

Uma ontologia para integrar dados de evasão escolar, desempenho acadêmico e distribuição de auxílios socioeconômicos: estudo de caso no IFMG

Eduardo Cardoso Melo^{1,*†} and Marcello Peixoto Bax^{1,†}

¹ Universidade Federal de Minas Gerais (UFMG), Av. Antônio Carlos, 6627, Belo Horizonte, MG, Brasil

Abstract

The massive use of computing technologies by organizations from the most varied fields of activity is linked, in particular, to the increasingly frequent emergence of environments that require extreme decision-making capacity from managers, combining quality and agility. In scenarios marked by decisions supported by data, the support of artifacts that contribute to integrating heterogeneous data sources in organizations, such as ontologies, is essential. This article presents how to develop a domain ontology to organize knowledge related to school dropout, academic performance and distribution of socioeconomic aid in a Federal Higher Education Institution, as well as integrating data from such constructs. The methodology used, with emphasis on the SABiO approach, provided the necessary support for the entire domain ontology development process. As main results, there was the integration of data from four heterogeneous sources in accordance with the specified ontology, allowing the recovery of information that answered the competence questions that indicate the coherence between the purpose of the ontology and the reality of its scope. Regarding the institutional data analyzed, the importance of the socioeconomic aid distributed for reducing school dropouts and increasing students' academic performance was evident.

Resumo

O emprego massivo de tecnologias computacionais por organizações dos mais variados ramos de atuação está ligado, em especial, com o surgimento cada vez mais frequente de ambientes que requerem dos gestores extrema capacidade decisória, mesclando qualidade e agilidade. Em cenários marcados por decisões subsidiadas por dados, é fundamental o apoio de artefatos que contribuam para integrar fontes heterogêneas de dados nas organizações, tais como as ontologias. Este artigo apresenta como elaborar uma ontologia de domínio para organizar o conhecimento relacionado com evasão escolar, desempenho acadêmico e distribuição de auxílios socioeconômicos em uma Instituição Federal de Ensino Superior, bem como integrar os dados de tais construtos. A metodologia empregada, com destaque para a abordagem SABiO, forneceu o suporte necessário para todo o processo de desenvolvimento da ontologia de domínio. Como principais resultados, houve a integração dos dados de quatro fontes heterogêneas em conformidade com a ontologia especificada, permitindo a recuperação de informações que responderam as questões de competência que indicam a coerência entre o propósito da ontologia e a realidade do seu escopo. Quanto aos dados institucionais analisados, ficou evidente a importância dos auxílios socioeconômicos distribuídos para a redução da evasão escolar e aumento do desempenho acadêmico dos estudantes.

Keywords

Ontologies, Knowledge Organization, Knowledge Representation

1. Introdução

O aumento da oferta de cursos e vagas nas Instituições Federais de Ensino Superior (IFES) ocorrido nas últimas duas décadas habilita a discussão de questões relacionadas com a gestão acadêmica dessas unidades educacionais, em especial aquelas que dizem respeito à evasão dos estudantes e o seu desempenho acadêmico. A partir do ingresso de indivíduos com perfis e realidades distintas das vigentes até então, novas variáveis passaram a influenciar na descontinuidade do curso e no

Proceedings of the 17th Seminar on Ontology Research in Brazil (ONTOBRAS 2024) and 8th Doctoral and Masters Consortium on Ontologies (WTDO 2024), Vitória, Brazil, October 07-10, 2024.

* Corresponding author.

† These authors contributed equally.

✉ eduardo.melo@ifmg.edu.br (E.C. Melo); bax.ufmg@gmail.com (M.P. Bax)

ORCID 0000-0002-1323-5859 (E.C. Melo); 0000-0003-0503-3031 (M.P. Bax)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

desempenho dos estudantes, tais como renda mensal familiar *per capita*, condições de moradia e transporte até o campus, recursos para aquisição de alimentos e materiais de estudo, dentre outras [1, 2, 3].

A concepção dos Institutos Federais de Ensino, Ciência e Tecnologia ocorreu neste momento de valorização do ensino público, aumentando a capilaridade e alcance da rede federal de educação profissional para os mais remotos locais do país [4]. A criação do Instituto Federal de Minas Gerais (IFMG) aconteceu em 2008, o mesmo possui em sua estrutura uma Reitoria sediada em Belo Horizonte e dezoito *campi* localizados, em sua maioria, na região central do estado. São ofertados cursos técnicos, graduações e pós-graduações, objetivando promover a verticalização do ensino para os mais de vinte mil estudantes matriculados [5, 6].

Uma das ações realizadas pelas IFES com o objetivo de possibilitar a manutenção dos estudantes no ambiente acadêmico e contribuir para o seu desempenho ao longo do curso, foi o incremento na oferta de programas de assistência estudantil para os mais vulneráveis tanto econômica quanto socialmente. O apoio de tais programas seria uma estratégia eficaz para a permanência discente e para a obtenção de resultados satisfatórios nas atividades previstas pelo curso [7].

Visando fortalecer em toda a rede federal as políticas de assistência estudantil, bem como oferecer condições para que os estudantes permaneçam na instituição até a conclusão do seu curso, o Governo Federal institucionalizou em 2010 o Programa Nacional de Assistência Estudantil (PNAES), indicando a necessidade de atenção para com moradia, alimentação, transporte e apoio pedagógico aos estudantes [8]. Desde então, as IFES passaram a contar com orientação formal para a atuação da assistência estudantil e dotação orçamentária específica [9].

Considerando que alguns dos objetivos do PNAES são a redução da evasão escolar e a melhora no desempenho acadêmico nas IFES, a existência de um monitoramento contínuo com formas de mensurar e analisar o seu impacto nesses aspectos seria benéfica para todos os envolvidos (IFES, Governo Federal, sociedade civil, dentre outros). Entretanto, esta ainda não é a realidade encontrada atualmente, conforme observado por [10, 11, 12]. Estudos que visam promover uma análise do PNAES muitas vezes o fazem sob o ponto de vista descritivo [13, 14, 15], buscando explorar como as ações assistenciais são conduzidas pela IFES, sem focar no alcance dos resultados em contraponto às propostas do Programa.

