UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

REPRESENTAÇÃO DO SUBDOMÍNIO DO REGISTRO DO DIPLOMA DIGITAL BRASILEIRO: PROPOSIÇÃO DE UMA ONTOLOGIA OPERACIONAL BASEADA EM PADRÕES DA WEB SEMÂNTICA

Renan Teixeira de Souza
Vitória – ES - Brasil
2026

RENAN TEIXEIRA DE SOUZA

REPRESENTAÇÃO DO SUBDOMÍNIO DO REGISTRO DO DIPLOMA DIGITAL BRASILEIRO: PROPOSIÇÃO DE UMA ONTOLOGIA OPERACIONAL BASEADA EM PADRÕES DA WEB SEMÂNTICA.

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Ciência da Informação.

Orientador: Henrique Monteiro Cristóvão

Coorientadora: Daniela Lucas da Silva Lemos

Área de concentração:
Informação, Sociedade e Cultura.

VITÓRIA
2026

Lista de Abreviaturas

ABox	Assertional Box
CAdES	CMS Advanced Electronic Signatures
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CC BY	Creative Commons Attribution
CCJE	Centro de Ciências Jurídicas e Econômicas
CEN	European Committee for Standardization
CEP	Código de Endereçamento Postal
CNPJ	Cadastro Nacional da Pessoa Jurídica
CONARQ	Conselho Nacional de Arquivos
CPF	Cadastro de Pessoas Físicas
CSV	Comma-Separated Values
CTS	CoreTrustSeal
DCAT	Data Catalog Vocabulary
DCMI	Dublin Core Metadata Initiative
DCTERMS	DCMI Metadata Terms — designação associada ao prefixo dcterms:
DMP	Data Management Plan
DOI	Digital Object Identifier
DSDP	Digital Student Data Portability
e-ARQ Brasil	Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos
ED	Educação — código de campo de aplicação do INPI
EFGD	Estratégia Federal de Governo Digital
ELM	European Learning Model
ELMO	Formato XML utilizado pelo EMREX para o intercâmbio de resultados acadêmicos
e-MEC	Sistema eletrônico de gestão dos processos de regulação da educação superior
EMREX	Solução para transferência eletrônica de registros acadêmicos entre instituições
ENGD	Estratégia Nacional de Governo Digital
FAIR	Findable, Accessible, Interoperable and Reusable
FOAF	Friend of a Friend
HEI	Higher Education Institution
HTTP	Hypertext Transfer Protocol
IBGE	Instituto Brasileiro de Geografia e Estatística
ID	Identificador
ICP-Brasil	Infraestrutura de Chaves Públicas Brasileira
IES	Instituição de Ensino Superior
IF	Informação — código de campo de aplicação do INPI
INPI	Instituto Nacional da Propriedade Industrial
IRI	Internationalized Resource Identifier
LOCN	Core Location Vocabulary
LOV	Linked Open Vocabularies
MEC	Ministério da Educação
NSF	Ramo do leiaute relativo à diplomação fora do Sistema Federal de Ensino
OWL	Web Ontology Language
PAdES	PDF Advanced Electronic Signatures
PPGCI	Programa de Pós-Graduação em Ciência da Informação
PROGRAD	Pró-Reitoria de Graduação
PROV-O	PROV Ontology
PTD	Plano de Transformação Digital
QC	Questão de competência
QR Code	Quick Response Code
RDC-Arq	Repositório Arquivístico Digital Confiável

RDF	Resource Description Framework
RDFS	RDF Schema
RG	Registro Geral
RNP	Rede Nacional de Ensino e Pesquisa
RVDD	Representação Visual do Diploma Digital
SABiO	Systematic Approach for Building Ontologies
SIGAD	Sistema Informatizado de Gestão Arquivística de Documentos
SOC	Sistema de Organização do Conhecimento
SPARQL	SPARQL Protocol and RDF Query Language
SPO	Subject, Predicate and Object
STI	Superintendência de Tecnologia da Informação
TBox	Terminological Box
TOVE	Toronto Virtual Enterprise
TTL	Extensão dos arquivos em formato Turtle
UF	Unidade da Federação
UFES	Universidade Federal do Espírito Santo
UFMG	Universidade Federal de Minas Gerais
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium
XAdES	XML Advanced Electronic Signatures
XML	Extensible Markup Language
XMLDSig	XML Digital Signature
XSD	XML Schema Definition

Resumo

A regulamentação do Diploma Digital Brasileiro, estabelecida pelas Portarias do Ministério da Educação nº 330/2018 e nº 554/2019, definiu requisitos comuns para a emissão e o registro desses documentos por meio de estruturas padronizadas em Extensible Markup Language (XML) e XML Schema Definition (XSD). Embora favoreça a interoperabilidade sintática, essa padronização não explicita, por si só, os significados e as relações conceituais necessários à interpretação e ao reúso automatizado dos dados acadêmicos por diferentes sistemas. Esta dissertação teve como objetivo propor uma ontologia operacional para representar o subdomínio do Registro do Diploma e favorecer sua interoperabilidade semântica, com base nos padrões da Web Semântica, na abordagem dos Dados Ligados e nos princípios Findable, Accessible, Interoperable and Reusable (FAIR), correspondentes à encontrabilidade, à acessibilidade, à interoperabilidade e ao reúso. A pesquisa apresentou natureza aplicada, abordagem qualitativa e caráter exploratório-descritivo, com análise documental e aplicação adaptada da metodologia OntoForInfoScience. As normas, os documentos técnicos e os esquemas XSD do Ministério da Educação fundamentaram a aquisição e a conceitualização do conhecimento. Os fundamentos da Organização e Representação do Conhecimento orientaram a delimitação do domínio, a definição dos conceitos, o controle terminológico e a explicitação das relações semânticas. Como resultado, foi desenvolvida a OntoDiplomaDigital, formalizada por meio de Resource Description Framework (RDF), RDF Schema (RDFS), Web Ontology Language (OWL) e Simple Knowledge Organization System (SKOS), com serialização em formato textual Turtle e alinhamento controlado a vocabulários externos. A prova de conceito utilizou 252 unidades tabulares com dados fictícios, transformadas em um grafo de instâncias composto por 25.553 triplas. Consultas formuladas em SPARQL Protocol and RDF Query Language (SPARQL) permitiram recuperar informações sobre diplomados, cursos, instituições, registros, assinaturas, mecanismos de validação e proveniência. Consultas federadas também demonstraram, em ambiente controlado, a possibilidade de relacionar recursos locais a dados externos da DBpedia e do Wikidata. Os resultados evidenciaram a coerência entre a especificação, a conceitualização, a formalização e a recuperação das informações previstas, sem constituir validação institucional, integração com sistemas oficiais ou avaliação com dados acadêmicos reais. Conclui-se que a OntoDiplomaDigital complementou a estrutura sintática dos esquemas oficiais com uma camada conceitual capaz de explicitar significados, relações e valores controlados. A pesquisa demonstrou a contribuição da Ciência da Informação para organizar semanticamente os dados do Registro do Diploma Digital e recuperar suas relações no conjunto sintético. A representação proposta pode apoiar o acesso contextualizado às informações e tornar mais explícitos os vínculos acadêmicos e administrativos, oferecendo uma base para aplicações educacionais, profissionais, científicas e de gestão pública, cujos benefícios dependem de avaliação institucional.

Palavras-chave: Diploma Digital; Ontologia; Web semântica; Interoperabilidade; Princípios FAIR

ABSTRACT

The regulation of the Brazilian Digital Diploma, established by Ministry of Education Ordinances No. 330/2018 and No. 554/2019, defined common requirements for issuing and registering these documents through standardized structures in Extensible Markup Language (XML) and XML Schema Definition (XSD). Although this standardization supports syntactic interoperability, it does not, by itself, make explicit the meanings and conceptual relationships required for different systems to interpret and automatically reuse academic data. This dissertation aimed to propose an operational ontology to represent the Diploma Registration subdomain and support its semantic interoperability, based on Semantic Web standards, the Linked Data approach, and the Findable, Accessible, Interoperable and Reusable (FAIR) principles. The research was applied in nature, adopted a qualitative approach, and was exploratory and descriptive, combining document analysis with an adapted application of the OntoForInfoScience methodology. Ministry of Education regulations, technical documents, and XSD schemas provided the basis for knowledge acquisition and conceptualization. The foundations of Knowledge Organization and Representation guided domain delimitation, concept definition, terminological control, and the specification of semantic relationships. As a result, OntoDiplomaDigital was developed and formalized using the Resource Description Framework (RDF), RDF Schema (RDFS), Web Ontology Language (OWL), and Simple Knowledge Organization System (SKOS), serialized in the Turtle textual format, and selectively aligned with external vocabularies. The proof of concept used 252 tabular records containing fictitious data, which were transformed into an instance graph comprising 25,553 triples. SPARQL queries retrieved information about diploma holders, academic programs, institutions, registrations, signatures, validation mechanisms, and provenance. Federated queries also demonstrated, in a controlled environment, the possibility of linking local resources to external data from DBpedia and Wikidata. The results showed coherence among the specification, conceptualization, formalization, and retrieval of the intended information, without constituting institutional validation, integration with official systems, or evaluation using real academic data. It was concluded that OntoDiplomaDigital complemented the syntactic structure of the official schemas with a conceptual layer capable of making meanings, relationships, and controlled values explicit. The research demonstrated the contribution of Information Science to the semantic organization of Digital Diploma Registration data and the retrieval of their relationships within the synthetic dataset. The proposed representation may support contextualized access to information and make academic and administrative connections more explicit, providing a basis for educational, professional, scientific, and public administration applications whose benefits require institutional evaluation.

Keywords: Digital Diploma. Ontology. Semantic interoperability. Information Science. Linked Data, FAIR principles.

Sumário

1 Introdução 10
1.1 Problema de Pesquisa 12
1.2 Objetivo Geral 15
1.3 Objetivos Específicos 15
1.4 Justificativa 16
1.5 Organização do trabalho 17
2 Estado da arte: 19
2.1 Trabalhos científicos correlatos 19
2.2 Iniciativas internacionais de interoperabilidade de credenciais acadêmicas 26
3 Fundamentação Teórica 28
3.1 Transformação Digital e Governança de dados 29
3.2 Organização e Representação do Conhecimento 32
3.3 Metadados 35
3.4 Princípios FAIR 39
3.5 Web Semântica e Dados Ligados 44
3.6 Ontologia Operacional 54
3.7 Metodologia para Criação de Ontologias 57
3.8 Padrões e Esquemas de Metadados para Interoperabilidade 61
3.8.1 XML e XSD enquanto Padrões de Dados 65
3.9 Diplomas Digitais de IES 70
3.9.1 Arquitetura modular dos esquemas XSD do diploma 71
3.9.2 Requisitos tecnológicos da assinatura eletrônica 79
3.9.3 O Acervo Acadêmico e o Contexto Institucional das IES 79
4 Procedimentos Metodológicos 83
4.1 Aplicação da metodologia OntoForInfoScience na construção da OntoDiplomaDigital. 84
4.1.1 Decisão ontológica, delimitação do subdomínio e especificação dos requisitos - Etapa preliminar da OntoForInfoScience 85
4.1.2 Especificação da ontologia - Etapa 1 da OntoForInfoScience 86
4.1.3 Aquisição e extração do conhecimento - Etapa 2 da OntoForInfoScience 87
4.1.4 Conceitualização do subdomínio - Etapa 3 da OntoForInfoScience 90
4.1.5 Fundamentação ontológica - Etapa 4 da OntoForInfoScience 92
4.1.6 Formalização da OntoDiplomaDigital - Etapa 5 da OntoForInfoScience 92
4.1.7 Avaliação da ontologia - Etapa 6 da OntoForInfoScience 95
4.1.8 Documentação da ontologia - Etapa 7 da OntoForInfoScience 97
4.1.9 Disponibilização da ontologia - Etapa 8 da OntoForInfoScience 98
4.2 Procedimentos complementares de implementação e validação computacional 99
4.2.1 Preparação e normalização da fonte tabular 101
4.2.2 Construção do mapeamento RDF no Ontotext Refine 102
4.2.3 Configuração e carga do repositório no GraphDB 104
4.2.4 Geração da ABox e validação no GraphDB 107
4.2.5 Prova de conceito 107
5 Resultados e Discussão 109
5.1 Resultado da Especificação da Ontologia – Etapa 1 da OntoForInfoScience 109
5.1.1 Identificação e caracterização da ontologia 110
5.1.2 Domínio e escopo 111
5.1.3 Tipo de ontologia 113
5.1.4 Grupos de usuários 115
5.1.5 Cenários de aplicação 116
5.2 Resultado da aquisição e extração do conhecimento – Etapa 2 da OntoForInfoScience 119
5.2.1 Fontes normativas e técnicas 120
5.2.2 Elementos extraídos e artefatos produzidos 124
5.3 Resultado da Conceitualização - Etapa 3 da OntoForInfoScience 134
5.3.1 Dicionário de conceitos, Tabela de Conceitos e Propriedades 135
5.3.2 Dicionário de Verbos 143
5.3.3 Modelos conceituais gráficos 149
5.4 Resultado da fundamentação ontológica – Etapa 4 da OntoForInfoScience 162
5.5 Resultado da Formalização da ontologia – Etapa 5 da OntoForInfoScience 164

5.5.1 Arquitetura formal implementada 164
5.5.2 Correspondência entre os passos metodológicos e a formalização realizada 166
5.5.3 Núcleo conceitual efetivamente implementado 168
5.5.4 Termos das classes principais 169
5.5.5 Propriedades e relações formalizadas 171
5.5.6 Axiomas e restrições lógicas 173
5.5.7 Instâncias, conceitos controlados e separação da ABox 175
5.6 Resultado da avaliação da ontologia – Etapa 6 da OntoForInfoScience 176
5.6.1 Resultados da validação semântica interna da OntoDiplomaDigital 177
5.6.2 Resultados da verificação formal e operacional da ontologia 178
5.6.3 Resultados da prova de conceito e da verificação operacional 180
5.7 Resultado da documentação da ontologia – Etapa 7 da OntoForInfoScience 189
5.7.1 Justificativa das escolhas semânticas 190
5.7.2 Consistência terminológica 190
5.7.3 Memória e evolução do projeto 191
5.7.4 Comunicação acadêmica da ontologia 191
5.7.5 Artefatos documentais consolidados 192
5.8 Resultado da disponibilização da ontologia – Etapa 8 da OntoForInfoScience 193
5.9 Discussão integrada dos resultados à luz da Ciência da Informação e campos interdisciplinares 193
6 Considerações finais 200
7 Referências 205
Apêndice A – Namespaces e prefixos utilizados **212**
Apêndice B – Consultas SPARQL e resultados da prova de conceito **214**

Índice de figuras:

Figura 1 - Mapa conceitual para conceitualizar a problemática de pesquisa 13
Figura 2 - Ciclo de vida dos dados ligados 46
Figura 3 - Etapas da metodologia OntoForInfoScience 58
Figura 4 - Representação da arquitetura dos arquivos XSD/XML 73
Figura 5 - Fluxo dos procedimentos computacionais da prova de conceito da OntoDiplomaDigital 98
Figura 6 - Recorte do mapeamento RDF no Ontotext Refine 102
Figura 7 - Prévia RDF de um registro sintético no Ontotext Refine 103
Figura 8 - Configuração do repositório diplomaDB no GraphDB 104
Figura 9 - Interface de importação de arquivos RDF no GraphDB 105
Figura 10 - Arquitetura dos arquivos XSD/XML - MEC 149
Figura 11 - Taxonomia da ontologia 151
Figura 12 - Encadeamento das relações do domínio Registro do Diploma 153
Figura 13 - Validação dos tipos XSD/XML - ramo normal 155
Figura 14 - Validação dos tipos XSD/XML - ramo judicial 156
Figura 15 - Validação do XSD/XML - ramo NSF 158
Figura 16 - Modelo conceitual de encadeamento das relações do diploma digital 160
Figura 17 — Grafo de informações recuperadas para um diplomado na prova de conceito 185
Figura 18 - Ligação entre a base sintética da OntoDiplomaDigital e o Wikidata por código IBGE 188

Índice de quadros

Quadro 1 - Expressões utilizadas no levantamento bibliográfico 20
Quadro 2 - Síntese das referências selecionadas para o estado da arte 21
Quadro 3 - Principais autores e contribuições teóricas 28
Quadro 4 - Tipo de metadados e suas finalidades 37
Quadro 5 - Princípios FAIR 41
Quadro 6 - Conceitos para Objetos Digitais FAIR e o ecossistema 42
Quadro 7 - Implementação - cultura, tecnologia e habilidades para a Prática FAIR 42
Quadro 8 - Incorporação e sustentabilidade: incentivos, métricas e investimentos 43
Quadro 9 - Elementos fundamentais da representação RDF 47
Quadro 10 - Representação de indivíduos em um trecho ABox 48
Quadro 11 - Princípios dos dados ligados 49
Quadro 12 - Representação de consulta SPARQL 51
Quadro 13 - Correspondência entre tripla RDF e consulta SPARQL 51
Quadro 14 - Categorias de recursos para representação e interoperabilidade 64

Quadro 15 - Trecho de código XSD/XML 68
Quadro 16 - Trecho do schema DiplomaDigital_v1.05.xsd: declaração do elemento raiz <Diploma> 74
Quadro 17 - Trecho de código XML-XSD do Diploma Digital - <xs:annotation> 75
Quadro 18 - Trecho de código XML-XSD do Diploma Digital - <xs:sequence> 76
Quadro 19 - Trecho de código XML-XSD do Diploma Digital - "required" 76
Quadro 20 - Trecho de código XML-XSD do Diploma Digital - "colapse 77
Quadro 21 - Trecho de código XML-XSD do Diploma Digital - <xs:group> 78
Quadro 22 - Etapas de preparação da fonte tabular 101
Quadro 23 - Especificação da OntoDiplomaDigital 109
Quadro 24 - Domínio da ontologia 111
Quadro 25 - Classificação da Ontologia 112
Quadro 26 - Grupo de usuários e beneficiários da ontologia 115
Quadro 27 - Cenário de aplicação da ontologia 116
Quadro 28 - Questões de competência e requisitos informacionais da OntoDiplomaDigital 117
Quadro 29 - Fontes primárias da ontologia 119
Quadro 30 - Arquivos XSD do pacote MEC v1.05 e sua utilização na ontologia 121
Quadro 31 - Artefatos produzidos da Ontologia 123
Quadro 32 - Glossário de conceitos da Etapa 2 da OntoForInfoScience 124
Quadro 33 - Glossário de verbos da Etapa 2 da metodologia OntoforInfoscience 126
Quadro 34 - Glossário de relações 128
Quadro 35- Propriedades de dados identificadas na Etapa 2 da metodologia OntoForInfoScience 131
Quadro 36 - Valores controlados identificados na Etapa 2 da metodologia OntoForInfoScience 132
Quadro 37 - Mecanismos de rastreabilidade declarados no TTL 133
Quadro 38 - Definição e origem dos conceitos do subdomínio do Registro do Diploma 135
Quadro 39 - Propriedades, atributos e regras de aplicação dos conceitos 138
Quadro 40 - Restrições e decisões de modelagem das classes da ontologia 141
Quadro 41 – Dicionário de verbos 143
Quadro 42 - Relações semânticas complementares 146
Quadro 43 - Arquitetura formal implementada na ontologia 164
Quadro 44 - Correspondência entre os passos da Etapa 5 e a formalização da OntoDiplomaDigital 166
Quadro 45 - Recorte do núcleo conceitual implementado na OntoDiplomaDigital 167
Quadro 46 - Grupos funcionais de propriedades formalizadas 171
Quadro 47 - Exemplos de restrições formalizadas e sua leitura em linguagem natural 173
Quadro 48 - Síntese dos resultados da avaliação, validação e verificação da OntoDiplomaDigital 175
Quadro 49 – Resultados da validação semântica interna da OntoDiplomaDigital 176
Quadro 50 – Resultados da verificação formal da Ontologia 177
Quadro 51 - Questões de competência e evidências da verificação operacional da OntoDiplomaDigital 181
Quadro 52 - Exemplo abreviado de consulta SPARQL para a QC1 183
Quadro 53 - Dados retornados pela consulta QC1 para o diploma sintético selecionado 184
Quadro 54 - Consulta federada para relacionar a naturalidade dos diplomados fictícios ao município no Wikidata 186
Quadro 55 — Dados retornados pela consulta federada ao Wikidata 187
Quadro 56 - Artefatos documentais consolidados 191

1 Introdução

A transformação digital do ensino superior brasileiro modificou os processos de produção, gestão, circulação e preservação dos documentos acadêmicos. Esse movimento integrou um processo mais amplo de digitalização da administração pública, orientado para a eficiência administrativa, a desburocratização, a inovação e a ampliação da oferta de serviços públicos digitais (Brasil, 2021). Entre as ações associadas a esse processo, destacou-se a consolidação do Portal Gov.br como plataforma unificada de acesso a informações e serviços públicos digitais, nos termos do Decreto nº 9.756/2019 (Brasil, 2019a). No âmbito acadêmico, uma das principais manifestações dessa transformação foi a instituição do Diploma Digital pelo Ministério da Educação, inicialmente por meio da Portaria MEC nº 330/2018 e, posteriormente, da Portaria MEC nº 554/2019 (Brasil, 2018; Brasil, 2019c).

A regulamentação estabeleceu requisitos comuns para a emissão e o registro do Diploma Digital, apoiados em documentos estruturados em *Extensible Markup Language* (XML) e XML Schema Definition (XSD). Esses esquemas definem elementos, tipos de dados, cardinalidades e regras aplicáveis às informações sobre instituições, cursos, diplomados e processos de emissão e registro. A versão 1.05, adotada como referência técnica nesta pesquisa, integrou o conjunto de atualizações disponibilizadas pelo MEC e forneceu a estrutura empregada na validação sintática dos documentos produzidos pelas Instituições de Ensino Superior (Brasil, 2020b; Brasil, 2022; W3C, 2012). A regulamentação também estabeleceu requisitos relacionados à integridade, à assinatura digital e ao carimbo de tempo (Brasil, 2025).

A adoção de estruturas comuns em XML/XSD favorece a interoperabilidade sintática, pois permite que diferentes sistemas reconheçam e processem documentos organizados segundo regras previamente estabelecidas. Essa padronização constitui uma condição importante para o intercâmbio automatizado de informações entre Instituições de Ensino Superior, plataformas governamentais e outros sistemas envolvidos na emissão e no registro do Diploma Digital. A compatibilidade estrutural, entretanto, não assegura que os dados intercambiados sejam interpretados com o mesmo significado nos diferentes ambientes informacionais. Enquanto a interoperabilidade sintática está relacionada à compatibilidade dos formatos e das estruturas, a interoperabilidade semântica requer a preservação dos significados associados aos elementos e às relações representadas (Riley, 2017; Laufer, 2015).

No campo da Ciência da Informação, esse problema aproxima-se dos fundamentos da Organização e Representação do Conhecimento, que orientam a identificação, a organização e a explicitação de conceitos, categorias e relações conforme os contextos e as finalidades dos sistemas de informação (Hjørland, 2016). Os metadados participam desse processo ao descrever características dos recursos informacionais por meio de propriedades e valores (Qin; Zeng, 2022). Aplicados ao Diploma Digital, esses fundamentos permitem compreender que o intercâmbio dos dados depende não apenas da transmissão de seus valores, mas também da representação dos recursos aos quais se referem e das relações estabelecidas entre diplomados, cursos, instituições, registros e mecanismos de validação.

Para essa finalidade, a Organização do Conhecimento dispõe de diferentes instrumentos destinados à estruturação conceitual e ao estabelecimento de relações entre os elementos de um domínio, com distintos níveis de formalização (Mazzocchi, 2018). Entre esses instrumentos, destacam-se as ontologias, cuja formalização permite utilizar a representação conceitual em aplicações computacionais.

As ontologias formalizam conceitos, propriedades, relações e restrições de determinado domínio. Entre as possibilidades existentes, as ontologias operacionais correspondem a representações implementadas e legíveis por máquinas, destinadas ao uso em sistemas e aplicações computacionais (Falbo, 2014). Essa característica mostrou-se pertinente ao problema investigado, que exigiu não apenas a conceitualização do subdomínio do Registro do Diploma, mas também uma estrutura capaz de ser materializada com dados e utilizada em consultas.

A operacionalização dessa representação encontrou suporte nos padrões da Web Semântica e na abordagem dos Dados Ligados. Essa operacionalização emprega Resource Description Framework (RDF) para representar dados em triplas e grafos, Web Ontology Language (OWL) para formalizar a ontologia e SPARQL Protocol and RDF Query Language (SPARQL) para consultar os dados representados em RDF. Em conjunto, essas tecnologias oferecem meios para identificar, representar, relacionar e consultar recursos em estruturas baseadas em grafos (Berners-Lee; Hendler; Lassila, 2001; W3C, 2013). Os Dados Ligados ampliam essa possibilidade ao favorecer o uso de identificadores globais e o estabelecimento de relações entre recursos mantidos em fontes distintas (Berners-Lee, 2006; Bizer; Heath; Berners-Lee, 2009). Os princípios FAIR complementam essa abordagem ao orientar a encontrabilidade, a acessibilidade, a interoperabilidade e o potencial de

reúso dos dados e metadados por pessoas e sistemas computacionais, sem pressupor que todas as informações sejam abertas ou irrestritamente acessíveis (Wilkinson et al., 2016).

A articulação entre a estrutura sintática estabelecida para o Diploma Digital e a necessidade de explicitar os conceitos e as relações do domínio delimitou o problema desta pesquisa. Embora os schemas XML/XSD forneçam uma base comum para a produção e a validação dos documentos, permanece a necessidade de uma representação semântica formal do subdomínio do Registro do Diploma que favoreça a interpretação, a recuperação e a integração dos dados. Essa questão é delimitada na seção seguinte.

1.1 Problema de Pesquisa

A instituição do Diploma Digital modificou os processos acadêmicos e documentais das Instituições de Ensino Superior brasileiras. Sua emissão, seu armazenamento e sua preservação em meio digital passaram a exigir mecanismos capazes de assegurar validade jurídica, autenticidade, integridade, rastreabilidade, privacidade e disponibilidade das informações (Sousa, 2023). Essa transformação não se restringiu à substituição do suporte físico, pois também ampliou a necessidade de intercâmbio de dados entre sistemas acadêmicos, plataformas governamentais e outros ambientes informacionais.

A regulamentação do Ministério da Educação estabeleceu diretrizes técnicas para a geração e a validação do Diploma Digital e promoveu a padronização estrutural dos documentos intercambiados (Brasil, 2019c; Brasil, 2020b). Conforme discutido na abertura deste capítulo, essa conformidade sintática não assegura, isoladamente, que os dados sejam interpretados de maneira uniforme pelos diferentes sistemas. A interoperabilidade semântica requer representações formais e compartilhadas que explicitem as conceitualizações do domínio (Guizzardi, 2020). No campo da Ciência da Informação, essa questão também possui dimensão terminológica, pois conceitos empregados sob diferentes perspectivas e contextos podem produzir ambiguidades e divergências de interpretação (Arakaki; Arakaki, 2020).

A dimensão prática do problema foi observada durante a experiência do pesquisador como Superintendente de Tecnologia da Informação da Universidade Federal do Espírito Santo, entre abril de 2020 e março de 2024, período no qual

ocorreu a implantação institucional do Diploma Digital. A adequação aos requisitos do Ministério da Educação exigiu a revisão de fluxos informacionais, a reorganização de cadastros acadêmicos e a realização de adaptações nos sistemas institucionais. Essa experiência foi utilizada como elemento de contextualização e apoio à identificação de pontos de atenção, posteriormente confrontados com as normas, os schemas oficiais e as demais fontes empregadas na pesquisa.

A necessidade de preservar o significado dos dados torna-se particularmente relevante quando as informações acadêmicas precisam ser relacionadas a recursos mantidos em diferentes ambientes. Consultas que envolvem instituições, cursos, registros, localização geográfica, natureza das organizações ou outros dados distribuídos dependem não apenas da existência e da compatibilidade estrutural desses dados, mas também da interpretação consistente das entidades e das relações representadas.

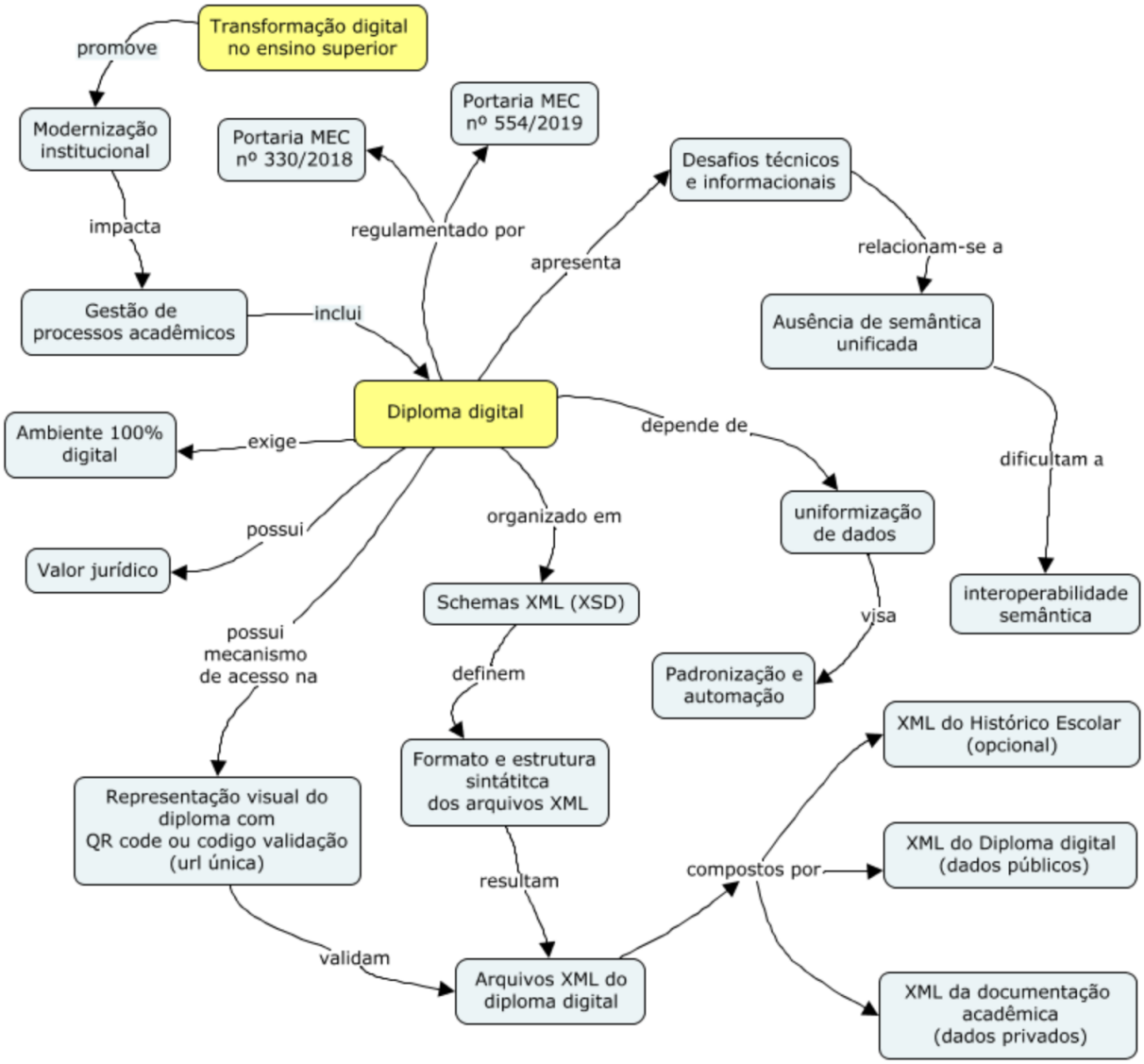
O Diploma Digital possui validade jurídica e utiliza documentos estruturados, mas essas características não o tornam automaticamente semanticamente interoperável. Termos como aluno, estudante, egresso e diplomado, por exemplo, podem ser considerados equivalentes em determinados contextos e representar categorias distintas em outros. Sem uma representação que explicite essas diferenças, a recuperação, a integração e o reúso das informações podem ser comprometidos.

O problema investigado consistiu, portanto, na ausência de uma representação ontológica operacional desenvolvida especificamente a partir do contexto normativo brasileiro e das estruturas que organizam os dados do subdomínio do Registro do Diploma Digital. Essa lacuna limita a possibilidade de representar e recuperar os dados desse subdomínio de forma semanticamente consistente em ambientes informacionais distribuídos. Diante desse problema, formulou-se a seguinte questão de pesquisa:

Como organizar os dados do subdomínio do Registro do Diploma Digital das Instituições de Ensino Superior brasileiras de modo a favorecer sua interoperabilidade semântica, acesso e recuperação?

Os principais componentes do problema de pesquisa foram sintetizados no mapa conceitual apresentado na Figura 1.

Figura 1 - Mapa conceitual para conceitualizar a problemática de pesquisa



Fonte: Elaborado pelo autor (2026).

O mapa conceitual relaciona o processo de transformação digital do ensino superior, a estrutura documental estabelecida para o Diploma Digital e as limitações decorrentes da ausência de uma representação conceitual compartilhada. O mapa conceitual também situa a Organização e Representação do Conhecimento, as ontologias e os padrões da Web Semântica como fundamentos empregados na pesquisa para tratar o problema identificado. Dessa forma, o mapa conecta o contexto normativo e institucional à solução teórico-metodológica investigada.

A questão formulada exigiu não apenas a conceitualização do domínio, mas uma representação formal que pudesse ser materializada com dados, processada computacionalmente e consultada. Essa exigência fundamentou a adoção de uma

ontologia operacional e orientou a definição dos objetivos apresentados nas seções seguintes.

1.2 Objetivo Geral

O objetivo geral desta pesquisa é propor uma ontologia operacional para organizar e representar o conhecimento do subdomínio do Registro do Diploma Digital Brasileiro, fundamentada nos princípios da Organização e Representação do Conhecimento e implementada por meio de padrões da Web Semântica, de modo a favorecer a interoperabilidade semântica e a recuperação da informação.

