro aspecto a ser considerado: dado às restrições nas camadas de segurança de extensões para navegadores, não é possível acessar arquivos locais. Portanto, a ativação do *plugin* só se dá em arquivos PDF disponibilizados na Internet.

Ressalta-se ainda que a classificação apresentada ppr DeClassIFY deve ser objeto de análise crítica por parte dos usuários e não recebida como definitiva. O *plugin* disponibilizado representa um artefato de software aplicando regras propostas em [Osorio 2024] para expressar uma taxonomia para pesquisa também apresentada neste mesmo documento. Assim, tanto quanto a taxonomia pode ser discutida em relação a sua amplitude e cobertura, o *plugin* implementado pode ser questionado em relação a sua precisão. No entanto, uma vez ciente do contexto em que a classificação se apresenta, o usuário pode utilizar as informações obtidas para promover suas próprias investigações.

7. Conclusão

Este artigo apresentou DeClassIFY, uma extensão para navegadores de Internet capaz classificar o(s) tipo(s) de pesquisa(s) envolvido(s) em relatos científicos, na forma de artigos. A taxonomia para essa classificação ([Osorio 2024]) apresenta 83 Descritores distribuídos entre 10 Categorias. No artigo, além de descrever o *plugin* e como deve ser utilizado, foi apresentada uma introdução à taxonomia utilizada e um relato de casos de uso. As formas de obter este *plugin* também se encontram relatadas.

O desenvolvimento do espírito crítico e científico é uma preocupação real na concepção dos projetos pedagógicos de cursos de graduação, e mesmo de pós-graduação, nas áreas de Exatas. Na Computação, em particular, a inclusão de tópicos relacionados ao Método Científico é promovida e usualmente materializadas nos documentos de definição dos cursos [Zorzo et al. 2017, Moraes et al. 2019]. Neste contexto, a ferramenta apresentada se insere como um recurso de apoio à coleta de material de referência e de auxílio a compreensão das diferentes formas de realizar a pesquisa em Computação.

Como atividade imediata futura pretende-se aplicar o uso de DeClassIFY em disciplinas de graduação. Os objetivos desta atividade são em número de dois. O primeiro é consolidar a taxonomia proposta em [Osorio 2024] e, eventualmente ajustando-a e adequar as regras implementadas no *plugin* para contemplar os ajustes. O segundo objetivo é o de oportunizar os estudantes a ampliar suas visões sobre a pesquisa em Computação e as oportunidades que a área oferece com apoio de uma ferramenta de apoio ao aprendizado.

Referências

- Google Chrome. <https://www.google.com/intl/pt-BR/chrome>. Acesso em 22 de março de 2024.
- IBM SPSS. <https://www.ibm.com/br-pt/spss>. Acesso em 22 de março de 2024.
- Kanban Flow. <https://kanbanflow.com>. Acesso em 22 de março de 2024.
- Mendeley. <https://www.mendeley.com>. Acesso em 22 de março de 2024.
- Mozilla pdf.js. <https://mozilla.github.io/pdf.js>. Acesso em 22 de março de 2024.
- NVivo. <https://lumivero.com/products/nvivo>. Acesso em 22 de março de 2024.
- Parsifal. <https://parsif.al>. Acesso em 22 de março de 2024.
- Protocols.io. <https://www.protocols.io>. Acesso em 22 de março de 2024.
- R package for qualitative data analysis,. <https://rqda.r-forge.r-project.org>. Acesso em 22 de março de 2024.
- STATA. <https://www.stata.com/>. Acesso em 22 de março de 2024.
- Taguette. <https://www.taguette.org>. Acesso em 22 de março de 2024.
- Trello. <https://trello.com>. Acesso em 22 de março de 2024.
- Zotero. <https://www.zotero.org>. Acesso em 22 de março de 2024.
- Araujo, R., Simão, A., Malucelli, A., Zorzo, A., Monteiro, J., and Chaimowicz, L. (2019). *Referenciais de Formação para os Cursos de Pós-Graduação Stricto Sensu em Computação*. Sociedade Brasileira de Computação (SBC). Sociedade Brasileira de Computação (SBC).
- Bispo Jr, E., Fonseca, L., and Santos, S. (2021). Reflexões e desafios sobre a formação na Ética em pesquisa na computação envolvendo humanos. In *Anais do XXIX Workshop sobre Educação em Computação*, pages 488–497, Porto Alegre, RS, Brasil. SBC.
- Ghezzi, C. (2020). What is research and why we do it. In *Being a Researcher*, pages 1–23. Springer.
- Hassani, H. (2017). Research methods in computer science: The challenges and issues. *arXiv preprint arXiv:1703.04080*.
- Holz, H. J., Applin, A., Haberman, B., Joyce, D., Purchase, H., Reed, C., and Others (2006). Research methods in computing: What are they, and how should we teach them? In *Working group reports on ITiCSE on Innovation and technology in computer science education*, pages 96–114.
- Kitchenham, B. and Charters, S. (2007a). Guidelines for performing systematic literature reviews in software engineering. 2.
- Kitchenham, B. and Charters, S. (2007b). Guidelines for performing systematic literature reviews in software engineering.

- Mello, R. and Mello, F. (2022). Ensino de metodologia de pesquisa qualitativa na graduação em computação: Uma proposta baseada em evidências. In *Anais do XXX Workshop sobre Educação em Computação*, pages 322–333, Porto Alegre, RS, Brasil. SBC.
- Morais, P., Rosa, J., Matos, E., Souza, M., Carneiro, L., and Souza, M. (2019). A formação do professorado da Área de ciência da computação no brasil: uma análise documental da pós-graduação stricto sensu. In *Anais do XXVII Workshop sobre Educação em Computação*, pages 291–304, Porto Alegre, RS, Brasil. SBC.
- Osorio, A. (2020). Uma proposta de protocolo de apoio à pesquisa em computação. Master's thesis, Universidade Federal de Pelotas.
- Osorio, A. (2024). *Metaciência da Computação: Um Estudo Sobre Taxonomia de Pesquisa*. PhD thesis, PPGC/UFPel. Defesa prevista para abril de 2024.
- Ramesh, V., Glass, R. L., Vessey, I., and Others (2004). Research in computer science: an empirical study. *Journal of systems and software*, 70(1-2):165–176.
- Tedre, M. (2007). Know your discipline: Teaching the philosophy of computer science. *Journal of Information Technology Education: Research*, 6(1):105–122.
- Tedre, M. (2011). Computing as a science: A survey of competing viewpoints. *Minds and Machines*, 21(3):361–387.
- Zorzo, A. F., Nunes, D., Matos, E., Steinmacher, I.; Leite, J., Araujo, R. M., Correia, R., and Martins, S. (2017). Referenciais de formação para os cursos de graduação em Computação. Sociedade Brasileira de Computação (SBC).