Dentre as razões para isso, [16, 17] destacam a dificuldade para integração de dados oriundos de fontes heterogêneas, fazendo com que a geração de novos conhecimentos fique prejudicada em função de um ambiente incapaz de fornecer os subsídios adequados para o entendimento de questões acadêmicas [18]. Trata-se de problema não exclusivo da área educacional pública, mas em um contexto marcado pela alta evasão escolar, o contorno da situação se delinea negativamente. A partir do momento que a integração dos dados é frágil ou o custo para alcançá-la é inviável, o gestor público pode ficar refém de informações que não contribuem satisfatoriamente para a tomada de decisões com qualidade [19, 20].

Este artigo descreve como elaborar uma ontologia de domínio para organizar formalmente o conhecimento sobre evasão escolar, desempenho acadêmico e distribuição de auxílios socioeconômicos no âmbito do IFMG, suportando a integração dos dados desses construtos para identificar eventuais relações entre eles.

A aplicabilidade desta pesquisa se justifica pela demanda por uma abordagem diferenciada para integrar os dados dos construtos envolvidos, ancorada na construção de uma ontologia que une semântica e formalmente os dados do domínio, bem como padroniza o entendimento dos conceitos e relações existentes neste contexto. Como ineditismo e diferencial, destaca-se a integração com base ontológica dos dados dos construtos, promovendo contribuições diretas para o aspecto da transdisciplinaridade existente na área de Ciência da Informação (CI).

2. Ontologias

Uma das principais definições sobre ontologia foi introduzida por [21] ao afirmar que se trata de um termo advindo da Filosofia com o intuito de indicar a “especificação explícita de uma

conceitualização”. Tal conceitualização refere-se a uma compreensão de mundo simplificada e abstrata a ser representada com um objetivo bem delineado, sendo verdadeiro o fato de que qualquer sistema baseado em conhecimento está, direta ou indiretamente, vinculado a alguma conceitualização.

A partir de uma extensão da conceitualização proposta por [21], [22] define ontologia como a “especificação formal de uma conceitualização compartilhada”. Esta formalidade pressupõe a indicação explícita de quais são os conceitos, propriedades, relações e restrições da ontologia, estabelecendo assim o conhecimento consensual advindo do entendimento provido pelos pares. Quanto à conceitualização, trata-se da noção de que a ontologia será, quando finalizada, um modelo abstrato de algum domínio ou fenômeno existente no que conhecemos por mundo real.

De acordo com [23], uma ontologia pode ser compreendida como um artefato composto de um vocabulário vinculado a determinada realidade que, quando utilizado por indivíduos ou sistemas eletrônicos, permite descrever o domínio em questão. Este artefato possui, ainda, a explicitação de fatos aceitos e relacionados com o sentido desejado para aqueles termos constantes do vocabulário. [24, 25] apresentam entendimento semelhante, reforçando a importância da característica das ontologias de criarem um vocabulário de termos com a devida especificação dos significados e apreciação conjunta de especialistas, o qual será compartilhado sob a forma de conhecimento.

Em uma conceitualização mais relacionada com a área de Ciência da Informação, [26, 27] afirmam que uma ontologia se refere ao entendimento semântico, por vários indivíduos, de um domínio específico, normalmente representada em uma estrutura formal baseada em linguagem lógica e contendo as devidas explicações semânticas sobre cada conceito relevante, suas propriedades e também acerca dos relacionamentos entre estes. [28, 29, 22] corroboram este entendimento, complementando que as ontologias possuem importante papel na integração e interoperabilidade semântica entre distintos sistemas de informação na *web*.

Embora diferentes propostas de classificações das ontologias tenham sido indicadas por variados autores ao longo dos últimos anos, existem duas tipificações mais comuns que são adequadas para complementar o entendimento teórico requerido pelo presente artigo. Essa classificação se origina dos estudos de [22, 30, 31]. Em uma ontologia de **fundamentação** ou de **alto nível** o objetivo é formalizar a descrição de classes gerais sem dependência direta com algum domínio, abordando conceitos abrangentes como tempo, objeto, espaço e matéria. Por outro lado, na ontologia de **domínio** busca-se explorar um domínio específico para descrever os seus conceitos e considerando que, normalmente, aqueles que fazem parte da estrutura da ontologia são próprios do domínio.

A construção de ontologias de domínio requer o suporte de uma ontologia de fundamentação apropriada, responsável por indicar os possíveis estados do mundo conforme os compromissos ontológicos fundamentais [32]. O alinhamento da ontologia de domínio com a ontologia de fundamentação permite, de acordo com o autor, especificar um subconjunto desses estados conforme sua admissibilidade dentro do escopo de um domínio particular. A UFO (*Unified Foundational Ontology*) é uma ontologia de fundamentação elaborada com o agrupamento de teorias das áreas de ontologia formal, linguística, lógica filosófica e psicologia cognitiva, sendo composta de pequenas teorias relacionadas, especialmente, com aspectos essenciais da modelagem conceitual [33, 34].

Conforme [35], as ontologias contribuem para a interoperabilidade e integração semântica de dados com a definição do significado dos termos e das relações entre esses termos dentro de um domínio específico, explicitando, assim, como eles devem ser interpretados por pessoas e máquinas. [36] reforçam a validade do uso de ontologias como referências semânticas na integração de dados que, mesmo obtidos em variadas origens e dotados de significados e formatos diferentes, são mapeados para uma estrutura uniforme que, por sua vez, padroniza o acesso por aplicações heterogêneas.

3. Metodologia

Esta pesquisa se caracteriza como exploratória, tendo como justificativa um estudo de caso realizado no âmbito do IFMG para organizar o conhecimento relacionado à evasão escolar, desempenho

acadêmico e distribuição de auxílios socioeconômicos no formato de uma ontologia. Este artefato permitiu integrar dados desses construtos e suportou a geração de novos conhecimentos acerca das relações entre eles, os quais podem servir como subsídios aos gestores acadêmicos na construção de políticas institucionais [37].

O processo de criação de ontologias com qualidade não é trivial, pelo contrário, os desafios para que o conhecimento seja organizado sob esta abordagem são inúmeros e, por vezes, complexos. Assim, utilizar metodologias que auxiliam o projeto de criação da ontologia é benéfico não apenas para os indivíduos envolvidos, mas também para o produto final a ser entregue.