1.3 Objetivos Específicos

I. Analisar os padrões normativos e técnicos estabelecidos pelo Ministério da Educação (MEC) para a implementação do diploma digital no Brasil, com foco nos elementos que compõem o subdomínio do Registro do Diploma;
II. Identificar ontologias, vocabulários controlados e esquemas de metadados relevantes ao contexto da representação interoperável dos dados de certificação acadêmica do diploma digital brasileiro;
III. Construir um modelo conceitual do subdomínio do Registro do Diploma, consolidando as entidades, relações e restrições essenciais à certificação acadêmica digital;
IV. Estabelecer correspondências entre os elementos dos XML/XSD, o modelo conceitual do subdomínio do Registro do Diploma e os termos selecionados para a ontologia operacional;
V. Implementar uma ontologia operacional baseada em padrões da Web Semântica compatíveis com os princípios da interoperabilidade semântica e dos dados ligados;
VI. Realizar uma prova de conceito para verificar a viabilidade de representação, transformação e recuperação semântica de dados sintéticos do Registro do Diploma e demonstrar, no conjunto de teste, uma integração federada baseada em identificador compartilhado.

A prova de conceito foi adotada como estratégia para verificar, em ambiente controlado e com dados sintéticos, a possibilidade de materialização e consulta da

representação proposta. Sua finalidade não foi validar a ontologia em contexto institucional ou avaliar seu desempenho em escala, mas demonstrar sua utilização operacional no recorte definido pela pesquisa.

1.4 Justificativa

A crescente digitalização dos processos acadêmicos brasileiros, impulsionada pela obrigatoriedade do Diploma Digital, evidencia a necessidade de mecanismos capazes de favorecer a interoperabilidade semântica entre os sistemas utilizados pelas Instituições de Ensino Superior (IES). Embora a regulamentação do Ministério da Educação tenha estabelecido padrões estruturais para a emissão e o registro desses documentos, permanecem desafios relacionados à representação uniforme dos dados, à integração entre diferentes sistemas e ao reúso das informações em ambientes digitais distribuídos. A compatibilidade estrutural entre sistemas não assegura, por si só, a preservação do significado dos dados intercambiados, uma vez que a interoperabilidade semântica requer que diferentes estruturas informacionais estabeleçam relações preservando o significado dos elementos representados (Guizzardi, 2020). As ontologias oferecem mecanismos formais para explicitar conceitos e relações de um domínio e estabelecer conceitualizações compartilhadas, contribuindo para reduzir ambiguidades e favorecer a interoperabilidade semântica (Guizzardi, 2020).

No campo da Ciência da Informação, essa perspectiva aproxima-se dos fundamentos da Organização e Representação do Conhecimento, que tratam da organização dos conceitos e das relações que estruturam determinado domínio do conhecimento (Capurro, 2003; Hjørland, 2016). A articulação entre esses fundamentos e as tecnologias da Web Semântica sustenta a escolha de uma ontologia operacional, pois a pesquisa requer não apenas explicitar a conceitualização do subdomínio do Registro do Diploma, mas também materializá-la com dados, processá-la computacionalmente e consultá-la.

Os princípios FAIR foram adotados para orientar a identificação, a descrição, a interoperabilidade e o reúso dos dados e dos artefatos ontológicos, sem pressupor a disponibilização irrestrita de informações acadêmicas sujeitas a condições legais e institucionais de acesso (Wilkinson et al., 2016; Sayão et al., 2021). A abordagem de Dados Ligados foi incorporada porque o problema envolve recursos mantidos em

ambientes distribuídos e requer identificadores e relações explícitas que permitam conectá-los e recuperá-los por meio dos padrões da Web Semântica. A integração federada, por sua vez, foi incluída na prova de conceito para verificar se um identificador compartilhado permitiria relacionar dados representados localmente a informações mantidas em uma fonte externa, sem exigir sua incorporação permanente ao repositório local.

A relevância da pesquisa decorre da convergência entre as dimensões científica, tecnológica e institucional. No âmbito científico, aproxima os fundamentos da Organização e Representação do Conhecimento das tecnologias da Web Semântica aplicadas ao domínio educacional e busca responder à lacuna identificada no Estado da Arte quanto à representação semântica do Diploma Digital por meio de uma ontologia operacional. No âmbito tecnológico, propõe uma estrutura formal e processável, baseada em padrões da Web Semântica e em Dados Ligados, para representar o subdomínio do Registro do Diploma e verificar possibilidades de recuperação e integração semântica. No plano institucional, a proposta pode oferecer subsídios para a organização e o reúso dos dados acadêmicos pelas Instituições de Ensino Superior e criar condições para sua futura integração com outros ecossistemas informacionais.

1.5 Organização do trabalho

Esta dissertação está organizada em seis capítulos, seguidos pelas referências e pelos apêndices. O Capítulo 1 contextualiza o Diploma Digital Brasileiro e apresenta o problema e a questão de pesquisa, os objetivos, a justificativa e a organização do trabalho. Nesse percurso, a questão de pesquisa delimita o problema investigado, orienta a definição dos objetivos e direciona o desenvolvimento da investigação.

O Capítulo 2 examina a produção científica relacionada às credenciais acadêmicas digitais, identifica as abordagens desenvolvidas e delimita a lacuna na qual esta pesquisa se insere. O Capítulo 3 reúne os referenciais conceituais necessários à compreensão do estudo, abrangendo a Organização e Representação do Conhecimento, os metadados, os princípios FAIR, a Web Semântica, os Dados Ligados, as ontologias, a interoperabilidade, os padrões XML e XSD e o Diploma Digital Brasileiro.

O Capítulo 4 apresenta o delineamento metodológico da pesquisa, a aplicação da metodologia OntoForInfoScience e os procedimentos complementares empregados na construção e na implementação da ontologia e no desenvolvimento da prova de conceito. O Capítulo 5 apresenta e discute os resultados da pesquisa, compreendendo os artefatos produzidos durante a construção da ontologia, sua formalização, avaliação, validação, documentação e disponibilização, bem como a materialização da prova de conceito e os resultados obtidos por meio das consultas SPARQL. A discussão relaciona esses resultados à Fundamentação Teórica e ao Estado da Arte, examinando a interoperabilidade semântica, a aplicação dos princípios FAIR, o atendimento aos objetivos da pesquisa, suas contribuições e suas limitações.

O Capítulo 6 apresenta as Considerações Finais, nas quais são sintetizados o percurso investigativo e os principais resultados, respondida a questão de pesquisa e retomado o atendimento aos objetivos propostos. O capítulo também sistematiza as contribuições e as limitações do estudo e indica possibilidades para trabalhos futuros. Por fim, as referências registram as fontes utilizadas, enquanto os apêndices reúnem o material complementar necessário à documentação, à rastreabilidade e à compreensão da pesquisa.

2 Estado da arte:

O Estado da Arte foi elaborado com a finalidade de mapear e analisar a produção científica relacionada às credenciais acadêmicas digitais e situar o problema investigado diante das abordagens já desenvolvidas. A análise concentrou-se em aspectos relacionados à circulação de registros acadêmicos, à interoperabilidade entre sistemas, à representação semântica, à portabilidade dos dados estudantis, à verificação de qualificações e à implementação institucional de documentos acadêmicos digitais.

O capítulo está organizado em duas seções complementares. A primeira apresenta os procedimentos de busca e seleção dos trabalhos científicos correlatos, bem como a análise de suas contribuições e limitações em relação ao problema investigado. A segunda examina iniciativas internacionais relacionadas ao intercâmbio e à representação de credenciais acadêmicas, consideradas como referências complementares para a compreensão do domínio.

2.1 Trabalhos científicos correlatos

Para a construção do Estado da Arte, realizou-se uma busca bibliográfica específica destinada a identificar estudos relacionados às credenciais acadêmicas digitais e a examinar suas contribuições para temas associados à interoperabilidade entre sistemas, à representação semântica, à integração de infraestruturas educacionais, à portabilidade dos dados acadêmicos, à verificação de qualificações e à implementação institucional do Diploma Digital.

O levantamento foi realizado com apoio do software Publish or Perish, ferramenta voltada à recuperação e à análise de citações acadêmicas em diferentes fontes de informação científica (Harzing, 2016). As consultas foram efetuadas entre 2 e 29 de março de 2026 no Google Scholar, no Portal de Periódicos CAPES, na Scopus, na OpenAlex e na Web of Science. Tratou-se de um levantamento focal, orientado ao problema específico da representação ontológica e da interoperabilidade do Diploma Digital. A estratégia de busca utilizou descritores em português e inglês relacionados a ontologia, interoperabilidade e diploma digital, conforme apresentado no Quadro 1.

Quadro 1 - Expressões utilizadas no levantamento bibliográfico

Idioma	Expressão de busca
Inglês	ontology AND interoperability AND digital diploma
Português	ontologia AND interoperabilidade AND diploma digital

Fonte: Elaborado pelo autor (2026).

As expressões em português e inglês integraram uma estratégia bilíngue de recuperação. Por esse motivo, os resultados foram contabilizados por fonte de informação, e não separadamente pelo idioma da expressão. As consultas recuperaram 335 registros: 247 no Google Scholar, 86 no Portal de Periódicos CAPES e dois na Scopus. A OpenAlex e a Web of Science não apresentaram resultados que atendessem à estratégia adotada.

Originalmente, o período de publicação abrangia cinco anos anteriores; contudo, com base na análise dos resultados, optou-se por adotar um intervalo de dez anos anteriores à conclusão do levantamento, correspondendo ao período de 2016 a 2026. Esse recorte foi aplicado durante a triagem dos registros, e não como restrição uniforme nas interfaces das fontes consultadas. A definição do período levou em conta a natureza recente das transformações tecnológicas e institucionais associadas às credenciais acadêmicas digitais, visando captar a produção contemporânea ao desenvolvimento das infraestruturas, dos padrões e das iniciativas analisados na pesquisa.

Os registros recuperados foram organizados no Zotero, ferramenta que permite a classificação e o agrupamento de itens em coleções, bem como a atribuição de palavras-chave, e submetidos à identificação de duplicatas. A comparação foi efetuada com base no *Digital Object Identifier* (DOI), no título normalizado e no ano de publicação, com conferência manual dos registros que apresentavam variações na descrição bibliográfica. Nove duplicatas foram excluídas, resultando em 326 registros para triagem.

A análise dos títulos e resumos identificou a pertinência dos trabalhos em relação a diplomas ou credenciais acadêmicas digitais e sua contribuição para pelo menos uma das dimensões investigadas: interoperabilidade entre sistemas, representação semântica, integração de infraestruturas educacionais, portabilidade de dados acadêmicos, verificação de qualificações ou implementação institucional do Diploma Digital. Nessa fase, 296 registros foram excluídos por não atenderem adequadamente ao objeto, à disponibilidade ou à problemática da pesquisa, restando 30 trabalhos para avaliação de elegibilidade por meio da leitura integral.

A leitura integral resultou na exclusão de 25 trabalhos. Foram excluídos os estudos situados fora do recorte temporal, aqueles que utilizavam a expressão “diploma digital” sem tratar do documento ou da credencial acadêmica digital, os trabalhos concentrados exclusivamente em aspectos de segurança ou autenticação sem contribuição para as dimensões analisadas e os estudos cuja abordagem se mostrava redundante em relação a outro trabalho mais diretamente relacionado ao problema investigado.

O corpus final foi composto por cinco estudos: Halloluwa e Silva (2025), El Chawra (2020), Kontzinos et al. (2020), Scheid (2022) e Silva e Zárate (2022). Quatro trabalhos foram publicados em inglês e um em português. A seleção reuniu contribuições relacionadas ao emprego de grafos de conhecimento em ambientes educacionais, à portabilidade de dados estudantis, à interoperabilidade entre infraestruturas heterogêneas, à representação semântica de qualificações e à implementação institucional do Diploma Digital Brasileiro. Esses estudos são apresentados e analisados no capítulo dedicado ao Estado da Arte. O Quadro 2 apresenta os trabalhos selecionados, seus respectivos títulos, anos e fontes de publicação.

Quadro 2 - Síntese das referências selecionadas para o estado da arte

Referência do levantamento bibliográfico
HALLOLUWA, Thilina; DE SILVA, Chameera. Blockchain in education technology: “a tool for challenges in education technology”: a review paper. Preprints, [S. l.], 3 mar. 2025. Preprint. DOI: https://doi.org/10.20944/preprints202503.0130.v1. Disponível em: https://www.preprints.org/manuscript/202503.0130/v1.
EL CHAWRA, T. Nevine. Implementing digital student data portability: experiences and lessons learned. 2020. Dissertação (Master of Arts) — University of Toronto, Toronto, 2020.
KONTZINOS, Christos; MARKAKI, Ourania; KOKKINAKOS, Panagiotis; KARAKOLIS, Vagelis; KAPSALIS, Panagiotis; PSARRAS, John. Exploring blockchain, semantics and decision support to optimise qualification certification, recruitment and competency management: an assessment of challenges, current practices and opportunities. International Journal on Advances in Intelligent Systems, [S. l.], v. 13, n. 3-4, p. 192-203, 2020. ISSN 1942-2679. Disponível em: https://personales.upv.es/thinkmind/dl/journals/intsys/intsys_v13_n34_2020/intsys_v13_n34_2020_3.pdf.
SCHEID, Eder John. An intent-based blockchain-agnostic interaction environment. 2022. 229 p. Tese (Doutorado) — University of Zurich, Zurich, 2022. DOI: https://doi.org/10.5167/uzh-220698. Disponível em: https://www.zora.uzh.ch/id/eprint/220698/.
SILVA, Patrícia Nascimento; ZÁRATE, Daniele Claudia Matta Fagundes. Desafios na implementação do Diploma Digital: um relato de experiência na UFMG. In: WORKSHOP DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO DAS INSTITUIÇÕES FEDERAIS DE ENSINO SUPERIOR DO BRASIL (WTICIFES), 14., 2022, Aracaju. Anais [...]. Aracaju: [s. n.], 2022. p. 48-52. Disponível em: https://hdl.handle.net/1843/61842.

Fonte: Elaborado pelo autor (2026)

Os estudos selecionados abordam a digitalização das credenciais e dos registros acadêmicos sob perspectivas distintas e complementares. Entre os aspectos identificados encontram-se a segurança e a confiança nos registros digitais, a portabilidade de dados estudantis, a integração entre infraestruturas tecnológicas, a verificação de qualificações, a utilização de recursos semânticos e as dificuldades institucionais decorrentes da implementação de documentos acadêmicos digitais. A tecnologia *blockchain* aparece de maneira recorrente em parte desses trabalhos como infraestrutura relacionada à integridade, ao registro e à verificação das informações, embora com finalidades e níveis de implementação distintos.

Halloluwa e Silva (2025) apresentam uma revisão sobre a utilização de *blockchain* na tecnologia educacional e propõem uma arquitetura de sistema de gestão da aprendizagem baseada na combinação entre *blockchain* e grafo de conhecimento. Na estrutura proposta, os estudantes constituem os nós centrais do grafo, aos quais são relacionados dados produzidos durante sua trajetória educacional, incluindo aulas, atividades, avaliações, resultados e certificados. A proposta contempla ainda a participação de diferentes atores do ecossistema educacional e a utilização da *blockchain* como componente da infraestrutura de registro e compartilhamento das informações.

A proposta de Halloluwa e Silva (2025) associa dados educacionais a uma estrutura baseada em grafo, possibilitando representar relações entre informações acadêmicas para além de seu armazenamento isolado. A arquitetura, entretanto, permanece em nível propositivo, e os próprios autores indicam possibilidades de desenvolvimento e aperfeiçoamento do grafo em trabalhos futuros. Além disso, o estudo não se destina à representação formal de um modelo normativo específico de credenciais acadêmicas.

El Chawra (2020) aborda o problema sob a perspectiva da *Digital Student Data Portability* (DSDP), investigando experiências de organizações que implementaram iniciativas destinadas à portabilidade de dados estudantis. A autora identifica a adoção de plataformas capazes de representar realizações acadêmicas por meio de registros portáveis, interoperáveis, comparáveis, autênticos, transparentes, compartilháveis e verificáveis. Entre os benefícios associados a essas iniciativas encontram-se o acesso, o compartilhamento, a empregabilidade, a prevenção de fraude, a privacidade e a portabilidade. A pesquisa também identifica valores, prioridades e o papel

desempenhado pelas organizações como fatores capazes de influenciar a adoção dessas iniciativas (El Chawra, 2020).

Essa perspectiva demonstra que a circulação de dados estudantis não se restringe à capacidade de transferir eletronicamente registros entre instituições. A portabilidade envolve também condições que permitam comparar, verificar e utilizar os dados recebidos. O estudo evidencia, dessa forma, dimensões técnicas e institucionais associadas à circulação de informações acadêmicas, sem propor uma ontologia para a representação dos registros estudantis.

Scheid (2022), por sua vez, concentra-se no problema da interoperabilidade entre plataformas *blockchain* heterogêneas. O autor identifica a diversidade de implementações, interfaces e modelos de dados, associada à escassez de padrões, como fator que dificulta a integração entre essas plataformas e favorece a formação de soluções não interoperáveis. A ausência de interoperabilidade pode resultar em silos de informação, fragmentação, redundância e inconsistência dos dados (Scheid, 2022).

Como resposta ao problema, Scheid (2022) desenvolve e avalia um ambiente de interação independente da plataforma *blockchain*, incluindo mecanismos destinados a abstrair detalhes técnicos das diferentes tecnologias e permitir a interação com múltiplas plataformas. A tese não se limita, portanto, à discussão conceitual da interoperabilidade, mas desenvolve e avalia componentes destinados a viabilizá-la. Embora não tenha como objeto específico as credenciais acadêmicas ou o Diploma Digital Brasileiro, o estudo evidencia problemas de interoperabilidade decorrentes da heterogeneidade das infraestruturas tecnológicas.

Kontzinos et al. (2020) aproximam-se mais diretamente da dimensão semântica ao apresentarem o projeto QualiChain, que articula *blockchain*, semântica, análise de dados e mecanismos de apoio à decisão para apoiar o armazenamento, o compartilhamento e a verificação de qualificações acadêmicas e profissionais. Os serviços básicos da plataforma são fundamentados em *blockchain* e semântica e abrangem o arquivamento, o armazenamento e a verificação de qualificações, a gestão de portfólios e a verificação de equivalência entre certificados.

A dimensão semântica ocupa papel relevante na arquitetura proposta. Kontzinos et al. (2020) associam o emprego de tecnologias semânticas à interoperabilidade dos dados e à possibilidade de estabelecer relações legíveis por máquina entre informações provenientes de fontes heterogêneas. O projeto considera

ainda a utilização e o reúso de ontologias e vocabulários para representar elementos relacionados a qualificações, competências, recrutamento, recursos humanos, *blockchain* e mecanismos de verificação.

O QualiChain demonstra, portanto, que mecanismos de representação semântica já são aplicados ao domínio das qualificações acadêmicas e profissionais, inclusive em situações que envolvem equivalência entre certificados e integração de informações oriundas de diferentes contextos. Seu escopo, entretanto, é amplo e contempla, além da certificação e da verificação de qualificações, processos relacionados ao recrutamento, à gestão de competências e ao apoio à decisão (Kontzinos et al., 2020).

Silva e Zárate (2022) aproximam a análise diretamente do contexto brasileiro ao relatarem a experiência de implementação do Diploma Digital na Universidade Federal de Minas Gerais (UFMG). O estudo descreve alterações realizadas no sistema acadêmico institucional, a integração com a solução desenvolvida pela Rede Nacional de Ensino e Pesquisa (RNP) e as adequações necessárias para atender às normas estabelecidas pelo Ministério da Educação.

Durante a implementação, foram identificadas situações nas quais os dados existentes no sistema acadêmico não correspondiam diretamente às estruturas e às regras exigidas pela solução. Entre os exemplos relatados encontram-se atividades sem docente cadastrado, formas de avaliação incompatíveis com campos obrigatórios, identificações repetidas de atividades, disciplinas de conteúdo variável e ocorrências acadêmicas para as quais não havia campos específicos. Algumas dessas questões demandaram esclarecimentos junto ao MEC e adaptações no sistema institucional (Silva; Zárate, 2022).

Silva e Zárate destacam ainda que, embora os desafios técnicos sejam relevantes, a definição dos conceitos e a adequação dos requisitos diante da diversidade dos currículos e dos formatos de registro constituem dificuldades significativas para as Instituições de Ensino Superior. O estudo registra também problemas relacionados às alterações sucessivas dos arquivos XSD e às indefinições existentes durante o processo de implantação.

A experiência relatada pela UFMG (Silva; Zárate, 2022) evidencia que a padronização estrutural em XML/XSD, embora necessária à implementação e à validação dos documentos, convive com dificuldades relacionadas à interpretação e à adequação dos conceitos acadêmicos às estruturas definidas pelo modelo. O estudo

fornece, portanto, evidências de questões conceituais e institucionais encontradas na implementação do Diploma Digital Brasileiro, embora não proponha uma formalização ontológica para tratá-las.

A comparação entre os estudos evidenciou contribuições distintas em relação ao problema investigado. Halloluwa e Silva (2025) associaram blockchain e grafo de conhecimento para organizar informações educacionais, sem se concentrarem em um modelo normativo específico de certificação acadêmica. El Chawra (2020) examinou condições técnicas e institucionais para a portabilidade de dados estudantis, sem propor uma ontologia para representar os registros transferidos. Scheid (2022) tratou da interoperabilidade entre plataformas blockchain heterogêneas, com foco na integração de infraestruturas e não nas credenciais acadêmicas. Kontzinos et al. (2020) incorporaram recursos semânticos à gestão e à verificação de qualificações, mas em um domínio amplo, que também abrange competências, recrutamento e apoio à decisão. Silva e Zárate (2022), por sua vez, evidenciaram desafios da implementação do Diploma Digital Brasileiro, sem formalizar uma ontologia destinada à representação semântica de seu subdomínio de registro.

Nesse conjunto, a OntoDiplomaDigital diferenciou-se por propor uma ontologia operacional delimitada ao Registro do Diploma, construída a partir do contexto normativo brasileiro e das estruturas modulares dos schemas XSD do Ministério da Educação. A proposta não substituiu soluções de portabilidade, blockchain ou infraestrutura institucional examinadas nos estudos correlatos; acrescentou uma camada de representação semântica destinada a explicitar os conceitos, as relações, os atributos, as restrições e os valores controlados necessários à organização e à recuperação dos dados do domínio. Assim, a lacuna identificada não correspondeu à ausência geral de ontologias ou tecnologias semânticas para credenciais acadêmicas, mas à inexistência, no corpus analisado, de uma modelagem operacional voltada especificamente ao Registro do Diploma Digital Brasileiro.

Essa delimitação orientou a construção da OntoDiplomaDigital como uma modelagem formal baseada nos padrões da Web Semântica, destinada a representar, interligar e favorecer o reúso dos dados do subdomínio do Registro do Diploma Brasileiro, à luz dos princípios FAIR e da abordagem dos Dados Ligados. A representação desses dados não pressupõe sua abertura irrestrita, devendo ser distinguidas as possibilidades de organização e recuperação das condições de acesso e utilização.

A compreensão dessa lacuna foi complementada pela análise de referências internacionais relacionadas ao intercâmbio e à representação de credenciais acadêmicas, apresentada na seção seguinte.

2.2 Iniciativas internacionais de interoperabilidade de credenciais acadêmicas

Além dos cinco estudos selecionados no Estado da Arte, foram consideradas, como referências complementares, três iniciativas europeias relacionadas ao intercâmbio e à representação de informações acadêmicas: o EMREX (European Student Mobility Data Exchange), o formato XML ELMO e o European Learning Model (ELM). Embora relacionadas, essas iniciativas possuem funções distintas e trajetórias próprias de desenvolvimento.

O EMREX é apresentado em sua documentação como uma solução para a transferência eletrônica de registros estudantis entre instituições europeias de ensino superior. O projeto, denominado *Field Trial on the Impact of Enabling Easy Mobility on Recognition of External Studies*, foi executado durante três anos, entre 2015 e 2017. Sua finalidade consistiu em testar formas de facilitar a administração da mobilidade estudantil e o reconhecimento de estudos realizados em outras instituições. O manual esclarece que o EMREX não pretendia desenvolver um sistema novo, pois a solução utilizava infraestruturas existentes, como serviços de autenticação, autorização e armazenamento de dados (EMREX, [s. d.]b).

O formato ELMO foi empregado pelo EMREX para estruturar os resultados acadêmicos intercambiados. A documentação técnica define o ELMO como um formato XML baseado no padrão CEN EN 15981:2011 EuroLMAI. Esse padrão descreve informações de avaliação presentes principalmente em diplomas, suplementos ao diploma e históricos escolares. O ELMO constitui, portanto, o formato utilizado para organizar os dados recuperados pelos componentes participantes do EMREX (EMREX, [s. d.]b; EMREX, [s. d.]c). As fontes oficiais consultadas não atribuem ao ELMO um período autônomo de desenvolvimento equivalente ao período do projeto EMREX.

O European Learning Model possui finalidade distinta. A Comissão Europeia o define como um modelo de dados multilíngue que fornece um vocabulário comum para a descrição da aprendizagem na Europa. A versão 3 foi disponibilizada em abril

de 2023, e seu evento oficial de lançamento ocorreu em 11 de maio do mesmo ano. A documentação informa também que o ELM v3 está mapeado e é interoperável com outros modelos, entre os quais o ELMO (EUROPEAN COMMISSION, [s. d.]a; EUROPEAN COMMISSION, 2023).

Na situação observada durante a realização desta pesquisa, o EMREX User Group apresenta-se como uma rede internacional independente dedicada à portabilidade dos dados estudantis. A documentação técnica do EMREX continua apresentando o ELMO como o formato XML utilizado no intercâmbio das informações de resultados acadêmicos. O ELM v3, por sua vez, é apresentado pela Comissão Europeia como uma versão estável do modelo, com suporte de longo prazo (EMREX, [s. d.]a; EMREX, [s. d.]c; EUROPEAN COMMISSION, [s. d.]a).

As referências examinadas não integraram o conjunto de trabalhos selecionados pelo levantamento bibliográfico, pois foram consultadas como fontes técnicas e institucionais complementares. Sua análise permitiu distinguir três recursos relacionados, mas funcionalmente diferentes: o EMREX, como infraestrutura de transferência de registros estudantis; o ELMO, como formato XML utilizado na estruturação dos dados intercambiados; e o European Learning Model, como modelo semântico destinado à descrição de informações sobre aprendizagem e qualificações.

Em conjunto, esses recursos evidenciaram que a interoperabilidade das credenciais acadêmicas pode envolver diferentes camadas, compreendendo infraestruturas de intercâmbio, formatos estruturados e modelos semânticos. Entretanto, eles não responderam diretamente à lacuna delimitada nesta pesquisa, referente à representação ontológica do subdomínio do Registro do Diploma Brasileiro com base nas normas e nos esquemas XSD do Ministério da Educação.

O Estado da Arte permitiu, assim, situar a OntoDiplomaDigital em relação aos estudos científicos e às referências internacionais examinadas. A construção da proposta exigiu, adicionalmente, a articulação dos fundamentos conceituais relacionados à organização da informação, aos metadados, à representação semântica, à interoperabilidade e ao Diploma Digital Brasileiro, apresentados no capítulo seguinte.

3 Fundamentação Teórica

A fundamentação teórica articula os conceitos de organização e representação do conhecimento, metadados, interoperabilidade e ontologias que sustentaram a construção da OntoDiplomaDigital.

O Quadro 3 a seguir sintetiza os principais referenciais mobilizados e suas contribuições para a fundamentação da pesquisa.

Quadro 3 - Principais autores e contribuições teóricas

Eixo conceitual	Principais referenciais	Contribuição para a fundamentação teórica
Transformação digital e governança de dados	Brasil (2021; 2024a; 2024b; 2024c)	Delimitação normativa do Governo Digital; distinção entre as estratégias nacional e federal; definição dos instrumentos de planejamento da transformação digital; e caracterização da governança, da gestão e da Infraestrutura Nacional de Dados.
Organização e Representação do Conhecimento	Mazzocchi (2018); Hjørland (2016)	Compreensão da organização do conhecimento como processo conceitual e contextual; explicitação das escolhas teóricas envolvidas na definição de conceitos, categorias e relações de representação.
Sistemas de Organização do Conhecimento	Zeng (2008); Barité (2011); Novak e Cañas (2010); Cañas et al. (2010)	Classificação dos Sistemas de Organização do Conhecimento segundo seus níveis de estrutura e expressividade; caracterização dos vocabulários controlados; e compreensão dos mapas conceituais como instrumentos de representação das relações entre conceitos.
Metadados	SAA (2005); Baca (2016); Arakaki e Arakaki (2020); Woodley (2016); Cardoso e Luz (2005); Lemos et al. (2022); Zeng e Qin (2022)	Definição das funções dos metadados; classificação de suas tipologias; discussão de seu ciclo de vida, qualidade e padronização; e delimitação de sua contribuição para a descrição, a gestão, a recuperação e o intercâmbio de recursos informacionais.
Princípios FAIR	Wilkinson et al. (2016); European Commission (2018); GO FAIR Initiative (2025)	Formulação e detalhamento dos princípios de encontrabilidade, acessibilidade, interoperabilidade e reutilização; caracterização da capacidade de processamento por máquinas; e compreensão das condições técnicas, informacionais e institucionais necessárias à implementação de ambientes FAIR.
Web Semântica e Dados Ligados	Berners-Lee, Hendler e Lassila (2001); Bizer, Heath e Berners-Lee (2009); Berners-Lee (2009); Heath e Bizer (2011); Laufer (2015)	Fundamentação da Web Semântica e da passagem de uma Web de documentos para uma Web de dados; apresentação dos princípios dos Dados Ligados; e compreensão do papel dos identificadores, das relações e dos padrões semânticos na publicação e integração de dados distribuídos.
Ontologias e interoperabilidade semântica	Gruber (1993); Borst (1997); Noy e McGuinness (2001); Falbo (2014); Guizzardi (2020)	Definição de ontologia como especificação explícita e formal de uma conceituação; distinção entre ontologias de referência e operacionais; discussão dos elementos da modelagem ontológica; e delimitação da

Eixo conceitual	Principais referenciais	Contribuição para a fundamentação teórica
		contribuição das ontologias para a interoperabilidade semântica.
Metodologias para criação de ontologias	Noy e McGuinness (2001); Mendonça (2015; 2021); Almeida (2024)	Sistematização das etapas de construção de ontologias; apresentação dos processos de especificação, aquisição do conhecimento, conceitualização, formalização, avaliação, documentação e disponibilização; e fundamentação da OntoForInfoScience.
Padrões, esquemas e vocabulários para interoperabilidade	Baca (2016); Riley (2017); Zeng e Qin (2022); Vandenbussche et al. (2016); W3C (2008; 2012)	Distinção entre interoperabilidade estrutural, sintática e semântica; caracterização de padrões e esquemas de metadados; explicação das funções do XML e do XSD; e discussão da reutilização de vocabulários semanticamente estruturados.
Diploma Digital e registros acadêmicos	Brasil (2018; 2019c; 2020b; 2025); Ribeiro et al. (2021); Santana Hüller, Gomes e Santos (2022)	Delimitação do contexto normativo e tecnológico do Diploma Digital Brasileiro; descrição dos requisitos de emissão, registro, estruturação, assinatura e preservação; e discussão dos desafios institucionais relacionados à integração dos sistemas e dos documentos acadêmicos.
Credenciais acadêmicas digitais em perspectiva internacional	EMREX (2020); European Commission ([s. d.]; 2025)	Apresentação de infraestruturas internacionais para intercâmbio de resultados acadêmicos; caracterização do formato ELMO, do European Learning Model e das credenciais digitais europeias; e articulação entre padronização estrutural, representação semântica e serviços de verificação.

Fonte: Elaborado pelo autor (2026)

3.1 Transformação Digital e Governança de dados

A transformação digital na Administração Pública não se limita à conversão de documentos ou procedimentos analógicos para formatos eletrônicos. Ela envolve mudanças nos processos organizacionais, na prestação dos serviços, na execução das políticas públicas e nas formas de interação entre o Estado e a sociedade. A transformação digital emprega tecnologias digitais para ampliar a integração dos serviços, melhorar o atendimento ao cidadão e produzir valor público, enquanto o governo digital corresponde à abordagem de gestão que orienta essa transformação nas organizações públicas. (Brasil, 2024a).

No Brasil, a Lei nº 14.129/2021 estabeleceu princípios, regras e instrumentos para o Governo Digital e para o aumento da eficiência pública. Entre suas diretrizes estão a desburocratização, a simplificação da relação entre o poder público e a sociedade, a disponibilização de serviços por meios digitais, a interoperabilidade de sistemas e a promoção do uso de dados pela Administração Pública. Esse marco

ampliou a compreensão da transformação digital ao vinculá-la à qualidade dos serviços, à transparência e à integração das informações governamentais (Brasil, 2021).

A Estratégia Nacional de Governo Digital (ENGD), instituída pelo Decreto nº 12.069/2024 para o período de 2024 a 2027, articula as estratégias de transformação digital da União, dos Estados, do Distrito Federal e dos Municípios. Seu alcance é federativo e busca orientar a oferta de serviços públicos inclusivos, eficazes, participativos e sustentáveis. O Decreto também relaciona a transformação digital à integração de serviços e políticas públicas, à transparência e à utilização de infraestruturas públicas digitais baseadas em padrões interoperáveis (Brasil, 2024a).

Em âmbito específico da Administração Pública Federal, o Decreto nº 12.198/2024 instituiu a Estratégia Federal de Governo Digital (EFGD) para o mesmo período. A EFGD sucede a estratégia anteriormente disciplinada pelo Decreto nº 10.332/2020, inicialmente destinada ao período de 2020 a 2022 e posteriormente prorrogada até 2023. Embora relacionadas, a ENGD e a EFGD possuem abrangências distintas: a primeira promove articulação entre os diferentes entes federativos, enquanto a segunda orienta os órgãos e as entidades da administração pública federal direta, autárquica e fundacional (Brasil, 2020a; Brasil, 2024a; Brasil, 2024b).

Para concretizar os objetivos da EFGD, os órgãos e as entidades federais devem elaborar instrumentos de planejamento, entre os quais o Plano de Transformação Digital (PTD). Esse plano deve contemplar, no mínimo, ações relacionadas aos serviços digitais e à melhoria da qualidade, à unificação dos canais digitais, à governança e à gestão de dados, bem como à segurança e à privacidade. Esses instrumentos caracterizam a transformação digital como um processo planejado e contínuo, e não como simples adoção isolada de ferramentas tecnológicas (Brasil, 2024b).

O mesmo Decreto instituiu a Infraestrutura Nacional de Dados (IND), compreendida como um conjunto de normas, políticas, arquiteturas, padrões, ferramentas tecnológicas e ativos informacionais voltados ao uso estratégico dos dados mantidos pelo Poder Executivo federal. Sua constituição evidencia que a circulação e o aproveitamento dos dados governamentais dependem da articulação entre recursos tecnológicos, padrões compartilhados, políticas institucionais e mecanismos de coordenação (Brasil, 2024b).

No âmbito da Infraestrutura Nacional de Dados, a governança de dados compreende a definição de diretrizes papéis, responsabilidades, padrões e práticas destinados a orientar, avaliar e monitorar a gestão dos dados sob responsabilidade das instituições. A gestão de dados, por sua vez, corresponde à execução das ações necessárias à utilização, à proteção, à manutenção e ao tratamento desses ativos. Os dois conceitos são complementares, mas não equivalentes: a governança estabelece direção, responsabilidades e mecanismos de acompanhamento, enquanto a gestão operacionaliza as decisões e os procedimentos definidos institucionalmente (Brasil, 2024c).

A implementação de programas de governança de dados pode ser organizada por meio de três instrumentos articulados: uma política interna, que estabelece diretrizes e responsabilidades; uma estratégia de dados, que define objetivos, iniciativas e metas; e um plano de implementação, que distribui ações, prazos, recursos e responsáveis. Essa organização demonstra que qualidade, segurança, integração e reúso dos dados dependem tanto de condições técnicas quanto de decisões institucionais capazes de sustentar sua gestão ao longo do tempo (Brasil, 2024c).

A interoperabilidade ocupa posição central nesse processo, pois serviços públicos integrados dependem da capacidade de diferentes órgãos e sistemas intercambiarem informações de maneira consistente. Essa capacidade não resulta apenas da conexão entre plataformas: requer padrões, identificadores, estruturas de dados e critérios comuns que possibilitem interpretar e utilizar as informações recebidas. A Infraestrutura Nacional de Dados incorpora essa perspectiva ao associar o uso estratégico dos dados à sua descoberta, ao acesso, à interoperabilidade, à privacidade, à segurança e à proteção (Brasil, 2024b; Riley, 2017).

No ensino superior, a instituição do Diploma Digital inseriu os processos de emissão e registro acadêmico nesse ambiente de transformação dos serviços e documentos públicos. A Portaria MEC nº 330/2018 estabeleceu a emissão de diplomas de graduação em formato digital, enquanto a Portaria MEC nº 554/2019 disciplinou requisitos relacionados à emissão, ao registro e à preservação desses documentos. A digitalização do diploma, contudo, não elimina a necessidade de organizar e representar consistentemente as informações que circulam entre Instituições de Ensino Superior, órgãos reguladores e plataformas governamentais (Brasil, 2018; Brasil, 2019c).

A governança estabelece responsabilidades, políticas e condições para a gestão dos dados, mas sua execução também depende de mecanismos capazes de estruturar conceitos, terminologias e relações. Quando informações provenientes de diferentes instituições precisam ser integradas, torna-se necessário considerar não somente seu formato, mas também a maneira como o conhecimento representado é organizado e interpretado. A Organização e Representação do Conhecimento oferece fundamentos para essa estruturação conceitual (Mazzocchi, 2018; Zeng; QIN, 2022).

3.2 Organização e Representação do Conhecimento

A Organização e Representação do Conhecimento compreende processos destinados à estruturação conceitual da informação, envolvendo a identificação, a seleção, a categorização e o estabelecimento de relações entre conceitos de determinado domínio. Esses processos não se limitam à organização física ou descritiva dos recursos informacionais, mas envolvem a construção de estruturas conceituais capazes de apoiar sua interpretação, recuperação e compartilhamento. Mazzocchi (2018) destaca que a Organização do Conhecimento busca representar os conceitos e as relações existentes em um domínio, permitindo sua organização sistemática em diferentes contextos informacionais.

A representação do conhecimento torna explícita essa organização ao expressar conceitos, categorias e relações de forma estruturada. Hjørland (2016) caracteriza os sistemas de organização do conhecimento como seleções de conceitos acompanhadas de relações semânticas e destaca que os processos de organização são orientados por conhecimentos e pressupostos teóricos. Na abordagem de análise de domínio, os critérios de organização precisam ser examinados em relação às teorias e às perspectivas presentes no campo considerado, incluindo possíveis divergências entre elas. A representação de um domínio expressa, portanto, escolhas sobre os elementos e as relações considerados relevantes, cuja compreensão exige explicitar os fundamentos dessas escolhas.

A crescente complexidade dos ambientes informacionais amplia essas exigências. A diversidade de fontes, a heterogeneidade terminológica e a produção contínua de dados demandam estruturas capazes de representar diferentes níveis de granularidade e múltiplos contextos de aplicação. Mazzocchi (2018) destaca que os ambientes contemporâneos de informação exigem instrumentos capazes de organizar

o conhecimento de maneira consistente, considerando diferentes perspectivas e necessidades de representação.

Organizar e representar conhecimento envolve, portanto, transformar informações dispersas em estruturas conceituais que permitam reconhecer entidades, estabelecer distinções, controlar ambiguidades e explicitar relações. Esse processo fundamenta a construção de diferentes instrumentos de organização do conhecimento, entre os quais se encontram os Sistemas de Organização do Conhecimento.

Os Sistemas de Organização do Conhecimento (SOCs) são instrumentos utilizados para estruturar e representar conhecimento de maneira sistemática, favorecendo a organização conceitual, o controle terminológico e a recuperação da informação (Mazzocchi, 2018). Esses sistemas apresentam diferentes níveis de complexidade e formalização, conforme as necessidades do domínio e os objetivos do ambiente informacional em que são empregados.

Zeng (2008) organiza os SOCs em quatro categorias principais: listas de termos, modelos baseados em metadados, sistemas de classificação e modelos relacionais. As listas de termos são utilizadas principalmente para controle e normalização terminológica; os modelos baseados em metadados acrescentam elementos estruturados para descrição de recursos; os sistemas de classificação organizam conceitos em categorias e hierarquias; e os modelos relacionais ampliam a capacidade de representar associações entre conceitos.

Essa classificação demonstra que os SOCs possuem diferentes capacidades expressivas. Alguns instrumentos concentram-se no controle de denominações, enquanto outros permitem representar hierarquias, associações e relações conceituais mais complexas. A escolha de um determinado sistema depende das características do domínio, do nível de detalhamento necessário e das finalidades informacionais pretendidas.

Barité (2011) amplia essa perspectiva ao tratar os Sistemas de Organização do Conhecimento como estruturas conceituais destinadas à representação e recuperação da informação. O autor apresenta uma visão abrangente desses sistemas, incluindo instrumentos como sistemas de classificação, vocabulários controlados, tesauros, ontologias e mapas conceituais, destacando que essas estruturas apresentam diferentes níveis de formalização e diferentes possibilidades de representação.

Além da organização terminológica, os SOCs contribuem para reduzir ambiguidades decorrentes da linguagem natural. Zeng (2008) destaca que esses sistemas auxiliam no controle de variações terminológicas, relações hierárquicas e associações conceituais, favorecendo maior consistência na representação e recuperação da informação.

Entre os instrumentos utilizados para representação do conhecimento, os mapas conceituais constituem uma forma gráfica de explicitar conceitos e relações entre elementos de um domínio. Novak e Cañas (2010) descrevem os mapas conceituais como estruturas formadas por conceitos, palavras de ligação e proposições, permitindo representar graficamente relações significativas entre conceitos. Embora apresentem menor grau de formalização computacional quando comparados às ontologias, esses instrumentos contribuem para a compreensão e organização inicial de domínios de conhecimento.

A elaboração de um SOC envolve escolhas teóricas e metodológicas relacionadas aos conceitos que serão representados, às categorias utilizadas e às relações estabelecidas. Mazzocchi (2018) destaca que toda organização do conhecimento resulta de decisões sobre como um domínio será compreendido e estruturado. Zeng (2008) complementa que a definição das categorias, níveis de abstração e relações deve estar alinhada aos objetivos do sistema de informação. Assim, um SOC representa uma determinada perspectiva sobre o conhecimento organizado, e não apenas uma estrutura técnica neutra.

No contexto dos ambientes digitais, a Organização do Conhecimento aproxima-se de tecnologias voltadas à representação formal e ao processamento computacional da informação. Pereira, Luz e Santarem Segundo (2019) relacionam essas práticas ao desenvolvimento e gestão de vocabulários controlados, construção de ontologias, uso de metadados e aplicação de tecnologias como RDF, URIs e SPARQL. Essa aproximação evidencia a articulação entre os fundamentos da Ciência da Informação e os mecanismos utilizados para representar conhecimento em ambientes digitais.

Nesse percurso, as ontologias representam instrumentos com maior grau de formalização, pois permitem representar conceitos, propriedades e relações de um domínio de maneira estruturada e interpretável por máquinas. Essa característica amplia as possibilidades de integração e reúso da informação, estabelecendo uma conexão entre os fundamentos da Organização do Conhecimento e as abordagens da Web Semântica, discutidas nas próximas seções.

3.3 Metadados

Os metadados constituem informações estruturadas associadas a recursos informacionais, utilizadas para identificá-los, descrevê-los, contextualizá-los, administrá-los e possibilitar sua recuperação e utilização em diferentes ambientes. Sua função ultrapassa a simples identificação de um objeto, pois envolve informações relacionadas ao conteúdo, à estrutura, à origem, às condições de acesso e às formas de gerenciamento dos recursos ao longo de seu uso (SAA, [s. d.]; Zeng; Qin, 2022).

A Society of American Archivists (SAA) define metadados como informações sobre dados que contribuem para a descoberta, a estruturação de objetos de dados e o apoio à administração e à preservação dos registros. Essa compreensão evidencia que os metadados podem atuar em diferentes níveis de granularidade e acompanhar os recursos informacionais em diferentes momentos de sua existência (SAA, [s. d.]). Riley (2017) amplia essa perspectiva ao caracterizar os metadados como informações estruturadas utilizadas para descrever, explicar, localizar ou facilitar a recuperação, o uso e a gestão de recursos informacionais.

Os metadados desempenham diferentes funções conforme a finalidade e o contexto em que são empregados. Zeng e Qin (2022) organizam essas funções em três categorias abrangentes: metadados descritivos, administrativos e estruturais. Os metadados descritivos são utilizados para identificar, selecionar e descobrir recursos, incluindo elementos como título, autoria, assunto e data. Os administrativos apoiam a gestão dos recursos, abrangendo informações técnicas, direitos, condições de acesso e aspectos relacionados à preservação. Os estruturais representam a organização interna de um recurso e as relações existentes entre suas partes, como ocorre em objetos digitais compostos por múltiplos componentes (Zeng; Qin, 2022).

Essas categorias não devem ser compreendidas como estruturas isoladas ou mutuamente excludentes. Diferentes classificações refletem perspectivas distintas sobre as funções desempenhadas pelos metadados. Gilliland (2016) apresenta uma abordagem funcional que distingue metadados administrativos, descritivos, de preservação, técnicos e de uso, enquanto Riley (2017) destaca os agrupamentos descritivos, administrativos e estruturais, incluindo aspectos técnicos, de preservação e direitos no âmbito administrativo. Essas abordagens demonstram que a classificação dos metadados depende das necessidades da comunidade, do domínio de aplicação e dos objetivos do sistema informacional.

Arakaki e Arakaki (2020), ao analisarem diferentes perspectivas terminológicas sobre metadados, apresentam categorias adicionais relacionadas à autenticação, proveniência, acesso e uso, meta-metadados, aspectos técnicos e estruturais, além das linguagens de marcação. Essas categorias evidenciam que os metadados podem registrar não apenas características do recurso, mas também informações sobre sua origem, responsabilidades envolvidas, condições de utilização e características do próprio registro de metadados.

O Quadro 4 sintetiza categorias de metadados identificadas na literatura e suas respectivas finalidades.

Quadro 4 - Tipo de metadados e suas finalidades

Tipo de metadado	Finalidade principal
Descritivo	Registra características destinadas à identificação, descoberta, seleção e recuperação dos recursos informacionais.
Administrativo	Apoia a gestão dos recursos, incluindo informações necessárias ao controle, administração e tomada de decisão.
De autenticação	Registra informações utilizadas para verificar identidade, integridade ou legitimidade do recurso.
De preservação	Reúne informações necessárias para manutenção, autenticidade e acesso ao recurso ao longo do tempo.
De direitos	Registra informações relacionadas à propriedade intelectual, licenças e condições jurídicas de utilização.
De acesso e uso	Indica condições de acesso e informações relacionadas à utilização do recurso.
De proveniência	Registra a origem do recurso, entidades responsáveis por sua produção e modificações realizadas.
Meta-metadados	Descreve características do próprio registro de metadados ou do conjunto de metadados.
Técnico	Registra características necessárias ao processamento, funcionamento ou reprodução do recurso em determinado ambiente tecnológico.
Estrutural	Representa a composição do recurso e as relações existentes entre seus componentes.
Linguagens de marcação	Integra elementos descritivos ou semânticos à estrutura do conteúdo, permitindo sua representação e interpretação.

Fonte: Adaptado de Riley (2017), Gilliland (2016) e Arakaki e Arakaki (2020)

A diversidade funcional dos metadados demonstra que sua utilização está relacionada não apenas à descrição dos recursos, mas também à sua gestão, preservação, autenticação, proveniência e reúso. Em ambientes digitais, diferentes processos podem gerar informações complementares sobre um mesmo recurso, formando camadas de descrição associadas às necessidades dos usuários, das instituições e dos sistemas responsáveis por sua administração (Gilliland, 2016).

A qualidade dos metadados depende da consistência das informações registradas, da adequação aos padrões utilizados e da aplicação de procedimentos que assegurem sua confiabilidade. Critérios como completude, consistência, conformidade e atualização contribuem para ampliar a capacidade de descoberta e

reutilização dos recursos informacionais. A adoção de padrões, regras de descrição e mecanismos de controle reduz divergências e favorece a manutenção da qualidade dos registros (Zeng; Qin, 2022).

A padronização constitui um elemento fundamental para o intercâmbio de informações entre sistemas. Padrões de metadados estabelecem conjuntos de elementos, estruturas e regras comuns para representação dos recursos, reduzindo diferenças entre descrições produzidas por diferentes comunidades. Cardoso e Luz (2005) destacam o papel dos padrões de metadados, como o Dublin Core, na organização dos objetos digitais e na possibilidade de coexistência com diferentes esquemas de representação. Os autores também ressaltam a importância de qualificadores e vocabulários controlados para ampliar a precisão da descrição dos recursos.

Entretanto, a existência de uma estrutura comum não garante, por si só, que diferentes sistemas atribuam o mesmo significado aos elementos representados. Dois modelos podem utilizar elementos estruturalmente equivalentes e ainda assim representar conceitos ou valores de maneira distinta. Riley (2017) observa que a interoperabilidade depende da capacidade dos sistemas reconhecerem e relacionarem informações provenientes de diferentes estruturas de representação. Woodley (2016) complementa essa perspectiva ao destacar que a interoperabilidade entre esquemas de metadados envolve aspectos sintáticos, estruturais e semânticos, sendo necessário analisar não apenas os elementos utilizados, mas também os significados associados aos valores e vocabulários empregados.

Essa distinção evidencia a diferença entre compatibilidade estrutural e interoperabilidade semântica. A utilização de padrões de descrição possibilita o intercâmbio de estruturas informacionais, mas a interpretação consistente dos dados requer mecanismos capazes de explicitar conceitos, relações e significados compartilhados. Nesse contexto, vocabulários controlados, sistemas de organização do conhecimento e ontologias complementam os metadados ao fornecerem mecanismos para representação formal dos conceitos e suas relações.

Lemos et al. (2022) destacam que a Ciência da Informação contribui para a organização e representação da informação por meio de recursos como padrões de metadados, vocabulários controlados e ontologias. Esses instrumentos atuam de forma complementar na descrição, organização e gestão dos recursos informacionais, ampliando as condições para seu uso e reúso em diferentes ambientes.

A relação entre metadados, interoperabilidade e reúso também se aproxima dos princípios FAIR (*Findable, Accessible, Interoperable and Reusable*). Wilkinson et al. (2016) estabelecem que dados e metadados devem ser descritos com informações suficientemente ricas, utilizar identificadores persistentes, empregar vocabulários compartilhados e registrar informações de proveniência para favorecer sua descoberta, acesso, integração e reutilização. Dessa forma, os metadados constituem uma base informacional para ambientes nos quais recursos precisam ser encontrados, interpretados e reutilizados por diferentes sistemas e comunidades.

Assim, os metadados representam uma camada fundamental para organização e intercâmbio de informações, mas sua contribuição para a interoperabilidade depende da existência de mecanismos capazes de superar limitações exclusivamente estruturais. Enquanto padrões de metadados estabelecem elementos e formatos compartilhados, mecanismos de representação semântica permitem explicitar conceitos e relações, favorecendo a interpretação consistente dos dados entre diferentes ambientes. Essa perspectiva fundamenta a abordagem da Web Semântica e das ontologias apresentada nas próximas seções.

3.4 Princípios FAIR

O aumento na produção e na circulação de dados digitais intensificou a demanda por práticas que promovam sua descoberta, acesso, integração e reutilização em diversos contextos informacionais. Wilkinson et al. (2016) consolidaram os princípios FAIR — Findable, Accessible, Interoperable and Reusable — como diretrizes para aprimorar a gestão, a descrição e a disponibilidade de dados e metadados. Tais princípios abrangem quatro atributos essenciais para objetos digitais: encontrabilidade, acessibilidade, interoperabilidade e reutilização, conforme detalhado no Quadro 5. Esses princípios não constituem uma norma tecnológica única, nem prescrevem uma forma exclusiva de implementação; tratam-se de orientações gerais que podem ser operacionalizadas por meio de diferentes padrões, tecnologias, políticas e práticas de gestão da informação, conforme os objetivos, os ambientes de uso e as comunidades responsáveis pelos dados.

Um aspecto central dos princípios FAIR é a capacidade de os objetos digitais serem identificados, interpretados e utilizados tanto por pessoas quanto por sistemas computacionais. Essa característica depende da existência de identificadores

persistentes, metadados estruturados, protocolos padronizados e linguagens formais capazes de representar informações de maneira compreensível para diferentes ambientes tecnológicos. A aplicação dos princípios FAIR, portanto, não se limita à disponibilização de dados, mas envolve a criação das condições necessárias para sua descoberta, integração e reutilização em novos contextos (Wilkinson et al., 2016).

A dimensão **Encontrável (*Findable*)** refere-se à capacidade de dados e metadados serem identificados e descobertos. Para isso, os objetos digitais devem possuir identificadores globalmente únicos e persistentes, ser descritos por metadados suficientemente ricos e estar registrados ou indexados em recursos que possibilitem sua localização. Além disso, os metadados devem conter explicitamente o identificador dos dados aos quais se referem, estabelecendo uma relação clara entre descrição e objeto digital (Wilkinson et al., 2016).

A dimensão **Acessível (*Accessible*)** está relacionada às condições pelas quais dados e metadados podem ser recuperados a partir de seus identificadores. Essa recuperação deve ocorrer por meio de protocolos de comunicação padronizados, abertos e amplamente implementáveis, admitindo mecanismos de autenticação e autorização quando necessários. O princípio também estabelece que os metadados devem permanecer disponíveis mesmo quando os dados correspondentes deixarem de estar acessíveis, preservando informações necessárias para sua identificação e compreensão (Wilkinson et al., 2016).

A dimensão **Interoperável (*Interoperable*)** exige que dados e metadados sejam representados utilizando linguagens formais, acessíveis, compartilhadas e aplicáveis à representação do conhecimento. Também pressupõe o uso de vocabulários compatíveis com os princípios FAIR e a existência de referências qualificadas entre objetos digitais. Essas condições permitem relacionar informações provenientes de diferentes fontes e favorecem sua interpretação consistente por pessoas e sistemas computacionais (Wilkinson et al., 2016).

A dimensão **Reutilizável (*Reusable*)** depende da existência de descrições suficientemente detalhadas para que os dados possam ser compreendidos e utilizados em diferentes contextos. Para isso, os objetos digitais devem apresentar informações sobre condições de uso, licenciamento, proveniência e padrões reconhecidos pelas comunidades às quais se destinam. A reutilização não depende apenas da disponibilidade técnica do recurso, mas também da existência de

informações que permitam avaliar sua origem, contexto e adequação para novas finalidades (Wilkinson et al., 2016).

Quadro 5 - Princípios FAIR

Princípio	Condições fundamentais
Encontrável — Findable	Os dados e metadados recebem identificadores globalmente únicos e persistentes; são descritos por metadados ricos; os metadados incluem explicitamente o identificador dos dados; e são registrados ou indexados em recurso pesquisável.
Acessível — Accessible	Os dados e metadados podem ser recuperados por seus identificadores mediante protocolo de comunicação padronizado; o protocolo é aberto, gratuito e universalmente implementável; admite autenticação e autorização quando necessário; e os metadados permanecem acessíveis mesmo quando os dados não estão mais disponíveis.
Interoperável — Interoperable	Os dados e metadados utilizam linguagem formal, acessível, compartilhada e amplamente aplicável à representação do conhecimento; empregam vocabulários que seguem os princípios FAIR; e apresentam referências qualificadas a outros dados ou metadados.
Reutilizável — Reusable	Os dados e metadados são descritos com atributos relevantes; apresentam licença de uso clara e acessível; contêm informações detalhadas de proveniência; e observam padrões pertinentes ao domínio ou à comunidade.

Fonte: Adaptado de Wilkinson et al. (2016).

A conformidade com os princípios FAIR não significa necessariamente que os dados devam estar disponíveis de forma irrestrita. O acesso aos recursos digitais pode estar sujeito a limitações legais, éticas, institucionais, comerciais ou relacionadas à privacidade. Nesses casos, os princípios FAIR estabelecem que as condições de acesso devem ser explicitadas e que os metadados associados aos recursos devem permanecer disponíveis para possibilitar sua descoberta, avaliação e eventual reutilização. Dessa forma, a abertura constitui uma possibilidade desejável quando compatível com as restrições existentes, mas não representa uma condição obrigatória para que um objeto digital seja considerado FAIR (European Commission, 2018).

A implementação dos princípios FAIR ultrapassa a adoção isolada de tecnologias ou padrões. Ela envolve a articulação entre objetos digitais, identificadores, metadados, padrões, serviços, políticas institucionais, infraestrutura tecnológica e competências profissionais. O relatório *Turning FAIR into Reality* destaca que a construção de um ambiente FAIR depende da integração entre componentes técnicos e organizacionais capazes de sustentar a produção, descrição,

preservação, descoberta e reutilização dos objetos digitais (European Commission, 2018).

A Comissão Europeia (2018) organiza as recomendações para implementação do FAIR em três dimensões complementares detalhadas no Quadro 6. A primeira dimensão concentra-se na definição dos objetos digitais FAIR e do ecossistema necessário para sua produção e gestão. Nessa etapa, destacam-se elementos como identificadores persistentes, metadados essenciais, padrões comunitários, políticas de gestão e repositórios confiáveis.

Quadro 6 - Conceitos para Objetos Digitais FAIR e o ecossistema

Recomendação	Descrição
1 - Definir FAIR para implementação:	Incorporar conceitos relacionados à seleção, gestão de longo prazo, interoperabilidade, compartilhamento e relação entre FAIR e abertura dos dados, considerando o princípio de que os dados devem ser “o mais abertos possível e o mais restritos necessário”.
2 - Implementar um modelo para Objetos Digitais FAIR:	Garantir que objetos digitais FAIR possuam identificadores persistentes associados a metadados essenciais, incluindo informações de proveniência, licenciamento e padrões reconhecidos pela comunidade.
3 - Desenvolver componentes de um ecossistema FAIR:	Estabelecer políticas, planos de gestão de dados, identificadores, padrões, repositórios e fluxos de trabalho que sustentem a produção e o gerenciamento dos objetos digitais.

Fonte: Adaptado de European Commission (2018)

A segunda dimensão apresentada no Quadro 7, refere-se à implementação prática dos princípios FAIR e envolve cultura organizacional, infraestrutura tecnológica e desenvolvimento de competências. A adoção dessas práticas depende da participação integrada de pesquisadores, profissionais da informação, gestores e equipes técnicas responsáveis pela produção, organização e manutenção dos dados e metadados. Além disso, requer serviços especializados capazes de transformar princípios gerais em procedimentos aplicáveis a diferentes domínios científicos e institucionais (European Commission, 2018).

Quadro 7 - Implementação - cultura, tecnologia e habilidades para a Prática FAIR

Recomendação	Segunda dimensão: Implementar – cultura, tecnologia e habilidades para a prática FAIR
4 - Desenvolver frameworks de interoperabilidade para compartilhamento FAIR:	Apoiar comunidades de pesquisa na criação de frameworks que definam suas práticas de compartilhamento e promover padrões comuns para pesquisa interdisciplinar.

Recomendação	Segunda dimensão: Implementar – cultura, tecnologia e habilidades para a prática FAIR
5 - Garantir a gestão de dados por meio de planos de gerenciamento de dados:	Exigir que todos os projetos de pesquisa incluam um DMP como elemento central para todos os resultados de pesquisa, atualizado regularmente.
6 - Reconhecer e recompensar práticas FAIR	Incluir dados FAIR como um resultado de pesquisa central na avaliação de contribuições de pesquisa e progressão na carreira.
7 - Apoiar tecnologias semânticas:	Desenvolver, expandir e aplicar tecnologias semânticas dentro e entre disciplinas para a interoperabilidade.
8 - Facilitar o processamento automatizado:	Habilitar a interação máquina-máquina entre componentes FAIR em múltiplos níveis e disciplinas.
9 - Desenvolver frameworks de avaliação e certificação	Criar mecanismos para avaliar serviços e infraestruturas responsáveis pela gestão de objetos digitais FAIR.
10 - Profissionalizar funções relacionadas à ciência e gestão de dados	Desenvolver competências especializadas para apoiar produção, organização, preservação e reutilização de dados.

Fonte: Adaptado de European Commission (2018)..

A terceira dimensão trata da incorporação e da sustentabilidade das práticas FAIR conforme o Quadro 8. Sua consolidação depende de incentivos institucionais, modelos de avaliação, métricas adequadas e investimentos de longo prazo. Esses elementos favorecem a integração da gestão FAIR às atividades regulares das instituições e evitam que sua aplicação permaneça restrita a iniciativas pontuais ou dependentes de projetos temporários (European Commission, 2018).

Quadro 8 - Incorporação e sustentabilidade: incentivos, métricas e investimentos

Recomendação	Terceira dimensão: Incorporar e sustentar – incentivos, métricas e investimento
12 - Desenvolver métricas para Objetos Digitais FAIR:	Criar e implementar um conjunto de métricas para Objetos Digitais FAIR, guiadas pelas práticas comunitárias.
13 - Desenvolver métricas para certificar serviços FAIR:	Utilizar e adaptar frameworks de certificação existentes, como o CoreTrustSea (CTS)l, para avaliar todos os componentes do ecossistema FAIR.
14 - Fornecer financiamento estratégico e coordenado:	Adotar uma abordagem coordenada para apoiar infraestruturas e serviços centrais, baseada em investimentos existentes e ligada a esquemas de certificação e modelos de negócios sustentáveis.

Recomendação	Terceira dimensão: Incorporar e sustentar – incentivos, métricas e investimento
15 - Fornecer financiamento sustentável:	Garantir que o financiamento apoie a manutenção dos componentes do ecossistema FAIR em um nível de serviço profissional e sustentável.

Fonte: Adaptado de European Commission (2018)

Em conjunto, os princípios FAIR articulam dimensões informacionais, tecnológicas e institucionais. Identificadores persistentes e protocolos favorecem a localização e o acesso; metadados estruturados, linguagens formais e vocabulários compartilhados contribuem para a interoperabilidade; enquanto licenças, informações de proveniência e padrões comunitários estabelecem condições para a reutilização responsável dos recursos digitais. A implementação desses princípios depende, portanto, da relação entre a estruturação dos objetos digitais e o ambiente organizacional responsável por sua gestão (Wilkinson et al., 2016; European Commission, 2018).

Entre essas dimensões, a interoperabilidade apresenta especial relevância para esta pesquisa, pois evidencia a necessidade de representações formais capazes de explicitar conceitos, relações e significados compartilhados. O emprego de linguagens formais, vocabulários e referências qualificadas aproxima os princípios FAIR das abordagens da Web Semântica e dos Dados Ligados, que oferecem mecanismos para representar e relacionar objetos digitais de maneira processável por máquinas.

A interoperabilidade requer, portanto, não apenas estruturas compatíveis, mas também mecanismos capazes de explicitar o significado dos elementos representados. As ontologias contribuem para essa finalidade ao possibilitar a representação formal de conceitos, propriedades e relações de um domínio, estabelecendo uma base estruturada para a interpretação e o reúso dos dados.

3.5 Web Semântica e Dados Ligados

A semântica ocupa-se do significado e das relações que possibilitam interpretar palavras, expressões e conceitos. Almeida, Souza e Fonseca (2011) distinguem três dimensões relevantes para essa compreensão: a composicionalidade, referente à formação do significado pela combinação de elementos linguísticos; as propriedades semânticas, que expressam relações sistemáticas entre palavras e sentenças; e a

referência, por meio da qual as expressões são vinculadas aos objetos e conceitos que representam.

Com o desenvolvimento da Semântica Formal, tornou-se possível representar determinados aspectos do significado por meio de modelos lógicos. Essa formalização criou condições para que relações conceituais fossem expressas em estruturas processáveis por sistemas computacionais, embora a representação formal não elimine a complexidade linguística, contextual e epistemológica envolvida na produção do significado (Almeida; Souza; Fonseca, 2011).

A Web Semântica fundamenta-se na possibilidade de tornar explícitos os significados e as relações associados aos recursos disponíveis na Web. Berners-Lee, Hendler e Lassila (2001) a caracterizam como uma extensão da Web existente, na qual as informações recebem uma estrutura que permite a cooperação entre pessoas e agentes computacionais. Não se trata, portanto, de substituir a Web de documentos, mas de ampliar suas condições de integração, descoberta e processamento.

De acordo com Laufer (2015), esse desenvolvimento pode ser atribuído a três estágios distintos da Web: a Web de Documentos, caracterizada pela estrutura de páginas e hiperlinks; a Web Programável, fundamentada em serviços e interfaces de programação; e a Web de Dados, que se concentra na identificação, descrição e interligação dos recursos representados. As conexões deixam de se limitar a documentos, passando a abranger também os dados e as entidades descritas nesses documentos.

Os Dados Ligados, ou Linked Data, constituem um conjunto de princípios e práticas para publicar dados estruturados na Web de modo que os recursos possam ser identificados, acessados e relacionados. Sua proposta baseia-se no emprego de identificadores globais e na criação de ligações entre recursos pertencentes a diferentes fontes, permitindo que os dados permaneçam distribuídos e, ao mesmo tempo, integrem uma rede de relações explícitas (Berners-Lee, 2006; Bizer; Heath; Berners-Lee, 2009).

A publicação de Dados Ligados não se reduz à conversão de dados para um formato específico. Auer et al. (2013) apresentam o ciclo de vida dos Dados Ligados como um conjunto de atividades relacionadas à extração, autoria, armazenamento e consulta, interligação, enriquecimento, análise da qualidade, evolução e reparo dos dados, além de processos de pesquisa, navegação e exploração. Essas atividades não representam etapas isoladas ou necessariamente sequenciais, mas componentes

interdependentes que se apoiam mutuamente durante a publicação, manutenção e utilização dos dados na Web.

A Figura 2 - Ciclo de vida dos dados ligados sintetiza esse caráter contínuo do ciclo de vida dos Dados Ligados, evidenciando que sua produção envolve tanto atividades técnicas de transformação e armazenamento quanto processos de interligação, avaliação, manutenção e exploração dos dados (Auer et al., 2013).

Figura 2 - Ciclo de vida dos dados ligados



Fonte: Adaptado por Auer et al. (2011),

A representação dos recursos nesse ambiente é sustentada pelo Resource Description Framework (RDF), modelo que permite expressar informações em forma de grafos. No RDF, cada afirmação é representada por uma tripla composta por sujeito, predicado e objeto, estrutura também indicada pela sigla SPO, derivada dos termos subject, predicate e object (Cyganiak et al., 2014).

O sujeito identifica o recurso sobre o qual se realiza uma afirmação. O predicado expressa a propriedade ou a relação utilizada nessa afirmação, enquanto o objeto indica outro recurso ou um valor associado ao sujeito. Quando diferentes triplas compartilham recursos, forma-se um grafo cujas conexões podem ser percorridas e combinadas por aplicações computacionais (Cyganiak et al., 2014).

Sujeito, predicado e objeto designam posições na estrutura da tripla, enquanto recurso, propriedade, classe e literal indicam funções semânticas distintas. Um recurso corresponde a algo que pode ser identificado e descrito; uma propriedade expressa uma característica ou relação; uma classe reúne recursos que compartilham

determinada caracterização; e um literal representa um valor, como texto, número ou data (Cyganiak et al., 2014; Ramalho, 2015).