Neste sentido, utilizou-se a abordagem proposta por [38] denominada *Systematic Approach for Building Ontologies* (SABiO). Trata-se de um método sistematizado para organizar o processo de desenvolvimento de ontologias, composto por uma estrutura relativamente objetiva e adaptável no que tange às atividades e diretrizes para o seu emprego, podendo ser integrada com a UFO para a construção de ontologias aderentes a esta ontologia de fundamentação.

Observa-se na Figura 1 que a SABiO é composta de um processo macro (**Desenvolvimento**) constituído de fases próprias e vários processos de **Apoio**. O foco da metodologia é contribuir para a construção de ontologias de domínio, seja como uma ontologia de referência, vista sob a ótica de um tipo particular de modelo conceitual, ou como uma ontologia operacional, isto é, uma versão que permite a leitura e interpretação da ontologia por máquinas.

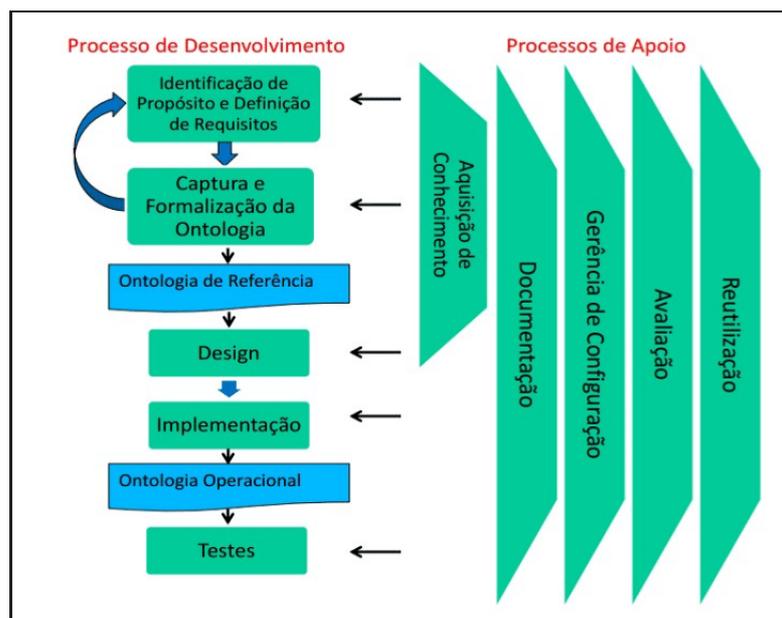


Figura 1: Processos e fases da Metodologia SABiO (adaptado e traduzido de [38]).

Na primeira fase (“Identificação de Propósito e Definição de Requisitos”) ocorreu a identificação do propósito da ontologia, bem como a eliciação dos seus requisitos e registro no formato de Questões de Competência (QC), responsáveis por delimitar o escopo da ontologia.

A segunda fase (“Captura e Formalização da Ontologia”) visava elaborar a ontologia de referência e envolveu: a) análise documental sobre as estruturas de armazenamento de dados do sistema gerenciador acadêmico do IFMG; b) definição formal dos conceitos envolvidos, com a realização de entrevistas não estruturadas em reuniões virtuais com duas especialistas de domínio que atuam no IFMG nas áreas de ensino e assistência estudantil, além de reaproveitamento conceitual da ontologia *Curriculum Course Syllabus Ontology* (CCSO)²; c) elaboração de um modelo gráfico no *software Visual*

² <https://vkreations.github.io/CCSO/>

Paradigm versão 16.2 contendo os elementos da ontologia e suas relações em conformidade com a linguagem *OntoUML* [34], um perfil que incorpora no Diagrama de Classes da UML estereótipos que fazem com que a ontologia resultante esteja alinhada com a UFO.

Com o objetivo de simplificar o entendimento acerca de como a presente pesquisa conduziu as ações requeridas para cumprir o proposto pelas fases Design, Implementação e Testes da SABiO, um conjunto de atividades foi definido e pode ser visualizado na Figura 2. Na sequência é feito o detalhamento sobre a execução de cada atividade.

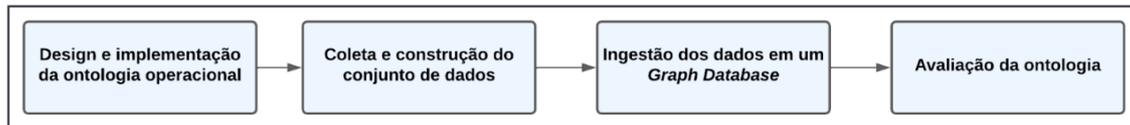


Figura 2: Atividades para a construção de um ambiente informacional integrado à ontologia

No escopo da primeira atividade (Design e implementação da ontologia operacional), houve a transformação da ontologia de referência em uma ontologia operacional no formato *OWL* (*Ontology Web Language*) e sintaxe *Turtle*, tarefa conduzida com o auxílio dos softwares *Visual Paradigm* (e *plugin* para *OntoUML*³) e *Protégé* (versão 5.5.0).

Na segunda atividade (Coleta e construção do conjunto de dados) é válido pontuar que, atualmente, a instituição dispõe de dois diferentes sistemas informatizados sem conexão direta ou interoperabilidade que gerenciam os dados demandados por esta pesquisa: um da assistência estudantil, específico para controle dos processos de concessão de auxílios socioeconômicos, e outro para gestão acadêmica. Ambos armazenam dados em repositórios do SGBD MySQL, motivo pelo qual criou-se um repositório próprio com esta tecnologia em ambiente local para receber os dados brutos selecionados dessas fontes.

Visando enriquecer esses registros obtidos junto à instituição, houve a integração com informações de repositórios externos: a Plataforma Nilo Peçanha e a central de dados abertos do INEP (especificamente os conjuntos com dados acerca do ENADE e do Censo da Educação Superior). O repositório próprio criado serviu como ambiente para execução de tarefas de pré-processamento nos dados brutos coletados, tais como a transformação e criação de atributos.