Essa distinção é necessária porque a palavra “objeto” pode ser utilizada em sentidos diferentes. Em RDF, objeto é a terceira posição da tripla e pode conter um recurso ou um literal. Na representação de ontologias com OWL, a expressão “propriedade de objeto” designa uma propriedade que relaciona recursos, enquanto uma propriedade de dados associa um recurso a um valor literal (W3C, [s. d.]; Ramalho, 2015).

O Quadro 9 - Elementos fundamentais da representação RDF organiza os elementos fundamentais da representação RDF, distinguindo as posições de uma tripla das estruturas que resultam da combinação entre diferentes afirmações (Cyganiak et al., 2014).

Quadro 9 - Elementos fundamentais da representação RDF

Elemento	Função	Posição ou uso na tripla RDF	Exemplo	Leitura didática
Sujeito	Indica o recurso sobre o qual se faz uma afirmação.	Primeira posição da tripla.	rec:diplomaCaso1	É o recurso que está sendo descrito.
Predicado	Expressa a relação ou característica atribuída ao sujeito.	Segunda posição da tripla; corresponde a uma propriedade.	ont:hasHolder	Indica como o sujeito se relaciona com o objeto.
Objeto	Representa aquilo que se relaciona ao sujeito ou o valor atribuído a ele.	Terceira posição da tripla; pode ser outro recurso ou um literal.	rec:diplomadoCaso1 ou "CPF-FICTICIO-0001"	Completa a afirmação feita sobre o sujeito.
Recurso	Representa uma entidade identificável no grafo.	Pode ocupar a posição de sujeito ou de objeto.	rec:diplomaCaso1, rec:diplomadoCaso1	É uma entidade específica representada no grafo.
Classe	Define um tipo ou categoria de recursos.	Pode aparecer como objeto de uma declaração com rdf:type, ou com sua forma abreviada a.	ont:DiplomaDigital, ont:Diplomado	Indica a categoria à qual o recurso pertence.
Propriedade	Define uma relação ou uma característica.	Ocupa a posição de predicado.	ont:hasHolder, ont:cpf	Estabelece o vínculo entre sujeito e objeto.
Propriedade de objeto	Relaciona um recurso a outro recurso.	O objeto é outro recurso identificável no grafo.	rec:diplomaCaso1 ont:hasHolder rec:diplomadoCaso1 .	O diploma do caso 1 tem como titular o diplomado do caso 1.

Elemento	Função	Posição ou uso na tripla RDF	Exemplo	Leitura didática
Propriedade de dados	Relaciona um recurso a um valor literal.	O objeto é um valor textual, numérico, temporal ou de outro tipo de dado.	rec:diplomadoCaso1 ont:cpf "CPF-FICTICIO-0001" .	O diplomado do caso 1 possui o CPF fictício indicado.
Literal	Representa um valor de dado.	Ocupa a posição de objeto.	"CPF-FICTICIO-0001", "Pessoa Fictícia-0001"	É o valor propriamente atribuído ao recurso.

Fonte: Adaptado de W3C (2006).

A aplicação desses elementos pode ser observada no Quadro 10 em um trecho da ABox[1] pública da prova de conceito, isto é, no conjunto de afirmações sobre indivíduos e relações instanciadas segundo a estrutura formal da ontologia.

Quadro 10 - Representação de indivíduos em um trecho ABox

```
rec:diplomaCaso1 a ont:DiplomaDigital ;
    ont:hasHolder rec:diplomadoCaso1 ;
    ont:hasProgram rec:cursoCaso1 ;
    ont:hasStatus vocab:Ativo .
rec:diplomadoCaso1 a ont:Diplomado ;
    ont:fullName "Pessoa Fictícia-0001" ;
    ont:cpf "CPF-FICTICIO-0001" ;
    ont:hasBirthPlace rec:localNaturalidadeBr3106200 .
```

Fonte: Elaborado pelo autor (2026), com base em W3C (2013) e na ABox pública da prova de conceito da OntoDiplomaDigital.

No primeiro conjunto de triplas, `rec:diplomaCaso1` é o sujeito das afirmações. A declaração a `ont:DiplomaDigital` indica que esse recurso pertence à classe `ont:DiplomaDigital`. As propriedades `ont:hasHolder`, `ont:hasProgram` e `ont:hasStatus` exercem a função de predicado, pois expressam relações entre o diploma e outros recursos do grafo. Seus objetos são, respectivamente, `rec:diplomadoCaso1`, `rec:cursoCaso1` e `vocab:Ativo`.

No segundo conjunto de triplas, rec:diplomadoCaso1 passa a ocupar a posição de sujeito. Isso demonstra que um recurso que aparece como objeto em uma afirmação pode tornar-se sujeito em outra, formando encadeamentos semânticos. A declaração a `ont:Diplomado` classifica o recurso como diplomado, enquanto

[1] Na tradição das Lógicas de Descrição, TBox (Terminological Box) refere-se ao componente terminológico de uma base de conhecimento, enquanto ABox (Assertional Box) corresponde ao componente assertional, formado por afirmações sobre indivíduos e relações entre eles (Baader et al., 2003). Esses conceitos são retomados e detalhados na seção dedicada às ontologias.

`ont:fullName` e `ont:cpf` registram valores literais associados a ele. Já ont:hasBirthPlace estabelece uma relação com outro recurso identificável, representando o local de naturalidade.

Esse encadeamento permite ampliar progressivamente o grafo sem reunir todas as informações em um único registro hierárquico. As relações tornam-se explícitas e podem ser combinadas com outras afirmações que utilizem os mesmos identificadores. Desse modo, a ABox da OntoDiplomaDigital demonstra como dados fictícios do Registro do Diploma podem ser representados como recursos, relações e valores literais, preservando a distinção entre entidades identificáveis e atributos descritivos (Cyganiak et al., 2014; Bizer; Heath; Berners-Lee, 2009).

A conexão entre grafos distribuídos depende da identificação inequívoca dos recursos e das propriedades. As especificações atuais do RDF utilizam *Internationalized Resource Identifiers (IRIs*), que ampliam os *Uniform Resource Identifiers (URIs)* ao admitir um conjunto mais abrangente de caracteres. Nos Dados Ligados, recomenda-se especialmente o emprego de identificadores baseados em HTTP, pois eles podem ser consultados por meio da infraestrutura da Web (Berners-Lee, 2006; Cyganiak et al., 2014).

Identificadores globais permitem diferenciar recursos que apresentam denominações semelhantes e estabelecer referências comuns entre conjuntos de dados distintos. Também possibilitam reutilizar propriedades e conceitos definidos em vocabulários externos, desde que sua semântica seja compatível com o contexto em que serão empregados (Heath; Bizer, 2011; Bizer; Heath; Berners-Lee, 2009).

Berners-Lee (2006) sistematiza os Dados Ligados em quatro princípios no Quadro 11: utilizar URIs para nomear recursos; empregar URIs HTTP para que esses recursos possam ser consultados; fornecer informações úteis por meio de padrões como RDF e SPARQL; e incluir ligações para outras URIs, favorecendo a descoberta de novos recursos.

Quadro 11 - Princípios dos dados ligados

Nº	Princípio	Finalidade
1	Utilizar URIs como nomes para recursos	Assegurar que recursos possam ser identificados de maneira global
2	Utilizar URIs HTTP para possibilitar o acesso	Permitir que os identificadores sejam consultados por meio da infraestrutura da Web

Nº	Princípio	Finalidade
3	Fornecer informações úteis utilizando padrões RDF e SPARQL	Possibilitar a descrição e a consulta estruturada dos recursos
4	Incluir ligações para outras URIs	Favorecer a descoberta de recursos e a conexão entre diferentes conjuntos de dados

Fonte: Adaptado de Berners-Lee (2006).

A disponibilização de dados na Web, sua abertura e sua vinculação semântica correspondem a condições distintas. Dados publicados na Web não são necessariamente abertos; da mesma forma, dados abertos não são automaticamente Dados Ligados, pois a abertura não garante o emprego de identificadores persistentes, representações estruturadas e relações semanticamente explícitas. Os Dados Abertos Ligados resultam da combinação entre condições jurídicas e técnicas de abertura e os princípios de identificação, descrição e conexão de recursos próprios dos Dados Ligados (Cristovão, 2026; Heath; Bizer, 2011)..

O Linked Open Data amplia as possibilidades de integração ao permitir que conjuntos de dados abertos permaneçam sob responsabilidade de suas fontes e sejam conectados por identificadores e relações explícitas. A rede resultante pode ser enriquecida à medida que novos recursos e ligações são publicados, sem exigir a centralização de todos os dados em uma única base (Bizer; Heath; Berners-Lee, 2009).

As informações representadas em RDF podem ser recuperadas por meio da SPARQL Protocol and RDF Query Language (SPARQL). A linguagem permite consultar grafos mediante padrões de triplas nos quais o sujeito, o predicado ou o objeto podem ser substituídos por variáveis. Os resultados correspondem às combinações de recursos e valores que satisfazem os padrões definidos na consulta (W3C, 2013).

Considere-se novamente a tripla extraída da ABox pública da prova de conceito:

```
rec:diplomaCaso1 ont:hasHolder rec:diplomadoCaso1 .
```

Para recuperar o recurso relacionado a rec:diplomaCaso1 pela propriedade `ont:hasHolder`, o objeto pode ser substituído pela variável `?titular` em uma consulta SPARQL representada no Quadro 12

Quadro 12 - Representação de consulta SPARQL

```
PREFIX rec: <https://w3id.org/ufes/ontodiplomadigital/registro#>
PREFIX ont: <https://w3id.org/ufes/ontodiplomadigital#>

SELECT ?titular
WHERE {
    rec:diplomaCaso1 ont:hasHolder ?titular .
}
```

Fonte: Elaborado pelo autor (2026), com base em W3C (2013) e na ABox pública da prova de conceito da OntoDiplomaDigital.

A cláusula PREFIX associa os prefixos `rec:` e `ont:` aos endereços utilizados na identificação dos recursos e das propriedades da *OntoDiplomaDigital*. A cláusula SELECT determina a variável que será apresentada no resultado, enquanto WHERE contém o padrão que deverá ser localizado no grafo. Como o sujeito `rec:diplomaCaso1` e o predicado `ont:hasHolder` permanecem definidos, a consulta recupera o recurso que ocupa a posição de objeto nas triplas correspondentes. Nesse exemplo, o resultado esperado é rec:diplomadoCaso1, que representa o titular fictício do diploma (W3C, 2013).

O Quadro 13 sintetiza a correspondência entre a afirmação RDF e o padrão utilizado na consulta SPARQL, evidenciando que a recuperação se baseia na própria estrutura relacional do grafo (W3C, 2013).

Quadro 13 - Correspondência entre tripla RDF e consulta SPARQL

Representação	Exemplo
Tripla RDF	rec:diplomaCaso1 ont:hasHolder rec:diplomadoCaso1 .
Padrão consultado	rec:diplomaCaso1 ont:hasHolder ?titular .
Variável	?titular
Resultado esperado	rec:diplomadoCaso1, recurso que representa o titular fictício do diploma.

Fonte: Elaborado pelo autor com base em W3C (2013).

Além de padrões simples, a SPARQL permite combinar diferentes padrões de triplas, aplicar filtros e agregações e consultar fontes distribuídas. Essas operações possibilitam recuperar relações que atravessam diferentes partes do grafo ou diferentes conjuntos de dados, desde que os recursos e as propriedades envolvidos sejam identificados de forma compatível (W3C, 2013; Cyganiak et al., 2014).

RDF, Dados Ligados e SPARQL desempenham funções complementares: o RDF fornece o modelo de representação; os princípios dos Dados Ligados orientam a identificação, a publicação e a conexão dos recursos; e a SPARQL oferece

mecanismos para consultar as afirmações presentes nos grafos. A combinação dessas funções cria condições técnicas para a integração e a exploração de dados distribuídos (Berners-Lee, 2006; Cyganiak et al., 2014; W3C, 2013).

Os Dados Ligados também podem contribuir para a implementação dos princípios FAIR, especialmente por meio de identificadores persistentes, metadados estruturados, vocabulários compartilhados e ligações entre recursos. Essa aproximação não torna os dois conceitos equivalentes: os princípios FAIR estabelecem condições para que dados sejam encontráveis, acessíveis, interoperáveis e reutilizáveis, enquanto os Dados Ligados oferecem práticas e tecnologias que podem favorecer parte dessas condições (Wilkinson et al., 2016; Berners-Lee, 2006).

A adoção de RDF, identificadores globais e padrões de consulta não assegura, isoladamente, que diferentes sistemas atribuam o mesmo significado aos dados. Zeng e Qin (2022) distinguem a interoperabilidade sintática, relacionada às estruturas e aos formatos utilizados no intercâmbio, da interoperabilidade semântica, dependente de definições, vocabulários e modelos conceituais capazes de preservar o significado. Laufer (2015) também ressalta que a integração semântica requer descrições que tornem os recursos compreensíveis além de sua organização técnica.

Para atender a essa dimensão semântica, RDF, RDF Schema (RDFS), OWL e SKOS cumprem funções distintas. RDF oferece o modelo de triplas; RDFS permite declarar classes, propriedades e hierarquias; OWL acrescenta recursos para a formalização de axiomas e restrições; e SKOS se destina à representação de conceitos e relações características dos sistemas de organização do conhecimento (Ramalho, 2015; Zeng; Qin, 2022).

No campo da Ciência da Informação, essas tecnologias aproximam a modelagem de dados e o enriquecimento semântico das práticas de Organização e Representação do Conhecimento. O desenvolvimento e a gestão de vocabulários, a aplicação de ontologias e a utilização de identificadores, RDF e SPARQL ampliam os espaços de atuação do profissional da informação em ambientes digitais interligados (Pereira; Luz; Santarém Segundo, 2019).

A interoperabilidade permanece, contudo, um processo que envolve dimensões técnicas, conceituais e institucionais. A existência de um grafo formalmente válido não elimina divergências terminológicas, políticas ou organizacionais entre os contextos de produção e utilização dos dados, tornando necessárias ações de documentação,

harmonização e mediação institucional (Lemos; Martins; Carmo, 2022; Zeng; Qin, 2022).

A Web Semântica e os Dados Ligados fornecem, portanto, uma base para identificar recursos, representar relações em grafos, estabelecer conexões entre fontes e recuperar informações por meio de consultas. A explicitação mais rigorosa dos conceitos, das propriedades e das restrições requer a utilização de ontologias (Berners-Lee; Hendler; Lassila, 2001; Ramalho, 2015).

3.6 Ontologia Operacional

O conceito de ontologia, quando aplicado à representação do conhecimento, está associado à explicitação de uma conceituação. Gruber (1993) define ontologia como uma especificação explícita de uma conceituação, entendendo a conceituação como uma visão abstrata e simplificada do domínio que se pretende representar.

Borst (1997) amplia essa formulação ao definir ontologia como uma especificação formal de uma conceituação compartilhada. O caráter formal indica que os conceitos e as relações devem ser expressos de maneira processável por sistemas computacionais, enquanto o caráter compartilhado remete a uma compreensão aceita por determinado grupo acerca do domínio representado. A ontologia não corresponde, portanto, a uma simples lista de termos, mas a uma estrutura conceitual na qual entidades, propriedades e relações recebem significados definidos (Borst, 1997).

O caráter explícito da ontologia é especialmente relevante quando diferentes pessoas, instituições ou sistemas precisam compartilhar uma compreensão do domínio. A coincidência entre os termos utilizados não assegura que os significados atribuídos sejam os mesmos, pois expressões semelhantes podem representar entidades ou relações conceitualmente distintas. A construção ontológica procura reduzir essa ambiguidade ao registrar as categorias reconhecidas e as condições segundo as quais se relacionam (Guizzardi, 2020).

Noy e McGuinness (2001) identificam entre as finalidades das ontologias o compartilhamento de uma compreensão comum da estrutura da informação, a possibilidade de reutilizar conhecimento de domínio, a explicitação de pressupostos e a separação entre o conhecimento conceitual e a lógica das aplicações. Os autores também observam que o desenvolvimento de uma ontologia é iterativo, pois as

definições iniciais podem ser revistas à medida que o domínio e os requisitos de representação são analisados.

As ontologias podem assumir diferentes funções conforme seus objetivos e níveis de aproximação com a implementação computacional. Falbo (2014) distingue as ontologias de referência, voltadas à representação conceitual do domínio, das ontologias operacionais, elaboradas para utilização efetiva por sistemas computacionais.

Uma ontologia de referência procura representar o domínio com clareza e fidelidade conceitual, mantendo independência em relação a tecnologias ou aplicações específicas. Sua função principal é favorecer a compreensão e a comunicação entre especialistas, servindo como modelo de referência para o exame das entidades, propriedades e relações consideradas relevantes (Falbo, 2014).

A ontologia operacional aproxima essa representação das condições necessárias ao processamento computacional. Nela, o conteúdo conceitual é expresso por meio de linguagens e estruturas formais que possibilitam representar classes, propriedades, relações e restrições, além de apoiar operações de consulta e integração de dados (Falbo, 2014; Ramalho, 2015).

Essa operacionalização não deve ser entendida como uma simples transcrição do modelo conceitual para determinada linguagem. As decisões relativas à tecnologia, à expressividade e às condições de processamento podem exigir adaptações, mas essas mudanças precisam preservar os compromissos conceituais que fundamentam a representação. O desafio consiste em compatibilizar fidelidade semântica e viabilidade computacional sem alterar indevidamente o significado das entidades e relações do domínio (Falbo, 2014).

No âmbito das Lógicas de Descrição, a estrutura de uma base de conhecimento pode ser compreendida pela distinção entre Terminological Box (TBox) e Assertional Box (ABox). A TBox reúne o componente terminológico, no qual são declarados conceitos, relações e axiomas gerais; a ABox contém as asserções referentes a indivíduos e às relações estabelecidas entre eles (Baader et al., 2003).

A TBox define a estrutura que permite a descrição dos recursos, enquanto a ABox registra ocorrências específicas representadas de acordo com essa estrutura. Essa distinção possibilita a separação do conhecimento geral sobre o domínio das afirmações relativas a entidades individualizadas, embora ambos os componentes

constituam a base de conhecimento utilizada para representação e raciocínio (Baader et al., 2003).).

Em uma ontologia operacional, classes e propriedades integram predominantemente a camada terminológica, enquanto os indivíduos e suas relações compõem a camada de asserções. Essa organização permite que diferentes conjuntos de indivíduos sejam descritos segundo um mesmo modelo e que as afirmações registradas sejam consultadas ou avaliadas à luz das definições estabelecidas na TBox (Baader et al., 2003; Falbo, 2014).

A separação entre classes, indivíduos e conceitos empregados em vocabulários controlados também precisa ser preservada. Classes representam categorias ontológicas; indivíduos correspondem a ocorrências particulares dessas categorias; e conceitos SKOS podem ser utilizados para organizar valores classificatórios e relações terminológicas sem que sejam automaticamente tratados como classes OWL (Ramalho, 2015).

A contribuição das ontologias para a interoperabilidade semântica reside na possibilidade de tornar explícito o significado das entidades e das relações representadas. Enquanto a compatibilidade estrutural permite que os dados sejam intercambiados, a representação ontológica oferece condições para que diferentes sistemas reconheçam os conceitos aos quais esses dados se referem e as relações estabelecidas entre eles (Guizzardi, 2020).

A formalização, contudo, não elimina automaticamente as diferenças de interpretação. Uma ontologia expressa determinados compromissos conceituais e precisa ser documentada, avaliada e compartilhada para que esses compromissos sejam compreendidos pelos usuários e sistemas envolvidos. A interoperabilidade depende tanto da estrutura formal quanto da adequação conceitual e da aceitação dos significados representados (Guizzardi, 2020; Noy; McGuinness, 2001).

As ontologias operacionais articulam, portanto, análise conceitual e representação computacional. Sua construção exige que o conhecimento do domínio seja adquirido, organizado, formalizado, avaliado e documentado segundo procedimentos capazes de manter a coerência entre os objetivos, o modelo conceitual e sua expressão formal (Falbo, 2014; Noy; McGuinness, 2001).

3.7 Metodologia para Criação de Ontologias

A construção de ontologias requer procedimentos sistemáticos que orientem a passagem do conhecimento sobre determinado domínio para uma representação conceitual formalizada. Esse processo envolve decisões relacionadas à delimitação do domínio, à identificação de conceitos, à definição de relações, à organização hierárquica, à formalização, à avaliação, à documentação e à disponibilização da ontologia. A adoção de uma metodologia permite organizar essas atividades, registrar as decisões tomadas e manter a coerência entre os objetivos estabelecidos e os artefatos produzidos.

As metodologias de construção de ontologias apresentam diferentes níveis de detalhamento, enfoques e formas de organização do processo. Entre as propostas consolidadas encontram-se TOVE, Methontology, Ontology Development 101 e NeOn, desenvolvidas principalmente nos campos da Engenharia do Conhecimento e da Ciência da Computação. Embora compartilhem atividades fundamentais, essas metodologias diferem quanto à sequência das etapas, aos artefatos previstos, ao tratamento do reúso e ao grau de detalhamento atribuído à conceitualização, à formalização e à avaliação (Mendonça, 2015).

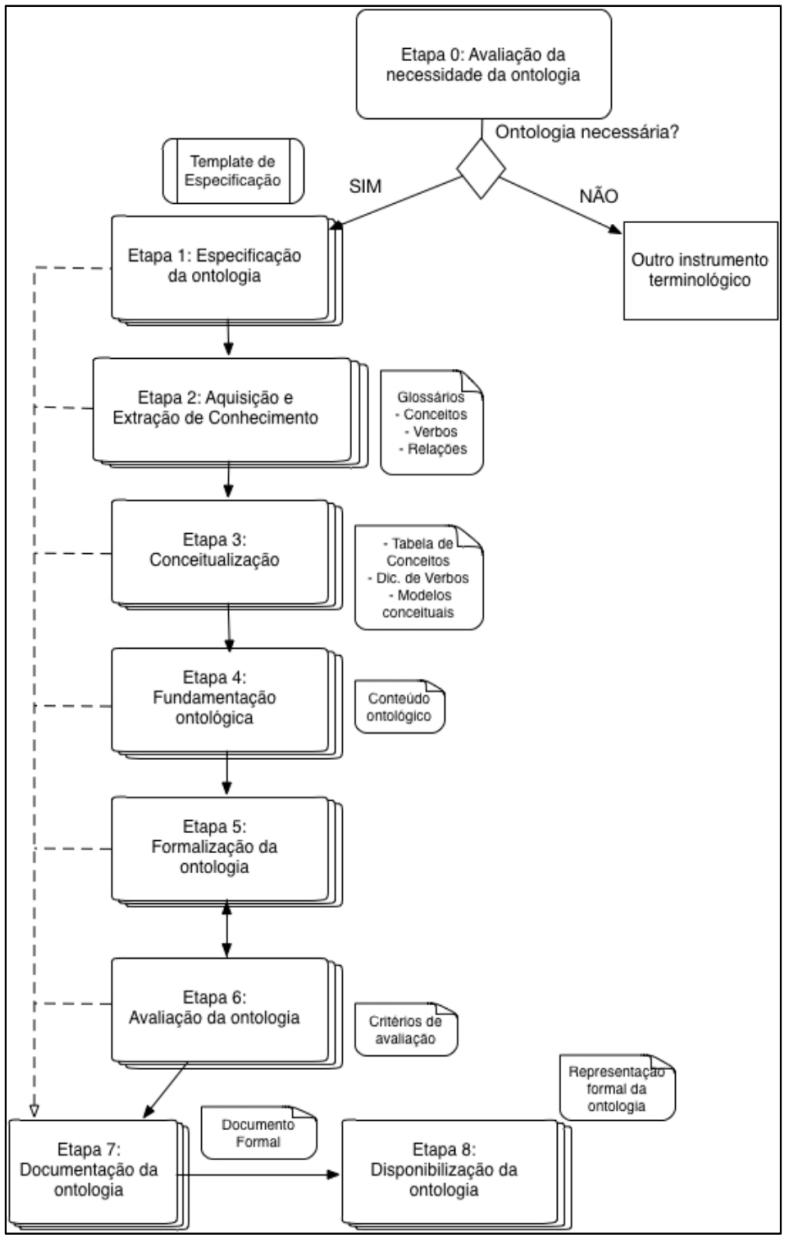
Não foi encontrada na literatura uma metodologia universalmente adequada a todos os projetos ontológicos. A escolha depende da finalidade da ontologia, das características do domínio, dos usuários previstos, das fontes disponíveis e do nível de formalização necessário. Em domínios informacionais, essa escolha também deve considerar a proximidade entre a construção de ontologias e atividades tradicionalmente relacionadas à Organização e Representação do Conhecimento, como análise de domínio, controle terminológico, elaboração de definições, identificação de relações conceituais e produção de instrumentos de representação.

Voltada a aproximar esse processo das práticas da Ciência da Informação, a OntoForInfoScience foi proposta por Mendonça (2015) como uma metodologia voltada à construção de ontologias na perspectiva da Ciência da Informação. Sua formulação procura tornar o processo de desenvolvimento ontológico mais compreensível para cientistas da informação, especialistas de domínio e profissionais ligados à Organização do Conhecimento. Para isso, detalha as atividades e os artefatos envolvidos na passagem entre a análise do domínio, a conceitualização e a representação formal.

A OntoForInfoScience incorpora contribuições de metodologias anteriores, entre elas Methontology, NeOn e Ontology Development 101, reorganizando-as em um processo orientado às necessidades da Ciência da Informação. A metodologia reconhece a ontologia não apenas como um artefato computacional, mas também como um sistema de representação do conhecimento cuja construção exige análise conceitual, controle terminológico, explicitação das relações e documentação das decisões adotadas (Mendonça, 2015).

O processo é composto por uma pré-etapa, também denominada etapa zero, e oito etapas subsequentes. Embora apresentadas segundo uma organização progressiva, essas etapas não devem ser compreendidas como uma sequência inteiramente rígida. A avaliação pode exigir o retorno à formalização ou à conceitualização, enquanto a documentação acompanha todo o processo. A Figura 3 representa essa estrutura e os principais artefatos associados às etapas.

Figura 3 - Etapas da metodologia OntoForInfoScience



Fonte: Mendonça (2015).

A **pré-etapa**, ou etapa zero, corresponde à avaliação da necessidade de construção da ontologia. Nesse momento, verifica-se se o problema identificado requer uma representação ontológica ou se pode ser atendido adequadamente por outros instrumentos, como glossários, taxonomias, tesauros, vocabulários controlados, esquemas de metadados ou modelos de dados. Essa avaliação evita a adoção de uma solução mais complexa do que aquela efetivamente necessária ao domínio e aos objetivos estabelecidos.

A **Etapa 1**, denominada especificação da ontologia, define os requisitos que orientarão seu desenvolvimento. Nela são estabelecidos aspectos como propósito,

domínio, escopo, usuários previstos, cenários de utilização, fontes de conhecimento e questões de competência. As questões de competência expressam perguntas que a ontologia deverá ser capaz de responder e funcionam como referência para a delimitação do conteúdo e para sua posterior avaliação. O preenchimento dessas informações resulta em um documento de especificação que orienta as etapas seguintes.

A Etapa 2 compreende a aquisição e a extração do conhecimento. Seu objetivo é reunir e examinar as fontes relevantes para a compreensão do domínio, incluindo literatura especializada, normas, documentos institucionais, bases de dados, sistemas existentes e, quando necessário, consultas a especialistas. A análise dessas fontes permite identificar os termos, conceitos e relações utilizados no domínio. Como resultado, a etapa produz artefatos terminológicos, como glossários de conceitos, verbos e relações, que subsidiam a etapa posterior de conceitualização.

A **Etapa 3** corresponde à conceitualização do domínio. Nessa etapa, o conhecimento adquirido é organizado em estruturas conceituais que tornam explícitos os conceitos, suas propriedades, relações, hierarquias e restrições. Entre os artefatos previstos encontram-se tabelas de conceitos e propriedades, dicionários de verbos e modelos conceituais gráficos. Esses instrumentos apoiam a análise do domínio e permitem avaliar a coerência da representação antes de sua codificação em uma linguagem formal.

A **Etapa 4** trata da fundamentação ontológica. Nela são examinadas as categorias e distinções necessárias para representar adequadamente a natureza das entidades do domínio. Quando pertinente aos objetivos do projeto, essa etapa pode incluir a análise ou a utilização de ontologias de fundamentação. Seu propósito é reduzir inconsistências conceituais e apoiar distinções como aquelas existentes entre classes e indivíduos, objetos e eventos, agentes e papéis ou partes e totalidades.

A **Etapa 5** consiste na formalização da ontologia. O modelo conceitual é convertido em uma representação formal e computacionalmente processável, mediante a definição de classes, propriedades, relações, axiomas e restrições. Essa passagem exige que as decisões conceituais elaboradas nas etapas anteriores sejam expressas de acordo com as características da linguagem de representação selecionada. A formalização não se limita à transcrição dos termos identificados, pois envolve decisões sobre a estrutura lógica e o comportamento semântico do modelo.

A **Etapa 6** corresponde à avaliação da ontologia. Nessa etapa, verifica-se se a representação atende aos objetivos e aos requisitos definidos na especificação, se mantém coerência conceitual e se apresenta consistência formal. A avaliação pode envolver a análise das questões de competência, a revisão por especialistas, a verificação dos conceitos e relações e a utilização de ferramentas computacionais. A relação entre formalização e avaliação é iterativa: problemas identificados durante a avaliação podem exigir alterações no modelo conceitual ou na representação formal.

A **Etapa 7** compreende a documentação da ontologia. Embora apareça como uma etapa específica, a documentação deve ser desenvolvida ao longo de todo o processo. Ela registra as fontes consultadas, as definições, as decisões de modelagem, os critérios utilizados, as mudanças realizadas e os artefatos produzidos. Esse registro favorece a rastreabilidade das decisões, a compreensão do modelo, sua manutenção e seu reúso por outros pesquisadores, profissionais ou sistemas.

A **Etapa 8** corresponde à disponibilização da ontologia. Essa etapa envolve a publicação dos artefatos produzidos em condições que permitam seu acesso, consulta, avaliação e reúso. A disponibilização pode compreender a representação formal da ontologia, sua documentação, os modelos conceituais, as definições terminológicas e outras informações necessárias à compreensão de seu escopo e de suas condições de utilização.

A OntoForInfoScience articula, portanto, atividades de análise de domínio, organização conceitual, formalização, avaliação e documentação em um processo direcionado à construção de ontologias. Sua contribuição para a Ciência da Informação reside especialmente na aproximação entre os procedimentos da engenharia ontológica e as práticas de Organização e Representação do Conhecimento. A aplicação dessa metodologia em uma investigação específica, assim como as adaptações decorrentes das características do domínio estudado, deve ser explicitada nos procedimentos metodológicos da pesquisa.

3.8 Padrões e Esquemas de Metadados para Interoperabilidade

A interoperabilidade corresponde à capacidade de sistemas distintos trocarem informações e utilizarem os conteúdos recebidos de maneira compatível com suas estruturas e finalidades. Essa capacidade não depende apenas da transmissão técnica dos dados, mas também da existência de acordos sobre sua organização e

seu significado, de modo que as informações compartilhadas possam ser interpretadas adequadamente nos diferentes ambientes em que circulam (Riley, 2017; Laufer, 2015).

Os padrões de metadados contribuem para esses acordos ao estabelecer elementos, definições e regras comuns para a descrição de recursos. Eles orientam quais características podem ser registradas, como os elementos devem ser denominados e quais condições devem ser observadas em sua utilização. Ao reduzir variações na produção das descrições, esses padrões favorecem a consistência dos registros e seu intercâmbio entre instituições e sistemas informacionais (Baca, 2016; Woodley, 2016).

Os esquemas de metadados constituem formas estruturadas de expressar os componentes de determinado padrão ou modelo descritivo. Um esquema pode definir elementos, atributos, tipos de valores, níveis de obrigatoriedade, possibilidades de repetição e relações entre os componentes da descrição. Dessa maneira, oferece uma estrutura compartilhada para organizar os metadados e verificar se os registros observam as regras estabelecidas para sua produção (Baca, 2016; Zeng; Qin, 2022).

Embora relacionados, padrão e esquema não são conceitos equivalentes. O padrão compreende o conjunto de princípios, elementos e convenções acordados para determinada finalidade, enquanto o esquema expressa esses componentes em uma estrutura organizada e, em alguns casos, processável computacionalmente. A aplicação de um padrão pode envolver um ou mais esquemas, além de orientações terminológicas, regras de conteúdo e documentos complementares necessários à sua interpretação (Baca, 2016; Woodley, 2016).

A adoção de estruturas compartilhadas favorece principalmente a interoperabilidade sintática, pois permite que diferentes sistemas reconheçam a organização, os elementos e os tipos de valores presentes nos dados intercambiados. A interoperabilidade semântica exige, adicionalmente, que os sistemas preservem e interpretem de maneira compatível o significado dos elementos e das relações representadas, o que depende de definições explícitas, modelos conceituais, vocabulários e ontologias (Laufer, 2015; Zeng; Qin, 2022).

A distinção entre os níveis sintático e semântico evidencia que a conformidade com um mesmo esquema não implica necessariamente em equivalência interpretativa. Dois sistemas podem gerar registros estruturalmente válidos e, ainda assim, atribuir significados distintos a elementos de nomes semelhantes. A

interoperabilidade semântica, portanto, requer correspondências conceituais que considerem o significado, o contexto e a função desempenhada por cada elemento na representação do domínio.. (Laufer, 2015; Zeng; Qin, 2022).

No ambiente da Web Semântica, linguagens formais oferecem mecanismos para representar dados e explicitar parte dessas estruturas conceituais. RDF fornece um modelo baseado em triplas; RDFS permite descrever classes, propriedades, domínios, alcances e hierarquias; e OWL oferece recursos para formalizar ontologias, axiomas e restrições. Essas linguagens estabelecem mecanismos representacionais, mas não determinam isoladamente quais conceitos e relações são adequados a determinado domínio (Ramalho, 2015; Zeng; Qin, 2022).