A coleta dos dados restringiu-se ao âmbito dos cursos de graduação presenciais ofertados pelos *campi* do IFMG, abrangendo as turmas de ingressantes entre o primeiro semestre letivo de 2018 e o segundo semestre letivo de 2022. Quanto aos auxílios socioeconômicos, foram analisados aqueles oferecidos sob a nomenclatura de “Bolsa Permanência (BP)”.

Em seguida, procedeu-se com o mapeamento entre os atributos do conjunto de dados e os conceitos da ontologia operacional elaborada anteriormente, habilitando a transformação e exportação dos dados armazenados no repositório intermediário para o formato requerido por uma plataforma de *Graph Database* (banco de dados em grafo).

Na terceira atividade (Ingestão dos dados em um *Graph Database*), os dados foram extraídos do repositório intermediário com a execução de consultas SQL e salvos em arquivos CSV (*Comma-Separated Values*) estruturados de acordo com os elementos integrantes da ontologia operacional. Optou-se pela escolha da plataforma *Neo4j* como ambiente para o gerenciamento (persistência e manipulação) do banco de dados em grafo, cuja versão para uso livre e sem custos disponibiliza todas as funcionalidades necessárias para o desenvolvimento das tarefas desta pesquisa, além de proporcionar alta performance na execução de ações (como leitura e escrita de dados) e ter baixa curva de aprendizado [39]. A importação dos arquivos CSV para o repositório em grafo demandou a execução de comandos específicos da linguagem *Cypher*, a qual é utilizada no *software Neo4j Browser*

³ <https://github.com/nemo-ufes/ontouml-lightweight-editor>

versão 5.12, gerando como resultado a criação de todas as instâncias e suas relações no projeto do *Neo4j*.

A atividade de avaliação da ontologia verificou a sua capacidade de responder as QCs elaboradas anteriormente, além da construção de uma matriz de rastreabilidade para verificar claramente aqueles elementos da ontologia (conceitos, relacionamentos ou propriedades) que participaram da resolução de cada QC, conforme proposto em [40]. Ressalta-se que o uso de médias simples nas comparações promovidas nas respostas às QCs foi precedido por uma análise que indicou se tratar de um conjunto de dados com distribuição simétrica.

4. Desenvolvimento

A primeira atividade executada definiu o propósito da ontologia: “Recuperar informações sobre a distribuição de auxílios socioeconômicos no formato de Bolsa Permanência em programas de assistência estudantil do IFMG para analisar sua contribuição na redução da evasão escolar e no aumento do desempenho acadêmico dos estudantes de cursos superiores presenciais”.

As QCs são apresentadas a seguir e buscaram explorar, ao máximo, aspectos das relações entre os três construtos analisados pela pesquisa:

- **QC1:** O desempenho acadêmico dos estudantes atendidos por auxílios socioeconômicos é superior ao daqueles que não são atendidos?
- **QC2:** Os estudantes que recebem auxílios socioeconômicos evadem menos proporcionalmente do que aqueles que não recebem?
- **QC3:** A proporção de estudantes formados é maior entre aqueles que recebem auxílios socioeconômicos?
- **QC4:** Existem variações no desempenho acadêmico e na taxa de evasão de acordo com o tipo de Bolsa Permanência destinada aos estudantes?

O levantamento das QCs contribuiu para a organização de uma visão mais otimizada sobre a abrangência da ontologia dentro do contexto dos três construtos. Para promover a captura e formalização da ontologia, foi preciso, inicialmente, especificar em linguagem natural as principais características dos seus componentes e relacionamentos. Para representar esta especificação, foi elaborado um modelo gráfico baseado na linguagem *OntoUML* com todos os conceitos definidos para a ontologia de referência e as relações entre eles. Tal construção foi realizada em várias iterações com a participação das especialistas de domínio no sentido de sugerir melhorias ou consensuar o entendimento sobre algum aspecto particular. A Figura 3 apresenta a versão final da ontologia de referência, denominada *Onto-Educacional*.

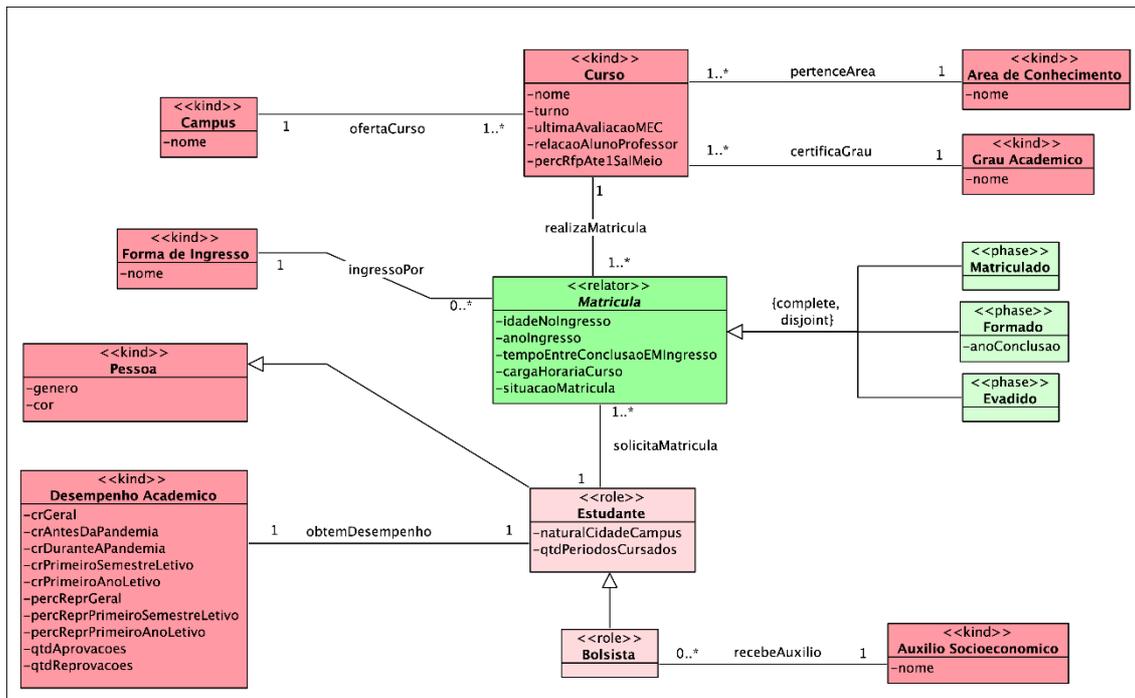


Figura 3: Ontologia de referência elaborada com OntoUML

É possível notar que a maior parte das classes que compõem a ontologia de referência foram estereotipadas como *kinds*, indicando que se referem a conceitos dotados do princípio uniforme de identidade, de acordo com a especificação da UFO. A classe *Curso* se relaciona diretamente com as classes *Campus* (para identificar o local de oferta), *Área de Conhecimento* (para caracterizar a área de estudos do curso) e *Grau Acadêmico* (para classificar o curso conforme a certificação final dos estudantes). A classe *Estudante* possui uma relação de especialização com a classe *Pessoa*, tendo recebido o estereótipo *role* devido ao fato deste conceito representar um papel executado pelo indivíduo dentro do contexto acadêmico em determinado momento e sob condições específicas.

A abstração dos dados que representam o desempenho acadêmico de um estudante levou à definição da classe *Desempenho Acadêmico*, estereotipada como *kind*. O estereótipo *role* também foi aplicado à classe *Bolsista*, pois se trata de um papel com aplicação dependente de outros fatores, como a aprovação em processo seletivo da assistência estudantil. Quando existe a instanciação de um bolsista, têm-se a definição de qual apoio o mesmo receberá, simbolizado pela classe *Auxílio Socioeconômico*. A relação entre um estudante e seu curso é mediada pelo relator *Matricula*, permitindo abstrair os dados próprios deste relacionamento, bem como integrá-lo com a classe *Forma de Ingresso*. As mudanças de estado de uma matrícula são representadas por três classes disjuntas e completas estereotipadas como *phase*, um tipo anti-rígido, apontando se ela está ativa (*Matriculado*) ou encerrada (*Formado* ou *Evadido*). Por motivo de simplificação do modelo, atributos como *genero* e *cor* foram definidos como propriedades ao invés de *qualities*.

A verificação de adequação do modelo às regras estipuladas pela UFO envolveu o uso de uma funcionalidade disponível no *plugin* da linguagem *OntoUML*, onde é analisada a coerência sintática dos elementos e suas relações com um conjunto de regras e restrições formais. Quando alguma inconsistência é identificada pelo processo de verificação, o *plugin* apresenta ao usuário detalhes sobre esse desvio de acordo com o seu nível de gravidade, erro (inconsistências ontológicas ou lógicas) ou aviso (problemas menos críticos). Após submeter a Onto-Educacional ao processo de verificação, nenhuma inconsistência foi encontrada.

No escopo da terceira fase, exportou-se o modelo criado em *OntoUML* para um arquivo com as especificações da UFO. Foi gerado um arquivo no formato *OWL* e sintaxe *Turtle* contendo tanto os elementos da ontologia quanto as definições sobre reutilização de recursos da UFO em tais elementos,

como *Kinds*, *Roles* e *Phases*. A Figura 4 contém um fragmento deste arquivo, no qual é possível observar a especificação da classe *Pessoa* integrada com recursos da *OWL* e da *UFO*.

```
onto-educacional:Pessoa rdf:type owl:Class, gufo:Kind, owl:NamedIndividual;  
rdfs:subClassOf gufo:FunctionalComplex;  
rdfs:label "Pessoa"@en.
```

Figura 4: Fragmento de código do arquivo com a ontologia exportada

A continuidade da transformação da ontologia de referência em ontologia operacional demandou a criação de um projeto no *software Protégé*, tendo como base o arquivo exportado em *OWL*. Na Figura 5 verifica-se a integração entre as classes próprias da Onto-Educacional e aquelas que compõem a UFO (quadro *Class hierarchy*), assim como a instanciação das classes de tipos (quadro *Individuals by type*).