Os vocabulários e as ontologias, por sua vez, disponibilizam termos semanticamente definidos que podem ser empregados na descrição e na integração de recursos. DCMI Metadata Terms, SKOS, FOAF, PROV-O, DCAT, Schema.org, LOCN e vCard RDF contemplam aspectos como descrição de recursos informacionais, organização de conceitos, representação de pessoas e organizações, proveniência, conjuntos de dados, localização e informações de contato (Baca, 2016; Woodley, 2016).

Linguagens formais, esquemas de metadados, vocabulários e ontologias desempenham, portanto, funções relacionadas, mas distintas. As linguagens fornecem mecanismos para estruturar a representação; os esquemas organizam elementos e regras descritivas; os vocabulários disponibilizam termos definidos para utilização compartilhada; e as ontologias explicitam conceitos, propriedades, relações e restrições de determinado domínio (Baca, 2016; Zeng; Qin, 2022).

O reúso de elementos provenientes de vocabulários e ontologias consolidados pode favorecer a interoperabilidade porque permite que diferentes modelos empreguem recursos previamente definidos e documentados. Esse reúso também pode reduzir a criação de termos locais desnecessários e ampliar as possibilidades de conexão com outros conjuntos de dados, desde que o significado original do elemento reutilizado seja preservado (Woodley, 2016; Zeng; Qin, 2022).

A semelhança entre denominações não constitui evidência suficiente para o reúso. A análise deve considerar a definição do elemento, a finalidade para a qual foi criado, os tipos de recursos aos quais se aplica e as relações que estabelece no vocabulário de origem. A utilização de um termo externo em contexto incompatível

pode introduzir ambiguidades e produzir inferências ou interpretações diferentes daquelas pretendidas pelo modelo local (Woodley, 2016; Zeng; Qin, 2022).

Quando existe correspondência conceitual suficiente, o elemento externo pode ser reutilizado diretamente. Quando o domínio exige uma relação mais específica, pode ser necessário criar um elemento local e estabelecer uma relação documentada com o recurso externo. Se não houver correspondência semanticamente segura, a definição de um elemento próprio é preferível à afirmação de uma equivalência inadequada (Woodley, 2016; Zeng; Qin, 2022).

A combinação de termos externos e locais não representa, por si mesma, uma limitação à interoperabilidade. Domínios especializados frequentemente exigem elementos próprios para representar suas particularidades, enquanto os conceitos mais gerais podem ser expressos por recursos compartilhados. A interoperabilidade depende da coerência dessa articulação e da documentação das decisões de reúso, especialização e alinhamento semântico (Baca, 2016; Zeng; Qin, 2022).

A descoberta de vocabulários potencialmente reutilizáveis pode ser apoiada por repositórios especializados. O Linked Open Vocabularies (LOV) reúne informações sobre vocabulários publicados na Web e permite examinar sua documentação, seu histórico de versões, sua estabilidade e suas relações com outros recursos semânticos (Vandenbussche et al., 2016).

Os repositórios auxiliam na localização e na avaliação preliminar dos vocabulários, mas não substituem a consulta às especificações oficiais. Antes de reutilizar uma classe, propriedade ou conceito, é necessário verificar sua definição, seu domínio, seu alcance, suas restrições e as recomendações estabelecidas pela organização responsável, pois a disponibilidade pública de um termo não assegura sua adequação a todos os contextos de representação (Woodley, 2016; Zeng; Qin, 2022).

O Quadro 14 organiza as categorias de recursos discutidas nesta seção e explicita as diferentes funções que desempenham na construção da interoperabilidade.

Quadro 14 - Categorias de recursos para representação e interoperabilidade

Categoria	Função conceitual	Exemplos
Padrões e termos de metadados	Estabelecer elementos e definições compartilhados para a descrição de recursos	DCMI Metadata Terms
Modelos de representação	Definir como os dados podem ser expressos e relacionados	RDF
Linguagens de descrição de esquemas	Representar classes, propriedades, domínios, alcances e hierarquias	RDFS
Linguagens de ontologias	Formalizar classes, propriedades, axiomas e restrições	OWL
Modelos para sistemas de organização do conhecimento	Representar conceitos, rótulos e relações semânticas	SKOS
Vocabulários e ontologias reutilizáveis	Disponibilizar termos para pessoas, organizações, proveniência, conjuntos de dados, localização e contato	FOAF, PROV-O, DCAT, Schema.org, LOCN e vCard RDF
Serviços de descoberta	Apoiar a localização e a avaliação preliminar de vocabulários publicados	LOV

Fonte: Elaborado pelo autor com base em Baca (2016), Woodley (2016), Vandenbussche et al. (2016) e Zeng e Qin (2022).

3.8.1 XML e XSD enquanto Padrões de Dados

A Extensible Markup Language (XML) é uma linguagem de marcação destinada à estruturação e ao intercâmbio de dados. Por ser extensível, permite definir elementos e atributos de acordo com as necessidades de determinado contexto, organizando-os em uma estrutura hierárquica que pode ser processada por diferentes aplicações computacionais (W3C, 2008).

Um documento XML é constituído por elementos organizados a partir de um único elemento raiz. Esses elementos podem conter textos, atributos e outros elementos, formando uma estrutura hierárquica semelhante a uma árvore. A relação entre os componentes é determinada por sua posição e seu nível de aninhamento no documento (W3C, 2008).

Para ser considerado bem-formado, o documento XML deve observar as regras sintáticas da linguagem, como possuir um único elemento raiz, apresentar correspondência entre as marcações de abertura e fechamento, preservar o

aninhamento correto dos elementos e utilizar adequadamente os valores dos atributos. A boa formação indica que o documento respeita a sintaxe XML, mas não demonstra sua conformidade com um modelo documental específico (W3C, 2008).

A Extensible Markup Language não estabelece previamente um conjunto fixo de elementos para todos os documentos. Os nomes, as estruturas e as regras aplicáveis a determinado conjunto de arquivos podem ser definidos separadamente por meio de um esquema. Entre as linguagens destinadas a essa finalidade encontra-se a XML Schema Definition Language (XSD), recomendada pelo W3C para descrever e validar estruturas documentais expressas em XML (W3C, 2008; W3C, 2012).

Um esquema XSD define quais elementos e atributos podem aparecer em determinada classe de documentos XML, como eles devem ser organizados e quais tipos de valores podem conter. O esquema também permite estabelecer regras relacionadas à ordem, à quantidade de ocorrências e às restrições aplicáveis aos conteúdos registrados nos documentos (W3C, 2012).

Documento XML e esquema XSD desempenham, portanto, funções distintas. O documento XML contém os valores correspondentes a uma ocorrência concreta, enquanto o XSD descreve a estrutura e as regras que essa categoria de documento deve observar. Um mesmo esquema pode ser utilizado para validar diferentes documentos XML produzidos segundo o modelo nele especificado (W3C, 2012).

O XSD diferencia tipos simples e tipos complexos. Os tipos simples representam valores que não contêm elementos ou atributos internos, como textos, números, datas e códigos. Os tipos complexos permitem reunir elementos e atributos em estruturas compostas, possibilitando representar registros formados por diferentes componentes relacionados (W3C, 2012).

Os tipos simples podem ser derivados de tipos predefinidos e receber restrições adicionais, denominadas facetas. Essas restrições podem delimitar comprimentos mínimos e máximos, intervalos numéricos, conjuntos de valores permitidos e padrões lexicais. O componente xs:pattern, por exemplo, permite definir uma expressão que o valor deverá observar para ser considerado válido (W3C, 2012).

A organização interna dos tipos complexos pode ser estabelecida por compositores. O componente xs:sequence determina que os elementos sejam apresentados na ordem declarada; xs:choice permite selecionar uma alternativa entre os elementos previstos; e xs:all admite a ocorrência dos componentes sem exigir a

mesma ordenação sequencial. A escolha do compositor determina como os elementos poderão ser combinados no documento XML (W3C, 2012).

As cardinalidades são expressas principalmente pelos atributos minOccurs e maxOccurs, que estabelecem as quantidades mínima e máxima de ocorrências de um elemento. Para os atributos XML, o componente use pode indicar se sua presença é obrigatória, opcional ou proibida. Esses mecanismos permitem controlar a ocorrência dos componentes e verificar se o documento apresenta a estrutura definida pelo esquema (W3C, 2012).

A validação consiste em verificar um documento XML em relação às declarações e restrições estabelecidas no XSD. Nesse processo, são examinados os elementos e atributos utilizados, sua ordem e cardinalidade, os tipos dos valores e as demais condições formalizadas no esquema. Um documento pode ser bem-formado segundo as regras gerais do XML e, ser inválido em relação ao XSD aplicável (W3C, 2008; W3C, 2012).

O emprego de um esquema comum favorece a interoperabilidade sintática porque permite que diferentes sistemas produzam e recebam documentos estruturados segundo a mesma gramática. A padronização dos nomes, tipos, posições e ocorrências dos componentes reduz divergências estruturais e possibilita a validação automatizada dos arquivos intercambiados (W3C, 2012; Zeng; Qin, 2022).

A conformidade sintática não assegura, entretanto, a interoperabilidade semântica. Dois sistemas podem processar documentos válidos segundo o mesmo XSD e ainda interpretar de maneira diferente os conceitos representados. O esquema controla a forma e parte das restrições aplicáveis aos dados, mas não formaliza necessariamente o significado das entidades, dos papéis e das relações do domínio em uma estrutura semântica compartilhada (Laufer, 2015; Zeng; Qin, 2022).

Essa limitação não significa que os elementos de um XSD sejam desprovidos de significado. Seus nomes, tipos, anotações, hierarquias e posições estruturais comunicam informações relevantes sobre o domínio. Contudo, parte desse significado permanece dependente da documentação técnica, das normas, das convenções institucionais e da interpretação atribuída pelos responsáveis pelos sistemas envolvidos (W3C, 2012; Laufer, 2015).

No Diploma Digital Brasileiro, o XML constitui o formato empregado para estruturar os documentos acadêmicos digitais, enquanto os esquemas XSD estabelecem os elementos, atributos, tipos, sequências, cardinalidades e restrições

que devem ser observados em sua produção. A regulamentação técnica do Ministério da Educação define a utilização dos leiautes e esquemas aplicáveis à emissão e ao registro dos diplomas de graduação em meio digital (Brasil, 2019c; Brasil, 2020b).

A adoção dos esquemas oficiais permite que documentos produzidos por diferentes Instituições de Ensino Superior sejam submetidos a regras comuns de estruturação e validação. Essa padronização favorece o intercâmbio técnico dos arquivos e a verificação de sua conformidade com os leiautes estabelecidos, sem eliminar a necessidade de acordos sobre o significado e a utilização institucional das informações registradas (Brasil, 2020b; Laufer, 2015).

O Quadro 15 apresenta um exemplo simplificado de componentes utilizados na definição de um esquema XSD, incluindo `xs:schema`, `xs:element`, `xs:attribute`, `xs:simpleType`, `xs:complexType`, `xs:sequence` e `xs:pattern`. Esses componentes exemplificam como elementos, atributos, tipos, estruturas compostas e restrições podem ser formalmente declarados em um esquema (W3C, 2012).

Quadro 15 - Trecho de código XSD/XML

```xml
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <!-- Declaração de um tipo simples customizado (simpleType) -->
  <xs:simpleType name="CodigoTipo">
    <xs:restriction base="xs:string">
      <!-- Exemplo de restrição: exatamente 6 caracteres alfanuméricos -->
      <xs:pattern value="[A-Za-z0-9]{6}"/>
    </xs:restriction>
  </xs:simpleType>

  <!-- Declaração de um tipo complexo para endereço -->
  <xs:complexType name="EnderecoType">
    <xs:sequence>
      <xs:element name="rua"     type="xs:string"/>
      <xs:element name="numero"  type="xs:integer"/>
      <xs:element name="cidade"  type="xs:string"/>
      <xs:element name="estado"  type="xs:string"/>
      <xs:element name="cep"     type="xs:string"/>
    </xs:sequence>
  </xs:complexType>

  <!-- Declaração de um elemento global do tipo complexo -->
  <xs:element name="pessoa">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="nome"       type="xs:string"/>
        <xs:element name="dataNasc"   type="xs:date"/>
        <xs:element name="endereco"   type="EnderecoType"/>
      </xs:sequence>
      <!-- Atributo associado à pessoa -->
      <xs:attribute name="id" type="CodigoTipo" use="required"/>
    </xs:complexType>
  </xs:element>

</xs:schema>
```

Fonte: Brasil, Ministério da Educação ([s. d.]).

No exemplo, `xs:schema` constitui o elemento raiz do esquema e declara o namespace da linguagem XML Schema. O componente `xs:simpleType` define um tipo simples sujeito a restrições, enquanto `xs:complexType` organiza uma estrutura composta por outros elementos ou atributos. O componente `xs:sequence` determina a ordem dos elementos internos, e as declarações `xs:element` e `xs:attribute` especificam os componentes admitidos no documento XML (W3C, 2012).

XML e XSD constituem mecanismos consolidados para a estruturação, intercâmbio e validação de documentos digitais. Sua contribuição reside na definição de uma sintaxe compartilhada e de regras verificáveis para a organização dos dados, enquanto a explicitação das relações conceituais demanda instrumentos complementares, tais como vocabulários e ontologias (W3C, 2012; Laufer, 2015; Zeng; Qin, 2022).

3.9 Diplomas Digitais de IES

A transformação digital dos serviços públicos brasileiros buscou ampliar o acesso da sociedade à administração pública por meio de canais digitais, com simplificação de procedimentos, integração de plataformas, segurança da informação e melhoria da qualidade dos serviços. Na administração pública federal direta, autárquica e fundacional, esse processo foi apoiado por infraestruturas e soluções compartilhadas, como identificação digital, avaliação de serviços e mecanismos de interoperabilidade. A trajetória iniciada com o levantamento dos serviços públicos federais consolidou uma política orientada à centralização do acesso no Gov.br e à reutilização de componentes tecnológicos pelos órgãos públicos.

No ensino superior, o Diploma Digital integrou esse movimento ao converter os processos de emissão e registro de diplomas em um fluxo documental nato-digital. Instituído pela Portaria MEC nº 330/2018 e regulamentado pela Portaria MEC nº 554/2019, o modelo foi detalhado por atos normativos e especificações técnicas que definiram requisitos de emissão, registro, assinatura, representação visual e preservação. Sua implantação não se limitou à substituição do papel por um arquivo eletrônico, pois exigiu das Instituições de Ensino Superior a revisão de procedimentos acadêmicos, cadastros, responsabilidades institucionais e sistemas de informação (Brasil, 2018; Brasil, 2019c).

A adoção ocorreu de forma progressiva. Desde 2021, as instituições puderam emitir o diploma de graduação como documento nato-digital, com a mesma validade jurídica atribuída ao diploma físico. A Portaria MEC nº 70/2025 atualizou a regulamentação e estabeleceu que as IES pertencentes ao Sistema Federal de Ensino passassem a emitir diplomas de graduação exclusivamente em formato digital a partir de 1º de julho de 2025. Esse percurso normativo consolidou o Diploma Digital como documento oficial e deslocou para as instituições a responsabilidade de adequar seus processos acadêmicos e tecnológicos às especificações definidas pelo MEC (Brasil, 2025).

O documento oficial é produzido em XML e reúne os dados acadêmicos, administrativos e técnicos associados ao diplomado, ao curso, às instituições emissora e registradora e ao registro do diploma. Sua validade jurídica está vinculada à assinatura digital com certificação da Infraestrutura de Chaves Públicas Brasileira e ao carimbo de tempo, que registra evidências sobre a existência e a integridade do

documento em determinado momento. A Representação Visual do Diploma Digital (RVDD) não substitui o XML assinado: ela funciona como forma de apresentação das informações e incorpora o código de validação e o QR Code que direcionam ao documento digital.

A verificabilidade pública constitui uma característica central dessa arquitetura. O acesso ao diploma pode ocorrer pelo endereço disponibilizado pela instituição, mediante o código de validação ou o QR Code presente na RVDD, enquanto o portal oficial do MEC permite visualizar documentos por meio do envio do arquivo XML. Esses mecanismos possibilitam conferir a conformidade e a integridade do documento sem depender da apresentação de uma cópia física.

A infraestrutura estabelecida pelo MEC assegura padronização documental, autenticidade, rastreabilidade e interoperabilidade sintática entre as IES. Os arquivos XML produzidos por diferentes instituições obedecem aos mesmos esquemas XSD e podem ser submetidos a procedimentos comuns de validação. Essa conformidade, entretanto, não explicita integralmente o significado das entidades nem as relações conceituais existentes entre diplomado, curso, instituição, registro, assinatura e validação. O Diploma Digital constitui, portanto, uma base documental estruturada e juridicamente confiável, mas sua integração com outros ecossistemas informacionais depende de uma camada semântica capaz de preservar e compartilhar o significado dos dados.

3.9.1 Arquitetura modular dos esquemas XSD do diploma

Os arquivos XSD do Diploma Digital foram definidos no âmbito da especificação técnica do Ministério da Educação, com base na Instrução Normativa MEC nº 01/2020, em uma organização modular e hierárquica. Essa arquitetura representada na Figura 4, em camadas, visa garantir reutilização de componentes, consistência sintática e validação padronizada dos diferentes documentos XML que compõem o ecossistema do Diploma Digital. A estrutura oficial do pacote XSD v1.05 pode ser compreendida em dois grupos principais: os arquivos principais, responsáveis por definir a estrutura dos documentos XML que serão gerados pelas Instituições de Ensino Superior, e os arquivos auxiliares, responsáveis por fornecer leiautes, tipos básicos, elementos reutilizáveis e componentes técnicos de assinatura digital.

Os arquivos principais representam o nível documental da especificação. Eles definem os documentos XML propriamente ditos, como o Diploma Digital, a

Documentação Acadêmica para Registro, o Histórico Escolar Digital, a Lista de Diplomas Anulados, o Arquivo de Fiscalização e o Currículo Escolar Digital. Em geral, esses arquivos principais funcionam como pontos de entrada da validação XML e incorporam ou referenciam os respectivos arquivos de leiaute, nos quais se encontra a descrição mais detalhada da estrutura interna de cada documento. O arquivo DiplomaDigital_v1.05.xsd define o esquema XML oficial do Diploma Digital, enquanto o arquivo leiauteDiplomaDigital_v1.05.xsd especifica a estrutura interna do documento, incluindo tipos, elementos, grupos e restrições.

No pacote XSD v1.05, os arquivos principais podem ser descritos da seguinte forma:

I. DiplomaDigital_v1.05.xsd: define a estrutura sintática do arquivo XML do Diploma Digital, documento público oficial, assinado digitalmente e verificável.
II. DocumentacaoAcademicaRegistroDiplomaDigital_v1.05.xsd: define a estrutura sintática da Documentação Acadêmica para Registro do Diploma Digital, reunindo dados e documentos de suporte ao processo de emissão e registro.
III. HistoricoEscolarDigital_v1.05.xsd: define a estrutura sintática do Histórico Escolar Digital, documento que registra a trajetória acadêmica do estudante.
IV. ListaDiplomasAnulados_v1.05.xsd: define a estrutura sintática da lista de diplomas com registro anulado pela Instituição de Ensino Superior registradora.
V. ArquivoFiscalizacao_v1.05.xsd: define a estrutura sintática do arquivo de fiscalização, utilizado para envio de informações ao Ministério da Educação quando solicitado.
VI. CurriculoEscolarDigital_v1.05.xsd: define a estrutura sintática do Currículo Escolar Digital.

Os arquivos auxiliares, por sua vez, funcionam como módulos de apoio à validação dos documentos principais. Eles reúnem definições comuns e padronizadas, evitando repetição de estruturas e assegurando uniformidade sintática em diferentes documentos XML. Esses arquivos são responsáveis, por exemplo, por tipos de dados reutilizáveis, padrões de identificação, campos como CPF, CNPJ, códigos, datas, nomes, estruturas de endereço e elementos de assinatura digital. A reutilização dessas definições contribui para a integridade técnica do pacote XSD e para a interoperabilidade sintática entre sistemas acadêmicos e plataformas governamentais.

No pacote XSD v1.05, os arquivos auxiliares podem ser descritos da seguinte forma:

I. tiposBasicos_v1.05.xsd: define tipos básicos reutilizáveis em diferentes arquivos da especificação, contribuindo para a padronização sintática de campos como identificadores, códigos, nomes, datas, CPF, CNPJ e demais estruturas recorrentes.

II. leiauteDiplomaDigital_v1.05.xsd: define o leiaute interno do XML do Diploma Digital, detalhando sua composição estrutural.

III. leiauteDocumentacaoAcademicaRegistroDiplomaDigital_v1.05.xsd: define o leiaute interno da Documentação Acadêmica para Registro do Diploma Digital.

IV. leiauteHistoricoEscolar_v1.05.xsd: define o leiaute interno do Histórico Escolar Digital.

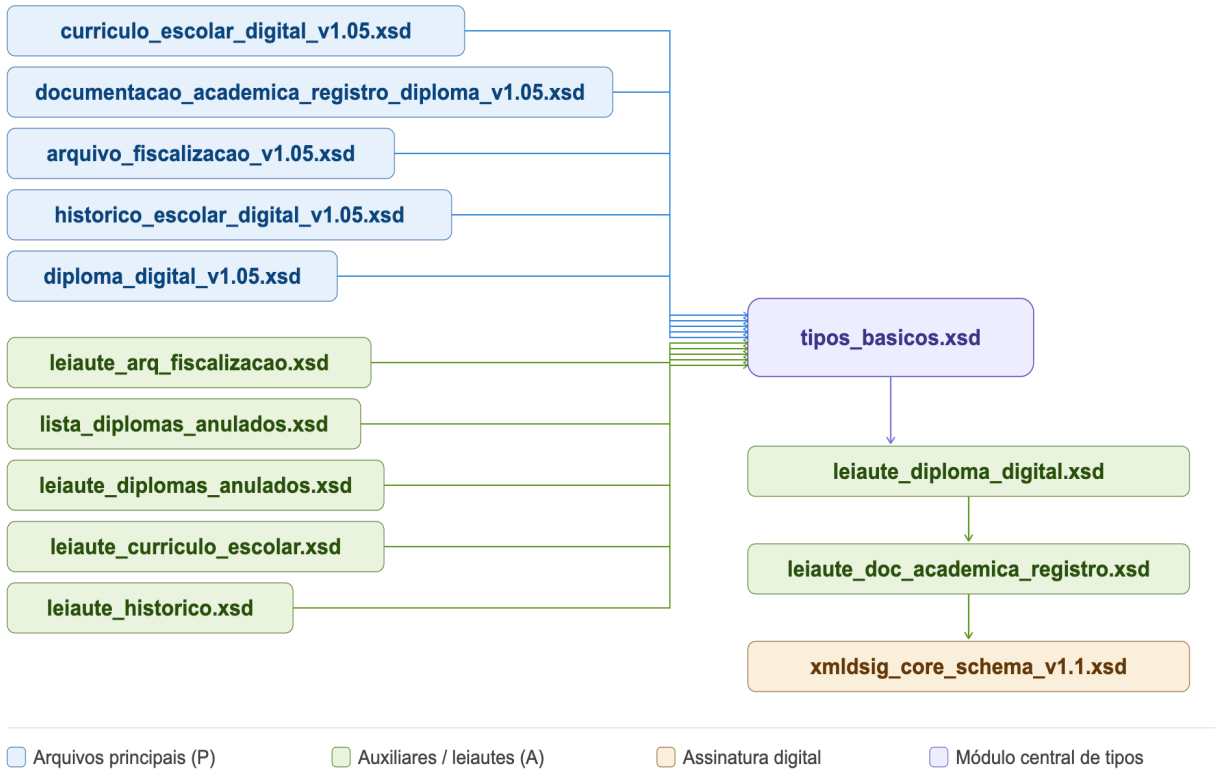
V. leiauteListaDiplomasAnulados_v1.05.xsd: define o leiaute interno da Lista de Diplomas Anulados.

VI. leiauteArquivoFiscalizacao_v1.05.xsd: define o leiaute interno do Arquivo de Fiscalização.

VII. leiauteCurriculoEscolar_v1.05.xsd: define o leiaute interno do Currículo Escolar Digital.

VIII. xmldsig-core-schema_v1.1.xsd: define os elementos relacionados à assinatura digital XML, permitindo representar a camada técnica de autenticação e integridade do documento.

Figura 4 - Representação da arquitetura dos arquivos XSD/XML



Fonte: Adaptado de Brasil, Ministério da Educação ([s. d.]).

Essa organização modular evidencia que os *schemas* XSD do Diploma Digital não atuam apenas como arquivos isolados de validação. Eles compõem uma arquitetura técnica interdependente, na qual os arquivos principais representam os documentos XML oficiais, enquanto os arquivos auxiliares fornecem estruturas reutilizáveis, leiautes internos, tipos básicos e elementos de assinatura digital. Para esta pesquisa, essa arquitetura é relevante porque constitui a principal fonte documental para a extração de conceitos, atributos, relações e restrições que, posteriormente, serão reinterpretados em uma camada semântica por meio da ontologia operacional.

A partir dessa compreensão geral, torna-se necessário observar como essa arquitetura se manifesta no próprio código XSD. O arquivo DiplomaDigital_v1.05.xsd é o ponto de entrada do XML oficial do Diploma Digital, pois declara o elemento raiz do documento e incorpora o leiaute responsável por sua estrutura interna. A análise desse trecho permite compreender como o MEC organiza tecnicamente o documento, define seus namespaces, referencia os arquivos auxiliares e estabelece os tipos complexos que orientam a validação do diploma em formato XML.

No trecho apresentado no Quadro 16 - Trecho do schema DiplomaDigital_v1.05.xsd: declaração do elemento raiz <Diploma>, observa-se que o declara o *namespace* da assinatura digital XML (xmldsig), *o* namespace padrão do XML Schema (xs) e o *namespace* próprio do Diploma Digital, vinculado ao endereço oficial do Ministério da Educação. Em seguida, o comando `<xs:include>` incorpora o arquivo leiauteDiplomaDigital_v1.05.xsd, responsável por definir a estrutura interna do diploma. Isso significa que o arquivo DiplomaDigital_v1.05.xsd funciona como o ponto de entrada do documento XML, enquanto o arquivo de leiaute especifica os tipos, elementos, grupos e restrições que compõem sua organização interna.

O elemento raiz <Diploma>, definido como do tipo complexo TDiploma, indica que o documento XML deve seguir uma estrutura previamente estabelecida no leiaute oficial. A anotação `<xs:annotation>`, por meio de `<xs:documentation>`, registra uma descrição textual do elemento, informando que se trata do Diploma Digital. Embora essa anotação não altere a validação técnica do XML, ela contribui para a documentação do schema e para a compreensão humana da estrutura. O *schema* atua, portanto, como estrutura sintática de validação, enquanto o leiaute associado detalha a composição do documento, garantindo padronização entre as Instituições de Ensino Superior.

Quadro 16 - Trecho do schema DiplomaDigital_v1.05.xsd: declaração do elemento raiz <Diploma>

```
<xs:schema
    xmlns:ds="https://www.w3.org/2000/09/xmldsig#"
    xmlns:xs="https://www.w3.org/2001/XMLSchema"
    xmlns="https://portal.mec.gov.br/diplomadigital/arquivos-em-xsd"
    targetNamespace="https://portal.mec.gov.br/diplomadigital/arquivos-em-xsd"
    elementFormDefault="qualified"
    attributeFormDefault="unqualified">

    <xs:include schemaLocation="leiauteDiplomaDigital_v1.05.xsd"/>

    <xs:element name="Diploma" type="TDiploma">
        <xs:annotation>
            <xs:documentation>
                Arquivo XML do Diploma Digital conforme leiaute v1.05
            </xs:documentation>
        </xs:annotation>
    </xs:element>
</xs:schema>
```

Fonte: Recorte do schema XSD do Diploma Digital, Brasil, Ministério da Educação ([s. d.]).

O arquivo XSD delimita de forma estruturada os componentes essenciais para a construção e validação do documento XML do diploma emitido por instituições de

ensino superior. Entre esses componentes destacam-se alguns elementos que são detalhados nos próximos parágrafos.

O elemento <Diploma> é declarado no Quadro 17 e conta com uma anotação <xs:annotation> para documentar seu propósito. A anotação não interfere na validação técnica, mas cumpre a função de explicitar, para humanos ou ferramentas de geração de documentação, que este elemento representa o “Arquivo XML do Diploma Digital conforme leiaute v1.05”. Esse recurso é útil para manutenção, compreensão e interoperabilidade do esquema, indicando contexto ou notas explicativas.

Quadro 17 - Trecho de código XML-XSD do Diploma Digital - `<xs:annotation>`

```
<xs:schema
    xmlns:ds="https://www.w3.org/2000/09/xmldsig#"
    xmlns:xs="https://www.w3.org/2001/XMLSchema"
    xmlns="https://portal.mec.gov.br/diplomadigital/arquivos-em-xsd"
    targetNamespace="https://portal.mec.gov.br/diplomadigital/arquivos-em-xsd"
    elementFormDefault="qualified"
    attributeFormDefault="unqualified">

    <xs:include schemaLocation="leiauteDiplomaDigital_v1.05.xsd"/>

    <xs:element name="Diploma" type="TDiploma">
        <xs:annotation>
            <xs:documentation>
                Arquivo XML do Diploma Digital conforme leiaute v1.05
            </xs:documentation>
        </xs:annotation>
    </xs:element>
</xs:schema>
```

Fonte: Recorte do schema XSD do Diploma Digital, Brasil, Ministério da Educação ([s. d.]).

Conforme apresentado no Quadro 18, a identificação do número do processo judicial e do nome do juiz é definida por um tipo complexo que exige a ocorrência dos elementos “NumeroProcessoJudicial” e “NomeJuiz” nessa ordem específica. O uso de <xs:sequence> estabelece explicitamente que a instância XML deve conter, primeiramente, o número do processo e, em seguida, o nome do juiz para que a validação seja aprovada. Tal estrutura garante consistência no layout e facilita a leitura e a extração dos dados por sistemas.

Quadro 18 - Trecho de código XML-XSD do Diploma Digital - <xs:sequence>

```
<xs:sequence>
    <xs:element name="NumeroProcessoJudicial" type="TNumeroProcessoJudicial"/>
    <xs:element name="NomeJuiz" type="TString"/>
    <xs:element minOccurs="0" name="Decisao" type="TString"/>
</xs:sequence>
```

Fonte: Recorte do schema XSD do Diploma Digital, Brasil, Ministério da Educação ([s. d.]).

No tipo complexo `TInfDiploma`, apresentado no Quadro 19 , declara-se o atributo `versao`, do tipo `TVersao`, cuja presença é obrigatória, conforme indicado por `use="required"`. A anotação `xs:documentation` identifica a versão do leiaute como v1.05. Entretanto, o trecho apresentado não contém a declaração `fixed="1.05"`; portanto, a obrigatoriedade refere-se à presença do atributo, e não à atribuição desse valor específico. Os valores admissíveis para o atributo dependem das restrições estabelecidas na definição do tipo `TVersao`. Esse atributo identifica a versão do leiaute utilizada no documento XML.

Quadro 19 - Trecho de código XML-XSD do Diploma Digital - "required"

```
<xs:attribute name="versao" type="TVersao" use="required">
    <xs:annotation>
        <xs:documentation>Versão do leiaute (v1.05)</xs:documentation>
    </xs:annotation>
</xs:attribute>
```

Fonte: Recorte do schema XSD do Diploma Digital, Brasil, Ministério da Educação ([s. d.]).

No Quadro 20, foi definido o tipo simples TCodigoValidacao, que restringe valores do tipo string, elimina espaços excedentes (colapso) e aplica uma expressão regular no atributo “pattern” para estabelecer o formato obrigatório. Esse tipo pode ser empregado na validação do “código de validação” do diploma digital, em conformidade com as normas vigentes, assegurando que o código atenda aos requisitos de tamanho e formato antes da aceitação da instância como válida.

Quadro 20 - Trecho de código XML-XSD do Diploma Digital - "colapse

```
<xs:simpleType name="TCodigoValidacao">
    <xs:annotation>
        <xs:documentation>Código de segurança é composto pelo: "e-MEC emissora" +
        </xs:documentation>
    </xs:annotation>
    <xs:restriction base="TString">
        <xs:whiteSpace value="collapse"/>
        <xs:pattern value="\d{1,}\.\d{1,}\.[a-f0-9]{12,}"/>
    </xs:restriction>
</xs:simpleType>
```

Fonte: Recorte do schema XSD do Diploma Digital, Brasil, Ministério da Educação ([s. d.]).

O trecho apresentado no Quadro 21 define o tipo complexo TDadosDiplomado, responsável por estruturar as informações do diplomado no schema XSD. No interior do elemento `<xs:sequence>`, cada `<xs:element name="...">` declara explicitamente um componente obrigatório da estrutura, associando-o a um tipo previamente definido (por exemplo, ID, Nacionalidade, CPF, DataNascimento). Já o elemento `<xs:group ref="GPessoa"/>` funciona como um grupo reutilizável de elementos, previamente declarado no *schema*, que encapsula um conjunto padronizado de informações relativas à pessoa (como nome, sexo, filiação, entre outros, dependendo da definição original do grupo). O uso de `<xs:group>` favorece a modularidade, evitando repetição de estruturas em diferentes partes do *schema* e garantindo consistência entre documentos XML. Por fim, o emprego de `<xs:choice>` estabelece uma escolha obrigatória entre dois elementos mutuamente exclusivos, RG ou OutroDocumentoIdentificacao, reforçando restrições semânticas da modelagem. Em conjunto, essas construções organizam o tipo complexo de forma clara, reaproveitável e semanticamente controlada, assegurando a padronização da representação dos dados do diplomado no Diploma Digital.

Quadro 21 - Trecho de código XML-XSD do Diploma Digital - <xs:group>

```
<xs:complexType name="TDadosDiplomado">
    <xs:annotation>
        <xs:documentation>Dados do Diplomado</xs:documentation>
    </xs:annotation>
    <xs:sequence>
        <xs:element name="ID" type="TId"/>
        <xs:group ref="GPressoa"/>
        <xs:element name="Nacionalidade" type="TNacionalidade"/>
        <xs:element name="Naturalidade" type="TNaturalidade"/>
        <xs:element name="CPF" type="TCpf"/>
        <xs:choice>
            <xs:element name="RG" type="TRg"/>
            <xs:element name="OutroDocumentoIdentificacao" type="TOutroDocumentoIdentificacao"/>
        </xs:choice>
        <xs:element name="DataNascimento" type="TData"/>
    </xs:sequence>
</xs:complexType>
```

Fonte: Recorte do *schema* XSD do Diploma Digital, Brasil, Ministério da Educação ([s. d.]).

3.9.2 Requisitos tecnológicos da assinatura eletrônica

Ribeiro, Almeida e Canedo (2021) abordam o Diploma Digital a partir dos requisitos tecnológicos relacionados à assinatura eletrônica dos documentos acadêmicos. Os autores apresentam um modelo arquitetural baseado em microsserviços e desenvolvem um protótipo em Java, com utilização da biblioteca XAdES4j, direcionado ao processo de emissão de diplomas em formato digital. A proposta evidencia que a transformação do diploma em documento nato-digital envolve não apenas a substituição do suporte físico, mas também a incorporação de mecanismos tecnológicos capazes de atender aos requisitos estabelecidos para sua emissão e assinatura.

A assinatura digital constitui um componente relevante da infraestrutura de confiança associada ao Diploma Digital. O emprego do padrão XAdES insere o documento acadêmico em um ambiente no qual a identificação do signatário e a integridade do conteúdo precisam ser tratadas por mecanismos próprios da certificação e da assinatura eletrônica. Ribeiro, Almeida e Canedo (2021) evidenciam a relação entre a emissão do diploma e os recursos tecnológicos empregados para assegurar as condições necessárias à verificação do documento no ambiente digital.

A abordagem dos autores também contribui para compreender o Diploma Digital como parte de um conjunto de relações entre documento, instituição, agentes responsáveis e mecanismos de autenticação. A assinatura não constitui apenas um elemento técnico acrescentado ao arquivo, mas integra o processo de produção e validação do documento acadêmico. O estudo de Ribeiro, Almeida e Canedo (2021)

oferece elementos para compreender a dimensão tecnológica e institucional da confiança no Diploma Digital e sua dependência de estruturas capazes de identificar os responsáveis pelos atos praticados no processo de emissão e registro.

3.9.3 O Acervo Acadêmico e o Contexto Institucional das IES

A digitalização da administração pública federal ampliou as exigências de transformação dos processos acadêmicos das IES. Esse movimento é balizado por marcos regulatórios complementares que exigem ações coordenadas: se, por um lado, as Portarias MEC nº 330/2018 e nº 554/2019 regulamentam a expedição do diploma de graduação em formato nato-digital estruturado em XML, por outro, a Portaria MEC nº 360/2022 determina a conversão integral do acervo acadêmico discente para o meio digital, vedando a produção de novos documentos em suporte físico. Sob a perspectiva da Arquivologia, essa conjuntura evidencia que o diploma digital não pode ser concebido como um artefato documental autônomo. Ele representa o cume indissociável de uma cadeia de custódia e de uma relação orgânica contínua que se inicia com a matrícula do discente e engloba todas as atividades de sua vida acadêmica. Portanto, a validade e a fidedignidade da certificação final dependem intrinsecamente da integridade e da guarda sistemática de todo o acervo que lhe serve de sustentação documental.

Entretanto, estudo de Santana Hüller, Gomes e Santos (2022) identificou indícios de fragmentação tecnológica e desarticulação entre sistemas no contexto investigado. No mapeamento realizado por Santana Hüller, Gomes e Santos (2022), os autores investigaram a existência de softwares registrados no Instituto Nacional da Propriedade Industrial (INPI) nas áreas de Educação (**ED**) e Informação (**IF**[2]), que abrangem subcampos epistemológicos centrais como graus e diplomas (ED-05), documentação (IF-02), arquivos (IF-04 e IF-06) e Sistemas de Informação (IF-07). Embora tenham identificado 21 programas de computador ativos para a gestão acadêmica, a pesquisa constatou que tais sistemas se limitavam a gerenciar atividades de registro transacional e rotinas administrativas correntes. Nenhum dos softwares mapeados apresentava funcionalidades voltadas ao gerenciamento

[2] ED e IF (Códigos de Campo de Aplicação do INPI): Refere-se às classificações adotadas pelo Instituto Nacional da Propriedade Industrial (INPI) para categorizar os registros de programas de computador no Brasil

eletrônico integrado de documentos do acervo ou à expedição do diploma digital em conformidade com as exigências técnicas do MEC.

Essa limitação observada no conjunto de softwares analisados também foi identificada entre as instituições participantes do levantamento realizado por Santana Hüller, Gomes e Santos (2022). Ao realizarem um levantamento empírico com representantes de 46 IES (obtendo 19,6% de adesão), Santana Hüller, Gomes e Santos (2022) observaram que nenhuma das instituições respondentes possuía um sistema que unificasse a guarda do acervo digital e a emissão do diploma. O estudo identificou apenas três instituições que se encontravam em fase de implementação de sistemas de diploma digital, contudo, estes operavam de maneira isolada e desalinhada em relação ao gerenciamento dos documentos do acervo discente. Sob a perspectiva da Ciência da Informação, esse desalinhamento sistêmico pode comprometer a fidedignidade e a adequada gestão do processo documental. A emissão de um diploma digital sem a correspondente integração com a base documental que o fundamenta também pode dificultar o atendimento aos requisitos de gestão, preservação, busca, acesso e recuperação previstos no Art. 46 da Portaria MEC nº 315/2018, que exige a manutenção de uma base de dados preservada e dotada de uma indexação que facilite a busca, o acesso e a recuperação das informações.

Por fim, a vulnerabilidade dessas soluções fragmentadas estende-se à dimensão da preservação e custódia continuada da informação pública. O mapeamento tecnológico de Santana Hüller, Gomes e Santos (2022) denunciou que os sistemas utilizados pelas IES carecem de conformidade com os requisitos mínimos estabelecidos pelo Conselho Nacional de Arquivos (CONARQ) no **e-ARQ Brasil**[3] para Sistemas Informatizados de Gestão Arquivística de Documentos (SIGAD[4]), além de não preverem parâmetros para a implantação de um Repositório Arquivístico Digital Confiável (RDC-Arq[5]), conforme as diretrizes da Resolução nº 43/2015 do CONARQ. Sem a estruturação de um RDC-Arq e a definição de políticas claras de preservação

[3] O e-ARQ Brasil é um modelo nacional de requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos (SIGAD), publicado pelo Conselho Nacional de Arquivos (CONARQ)

[4] SIGAD (Sistema Informatizado de Gestão Arquivística de Documentos): É a sigla para qualquer software que consiga gerenciar a vida de um documento (desde a sua criação, uso diário, até o arquivo definitivo ou descarte) seguindo as regras do e-ARQ Brasil

[5] RDC-Arq (Repositório Arquivístico Digital Confiável): Entende-se como "cofre digital ultra seguro". Um ambiente projetado com regras rígidas (segundo a Resolução nº 43/2015 do CONARQ) para garantir que um documento digital (como o XML do diploma) seja preservado por décadas.

digital, como migração, reprografia e atualização de versões, os diplomas e históricos correm riscos iminentes de obsolescência tecnológica e perda de autenticidade no longo prazo. Em vista do exposto, constata-se que o contexto institucional das universidades brasileiras demanda soluções informacionais que transcendam a limitação dos silos de dados, assegurando que a fidedignidade jurídica imediata da assinatura digital seja respaldada por uma rede de metadados integrada e semanticamente estruturada.

4 Procedimentos Metodológicos

A investigação adotou abordagem qualitativa, natureza aplicada e, quanto aos objetivos, caráter exploratório-descritivo. A abordagem qualitativa permitiu compreender a modelagem ontológica dos dados do Diploma Digital Brasileiro em suas dimensões informacional, institucional, normativa e tecnológica, considerando significados, práticas e relações que não poderiam ser reduzidos a variáveis exclusivamente quantificáveis. Conforme Creswell e Creswell (2021), essa abordagem mostra-se adequada à interpretação de fenômenos complexos em seus contextos de ocorrência. O caráter exploratório decorreu da incipiência dos estudos nacionais sobre ontologias operacionais aplicadas ao Diploma Digital, especialmente sob a perspectiva da Ciência da Informação e da interoperabilidade semântica. A dimensão descritiva orientou a identificação, a análise e a sistematização dos conceitos, entidades, propriedades, relações e restrições do domínio. Esse enquadramento mostrou-se compatível com Richardson e Peres (1999), para quem a pesquisa descritiva possibilita observar, registrar, analisar e ordenar fenômenos sem necessariamente manipulá-los.

A pesquisa combinou levantamento bibliográfico, análise documental e desenvolvimento de artefato ontológico. O levantamento bibliográfico subsidiou a fundamentação teórica e a construção do Estado da Arte, permitindo identificar estudos relacionados a credenciais acadêmicas digitais, interoperabilidade, representação semântica, portabilidade de dados estudantis e tecnologias aplicadas à verificação de qualificações. A análise documental concentrou-se nas normas, nos documentos técnicos e nos *schema*s XML/XSD associados ao Diploma Digital Brasileiro, utilizados como fontes para a delimitação do subdomínio e para a identificação dos elementos informacionais relevantes. O desenvolvimento do artefato foi conduzido conforme o percurso metodológico da OntoForInfoScience, cujas etapas orientaram a decisão ontológica, a aquisição do conhecimento, a conceitualização, a formalização, a avaliação e a documentação da ontologia.

Esse delineamento permitiu articular os fundamentos teóricos da Organização e Representação do Conhecimento, os requisitos normativos e técnicos do Diploma Digital e os procedimentos próprios da engenharia ontológica. A descrição detalhada das fontes consultadas, dos critérios de seleção bibliográfica e das etapas de

construção da ontologia é apresentada nas subseções seguintes, de modo a explicitar o percurso metodológico adotado e favorecer a reprodutibilidade da pesquisa.

4.1 Aplicação da metodologia OntoForInfoScience na construção da OntoDiplomaDigital.

A construção da OntoDiplomaDigital foi conduzida como um processo de Organização e Representação do Conhecimento aplicado ao subdomínio do Registro do Diploma no contexto do Diploma Digital Brasileiro. A decisão de desenvolver uma ontologia decorreu da análise do problema de pesquisa e da avaliação das características dos instrumentos de representação disponíveis. Verificou-se que os instrumentos terminológicos tradicionais, embora adequados à organização e ao controle de termos, não permitiam representar, com o nível de formalidade requerido, as entidades, propriedades, relações e restrições necessárias ao processamento computacional do domínio. De modo semelhante, os arquivos XML e os schemas XSD publicados pelo Ministério da Educação asseguravam a estruturação dos documentos, a tipagem dos elementos e a validação de sua conformidade sintática, mas não constituíam, isoladamente, uma conceitualização formal compartilhada nem explicitavam as relações semânticas necessárias à integração e ao reúso dos dados.

Em conformidade com a distinção estabelecida por Gil (2008, p. 182-183), esta seção apresenta os procedimentos adotados [...]. O capítulo metodológico concentra-se na descrição das operações realizadas e dos critérios que orientaram sua execução, enquanto os artefatos produzidos, as evidências obtidas e sua interpretação são reservados, respectivamente, aos capítulos de Resultados e de Discussão.

Para estruturar esse percurso, adotou-se a metodologia OntoForInfoScience, proposta por Mendonça (2015) para orientar a construção de ontologias no campo da Ciência da Informação. A metodologia compreende uma etapa preliminar, destinada à avaliação da necessidade de desenvolvimento da ontologia, e oito etapas subsequentes: especificação; aquisição e extração do conhecimento; conceitualização; fundamentação ontológica; formalização; avaliação; documentação; e disponibilização. Sua aplicação nesta pesquisa não correspondeu à reprodução linear ou mecânica de um roteiro previamente estabelecido. As etapas foram interpretadas e adaptadas às características de uma ontologia operacional de

domínio, à delimitação do Registro do Diploma, às fontes normativas e técnicas do Ministério da Educação e aos padrões de representação da Web Semântica empregados na implementação da OntoDiplomaDigital.

O desenvolvimento foi realizado de forma iterativa e incremental, com sucessivos ciclos de análise, modelagem, formalização e avaliação. Inconsistências terminológicas, lacunas conceituais, relações insuficientemente definidas ou incompatibilidades identificadas durante as verificações foram encaminhadas às etapas de aquisição do conhecimento, conceitualização ou formalização, de acordo com a natureza do problema observado. Após cada alteração relevante, foram repetidas as verificações pertinentes aos glossários, aos dicionários de conceitos e relações, aos modelos conceituais e ao arquivo formalizado em Turtle. Essa dinâmica operacionalizou a relação bidirecional entre formalização e avaliação prevista pela OntoForInfoScience e possibilitou o refinamento progressivo da ontologia em relação à coerência conceitual, à consistência lógica, à rastreabilidade das decisões de modelagem e à compatibilidade com os requisitos definidos para o domínio.

4.1.1 Decisão ontológica, delimitação do subdomínio e especificação dos requisitos - Etapa preliminar da OntoForInfoScience

Esta subseção corresponde à etapa preliminar da OntoForInfoScience, denominada Etapa zero, na qual foram definidos o tipo de instrumento de representação do conhecimento a ser adotado, o recorte do domínio investigado e os requisitos iniciais da ontologia. Essa etapa teve a função de estabelecer as bases metodológicas para as fases posteriores de aquisição do conhecimento, conceitualização, formalização, avaliação e documentação, evitando que a construção da ontologia se iniciasse sem delimitação explícita de escopo e finalidade.

A decisão ontológica foi realizada mediante a comparação entre os requisitos definidos para a pesquisa e as capacidades representacionais de taxonomias, tesauros, vocabulários controlados e ontologias. Considerou-se que os três primeiros instrumentos seriam adequados caso a finalidade se restringisse à padronização terminológica, à organização hierárquica de conceitos, à indexação ou à recuperação de documentos. A ontologia foi selecionada porque o problema investigado exigia a representação de recursos identificáveis, propriedades, diferentes tipos de relações, restrições e alinhamentos semânticos, além da possibilidade de materializar dados, realizar inferências e formular consultas sobre o domínio. Essa decisão seguiu o

critério proposto por Mendonça (2015, p. 185-188), segundo o qual a escolha do instrumento de representação deve corresponder à natureza do problema e ao grau de formalidade requerido.

4.1.2 Especificação da ontologia - Etapa 1 da OntoForInfoScience

A Etapa 1 da OntoForInfoScience corresponde à especificação da ontologia. Após a decisão pela construção de uma ontologia operacional, essa etapa teve a função de definir os parâmetros iniciais do artefato a ser desenvolvido, incluindo domínio, propósito, usuários previstos, cenários de aplicação, grau de formalidade e questões de competência. Essa especificação estabeleceu o vínculo entre o problema de pesquisa e as etapas posteriores da metodologia, orientando quais elementos deveriam ser investigados, modelados, formalizados e avaliados.

Para subsidiar essa especificação, foram considerados os documentos normativos e técnicos relacionados ao Diploma Digital Brasileiro, especialmente os arquivos associados ao Diploma Digital, à Documentação Acadêmica para Registro e ao Histórico Escolar Digital. A análise inicial considerou a extensão do conjunto documental, o tempo disponível para o desenvolvimento da pesquisa, a pertinência de cada componente aos objetivos específicos e sua contribuição para responder às questões de competência. Esse procedimento permitiu distinguir os elementos centrais do domínio das estruturas auxiliares destinadas predominantemente à validação sintática dos documentos XML.

Com base nesses critérios, o escopo da ontologia foi delimitado ao subdomínio do Registro do Diploma, entendido como o núcleo conceitual relacionado à certificação pública da conclusão de um curso superior. O recorte abrangeu as entidades e relações necessárias à identificação do diploma e de seu titular, do curso concluído, das instituições emissora e registradora, do ato de registro, dos agentes responsáveis e assinantes, dos mecanismos de autenticidade e validação e dos documentos diretamente associados ao fluxo. Os dados privados constantes da documentação acadêmica e a granularidade curricular própria do histórico escolar digital permaneceram fora do núcleo operacional, exceto quando sua referência se mostrou necessária para representar a proveniência, a sustentação documental ou o encadeamento informacional do diploma.

As questões de competência foram formuladas a partir da análise inicial das fontes normativas e técnicas e das necessidades de representação identificadas no

subdomínio do Registro do Diploma. Sua formulação consistiu em expressar, na forma de perguntas, quais informações e relações a OntoDiplomaDigital deveria permitir representar e recuperar. Foram contempladas a identificação das instituições emissora e registradora; os elementos de registro, assinatura e validação; a recuperação de diplomas por instituição, período e modalidade de oferta; a relação entre titular, curso, grau e título; os valores controlados; a proveniência do registro; e a representação da relação de acesso por plataforma governamental. As sete questões de competência resultantes são apresentadas na Seção 5.1.5, no Quadro 28.

Essas questões foram tratadas como requisitos informacionais da ontologia e orientaram a seleção dos conceitos, a definição das propriedades e relações e a posterior elaboração das consultas SPARQL. O critério de inclusão de cada elemento considerou sua pertinência ao subdomínio delimitado e sua contribuição para responder às perguntas formuladas, evitando a transformação indiscriminada dos componentes intermediários dos XSDs em entidades ontológicas. Dessa forma, as questões estabeleceram uma referência comum entre a especificação, a construção e a avaliação operacional da OntoDiplomaDigital. Sua operacionalização por meio de consultas SPARQL e os resultados obtidos na prova de conceito, realizada em ambiente controlado com dados sintéticos, são apresentados na Seção 5.6.3.

4.1.3 Aquisição e extração do conhecimento - Etapa 2 da OntoForInfoScience

A Etapa 2 da OntoForInfoScience correspondeu à aquisição e à extração do conhecimento necessário à construção da OntoDiplomaDigital. Seu ponto de partida foi o estudo do Diploma Digital Brasileiro apresentado na Seção 3.9, que examinou o contexto normativo e institucional, a arquitetura modular dos esquemas XSD e os requisitos tecnológicos associados à assinatura eletrônica. Essa compreensão inicial do domínio subsidiou a análise sistemática das fontes para identificar termos, conceitos, atributos, relações e restrições associados ao subdomínio do Registro do Diploma. A aquisição foi conduzida a partir de quatro conjuntos de fontes: normativas e técnicas, bibliográficas, empíricas/contextuais e vocabulários externos utilizados como referência para posterior alinhamento semântico.

As fontes normativas e técnicas compreenderam os atos oficiais que instituíram e regulamentaram o Diploma Digital Brasileiro, especialmente a Portaria MEC nº 330/2018, a Portaria MEC nº 554/2019, a Instrução Normativa MEC nº 01/2020, a Nota

Técnica MEC nº 13/2019 (Brasil, 2019b) e a Portaria MEC nº 70/2025. Esses documentos foram analisados para identificar responsabilidades institucionais, requisitos de emissão e registro, condições de validade jurídica, elementos de assinatura digital, preservação documental e regras gerais aplicáveis ao Diploma Digital. A análise normativa permitiu delimitar quais elementos pertenciam ao núcleo conceitual do Registro do Diploma e quais correspondiam a requisitos jurídicos, administrativos ou técnicos complementares.

A documentação técnica e os schemas XSD disponibilizados pelo Ministério da Educação foram examinados como fontes primárias para a extração dos elementos estruturais do domínio. Foram considerados os arquivos associados ao Diploma Digital, à Documentação Acadêmica para Registro e ao Histórico Escolar Digital, bem como tipos básicos, leiautes auxiliares, grupos, tipos simples e complexos, cardinalidades, enumerações, alternativas estruturais, dependências entre schemas e componentes de assinatura digital. Essa análise permitiu identificar candidatos a classes, propriedades, atributos, valores controlados e restrições, distinguindo os elementos efetivamente relevantes ao subdomínio do Registro do Diploma das estruturas voltadas predominantemente à validação sintática dos documentos XML.

As fontes bibliográficas foram utilizadas para apoiar a interpretação conceitual dos elementos extraídos e para orientar a organização do conhecimento no domínio investigado. Foram consideradas obras relacionadas ao Diploma Digital, às credenciais acadêmicas digitais, à interoperabilidade, à representação semântica, aos metadados, aos princípios FAIR, aos Dados Ligados e à engenharia ontológica. Essas fontes não substituíram os atos normativos e os schemas XSD como base primária do domínio, mas contribuíram para interpretar os conceitos, justificar decisões de representação e relacionar a construção da ontologia aos fundamentos da Ciência da Informação e da Web Semântica.

Também foram considerados vocabulários e modelos externos como fontes de referência para o alinhamento técnico-semântico posterior. A análise desses recursos teve caráter seletivo e foi orientada pela aderência ao domínio, pela estabilidade dos identificadores, pela capacidade de representar pessoas, organizações, documentos, cursos, endereços, conceitos controlados e relações de proveniência, e pela compatibilidade com padrões da Web Semântica. Nessa etapa, os vocabulários externos foram tratados como fontes de comparação e apoio, cabendo às etapas seguintes decidir quais alinhamentos seriam efetivamente incorporados à ontologia.

As fontes empíricas/contextuais incluíram a experiência institucional pregressa do pesquisador, decorrente de sua atuação na gestão de tecnologia da informação e no acompanhamento da implantação do sistema acadêmico e do Diploma Digital na Universidade Federal do Espírito Santo. Essa experiência foi respaldada por registros de atividades técnicas e gerenciais realizadas durante o processo de implantação, envolvendo a Superintendência de Tecnologia da Informação (STI), a Pró-Reitoria de Graduação (PROGRAD), fornecedores do sistema acadêmico e outros agentes institucionais. Os registros abrangeram atividades relacionadas à adesão ao Diploma Digital, adequações no sistema acadêmico, preparação de dados institucionais, validação da Representação Visual do Diploma Digital (RVDD) e fluxos de emissão e registro.

Essa experiência foi utilizada exclusivamente como elemento contextual para a compreensão do domínio, permanecendo as decisões de aquisição, conceitualização e formalização do conhecimento fundamentadas nas fontes normativas e técnicas oficiais, especialmente nos atos normativos e schemas XSD disponibilizados pelo MEC. Os registros institucionais anteriores à pesquisa não foram tratados como entrevistas ou procedimentos formais de coleta de dados, nem utilizados como mecanismo de validação externa da OntoDiplomaDigital.

A extração do conhecimento foi realizada pela identificação de substantivos, expressões nominais, verbos, expressões relacionais, atributos e restrições recorrentes nas fontes analisadas. Esses elementos foram organizados inicialmente como candidatos terminológicos, sem pressupor sua conversão automática em classes, propriedades ou indivíduos da ontologia. Como produto da Etapa 2, foram elaborados glossários preliminares de conceitos, verbos e relações, acompanhados de registros de proveniência terminológica. Os elementos extraídos e os artefatos produzidos são apresentados na Seção 5.2.2. Esses artefatos serviram como insumos para a etapa seguinte de conceitualização, na qual os termos extraídos foram definidos, agrupados, especializados, relacionados e avaliados quanto à sua pertinência para o modelo ontológico.

4.1.4 Conceitualização do subdomínio - Etapa 3 da OntoForInfoScience

A Etapa 3 da OntoForInfoScience correspondeu à conceitualização do subdomínio do Registro do Diploma. Nessa etapa, os termos e estruturas identificados durante a aquisição do conhecimento foram organizados em uma representação

intermediária do domínio, anterior à formalização em linguagem ontológica. O objetivo foi transformar os elementos extraídos das fontes normativas, técnicas, bibliográficas e contextuais em conceitos definidos, relações interpretadas e modelos conceituais capazes de orientar a posterior formalização da OntoDiplomaDigital.

A conceitualização foi conduzida em três frentes complementares. A primeira consistiu na elaboração do dicionário de conceitos e da tabela de conceitos e propriedades. Os candidatos extraídos dos arquivos XML/XSD, dos atos normativos e das demais fontes analisadas foram submetidos à comparação terminológica, à definição de significado, à identificação de variantes, à classificação conceitual e à análise de pertinência em relação ao escopo do Registro do Diploma. Para cada elemento, foram registrados a denominação adotada, a definição de trabalho, a origem documental ou decisão de modelagem, os possíveis sinônimos, os atributos associados, as relações centrais, as restrições de uso e as questões ainda dependentes de decisão nas etapas posteriores.

A segunda frente correspondeu à elaboração do dicionário de verbos. As expressões relacionais identificadas nas fontes foram analisadas para verificar se representavam vínculos semanticamente relevantes entre recursos do domínio. Para cada relação candidata, foram registrados a expressão verbal, o recurso de origem, o recurso de destino, a possível relação inversa, a fonte de extração e exemplos de leitura. Esse procedimento permitiu diferenciar relações hierárquicas, associações, atributos e valores controlados, além de preparar a posterior definição de propriedades ontológicas. A taxonomia preliminar foi construída separadamente, com base na pergunta “é um tipo de?”, para distinguir relações de generalização e especialização das demais relações associativas.

A terceira frente envolveu a construção dos modelos conceituais gráficos. Inicialmente, os modelos foram desenvolvidos no Miro[6], como apoio à organização visual dos agentes, documentos, recursos acadêmicos, fluxos informacionais e relações associados ao Registro do Diploma. Essa representação permitiu comparar, agrupar e reposicionar os elementos identificados durante a aquisição do conhecimento. Após os ciclos de análise, as estruturas selecionadas foram transcritas

[6] Miro é uma plataforma digital de colaboração visual baseada em canvas, utilizada para organizar quadros, diagramas, fluxos e mapas conceituais. Disponível em: https://miro.com/.

para PlantUML[7], utilizado como fonte textual dos modelos conceituais e diagramas. Essa representação facilitou a revisão das denominações, das ligações, das direções de leitura, das cardinalidades indicativas e das correspondências com os artefatos de origem.

A passagem da estrutura documental para o modelo conceitual foi orientada por critérios interpretativos. Um complexType nomeado foi considerado candidato a entidade ou classe quando representava um recurso reconhecível no subdomínio, com identidade própria, características descritivas e participação em relações. Nos modelos voltados à rastreabilidade, preservou-se o nome técnico do tipo; nos casos de complexType anônimo, adotou-se como referência a denominação do elemento que continha a estrutura. A existência de um tipo complexo, entretanto, não determinou automaticamente sua transformação em classe, pois também foram considerados seu significado no domínio, sua pertinência ao recorte e sua contribuição para as questões de competência.

Os simpleType, os atributos e os elementos de conteúdo simples foram tratados, em regra, como candidatos a características descritivas expressas por literais. Enumerações ou listas de valores com função terminológica estável e reutilizável foram examinadas como candidatas a vocabulários controlados. Durante a revisão conceitual, as regras de preenchimento dos XSDs e suas estruturas documentais foram tratadas como evidências para análise, e não como elementos de conversão automática em axiomas ou classes. Domínios, alcances, cardinalidades, características lógicas e requisitos relativos à qualidade da ABox foram reservados para exame nas etapas posteriores de formalização, materialização e validação.

Os produtos consolidados dessa etapa são apresentados no Capítulo 5, especialmente nas subseções relativas ao dicionário de conceitos, à tabela de conceitos e propriedades, ao dicionário de verbos e aos modelos conceituais gráficos. Nesta seção, descreve-se o percurso metodológico adotado para sua obtenção, evidenciando como os elementos extraídos na Etapa 2 foram interpretados, organizados e preparados para a formalização ontológica.

4.1.5 Fundamentação ontológica - Etapa 4 da OntoForInfoScience

[7] PlantUML é uma ferramenta de código aberto que permite gerar diagramas UML e outros diagramas a partir de descrições textuais. Disponível em: https://plantuml.com/.

A Etapa 4 da OntoForInfoScience recomenda explicitar a fundamentação ontológica e considerar o emprego de ontologias de alto nível. Mendonça (2015, p. 209–214) admite, entretanto, que ontologias de menor porte, ontologias leves de domínio ou subontologias podem demandar uma aplicação adaptada dessa etapa. Considerando que a OntoDiplomaDigital foi concebida como uma ontologia operacional de domínio, delimitada ao subdomínio do Registro do Diploma e orientada à interoperabilidade semântica de um recorte documental específico, não foi adotada uma ontologia de fundamentação como estrutura superior obrigatória. Em seu lugar, empregou-se uma estratégia de reúso técnico-semântico baseada na análise de vocabulários e modelos consolidados, entre eles RDF, RDFS, OWL, SKOS, DCMI Metadata Terms, PROV-O, FOAF, DCAT, Schema.org, LOCN e vCard RDF. Portais agregadores foram utilizados apenas como mecanismos de descoberta, enquanto as definições, os domínios, os alcances e as condições de uso dos termos candidatos foram conferidos nas respectivas especificações oficiais.

A adoção de correspondências externas foi condicionada à compatibilidade entre o significado atribuído ao elemento local e o significado, o domínio, o alcance e o uso previsto para o termo reutilizado. A relação rdfs:subPropertyOf foi empregada quando a propriedade local podia ser interpretada como especialização de uma propriedade mais ampla, preservando-se a compatibilidade semântica entre ambas; relações de equivalência mais fortes não foram utilizadas na ausência de evidência suficiente para sustentar identidade conceitual. Quando não foi identificado termo externo adequado ao significado requerido pelo modelo, manteve-se a classe ou a propriedade local, acompanhada de definição, fonte e justificativa de modelagem.

4.1.6 Formalização da OntoDiplomaDigital - Etapa 5 da OntoForInfoScience

A formalização converteu os artefatos conceituais em uma representação computável, consolidada no arquivo canônico OntoDiplomaDigitalregistrodiploma.ttl[8]. Utilizaram-se RDF e RDFS para a estrutura básica do grafo e da hierarquia, OWL para classes, propriedades e restrições, SKOS para os vocabulários controlados e Turtle como formato de serialização. O desenvolvimento ocorreu de forma incremental, com revisão das declarações após cada ciclo de alterações.

[8] Os arquivos da ontologia em formato Turtle (.ttl), incluindo a TBox canônica e a ABox utilizada na prova de conceito, encontram-se disponibilizados no repositório GitHub do projeto. A relação dos arquivos, sua organização e os respectivos endereços de acesso são apresentados no Apêndice B.

Para compreender como essa ontologia foi organizada, é necessário compreender os conceitos de TBox e ABox. No presente projeto, a TBox corresponde principalmente ao arquivo canônico OntoDiplomaDigital_registro_diploma.ttl. Esse arquivo contém o modelo da ontologia: nele são definidas classes como DiplomaDigital, Diplomado, Curso e InstituicaoEnsinoSuperior, além das propriedades, relações e restrições que estabelecem como esses elementos podem ser descritos e relacionados. A TBox informa, por exemplo, que um diploma pode ter um titular e que esse titular deve ser interpretado como um diplomado, mas não apresenta os dados de um diplomado específico.

A ABox corresponde à aplicação desse modelo aos indivíduos e aos seus dados. No projeto, ela foi mantida em arquivo separado, denominado ontologia_operacional_individuos_ficticios_registro_diploma.ttl. Esse arquivo utiliza as classes e propriedades definidas na TBox para representar casos concretos da prova de conceito: um determinado diploma, o diplomado ao qual ele pertence, o curso realizado, as instituições emissora e registradora, o registro e as demais informações relacionadas. Os dados empregados são representativos e foram preparados para a validação, não devendo ser confundidos com a exposição de dados pessoais reais. Em termos simples, a TBox é o arquivo que define como os dados do Diploma Digital devem ser organizados; a ABox é o arquivo que usa essa organização para expressar os dados dos diplomas e dos diplomados representados na prova de conceito. Quando os dois arquivos são reunidos no GraphDB, o modelo e seus indivíduos formam o grafo de conhecimento operacional.

O procedimento seguiu os sete passos da Etapa 5 da OntoForInfoScience descritos por Mendonça (2015, p. 214-244):

1. **Construção da taxonomia geral:** Os conceitos dos modelos foram examinados pela pergunta “é um tipo de?”. A relação rdfs:subClassOf foi utilizada apenas quando havia especialização genuína. A taxonomia foi reservada à hierarquia de classes; relações, atributos e valores controlados não foram apresentados como ramos taxonômicos.
2. **Definição das propriedades descritivas das classes:** As classes receberam identificadores, rótulos, definições, comentários, notas de escopo, sinônimos quando documentados, exemplos, fontes e justificativas. Essas anotações registraram o significado e o uso pretendido, sem serem confundidas com axiomas lógicos.

3. **Criação das definições formais:** Foram avaliadas relações de subclasse, disjunções, domínios, alcances, cardinalidades e restrições OWL. As restrições somente foram introduzidas quando sustentadas pelo recorte do domínio e pelas fontes. Condições necessárias expressas por rdfs:subClassOf foram distinguidas de definições necessárias e suficientes, que exigiriam axiomas de equivalência devidamente justificados.
4. **Definição das propriedades de dados:** Os atributos literais receberam domínio quando a inferência resultante era desejável e alcance compatível com o tipo de dado. A funcionalidade foi declarada apenas quando o modelo exigia no máximo um valor por sujeito. A origem XSD foi registrada por caminho, elemento e tipo sempre que disponível.
5. **Criação das instâncias:** A TBox foi mantida separada da ABox operacional. No arquivo ontológico permaneceram os indivíduos metodológicos e os conceitos controlados necessários à documentação do modelo; os indivíduos correspondentes aos dados do domínio foram gerados em artefato próprio durante a prova de conceito.
6. **Especificação das relações ontológicas:** As relações foram reutilizadas de vocabulários externos ou criadas localmente a partir do dicionário de verbos e dos modelos conceituais. A escolha considerou significado, direção, identidade dos recursos participantes e necessidade de navegação no grafo.
7. **Definição das propriedades das relações:** Para cada propriedade de objeto foram avaliados rótulo, definição, domínio, alcance, inversa, superpropriedade, funcionalidade, fonte e justificativa. Características lógicas adicionais não foram declaradas sem evidência de que se aplicavam universalmente à relação.