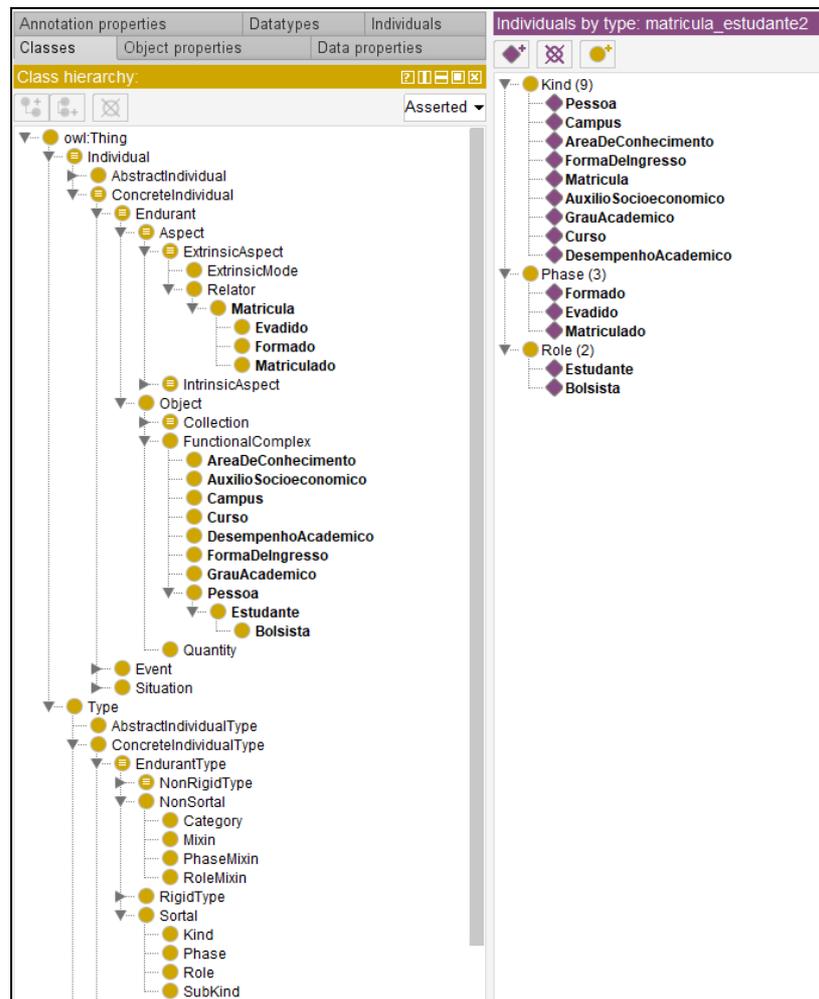


Figura 5: Integração das classes da Onto-Educacional com a UFO

Encerradas as atividades de design e implementação da ontologia operacional, a versão final foi publicada em *OWL* no repositório virtual do projeto e passou-se à atividade de construção do conjunto de dados a partir da obtenção de registros nos ambientes informacionais do IFMG, da Plataforma Nilo Peçanha e do INEP. O conjunto de dados mapeado conforme a ontologia operacional ficou com 10.263 registros compostos de 32 atributos (11 originais e 21 criados).

Na sequência, houve a ingestão dos registros em um banco de dados no formato de grafos, no caso o *Neo4j*. Os registros do repositório intermediário foram exportados no formato CSV e, com a execução de comandos na linguagem *Cypher*, importados no *Neo4j* em um repositório específico.

A primeira forma de avaliação da ontologia refere-se à execução de consultas escritas na linguagem *Cypher* para responder às QCs por meio da recuperação das informações armazenadas no repositório em grafo do *Neo4j*. A seguir são apresentados os resultados obtidos em cada QC. A título de exemplificação, apenas na primeira QC é exposto o comando *Cypher* utilizado para recuperação das informações.

- **QC1: O desempenho acadêmico dos estudantes atendidos por auxílios socioeconômicos é superior ao daqueles que não são atendidos?**

Consulta escrita em *Cypher*:

```
MATCH (bol:Bolsista)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH COUNT(bol) AS qtd, AVG(des.crGeral) AS cr_geral_medio, 'Bolsistas' AS tipo
RETURN tipo, qtd AS qtd, ROUND(cr_geral_medio, 2) AS cr_geral_medio
UNION ALL
MATCH (est:Estudante)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH COUNT(est) AS qtd, AVG(des.crGeral) AS cr_geral_medio, 'Não bolsistas' AS tipo
RETURN tipo, qtd AS qtd, ROUND(cr_geral_medio, 2) AS cr_geral_medio;
```

O indicador que mediu o desempenho acadêmico nesta pesquisa foi o coeficiente de rendimento (CR) do estudante obtido ao longo do curso. Esta QC envolveu a análise dos 10.263 estudantes do conjunto de dados, sendo que 2.142 (20,9%) são bolsistas e 8.121 (79,1%) não receberam nenhum tipo de auxílio socioeconômico. A média do CR dos estudantes bolsistas é de 58,99, superior à média dos não bolsistas que é de 40,15. Trata-se de uma considerável diferença (18,84 pontos ou 32% superior), indicando que, no contexto desta pesquisa, os estudantes que recebem auxílios possuem desempenho acadêmico superior. Os resultados da validação estatística (Teste T de Student) comprovaram a significância desta diferença.

- **QC2: Os estudantes que recebem auxílios socioeconômicos evadem menos proporcionalmente do que aqueles que não recebem?**

Analisando a Tabela 1, tem-se que a proporção de estudantes bolsistas que evadiram (26,9%) é menor do que entre os não bolsistas (47,5%), permitindo inferir a contribuição positiva do recebimento dos auxílios neste indicador. Quando investigada a diferença entre a proporção dos dois grupos, a evasão dos bolsistas é 43,5% menor do que aquela dos não bolsistas.

Tabela 1

Resultados da consulta para a QC2

Grupo	Estudantes	Evadidos	Proporção
Bolsistas	2.142	577	26,9%
Não bolsistas	8.121	3.857	47,5%

- **QC3: A proporção de estudantes formados é maior entre aqueles que recebem auxílios socioeconômicos?**

O resultado da terceira QC fornece indícios similares aos obtidos nas questões anteriores sobre o papel dos auxílios socioeconômicos no apoio aos estudantes do IFMG. Verifica-se na Tabela 2 que a proporção de estudantes bolsistas que formaram (11,2%) é duas vezes superior à proporção dos não bolsistas (5,3%).

Tabela 2

Resultados da consulta para a QC3

Grupo	Estudantes	Formados	Proporção
Bolsistas	2.142	239	11,2%
Não bolsistas	8.121	431	5,3%

- **QC4: Existem variações no desempenho acadêmico e na taxa de evasão de acordo com o tipo de Bolsa Permanência destinada aos estudantes?**

A distribuição da Bolsa Permanência (BP) no IFMG é materializada com o pagamento de dez parcelas anuais de um auxílio que pode ser de quatro tipos: BP1 (R\$400,00), BP2 (R\$300,00), BP3 (R\$200,00) e BP4 (R\$150,00). De acordo com a Tabela 3, BP4 é o tipo de bolsa mais distribuído aos estudantes, seguido por BP3, BP2 e, por fim, BP1, em uma ordem inversamente proporcional ao valor pago por cada uma delas. Esta sequência, contudo, não se repete ao ser analisada a taxa de evasão por tipo de bolsa; enquanto a menor taxa foi observada entre os estudantes que recebem a BP4, taxas próximas e superiores caracterizam a evasão de quem é beneficiado com os outros três tipos de bolsa.

Neste sentido, os dados indicam que 60% dos bolsistas que evadiram não conseguiram se manter no curso mesmo recebendo valores maiores de auxílio socioeconômico. Quando o coeficiente de rendimento acadêmico é apreciado, observa-se uma situação na qual os valores das bolsas diminuem (BP1 para BP4) enquanto as médias aumentam (46,58 para 61,67). Uma hipótese para que a menor média advenha justamente dos estudantes beneficiados com o tipo de bolsa de maior valor (BP1) é que se trata de indivíduos com alta vulnerabilidade socioeconômica que convivem em contextos familiares marcados por situações que dificultam a rotinas de estudos.