A criação dos conceitos SKOS seguiu procedimento específico. Primeiro, identificaram-se valores cujo significado deveria permanecer estável e compartilhado, evitando variações textuais livres. Em seguida, esses valores foram agrupados em esquema de conceitos e receberam *Internationalized Resource Identifiers* (IRIs) persistentes, skos:prefLabel, variantes terminológicas quando comprovadas, definição, nota de escopo, fonte e vínculo com o esquema. Os conceitos concretos foram tratados como instâncias de skos:Concept e das categorias locais correspondentes. Sua ligação com diplomas, cursos ou assinaturas foi realizada por propriedades de objeto especializadas da propriedade abstrata hasControlledValue, e não por propriedades de dados.

O arquivo TTL foi mantido como fonte canônica da formalização. As tabelas, glossários e diagramas derivados foram regenerados a partir dele ou comparados com ele, evitando que representações secundárias passassem a governar a ontologia. As alterações foram acompanhadas por registros de decisões e por verificações de impacto sobre TBox, ABox, consultas e documentação. Os procedimentos computacionais complementares à formalização são detalhados na Seção 4.2, que apresenta a preparação dos dados sintéticos, o mapeamento RDF no Ontotext Refine[9], a configuração e a carga do repositório no GraphDB[10] e a integração entre essas ferramentas. Esses procedimentos viabilizaram a materialização da ABox e sua integração com a TBox para a execução da prova de conceito.

4.1.7 Avaliação da ontologia - Etapa 6 da OntoForInfoScience

A Etapa 6 da OntoForInfoScience correspondeu à avaliação da OntoDiplomaDigital. Essa etapa foi definida antes da apresentação dos resultados e teve por finalidade examinar se a ontologia atendia aos objetivos da pesquisa, às questões de competência e aos requisitos definidos para o subdomínio do Registro do Diploma. A avaliação foi organizada em três dimensões complementares: avaliação, validação e verificação. A avaliação considerou a adequação global da ontologia aos objetivos e ao recorte da pesquisa. A validação teve caráter interno e documental, sendo realizada pela confrontação da representação construída com as fontes normativas, técnicas e conceituais utilizadas na modelagem e com os requisitos definidos na especificação. Não foi realizada validação externa por especialistas do domínio ou por instituições usuárias do Diploma Digital. A verificação incidiu sobre a correção sintática, estrutural e operacional da implementação. Essa distinção segue Mendonça (2015, p. 244-250), para quem a validação se refere à adequação da ontologia ao domínio representado, enquanto a verificação examina se sua construção foi realizada corretamente.

Foram adotados como parâmetros de avaliação a aderência ao subdomínio delimitado, a cobertura das questões de competência, a coerência entre classes,

[9] Ontotext Refine: Ferramenta desenvolvida pela Ontotext para limpeza, transformação e mapeamento de dados estruturados para RDF, com integração ao GraphDB. Disponível em: https://platform.ontotext.com/ontorefine/. Acesso em: 1 out. 2026.
[10] GraphDB: Sistema de gerenciamento de bancos de dados em grafo com suporte a RDF e SPARQL, utilizado para armazenamento e consulta de dados semânticos. Disponível em: https://graphdb.ontotext.com/documentation/10.1/. Acesso em: 1 out. 2026.

propriedades, indivíduos e conceitos controlados, a consistência de domínios, alcances e restrições, a rastreabilidade até as fontes e a processabilidade do artefato por ferramentas compatíveis com os padrões da Web Semântica. Esses parâmetros orientaram a análise da TBox, dos vocabulários controlados, da ABox utilizada na prova de conceito e das consultas SPARQL elaboradas a partir das questões de competência.

Para examinar a ontologia em uso, foi definida uma prova de conceito baseada na aplicação da TBox OntoDiplomaDigital_registro_diploma.ttl a uma ABox composta por dados sintéticos. A fonte de dados foi preparada em formato tabular e mapeada no Ontotext Refine para as classes, propriedades de objeto, propriedades de dados e conceitos controlados previamente definidos na TBox. Esse mapeamento incluiu regras para geração de IRIs, criação de indivíduos e estabelecimento de relações entre recursos. O procedimento foi tratado como instanciação de um modelo já formalizado, não como mecanismo para redefinir retroativamente a estrutura da ontologia.

A TBox e a ABox foram carregadas no GraphDB, que constituiu o ambiente de execução das consultas SPARQL vinculadas às questões de competência. A recuperação dos recursos, propriedades e relações esperados foi adotada como critério de aceitação da prova de conceito. O protocolo de verificação incluiu a análise sintática dos arquivos Turtle, a conferência de prefixos e declarações, a inspeção de classes, propriedades, domínios, alcances, restrições e relações inversas, a comparação entre os artefatos canônicos e a recarga controlada no repositório.

O protocolo de validação funcional incluiu a execução das consultas vinculadas às questões de competência, a verificação da presença dos recursos e relações esperados, consultas negativas destinadas à identificação de representações residuais, testes de regressão após alterações na TBox ou na ABox e a conferência da correspondência semântica entre sujeito, propriedade e objeto. A comparação entre representações RDF não foi condicionada exclusivamente à igualdade da contagem global de triplas, uma vez que inferências, relações inversas, tipagens e anotações podem produzir grafos ou serializações com quantitativos distintos sem comprometer a representação requerida. O critério principal de avaliação foi o atendimento aos requisitos semânticos definidos para o recorte da pesquisa e às questões de competência estabelecidas na especificação da ontologia.

4.1.8 Documentação da ontologia - Etapa 7 da OntoForInfoScience

A Etapa 7 da OntoForInfoScience correspondeu à documentação da OntoDiplomaDigital. Embora apareça como etapa específica da metodologia, a documentação foi conduzida de forma transversal ao desenvolvimento da ontologia, acompanhando as decisões tomadas desde a especificação inicial até a avaliação da prova de conceito. Essa organização permitiu registrar as fontes consultadas, os critérios de modelagem, as alterações realizadas, os artefatos produzidos e as justificativas das escolhas semânticas adotadas ao longo da pesquisa.

Foram preservados os arquivos de origem, os glossários, os dicionários, os modelos conceituais, os arquivos PlantUML, o código Turtle, a especificação semântica, os registros de decisões, os históricos de processamento, as consultas SPARQL e os registros de validação. A documentação foi organizada para manter a relação entre os schemas XSD do Ministério da Educação, as interpretações conceituais realizadas na pesquisa e os elementos formalizados na ontologia.

A rastreabilidade foi estruturada em dois sentidos complementares. No primeiro, classes, propriedades, restrições e valores controlados foram relacionados, sempre que possível, aos elementos, tipos e caminhos correspondentes nos XSDs e nos documentos de origem. No segundo, os artefatos derivados foram vinculados ao TTL canônico e às decisões de modelagem vigentes. No arquivo Turtle, essa rastreabilidade foi registrada por meio de propriedades e anotações destinadas a documentar fonte, escopo, justificativa de modelagem, correspondência com a estrutura XSD e notas de uso.

A documentação também teve a função de preservar a memória do projeto. Mudanças em classes, propriedades, restrições ou dados sintéticos foram acompanhadas da revisão dos artefatos afetados e da execução de verificações correspondentes. Fragilidades, lacunas e ambiguidades identificadas nas fontes oficiais foram registradas juntamente com as decisões adotadas para sua representação. Esse procedimento reduziu a dependência do conhecimento tácito do pesquisador e permitiu que o percurso de construção da ontologia pudesse ser auditado, revisado e reutilizado em atualizações futuras. A documentação consolidada desse percurso é apresentada na na Seção 5.7, e complementada pelos artefatos computacionais, pelas orientações de reprodução e pelos relatórios de verificação

disponibilizados no repositório público da pesquisa, cujo acesso é indicado no Apêndice B.

A comunicação acadêmica da ontologia foi realizada por meio da dissertação, dos quadros explicativos, dos modelos conceituais gráficos, dos trechos de código Turtle, dos exemplos de leitura das triplas e dos apêndices. Esses recursos traduziram a estrutura formal da ontologia para uma forma compreensível a pesquisadores, profissionais da informação, gestores acadêmicos e leitores não especializados em engenharia de ontologias, sem substituir o arquivo computacional processável que formaliza o modelo.

4.1.9 Disponibilização da ontologia - Etapa 8 da OntoForInfoScience

No âmbito desta pesquisa, a Etapa 8 da OntoForInfoScience foi interpretada como disponibilização acadêmica dos artefatos produzidos, e não como implantação operacional da ontologia em ambiente público de produção. Essa decisão decorreu do escopo da dissertação, que teve por finalidade propor, formalizar e demonstrar a OntoDiplomaDigital por meio de prova de conceito, sem assumir a manutenção de um endpoint SPARQL permanente, a publicação institucional de URIs dereferenciáveis ou uma política continuada de governança em rede. Assim, a disponibilização foi planejada por meio de repositório público no GitHub[11], no qual foram reunidos o arquivo Turtle da ontologia, a ABox com instâncias fictícias, os arquivos tabulares, as consultas SPARQL, as consultas federadas demonstrativas, os registros de validação e a documentação de reprodução da prova de conceito.

A preparação para disponibilização considerou requisitos de representação em formato lógico processável, atribuição de IRI, inclusão de metadados, documentação de versão, indicação de licença e definição de mecanismos de acesso. Para esse fim, foram utilizados Turtle, vocabulários padronizados e identificadores persistentes como requisitos técnicos compatíveis com a publicação e o reúso em ambientes de Web Semântica e Dados Ligados.

As URIs e referências a endereços eletrônicos presentes no arquivo TTL foram mantidas como representação do cenário de publicação da ontologia. Esses elementos demonstram como recursos, metadados e artefatos podem ser identificados em ambiente Web, mas não devem ser interpretados como evidência de

[11] O repositório público da prova de conceito da OntoDiplomaDigital encontra-se disponível em: https://github.com/renanteixeirasouza/ontodiplomadigital-proof-of-concept/releases/tag/v1.0.

que todos os recursos estejam efetivamente publicados, resolvíveis ou disponíveis em infraestrutura operacional. A disponibilização pública definitiva, incluindo hospedagem, resolução persistente das URIs, controle de versões e política de manutenção, permanece como etapa posterior ao trabalho.

Com essa delimitação, a Etapa 8 encerra o percurso metodológico da OntoForInfoScience nesta pesquisa sem ampliar o alcance da prova de conceito. A dissertação entrega a ontologia formalizada, documentada e preparada para disponibilização futura, mas não afirma implantação institucional, publicação pública operacional ou validação externa em ambiente de produção.

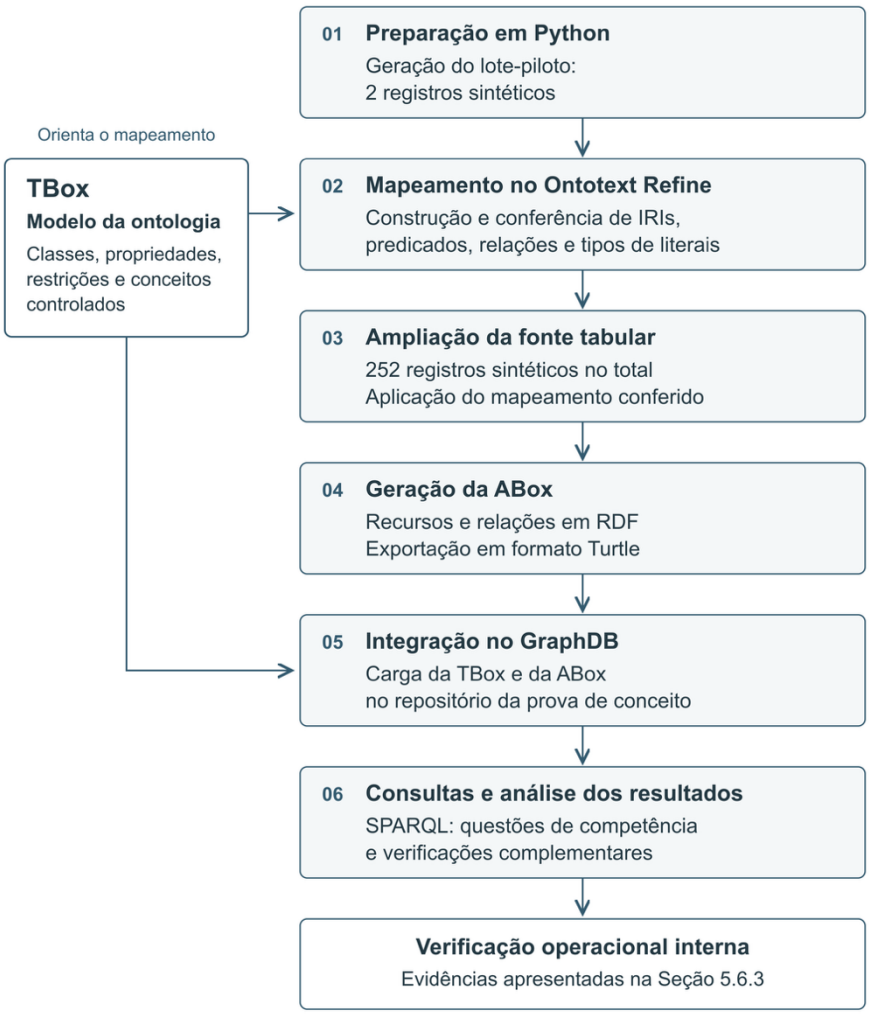
4.2 Procedimentos complementares de implementação e validação computacional

Esta seção foi incluída para facilitar a compreensão, inclusive para leitores não especializados, da transição entre a modelagem conceitual da ontologia e sua implementação computacional. A OntoForInfoScience orientou a delimitação do domínio, a organização do conhecimento e a formalização da TBox, definida no arquivo canônico OntoDiplomaDigital_registro_diploma.ttl, porém não especifica os procedimentos tecnológicos necessários para a transformação dos dados em instâncias RDF nem para sua consulta em um ambiente de grafos.

Por essa razão, foram utilizados procedimentos complementares: o Ontotext Refine realizou o mapeamento da fonte tabular para os elementos previamente definidos na TBox, gerando a representação RDF correspondente à ABox da prova de conceito, enquanto o GraphDB integrou TBox e ABox e possibilitou a execução das consultas SPARQL. Essa divisão de funções permitiu que a OntoForInfoScience sustentasse a construção conceitual da ontologia, enquanto as ferramentas computacionais viabilizaram sua materialização e exploração operacional.

A Figura 5 sintetiza os procedimentos computacionais empregados na prova de conceito, desde a preparação dos dados sintéticos e a construção do mapeamento RDF até a integração da TBox e da ABox no GraphDB e a execução das consultas SPARQL. A TBox forneceu a referência para o mapeamento, enquanto a ABox reuniu os recursos e as relações gerados a partir da fonte tabular.

Figura 5 - Fluxo dos procedimentos computacionais da prova de conceito da OntoDiplomaDigital



Fonte: Elaborado pelo autor (2026).

O fluxo mostra como dados organizados em uma tabela passam a compor uma representação com significado explícito. A TBox define os conceitos e as relações do domínio, como diploma, diplomado e instituição. Com base nesse modelo, o Ontotext Refine transforma os dados sintéticos da tabela em recursos relacionados, formando a ABox. No GraphDB, modelo e dados são reunidos para consulta. As consultas SPARQL verificam se as informações e relações previstas podem ser recuperadas. Esse processo ilustra uma base da interoperabilidade semântica: representar dados por conceitos e relações compartilhados, permitindo interpretar seu significado além da estrutura da tabela.

4.2.1 Preparação e normalização da fonte tabular

A fonte tabular utilizada na prova de conceito foi produzida por meio de código desenvolvido em Python[12], tomando como referência as classes, as propriedades, os tipos de dados e os conceitos controlados definidos na TBox canônica da OntoDiplomaDigital. A geração foi organizada em duas fases. Inicialmente, foram produzidos dois registros integralmente sintéticos, empregados como lote-piloto para verificar a correspondência entre os campos tabulares e os elementos previstos na ontologia antes da aplicação do procedimento ao conjunto ampliado.

Os registros-piloto foram importados no Ontotext Refine para a construção e a conferência do mapeamento RDF. Nessa etapa, foram examinadas a formação dos Identificadores Internacionalizados de Recursos (Internationalized Resource Identifiers — IRIs), a correspondência entre os predicados e as propriedades da TBox, a distinção entre propriedades de dados e propriedades de objeto, a associação aos conceitos SKOS, a tipagem dos literais e o tratamento dos campos sem valor. Essa conferência teve caráter operacional e destinou-se a verificar a adequação do mapeamento à estrutura formal definida na ontologia.

Após a conferência do lote-piloto, o mesmo código foi empregado para gerar um conjunto ampliado de registros sintéticos segundo a estrutura tabular e as regras definidas para a simulação. Nenhum dos registros foi derivado de dados acadêmicos de pessoas naturais; por essa razão, a fonte utilizada na prova de conceito foi caracterizada como integralmente sintética, e não como anonimizada. Cada linha da tabela foi concebida como unidade de entrada para o mapeamento e reuniu dados capazes de originar diferentes recursos RDF relacionados, como diploma, diplomado, curso, instituições, registro, assinaturas e código de validação, de modo que a quantidade de registros tabulares não corresponde diretamente à quantidade de indivíduos posteriormente materializados na ABox.

O Quadro 22 sintetiza as etapas de preparação da fonte tabular, desde a geração do lote-piloto e a conferência do mapeamento até a produção do conjunto ampliado destinado à materialização RDF. Os quantitativos correspondentes à fonte

[12] Python é uma linguagem de programação de alto nível, de código aberto e de propósito geral, amplamente utilizada em automação, processamento e análise de dados, desenvolvimento científico e implementação de aplicações computacionais. Disponível em: https://www.python.org/. Acesso em: 14 ago. 2026.

efetivamente produzida e aos recursos resultantes da materialização são apresentados no capítulo de Resultados.

Quadro 22 - Etapas de preparação da fonte tabular

Etapa	Atividades principais
1. Modelagem conceitual	Delimitação do domínio e dos requisitos; definição das questões de competência. Saída: modelo conceitual do Registro do Diploma.
2. Construção da TBox	Formalização de classes, propriedades, hierarquias, domínios, alcances, restrições e anotações; geração do arquivo canônico em Turtle e carga no GraphDB. Saída: OntoDiplomaDigital_registro_diploma.ttl (TBox).
3. Preparação do ambiente e da fonte de dados	Geração da fonte sintética em Python, composta por 252 registros; preparação e alinhamento dos dados para o mapeamento no Ontotext Refine. Saída: ambiente de mapeamento preparado e fonte alinhada à TBox.
4. Geração da ABox	Mapeamento dos campos para classes, propriedades e conceitos; criação de indivíduos e relações conforme a TBox; geração das triplas RDF e materialização em Turtle. Saída: ABox com 252 registros instanciados.
5. Integração TBox + ABox	Carga da ABox no GraphDB e integração com a TBox para composição do grafo. Saída: grafo de conhecimento integrado no GraphDB.
6. Consulta e exploração	Execução de consultas SPARQL; validação de instâncias e relações; recuperação de informações associadas às questões de competência. Saída: informações recuperadas e validadas.

Fonte: Elaborado pelo autor (2026)

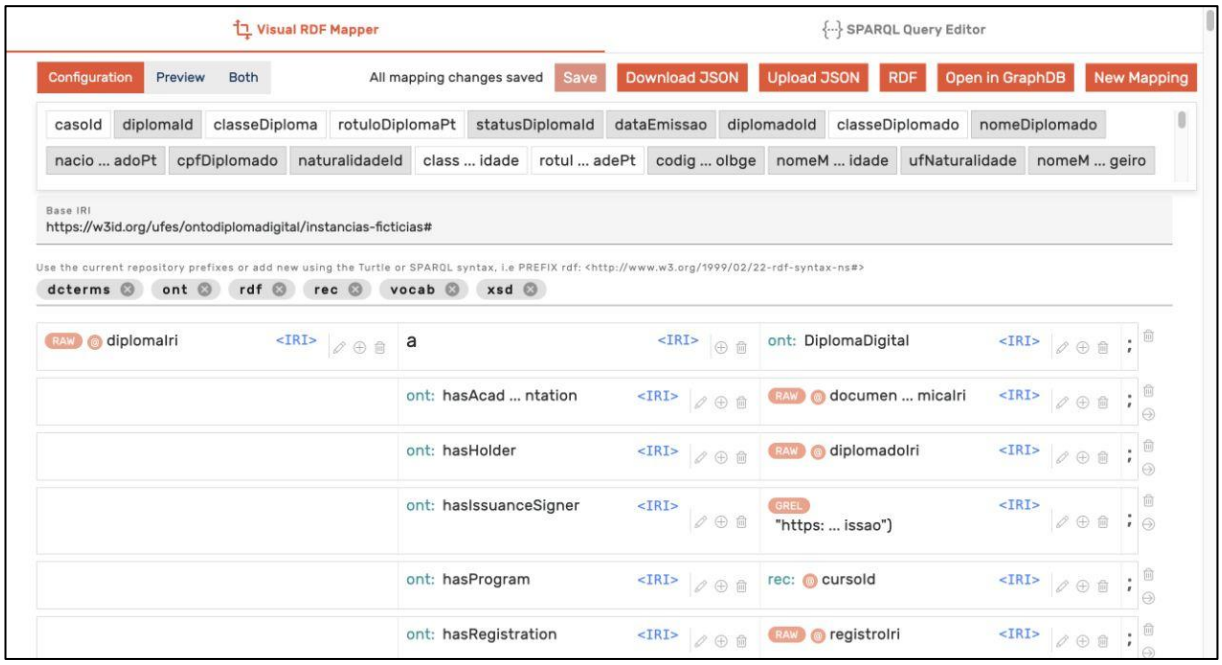
4.2.2 Construção do mapeamento RDF no Ontotext Refine

O mapeamento RDF foi construído no Ontotext Refine a partir dos registros sintéticos do lote-piloto. A TBox canônica da OntoDiplomaDigital, previamente carregada no GraphDB, forneceu as classes, propriedades e conceitos controlados utilizados como referência. No mapeamento, cada campo da fonte tabular foi associado ao elemento correspondente da ontologia: campos descritivos foram vinculados a propriedades de dados, vínculos entre entidades a propriedades de objeto e valores classificatórios a conceitos SKOS. Também foram definidas regras para a construção das IRIs e para a tipagem dos valores.

A Figura 6 apresenta um recorte do mapeamento visual no Ontotext Refine. Nessa interface, os campos da fonte tabular são associados às classes e propriedades

da OntoDiplomaDigital. Por exemplo, o campo que identifica o diplomado é utilizado para estabelecer a relação *hasHolder* (“tem titular”), vinculando o recurso que representa o diploma ao recurso que representa seu titular.

Figura 6 - Recorte do mapeamento RDF no Ontotext Refine

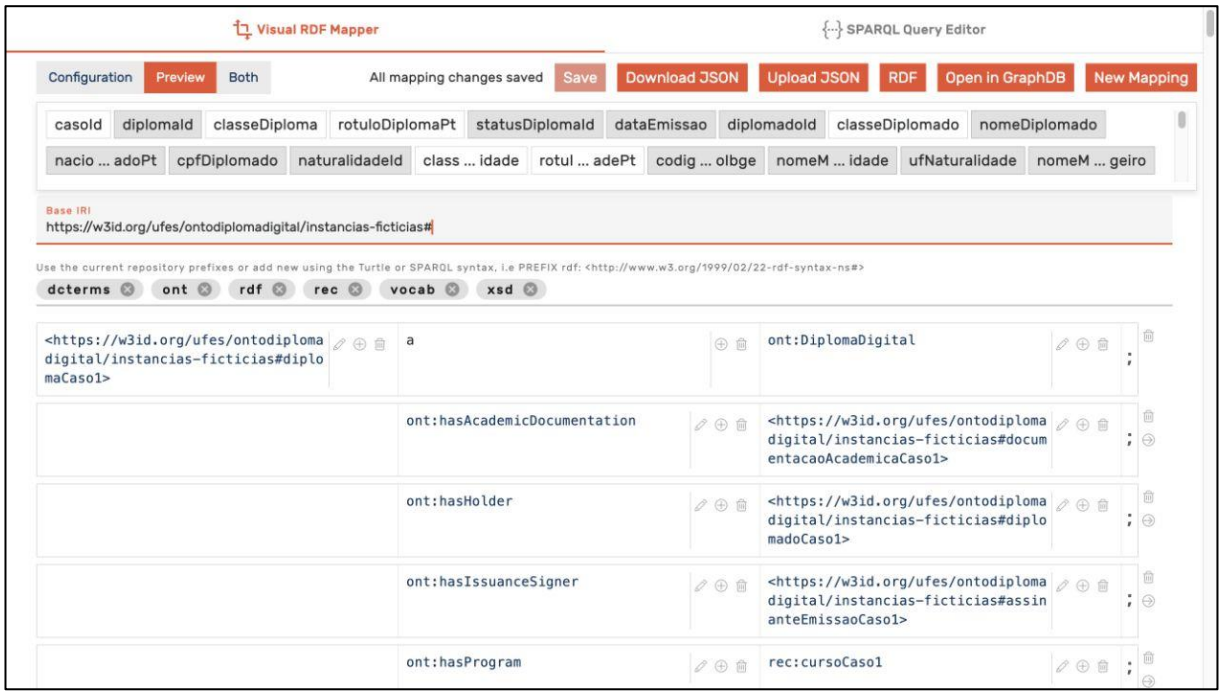


Fonte: Elaborado pelo autor (2026), a partir do ambiente da prova de conceito

A visualização prévia do RDF no Ontotext Refine foi utilizada para conferir a formação dos identificadores, os predicados selecionados, os tipos dos objetos e o tratamento dos campos sem valor. Após essa conferência, o mesmo mapeamento foi aplicado à fonte tabular ampliada da prova de conceito. A execução desse mapeamento produziu uma representação RDF serializada em Turtle contendo as instâncias e suas relações. Essa representação RDF corresponde à ABox da prova de conceito, posteriormente carregada no GraphDB em conjunto com a TBox. Os quantitativos da ABox e os resultados das consultas SPARQL são apresentados no capítulo de Resultados.

A Figura 7 apresenta a prévia RDF correspondente ao mapeamento. Enquanto a configuração define as regras de transformação, a prévia mostra os identificadores e as relações produzidos a partir de um registro sintético. Essa visualização permitiu conferir se o diploma estava associado aos recursos previstos, como seu titular, sua documentação acadêmica e seu curso, antes da exportação da ABox.

Figura 7 - Prévia RDF de um registro sintético no Ontotext Refine



Fonte: Elaborado pelo autor (2026), a partir do ambiente da prova de conceito.

4.2.3 Configuração e carga do repositório no GraphDB

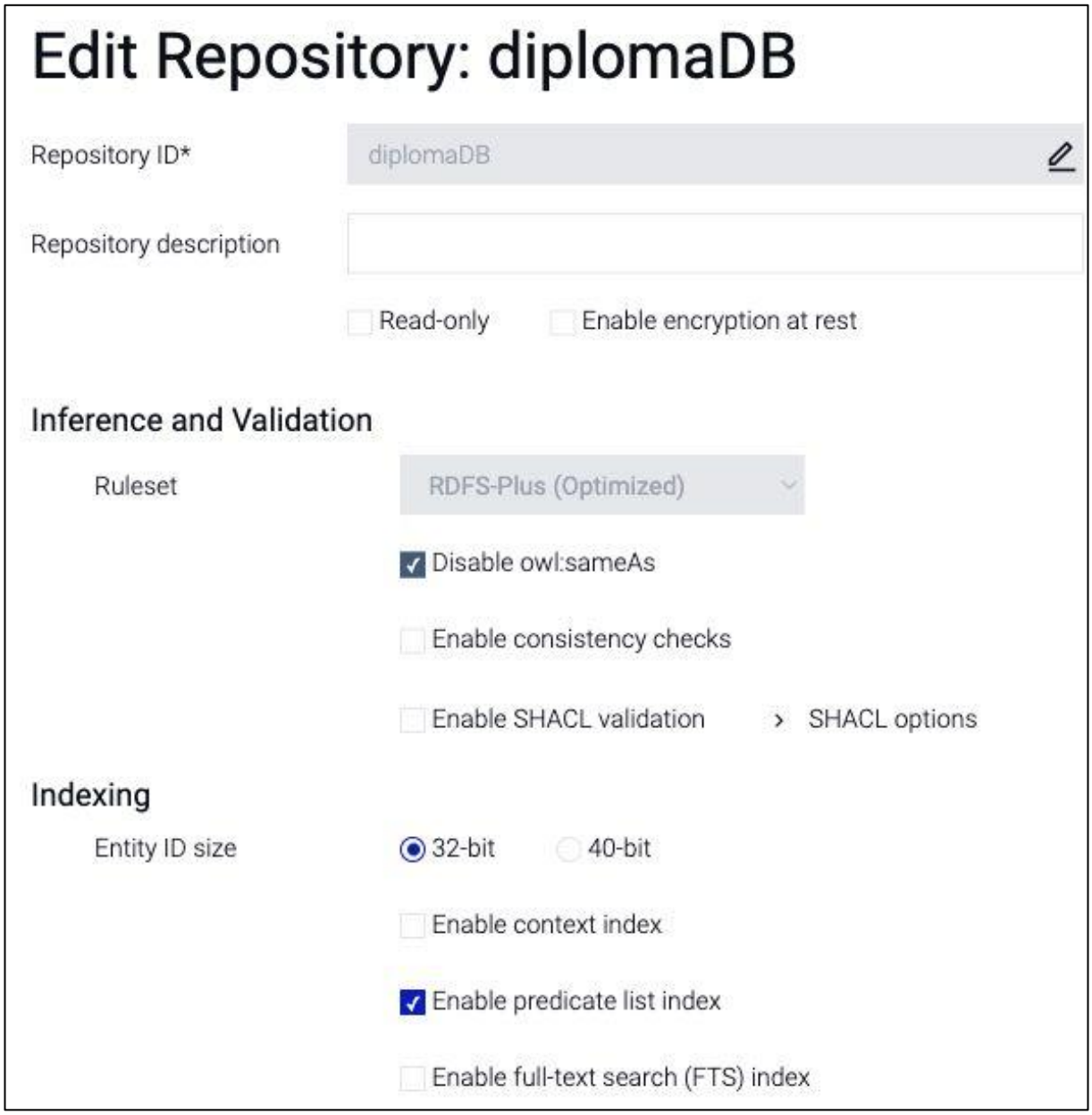
Os procedimentos adotados apoiaram-se nos estudos dirigidos e nos materiais produzidos durante as disciplinas do Programa de Pós-Graduação em Ciência da Informação, especialmente nos conteúdos relacionados a dados ligados, ontologias operacionais, organização de repositórios RDF e consultas SPARQL (CRISTÓVÃO, 2026). Esses referenciais complementaram a orientação conceitual da OntoForInfoScience quanto à implementação computacional dos artefatos e à preparação do ambiente destinado à prova de conceito.

Foi criado um repositório específico para a OntoDiplomaDigital, organizado para receber separadamente os arquivos correspondentes à TBox e à ABox. A TBox constituiu a camada terminológica e axiomatizada do domínio, reunindo classes, propriedades, hierarquias, restrições e vocabulários controlados, enquanto a ABox correspondeu à camada destinada às instâncias e às afirmações derivadas da materialização dos dados. A separação entre esses componentes foi adotada para permitir o controle independente dos artefatos e facilitar eventuais atualizações da estrutura ontológica ou dos dados instanciados. Como parte da preparação do repositório, foram definidos procedimentos de conferência da integridade sintática dos arquivos, do reconhecimento dos namespaces, da presença das classes e

propriedades requeridas e da vinculação dos indivíduos aos elementos previstos na ontologia.

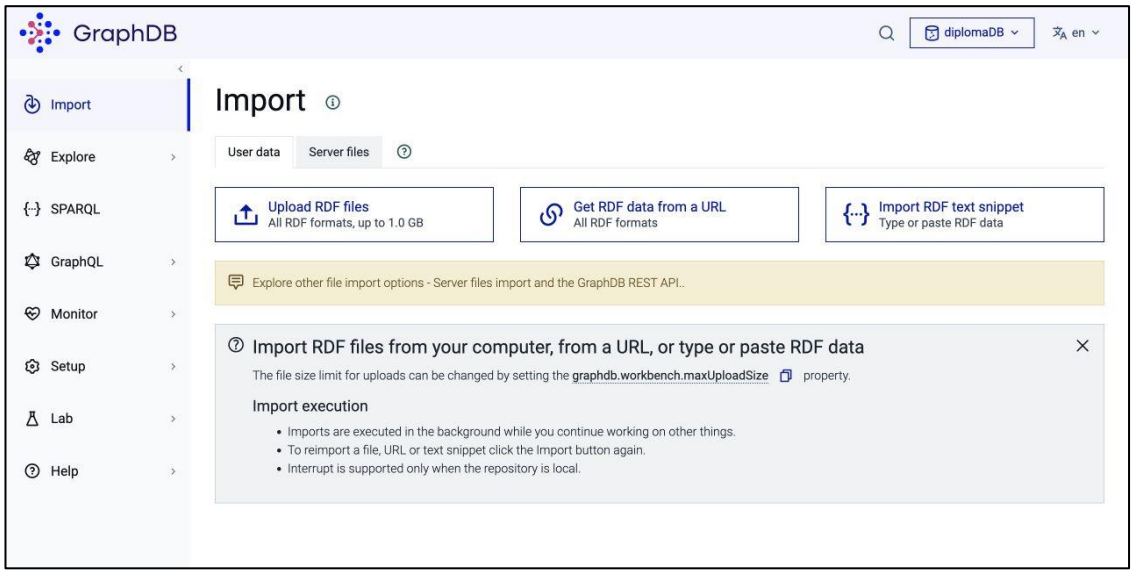
A Figura 8 apresenta a configuração do repositório diplomaDB no GraphDB, incluindo os parâmetros de inferência, indexação e consulta. O repositório constituiu o ambiente de armazenamento e integração do modelo ontológico e dos dados sintéticos utilizados na prova de conceito.