Tabela 3

Resultados da consulta para a QC4

Tipo de bolsa	Bolsistas	Evadidos	Taxa de evasão	CR geral médio
BP1	61	18	29,5%	46,58
BP2	347	107	30,8%	55,57
BP3	693	218	31,5%	57,77
BP4	1.041	234	22,5%	61,67

Como complemento à verificação da ontologia, a criação de uma matriz de rastreabilidade (Tabela 4) vinculou a cada QC os elementos da ontologia utilizados na sua resposta. A classe *Pessoa* não foi instanciada diretamente no repositório, apenas por meio de sua especialização via classe *Estudante*, por isso não foi listada na matriz de rastreabilidade. Houve baixo emprego das propriedades das classes nas QCs em função da delimitação de tamanho do artigo.

Tabela 4

Matriz de rastreabilidade das QCs da ontologia

QC	Conceitos, relações ou propriedades
1	Bolsista, obtemDesempenho, DesempenhoAcademico, crGeral, Estudante

2	Matricula, confirmaMatricula, Bolsista, obtemDesempenho, DesempenhoAcademico, situacaoMatricula, Estudante
3	Matricula, confirmaMatricula, Bolsista, obtemDesempenho, DesempenhoAcademico, situacaoMatricula, Estudante
4	Bolsista, recebeAuxilio, AuxilioSocioeconomico, obtemDesempenho, DesempenhoAcademico, crGeral, nome, situacaoMatricula

A matriz de rastreabilidade mostra como as QCs estão alinhadas com a ontologia e cobrem, adequadamente, o domínio envolvido.

5. Considerações finais

Apresentou-se neste artigo um estudo envolvendo diferentes áreas do conhecimento, como Educação, Ciência da Computação e Ciência da Informação, para, com o suporte de uma abordagem ontológica, organizar formalmente os conceitos relacionados com evasão escolar, desempenho acadêmico e distribuição de auxílios socioeconômicos. A ontologia elaborada foi a base para a construção de um ambiente tecnológico de recuperação de informações capaz de integrar os dados advindos de fontes heterogêneas e permitir a geração de conhecimentos englobando as relações entre esses três construtos.

Partindo de um alinhamento com o proposto pela Metodologia SABiO, vários artefatos elaborados compõem o conjunto de produtos resultantes deste estudo, tais como a ontologia de referência produzida em coerência com a ontologia de fundamentação UFO, a ontologia operacional que tornou possível a leitura e utilização dos dados reais do IFMG a nível de máquina e o repositório no formato de grafos contendo as instâncias e suas relações. Dada a capacidade organizadora e expressiva das ontologias, tem-se como principal resultado um modelo formal para auxiliar na padronização dos conceitos envolvidos no contexto em questão, bem como ser reutilizado por outras IFES como sustentação para integrar seus próprios dados.

É válido ressaltar que o desenvolvimento do ambiente tecnológico que suportou a ingestão dos dados organizados pela ontologia foi de grande valia para que a transformação das QCs em consultas escritas com uma sintaxe específica produzisse resultados tangíveis e interpretáveis e, adicionalmente, colaborasse para verificar a adequação da ontologia em relação a seus requisitos.

Quanto aos resultados específicos obtidos com os registros do IFMG, merece destaque o impacto positivo da distribuição de auxílios socioeconômicos tanto na evasão escolar quanto no desempenho acadêmico e na proporção de formados, demonstrando a importância de se aprimorar a concessão de tais benefícios para a comunidade acadêmica com vulnerabilidades socioeconômicas. Tais resultados foram encaminhados à gestão da instituição para análise e aprimoramento dos processos internos do setor de assistência estudantil.

Uma limitação deste estudo consiste na análise somente de cursos de graduação, embora o IFMG ofereça diversos cursos técnicos em suas unidades. Eventuais trabalhos futuros que busquem ampliar as possibilidades de emprego prático da Onto-Educacional podem envolver os dados desses cursos.

Agradecimentos

Agradecemos ao Instituto Federal de Minas Gerais (IFMG) pelo suporte à elaboração deste artigo a partir da anuência para a utilização dos dados institucionais. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) e do Programa de Pós-Graduação em Gestão & Organização do Conhecimento (PPGGOC) da Universidade Federal de Minas Gerais (UFMG).

Referências

- [1] Nonato BF, et al. Mudanças no perfil dos estudantes da UFMG: desafios para a prática docente, *Revista Docência do Ensino Superior* (2020); 10; 1–21.
- [2] Ristoff D. O novo perfil do campus brasileiro: uma análise do perfil socioeconômico do estudante de graduação, *Avaliação* (2014); 19; 3.
- [3] Senkevics AS, Mello UM. O perfil discente das universidades federais mudou pós-Lei de Cotas? *Cadernos de Pesquisa* (2019); 49; 184-208.
- [4] Brasil. Ministério da Educação. Instituições da rede federal, <http://portal.mec.gov.br/rede-federal-inicial/instituicoes> (accessed 19 May 2024).
- [5] Brasil. Ministério da Educação. Plataforma Nilo Peçanha, <https://www.gov.br/mec/pt-br/pnp> (accessed 02 Apr 2024).
- [6] Instituto Federal de Minas Gerais, IFMG. Histórico e missão, <https://www.ifmg.edu.br/portal/sobre-o-ifmg/historico-e-missao> (accessed 10 Apr 2024).
- [7] Santos EHM. As configurações da assistência estudantil no Brasil: o cenário do PNAES, https://bdt.d.ibict.br/vufind/Record/UFSP_cd5000bd2db156fb8d691bbf28db713f (2020, accessed 05 Mar 2024).
- [8] Brasil. Decreto N° 7.234, de 19 de julho de 2010. Dispõe sobre o Programa Nacional de Assistência Estudantil (PNAES), http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/decreto/d7234.htm (accessed 18 May 2024).
- [9] Pinheiro EB. A trajetória da assistência estudantil no contexto da política de educação: educação superior, assistência estudantil e o PNAES, *Anais do XVI Encontro Nacional de Pesquisadores em Serviço Social* (2018); 16; 1-21.
- [10] Brito ACT, Costa AKT, Almeida VL. Gestão universitária: políticas comparadas de monitoramento do programa bolsa permanência em universidades federais brasileiras, *Revista Ibero-Americana de Estudos em Educação* (2019); 1862-1875.
- [11] Fava HL, Hall RJ, Cintra RF. Indicadores na assistência estudantil: o que a literatura diz, *Jornal de Políticas Educacionais* (2022); 16; 1.
- [12] Santos DG, Siqueira PHL, Diniz DM. Assistência estudantil e PNAES: um estudo bibliométrico sobre a produção científica recente, *Práticas em Gestão Pública Universitária* (2021); 5; 28-53.
- [13] Eloi SSS, et al. O Programa Nacional de Assistência Estudantil–PNAES na Universidade Federal de Viçosa (UFV): uma avaliação na percepção dos beneficiários, *Revista Gestão Universitária na América Latina-GUAL* (2019); 12; 3; 106-129.
- [14] Freitas FCL. Avaliação do Programa Nacional de Assistência Estudantil na UFC: a permanência discente como direito à educação dos estudantes do campus de Sobral, <https://repositorio.ufc.br/handle/riufc/58080> (2021, accessed 27 Feb 2024).
- [15] Recktenvald M, Mattei L, Pereira VA. Avaliando o Programa Nacional de Assistência Estudantil (PNAES) sob a ótica das epistemologias, *Revista da Avaliação da Educação Superior* (2018); 23; 405-423.
- [16] Jorge EMFJ, et al. Arquitetura da informação analítica para integração de dados da pesquisa e pós-graduação: um estudo de caso da Universidade do Estado da Bahia, *Informação & Informação* (2020); 25; 115-140.
- [17] Khamdamov U, et al. Models of integration of higher education management information systems, *International Conference on Information Science and Communications Technologies* (2021), <https://ieeexplore.ieee.org/abstract/document/9670171> (accessed 10 Apr 2024).
- [18] Rodzi NAH, Othman, MS, Yusuf, LM. Significance of data integration and ETL in business intelligence framework for higher education, *International Conference on Science in Information Technology* (2015); 181-186.
- [19] Sanna FA. A importância da integração de dados para a qualidade da decisão do gestor público, <https://repositorio.enap.gov.br/handle/1/1940> (2014, accessed 21 May 2024).
- [20] Gruber TR. Toward principles for the design of ontologies used for knowledge sharing, *International Journal Human-Computer Studies* (1995); 43; 5; 907-928.

- [21] Borst WN. Construction of engineering ontologies for knowledge sharing and reuse, <https://research.utwente.nl/en/publications/construction-of-engineering-ontologies-for-knowledge-sharing-and-> (1997, accessed 05 Feb 2024).
- [22] Guarino N. Understanding, building and using ontologies, *International journal of human-computer studies* (1997); 46; 2-3; 293-310.
- [23] Uschold M, Gruninger M. Ontologies: principles, methods and applications, *The knowledge engineering review* (1996); 11; 2; 93-136.
- [24] DeGiacomo G, et al. Using ontologies for semantic data integration. In: Flesca, S., Greco, S., Masciari, E., Saccà, D. (eds) *A comprehensive guide through the italian database research over the last 25 years*, *Studies in Big Data* (2018); 31.
- [25] Estevão JSB, Strauhs FR. Proposta de uma ontologia como modelo de referência no domínio da memória organizacional histórica, *Perspectivas em Ciência da Informação* (2013); 18; 4; 35-53.
- [26] Jacob EK. Ontologies and the semantic web, *Bulletin of the American Society for Information Science and Technology* (2003); 29; 4; 19-39.
- [27] Breitman KK, Felicíssimo CH, Casanova MA. CATO-A lightweight ontology alignment tool. <http://www-di.inf.puc-rio.br/~casanova/Publications/Papers/2005-Papers/2005-CAISE-Breitman.pdf> (2005, accessed 20 Feb 2024).
- [28] Campos JG, et al. Estado da arte sobre arquiteturas de sistemas para integração de dados, <https://biblioteca.ibp.org.br/riooilegas/pt-BR/search/39610?exp=> (2020, accessed 02 Mar 2024).
- [29] Gonçalves JEA. Método ágil de integração semântica de dados científicos baseado em ontologias, <https://repositorio.ufmg.br/handle/1843/34013> (2020, accessed 21 Feb 2024).
- [30] Guizzardi G. Ontological foundations for structural conceptual model, *CTIT-Centre for Telematics and Information Technology*, <https://research.utwente.nl/en/publications/ontological-foundations-for-structural-conceptual-models> (2005, accessed 18 Apr 2024).
- [31] Haav HM, Lubi TL. A survey of concept-based information retrieval tools on the web, *Proceedings of the 5th East-European Conference ADBIS* (2001); 29-41.
- [32] Guizzardi G. On ontology, ontologies, conceptualizations, modeling languages, (Meta) models, *Frontiers in Artificial Intelligence and Applications, Databases and Information Systems IV* (2007); 18-39.
- [33] Guizzardi G, Falbo R, Guizzardi R. A importância de ontologias de fundamentação para a engenharia de ontologias de domínio: o caso do domínio de Processos de Software, *IEEE Latin America Transactions* (2008); 6; 3; 244-251.
- [34] Guizzardi G, et al. Towards ontological foundations for conceptual modeling: The unified foundational ontology (UFO) story, *Applied ontology* (2015); 10; 3-4; 259-271.
- [35] Guarino N. Formal Ontology and information systems, *Proceedings of FOIS* (1998); 3-15.
- [36] Britto GC, Ruy FB, Azevedo CLB. Um ambiente para integração de dados abertos relativos à despesa pública. *Ontobras* (2020); 13; 1-14.
- [37] Gil AC. Como elaborar projetos de pesquisa. São Paulo: Atlas, 2017.
- [38] Falbo R. SABiO: Systematic Approach for Building Ontologies, http://ceur-ws.org/Vol-1301/ontocomodise2014_2.pdf (2014, accessed 20 Jan 2024).
- [39] Scifo E. *Graph data science with Neo4j*. Packt Publishing, 2023.
- [40] Silveira PDN et al. Uma ontologia de referência para arquiteturas pedagógicas. *Anais do XXXII Simpósio Brasileiro de Informática na Educação* (2021); 32; 24-34.