Figura 8 - Configuração do repositório diplomaDB no GraphDB



Fonte: Captura de tela realizada pelo autor no GraphDB (2026)..

Figura 9 - Interface de importação de arquivos RDF no GraphDB



Fonte: Captura de tela realizada pelo autor no GraphDB (2026

A Figura 9 apresenta a interface de importação de arquivos RDF no GraphDB. Essa interface permite selecionar arquivos, como os documentos Turtle correspondentes à TBox e à ABox, para carregamento no repositório ativo. A imagem ilustra o mecanismo de carga, sem representar o registro de uma importação concluída.

Com o repositório configurado, o GraphDB foi utilizado como ambiente para integrar a TBox e a ABox e executar as consultas SPARQL previstas no protocolo de avaliação. As consultas foram estruturadas a partir das questões de competência, com o objetivo de examinar a recuperação de informações e os caminhos semânticos estabelecidos entre diplomas, diplomados, cursos, instituições, registros e os demais recursos representados. O ambiente também foi empregado para inspeção do grafo e conferência das relações produzidas pelo mapeamento, permitindo verificar eventuais divergências entre a estrutura formalizada e os dados instanciados. Os resultados dessas verificações, bem como os retornos obtidos nas consultas, são apresentados no capítulo de Resultados.

A integração entre GraphDB e Ontotext Refine foi estabelecida com o objetivo de mapear os dados tabulares para RDF utilizando classes e propriedades definidas pela TBox. O GraphDB forneceu a estrutura ontológica, enquanto o Ontotext Refine aplicou as regras de transformação, referenciando as classes e propriedades da ontologia. A representação RDF foi minuciosamente examinada antes da exportação para garantir a correspondência e o tratamento adequado dos dados.

4.2.4 Geração da ABox e validação no GraphDB

Concluída a conferência operacional do mapeamento com o lote-piloto, a configuração definida no Ontotext Refine foi aplicada à fonte tabular ampliada destinada à prova de conceito. O processamento seguiu as regras previamente estabelecidas para a construção dos recursos RDF, incluindo a geração das IRIs, a atribuição de propriedades de dados e de objeto, a vinculação aos conceitos controlados e a tipagem dos literais. A execução do mapeamento produziu uma representação RDF, posteriormente exportada em formato Turtle, correspondente à ABox utilizada na etapa de verificação computacional.

A ABox foi carregada no GraphDB em conjunto com a TBox canônica da OntoDiplomaDigital, constituindo o conjunto de dados utilizado nas consultas SPARQL definidas a partir das questões de competência e dos controles de regressão. As consultas tiveram como finalidade verificar a recuperação dos recursos e das relações previstos no modelo no ambiente controlado da prova de conceito. Os resultados das verificações realizadas são apresentados na seção 5.6.3.

4.2.5 Prova de conceito

A prova de conceito foi delimitada como verificação operacional interna, realizada com base artificial. Por essa razão, seus resultados não foram tratados como validação externa, teste com dados institucionais reais ou avaliação de desempenho em escala nacional.

Para essa verificação, foi preparada uma base tabular artificial, gerada especificamente para a pesquisa, sem utilização de dados reais de estudantes ou de instituições. A construção ocorreu em duas etapas. Primeiro, foram produzidos dois registros para a conferência inicial do mapeamento RDF no Ontotext Refine. Esse lote-piloto foi utilizado para verificar a formação das IRIs, a associação entre colunas e predicados, os tipos de objetos, os literais e as transformações aplicadas aos dados.

Após essa conferência, foram gerados 250 registros adicionais, totalizando 252 registros na fonte tabular utilizada na prova de conceito. Esse quantitativo não teve finalidade amostral ou estatística; foi definido apenas para verificar a aplicação repetida do mapeamento sobre um conjunto ampliado de entradas. Cada registro da planilha representou uma unidade de entrada do mapeamento, a partir da qual foram produzidos diferentes recursos e relações em RDF.

O mapeamento RDF conferido no lote-piloto foi aplicado ao conjunto completo no Ontotext Refine, gerando um arquivo RDF serializado em Turtle. Esse arquivo constituiu a ABox da prova de conceito. Em seguida, a ABox foi carregada no GraphDB juntamente com a TBox da OntoDiplomaDigital, permitindo a execução das consultas SPARQL elaboradas a partir das sete questões de competência. Definidas como requisitos informacionais na etapa de especificação, conforme descrito na Seção 4.1.2, essas questões orientaram a construção da ontologia e a verificação da recuperação dos recursos e das relações previstos no modelo. Sua apresentação encontra-se na Seção 5.1.5, no Quadro 28, e os resultados das consultas executadas com dados sintéticos são apresentados na Seção 5.6.3.

5 Resultados e Discussão

Este capítulo apresenta e discute os resultados obtidos no desenvolvimento da OntoDiplomaDigital. A organização acompanha as etapas da OntoForInfoScience adotadas na pesquisa, de modo que cada seção corresponde a uma fase do percurso metodológico e evidencia os produtos gerados, as decisões tomadas e sua relação com os objetivos da dissertação.

A discussão não é concentrada apenas ao final do capítulo. Sempre que necessário, os resultados são analisados no próprio ponto em que são apresentados, permitindo explicitar sua função no processo de construção da ontologia, sua relação com as fontes utilizadas e sua contribuição para a representação semântica do subdomínio do Registro do Diploma Digital.

Ao final, a seção Discussão Integrada dos Resultados retoma os achados de maneira transversal, articulando as etapas do processo, os fundamentos teóricos, o Estado da Arte, os princípios FAIR e as contribuições da ontologia para a interoperabilidade semântica. Essa seção não tem a função de repetir isoladamente os resultados de cada etapa, mas de interpretar o conjunto do percurso desenvolvido.

Para apoiar a leitura dos termos e das representações empregados ao longo deste capítulo, os namespaces e seus respectivos prefixos utilizados na dissertação são apresentados de forma consolidada no Apêndice A. O Apêndice B reúne as orientações de acesso aos artefatos computacionais e documentais da prova de conceito, permitindo consultar as evidências que sustentam os resultados apresentados nas seções seguintes.

5.1 Resultado da Especificação da Ontologia – Etapa 1 da OntoForInfoScience

Esta seção apresenta a especificação da OntoDiplomaDigital. As subseções detalham os resultados informacionais relativos à identificação e à caracterização da ontologia, abrangendo domínio e escopo, tipologia, grupos de usuários, cenários de aplicação, questões de competência e alinhamento aos objetivos específicos da pesquisa.

A especificação teve como finalidade definir os requisitos iniciais da OntoDiplomaDigital e orientar as etapas subsequentes da metodologia. Por esse motivo, os resultados aqui apresentados não se limitam à descrição administrativa do

artefato, mas refletem as decisões que delimitaram o que deveria ser modelado, formalizado e, posteriormente, avaliado.

5.1.1 Identificação e caracterização da ontologia

A especificação resultou na identificação inicial da OntoDiplomaDigital como uma ontologia operacional de domínio, delimitada ao subdomínio do Registro do Diploma Digital Brasileiro. Essa caracterização definiu a identidade do artefato, sua finalidade, sua URI de referência, o namespace adotado, os metadados descritivos, o idioma principal, a licença e os padrões de representação previstos para sua formalização. O Quadro 23 sintetiza esses elementos e indica, quando aplicável, sua correspondência com os termos utilizados posteriormente no arquivo TTL.

Quadro 23 - Especificação da OntoDiplomaDigital

Campo	Conteúdo	Termo correspondente no TTL canônico
Nome adotado	OntoDiplomaDigital	dcterms:title e rdfs:label
URI da ontologia	https://w3id.org/ufes/ontodiplomadigital	Recurso declarado como owl:Ontology
Tipo de recurso	Ontologia operacional de domínio	rdf:type owl:Ontology e dcterms:type
URI base da ontologia	https://w3id.org/ufes/ontodiplomadigital	owl:Ontology
Prefixo preferencial	odd	vann:preferredNamespacePrefix
URI do namespace	https://w3id.org/ufes/ontodiplomadigital#	vann:preferredNamespaceUri
Descrição	Modelo organizado do Diploma Digital Brasileiro, destinado a apoiar a compreensão, a consulta, a manutenção, o reúso e a evolução do domínio.	dcterms:description
Assunto	Diploma Digital Brasileiro; Ontologias; Web Semântica; Dados Abertos Ligados	dcterms:subject
Autor principal	Renan Teixeira de Souza	dcterms:creator
Instituição responsável	Programa de Pós-Graduação em Ciência da Informação (PPGCI/UFES)	dcterms:publisher
Data de criação	20 de abril de 2026	dcterms:created
Última modificação	23 de julho de 2026	dcterms:modified
Versão atual	OntoDiplomaDigital_registro_diploma.ttl	owl:versionInfo
IRI da versão	https://w3id.org/ufes/ontodiplomadigital/	owl:versionIRI
Idiomas	Português (@pt) como idioma principal, com rótulos, comentários e descrições complementares em inglês (@en).	dcterms:language, rdfs:label, rdfs:comment e demais literais com marcação linguística
Licença	Creative Commons Attribution 4.0 International (CC BY 4.0)	dcterms:license

Campo	Conteúdo	Termo correspondente no TTL canônico
Fontes documentais	Portal do Diploma Digital do MEC, schemas XSD v1.05 e documentação normativa utilizada na modelagem	dcterms:source
Referenciais de conformidade	Schemas XSD oficiais do Diploma Digital Brasileiro e respectivas especificações técnicas	dcterms:conformsTo
Fundamentação metodológica	Ontologia concebida como especificação formal e explícita de uma conceitualização compartilhada, fundamentada em Gruber (1993), Borst (1997) e desenvolvida segundo a OntoForInfoScience.	odd:methodologicalFoundation e dcterms:bibliographicCitation
Formato de representação	Ontologia desenvolvida em OWL 2 sobre RDF, serializada em Turtle (.ttl), com utilização complementar de RDFS e SKOS.	dcterms:format

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience (MENDONÇA, 2015)

Os dados apresentados no Quadro 23 cumprem a função de tornar explícita a identidade informacional da ontologia antes de sua conceitualização e formalização. A definição de uma URI persistente, de um namespace próprio e de metadados descritivos cria condições para que a OntoDiplomaDigital seja identificada, citada, documentada e reutilizada de maneira consistente. A distinção entre a URI da ontologia e a IRI da versão também contribui para separar a identidade estável do recurso de seus estados versionados.

A caracterização como ontologia operacional de domínio delimita o alcance do artefato: a proposta não busca representar todo o ecossistema do Diploma Digital, mas estruturar semanticamente o subdomínio do Registro do Diploma. Essa decisão torna o escopo mais controlável e compatível com os objetivos da pesquisa, pois concentra a modelagem nos elementos necessários à representação, consulta e interoperabilidade semântica dos dados do registro. Sua utilização operacional foi examinada com dados sintéticos, sem pressupor a disponibilização pública dos dados pessoais representados pelo modelo.

5.1.2 Domínio e escopo

A especificação do domínio definiu que a OntoDiplomaDigital se concentra nos diplomas digitais de Instituições de Ensino Superior brasileiras, conforme o arcabouço normativo estabelecido pelo Ministério da Educação. Dentro desse domínio geral, o escopo foi delimitado ao subdomínio do Registro do Diploma, abrangendo os

recursos, agentes, atributos e relações necessários à representação do diplomado, do curso concluído, das instituições envolvidas, do ato de registro e dos mecanismos de autenticidade e validação. O Quadro 23 apresenta a síntese dessa delimitação.

Quadro 24 - Domínio da ontologia

Dimensão	Conteúdo
Domínio geral	Diplomas digitais de Instituições de Ensino Superior brasileiras, conforme o arcabouço normativo estabelecido pelo Ministério da Educação.
Escopo de cobertura	Subdomínio do Registro do Diploma, abrangendo o diplomado, o curso concluído, o diploma emitido, as instituições envolvidas, o registro acadêmico e os mecanismos de autenticidade e validação. A cobertura conceitual desses elementos não determina suas condições de acesso.
Fontes estruturais do escopo	O pacote de schemas XSD v1.05 do MEC foi utilizado como fonte técnica de análise. A modelagem concentrou-se principalmente no DiplomaDigital_v1.05.xsd e no leiauteDiplomaDigital_v1.05.xsd, com consulta complementar aos schemas da Documentação Acadêmica, do Histórico Escolar Digital, dos tipos básicos e da assinatura XML.
Limites de cobertura	Não foram modelados integralmente os dados privados da Documentação Acadêmica, a granularidade curricular do Histórico Escolar Digital, os certificados de pós-graduação, os diplomas internacionais, a Lista de Diplomas Anulados, o Arquivo de Fiscalização e o Currículo Escolar Digital. Esses componentes permaneceram como possibilidades de extensão futura.
Granularidade representada	Incluiu entidades centrais, como Diplomado, Diploma Digital, Instituição de Ensino Superior, Curso e Registro do Diploma, além de entidades e estruturas complementares relacionadas à autenticidade, à validação e à rastreabilidade, como Assinatura Digital, Carimbo de Tempo, Código de Validação, responsáveis e assinantes institucionais.
Campo	Conteúdo
Domínio	Diplomas digitais de Instituições de Ensino Superior (IES) brasileiras, conforme regulamentação do Ministério da Educação (Portaria MEC nº 330/2018, nº 554/2019 e IN nº 01/2020).
Ponto de partida do escopo	Entidades formalizadas nos schemas XSD v1.05 do MEC: Diploma Digital (DiplomaDigital_v1.05.xsd), Documentação Acadêmica (DocumentacaoAcademicaRegistroDiplomaDigital_v1.05.xsd) e Histórico Escolar Digital (HistoricoEscolarDigital_v1.05.xsd).
Limite do escopo (o que não cobre)	Certificados de pós-graduação, diplomas internacionais, históricos de instituições privadas não aderentes ao sistema MEC, e documentos complementares (ArquivoFiscalizacao, ListaDiplomasAnulados, CurriculoEscolar), que poderão ser incorporados em versões futuras.
Granularidade	Inclui entidades de nível macro (Diplomado, Instituição, Curso, Diploma) e micro (Assinatura Digital, Carimbo de Tempo, ICP-Brasil, Código de Validação, Livro de Registro).

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience (MENDONÇA, 2015)

A delimitação apresentada no Quadro 24 evidencia que a ontologia não teve como finalidade reproduzir integralmente todos os documentos e schemas do Diploma Digital Brasileiro. O recorte adotado concentrou-se nos elementos do registro necessários à representação formal do diploma e à recuperação das relações previstas nas questões de competência. A seleção decorreu de sua pertinência ao problema de pesquisa e não de uma classificação de todos esses elementos como dados de acesso público. Essa decisão também permitiu distinguir os elementos centrais do domínio daqueles que permaneceram como possibilidades de extensão futura, como dados privados da Documentação Acadêmica, granularidade curricular do Histórico Escolar Digital e documentos complementares previstos no ecossistema normativo.

As normas, a documentação técnica e os schemas XSD utilizados como fontes da modelagem são de acesso público. Essa condição não se estende automaticamente aos valores registrados em documentos acadêmicos individuais, que podem incluir dados pessoais de diplomados e agentes institucionais. A prova de conceito empregou exclusivamente dados sintéticos. A presença de classes e propriedades na ontologia define possibilidades de representação, sem estabelecer autorização para divulgar ou reutilizar dados reais. Uma aplicação institucional deverá definir as condições de acesso e tratamento das informações.

5.1.3 Tipo de ontologia

A classificação registrada na especificação caracterizou a OntoDiplomaDigital como uma ontologia operacional de domínio. Quanto à abrangência, o artefato foi delimitado ao domínio dos diplomas digitais de Instituições de Ensino Superior brasileiras, com recorte específico no subdomínio do Registro do Diploma. Quanto ao propósito, foi concebido para representar recursos, agentes, atributos e relações de forma processável por máquinas, permitindo consultas e validações sobre os dados representados. O Quadro 25 sintetiza essa classificação e os padrões previstos para a formalização da ontologia.

Quadro 25 - Classificação da Ontologia

Dimensão	Classificação
Quanto ao nível	Ontologia de domínio, por representar conceitos, relações e restrições pertencentes ao domínio dos diplomas digitais das Instituições de

Dimensão	Classificação
	Ensino Superior brasileiras, com escopo delimitado ao Registro do Diploma.
Quanto ao propósito	Ontologia operacional, destinada à implementação computacional, à representação processável por máquinas, ao apoio à interoperabilidade semântica e à recuperação de informações".
Quanto ao grau de formalidade	Alto grau de formalidade no núcleo ontológico, formalizado em OWL por meio de classes, propriedades, hierarquias, restrições e axiomas. A camada terminológica complementar apresentou formalidade mais leve, utilizando SKOS para representar conceitos controlados e suas relações vocabulares.
Modelo de representação	RDF, utilizado como modelo de dados baseado em triplas para representar os recursos e as relações da ontologia.
Linguagem de representação do núcleo	OWL, empregada na definição formal de classes, propriedades de objeto, propriedades de dados, hierarquias, restrições e demais axiomas ontológicos.
Linguagem de representação complementar	SKOS, empregado na organização dos vocabulários controlados relativos a grau conferido, título conferido, modalidade de oferta, situação do diploma, formato de assinatura e papel da assinatura.
Formato de serialização	Turtle (TTL), adotado para codificação, leitura, documentação, publicação e compartilhamento da ontologia.

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience (MENDONÇA, 2015)

A classificação apresentada delimita a função da OntoDiplomaDigital dentro da pesquisa. A ontologia não automatiza a emissão ou o registro de diplomas, nem substitui os sistemas institucionais responsáveis por esses processos. Sua contribuição está na construção de uma camada semântica capaz de representar os dados do Registro do Diploma de maneira formal, consultável e alinhada a padrões da Web Semântica.

Essa delimitação também evita atribuir ao artefato resultados que não foram objeto da prova de conceito. A OntoDiplomaDigital não modelou integralmente trajetórias acadêmicas, indicadores de inclusão ou toda a granularidade curricular do Histórico Escolar Digital. O foco permaneceu no núcleo informacional público do Registro do Diploma, o que tornou a ontologia compatível com os objetivos da pesquisa e com o escopo definido na Etapa 1 da metodologia OntoForInfoScience.

5.1.4 Grupos de usuários

A especificação identificou os grupos de usuários e beneficiários da OntoDiplomaDigital a partir de sua relação com os processos de emissão, registro, consulta, validação e reúso dos dados do Diploma Digital. Foram considerados usuários diretos os atores institucionais, acadêmicos e computacionais que podem utilizar a ontologia para representar, processar, consultar ou analisar os dados do subdomínio. O cidadão diplomado foi caracterizado como beneficiário da interoperabilidade e da verificabilidade dos dados, pois tende a acessar os resultados por meio de sistemas, portais ou aplicações intermediárias. O Quadro 26 apresenta os grupos considerados e sua finalidade informacional.

Quadro 26 - Grupo de usuários e beneficiários da ontologia

Grupo de usuários ou beneficiários	Perfil e finalidade informacional
Gestores de tecnologia da informação e de registros acadêmicos das IES	Representar e intercambiar, de forma semanticamente consistente, dados relacionados ao diplomado, ao curso, à emissão, ao registro e à validação do diploma digital, apoiando a integração entre sistemas acadêmicos e serviços externos.
Pesquisadores em Ciência da Informação e áreas correlatas	Utilizar a ontologia como artefato de organização e representação do conhecimento em estudos sobre interoperabilidade semântica, metadados, dados ligados, ontologias operacionais e governança da informação acadêmica.
Gestores de políticas públicas educacionais e órgãos reguladores	Recuperar e relacionar dados estruturados sobre diplomas, cursos e instituições para apoiar atividades de acompanhamento, regulação e produção de informações agregadas sobre a certificação acadêmica.
Sistemas e agentes computacionais	Processar os recursos representados em RDF, realizar consultas SPARQL e recuperar relações entre diplomados, diplomas, cursos, instituições, registros e mecanismos de validação.
Cidadãos diplomados	Beneficiários da interoperabilidade, da verificabilidade e da portabilidade dos dados de certificação acadêmica, acessados por meio de portais, sistemas institucionais ou outras aplicações que utilizem a ontologia.

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience (MENDONÇA, 2015)

A distinção entre usuários diretos e beneficiários delimita o alcance prático da OntoDiplomaDigital. A ontologia foi projetada principalmente para apoiar sistemas, gestores, pesquisadores e órgãos institucionais na representação e consulta dos dados, não para ser manipulada diretamente pelo cidadão diplomado. Ainda assim, a melhoria da interoperabilidade semântica pode beneficiar o diplomado ao favorecer mecanismos de validação, portabilidade e acesso confiável às informações de sua certificação acadêmica.

5.1.5 Cenários de aplicação

Com base na metodologia OntoForInfoScience, foram definidos cinco cenários de aplicação para relacionar o propósito da OntoDiplomaDigital às necessidades informacionais dos usuários e beneficiários identificados na especificação do Quadro 27. Os cenários contemplaram situações de uso do subdomínio do Registro do

Diploma, incluindo consulta e recuperação de informações, verificação de dados associados à autenticidade do diploma e integração semântica com sistemas e plataformas governamentais.

Quadro 27 - Cenário de aplicação da ontologia

Cenário	Descrição
Verificação automática de autenticidade	Um sistema externo, como empresa, órgão público ou instituição estrangeira, consulta a ontologia via SPARQL para verificar se um diploma foi emitido por IES credenciada, possui informações associadas à assinatura ICP-Brasil.
Integração com Gov.br	Dados acadêmicos do diploma poderiam ser publicados como dados ligados, permitindo que plataformas governamentais consultem automaticamente informações de formação do cidadão, sem necessidade de upload manual de documentos.
Rastreabilidade do ciclo de vida do diploma	A ontologia registra eventos associados ao diploma, como emissão, registro, assinatura, validação e eventual anulação, com proveniência documentada.
Mobilidade acadêmica internacional	Instituições estrangeiras consultam a ontologia para reconhecimento do grau conferido, área de formação, carga horária e instituição emissora, reduzindo barreiras informacionais na revalidação.
Formulação de políticas públicas	Gestores consultam grafos de conhecimento para analisar distribuição de diplomados por município, curso, IES, grau, modalidade e período, subsidiando ações afirmativas e investimentos educacionais.

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience (MENDONÇA, 2015)

Os cenários apresentados no Quadro 27 funcionaram como critérios de orientação para o desenvolvimento da ontologia. Eles ajudaram a delimitar quais informações deveriam ser representadas para que o modelo pudesse responder a consultas sobre diplomados, cursos, instituições, registros e mecanismos de autenticidade. Ao mesmo tempo, a formulação desses cenários evidencia os limites da pesquisa: a OntoDiplomaDigital cria uma estrutura semântica compatível com essas aplicações, mas não comprova, nesta etapa, sua implantação efetiva em plataformas governamentais, sistemas institucionais ou processos internacionais de mobilidade acadêmica.

Como resultado da Etapa 1 da OntoForInfoScience, foi consolidado o Documento de Especificação da OntoDiplomaDigital, reunindo domínio, escopo, tipo de ontologia, grupos de usuários, cenários de aplicação e questões de competência. Nessa etapa, também foram definidos os requisitos gerais de representação computacional, indicando a adoção de padrões da Web Semântica, como OWL, RDF,

SKOS, TTL e SPARQL, sem ainda descrever a materialização formal desses elementos no artefato ontológico.

A arquitetura formal efetivamente implementada é apresentada na Etapa 5, correspondente à formalização da ontologia, pois nela são discutidos os elementos declarados no TTL, as classes, propriedades, restrições, vocabulários controlados, metadados e mecanismos de rastreabilidade formalizados na OntoDiplomaDigital.

Questões de competência da OntoDiplomaDigitalA especificação consolidou sete questões de competência destinadas a representar as necessidades informacionais que a OntoDiplomaDigital deveria atender. Essas questões estabeleceram a ligação entre o problema de pesquisa, os objetivos específicos e os requisitos de consulta da ontologia, abrangendo a identificação de instituições emissoras e registradoras, a recuperação de elementos de registro e validação, a consulta de diplomas por instituição, período e modalidade, a relação entre diplomado, curso, grau e título, o uso de valores controlados, a proveniência do registro e a representação de acesso por plataforma governamental no conjunto sintético. A correspondência entre as questões de competência e os objetivos específicos está apresentada no Quadro 28.

Quadro 28 - Questões de competência e requisitos informacionais da OntoDiplomaDigital

Nº	Questões de competência	Requisito de representação e recuperação
QC1	Quais são as IES emissoras e a IES registradoras de determinado diploma digital?	Distinguir e recuperar os papéis de instituição emissora e registradora associados a um diploma
QC2	Quais elementos representados permitem recuperar o registro, a assinatura digital, o carimbo do tempo e o código de validação de determinado diploma?	Analisar os padrões normativos e técnicos, estabelecer as correspondências para a ontologia operacional e realizar a prova de conceito.
QC3	Quais diplomas foram emitidos por determinada IES, em período definido, segundo a modalidade de oferta do curso?	Implementar a ontologia e realizar a prova de conceito.
QC4	Quem é o titular de determinado diploma, qual curso concluiu e quais grau e título foram conferidos?	Construir o modelo conceitual, implementar a ontologia e realizar a prova de conceito.
QC5	Quais valores controlados qualificam grau, título, modalidade, situação do	Identificar os recursos semânticos relevantes, mapear e implementar os valores

Nº	Questões de competência	Requisito de representação e recuperação
	diploma e formato ou papel da assinatura?	controlados e verificá-los na prova de conceito.
QC6	Qual é a proveniência do registro, incluindo IES registradora, agentes, data e processo administrativo?	Estabelecer as correspondências para a ontologia operacional, implementar a ontologia e realizar a prova de conceito.
QC7	Por meio de qual plataforma governamental o diploma de determinado cidadão, identificado por CPF, é representado como acessível no conjunto sintético?	Implementar a ontologia e verificar a relação de acesso na prova de conceito.

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience (MENDONÇA, 2015)

As questões de competência funcionaram como requisitos informacionais da OntoDiplomaDigital e estabeleceram a relação entre os objetivos específicos da pesquisa e as informações que a ontologia deveria permitir recuperar. Conforme sistematizado no Quadro 28, elas orientaram a seleção dos conceitos, das propriedades e das relações desenvolvidos nas etapas posteriores da OntoForInfoScience.

Na prova de conceito, as sete questões foram operacionalizadas por meio de consultas SPARQLexecutadas sobre a ABox no GraphDB. Também foram realizadas consultas complementares para examinar caminhos semânticos, relações entre recursos e aspectos de consistência que não eram abrangidos diretamente pelas consultas principais. Essas consultas adicionais não constituíram novas questões de competência, mas forneceram evidências auxiliares para a verificação operacional da ontologia. O conjunto de consultas está disponível no repositório público da pesquisa, indicado no Apêndice B, e seus resultados são apresentados na seção 5.6.3.

O Documento de Especificação estabeleceu os requisitos de referência utilizados nas etapas posteriores de aquisição do conhecimento, conceitualização, fundamentação ontológica, formalização, avaliação e verificação operacional. A arquitetura formal implementada é apresentada na seção 5.5.1, no Quadro 43.

5.2 Resultado da aquisição e extração do conhecimento – Etapa 2 da OntoForInfoScience

A seção 5.2 apresenta os resultados da Etapa 2 da OntoForInfoScience, correspondente à aquisição e extração do conhecimento. A etapa reuniu fontes

normativas e técnicas, complementadas por referências bibliográficas e informações contextuais, para identificar elementos relevantes ao subdomínio do Registro do Diploma. A análise permitiu reconhecer definições institucionais, estruturas documentais, termos especializados, atributos, relações e categorias terminológicas que subsidiaram a construção da OntoDiplomaDigital.

Os resultados são apresentados em duas subseções. A seção 5.2.1 reúne as fontes normativas e técnicas que fundamentaram a análise do domínio; a seção 5.2.2 apresenta os elementos extraídos e os artefatos produzidos. Conforme a OntoForInfoScience, os resultados formais da Etapa 2 consistiram em três glossários: de conceitos, de verbos e de relações, apresentados, respectivamente, nos Quadros 27, 28 e 29. As listas de atributos e de valores controlados funcionaram como registros auxiliares da extração, enquanto a rastreabilidade das fontes sustentou a documentação das decisões desenvolvidas nas etapas posteriores.

5.2.1 Fontes normativas e técnicas

A aquisição do conhecimento teve como ponto de partida dois conjuntos complementares de fontes: as fontes normativas e as fontes técnicas. As fontes normativas compreenderam os atos oficiais que instituem, regulamentam e orientam o Diploma Digital Brasileiro, como portarias, instruções normativas, decretos e notas técnicas. Elas foram utilizadas para identificar responsabilidades institucionais, requisitos de validade jurídica, papéis dos agentes e obrigações relacionadas à emissão, ao registro, à preservação e à validação dos diplomas digitais.

As fontes técnicas corresponderam aos schemas XSD e aos documentos estruturais do pacote técnico do Ministério da Educação. Esses arquivos não definem o sentido jurídico do Diploma Digital, mas descrevem sua organização sintática em XML, incluindo elementos, tipos de dados, atributos, agrupamentos, restrições, dependências entre schemas e componentes de assinatura digital. O Quadro 29 apresenta os atos normativos, documentos técnicos e schemas utilizados como fontes primárias da análise do domínio.

Quadro 29 - Fontes primárias da ontologia

Fonte	Conteúdo incorporado à análise do domínio
Portaria MEC nº 330/2018	Instituição do diploma digital e identificação das responsabilidades das Instituições de Ensino Superior no processo de emissão e registro.

Fonte	Conteúdo incorporado à análise do domínio
Portaria MEC nº 554/2019	Requisitos gerais de emissão, registro, preservação, assinatura e validade jurídica do diploma digital.
Instrução Normativa MEC nº 01/2020	Estruturação técnica dos documentos XML e definição do pacote de schemas XSD utilizado como referência para a análise.
Nota Técnica MEC nº 13/2019 (Brasil, 2019b)	Orientações técnicas relacionadas à implantação, à assinatura digital e à conformidade dos documentos eletrônicos.
Portaria MEC nº 70/2025	Atualização normativa relativa à validade jurídica, à assinatura digital e ao carimbo de tempo, utilizada como fonte complementar.
Decreto nº 9.756/2019	Referência ao ambiente de integração de serviços públicos digitais por meio do portal Gov.br.
Decreto nº 10.332/2020	Diretrizes de governo digital e interoperabilidade entre sistemas da Administração Pública Federal.
Pacote de schemas XSD v1.05	Fonte técnica para identificação de elementos, tipos, atributos, estruturas e dependências do diploma digital.
DiplomaDigital_v1.05.xsd e leiauteDiplomaDigital_v1.05.xsd	Fontes centrais para a extração dos elementos do Diploma Digital, do diplomado, do curso, das instituições, do registro e das assinaturas.
tiposBasicos_v1.05.xsd e xmldsig-core-schema_v1.1.xsd	Fontes complementares para identificação de tipos reutilizáveis, identificadores e elementos técnicos de assinatura digital.
Schemas da Documentação Acadêmica e do Histórico Escolar Digital	Fontes contextuais utilizadas para compreender as relações documentais do ecossistema, sem incorporação integral de sua granularidade ao escopo da ontologia.

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience (MENDONÇA, 2015)

No conjunto das fontes técnicas, os *schemas* XSD v1.05 foram analisados como arquivos principais e auxiliares. Os arquivos principais estruturam documentos completos do ecossistema do Diploma Digital, como o Diploma Digital, a Documentação Acadêmica para Registro, o Histórico Escolar Digital, a Lista de Diplomas Anulados, o Arquivo de Fiscalização e o Currículo Escolar Digital. Os arquivos auxiliares, por sua vez, definem leiautes internos, tipos reutilizáveis, valores controlados, padrões, restrições e elementos de assinatura digital utilizados pelos documentos principais. O Quadro 30 apresenta os arquivos examinados e sua utilização na OntoDiplomaDigital

.

Quadro 30 - Arquivos XSD do pacote MEC v1.05 e sua utilização na ontologia

Arquivo XSD / Categoria	Função no pacote MEC	Utilização na OntoDiplomaDigital
Principal: DiplomaDigital_v1.05.xsd	Estrutura o XML oficial do diploma.	Delimita o recurso DiplomaDigital e o núcleo do domínio modelado.
Principal: DocumentacaoAcademicaRegistroDiplomaDigital_v1.05.xsd	Estrutura a documentação acadêmica destinada à emissão e ao registro.	Referência para DocumentacaoAcademica, proveniência e suporte documental do registro.
Principal: HistoricoEscolarDigital_v1.05.xsd	Estrutura o Histórico Escolar Digital como documento próprio.	Referência para HistoricoEscolarDigital e para a delimitação da trajetória acadêmica mantida fora do núcleo operacional.
Principal: ListaDiplomasAnulados_v1.05.xsd	Estrutura o arquivo de diplomas com registro anulado.	Referência contextual para situações de anulação e estados do diploma.
Principal: ArquivoFiscalizacao_v1.05.xsd	Estrutura o arquivo de fiscalização solicitado pelo MEC às IES.	Delimita processos de fiscalização externos ao núcleo ontológico.
Principal: CurriculoEscolarDigital_v1.05.xsd	Estrutura o Currículo Escolar Digital.	Delimita informações curriculares não incorporadas integralmente ao núcleo da ontologia.
Auxiliar: tiposBasicos_v1.05.xsd	Reúne tipos simples e complexos, grupos, enumerações, padrões e restrições reutilizáveis.	Fonte de tipos, cardinalidades, grupos estruturais, valores controlados e restrições empregados na conceitualização e formalização.
Auxiliar: leiauteDiplomaDigital_v1.05.xsd	Define as estruturas internas do diploma, diplomado, curso, emissão e registro.	Principal fonte estrutural do subdomínio do Registro do Diploma, originando classes, propriedades, restrições e elementos de rastreabilidade XSD.
Auxiliar: leiauteDocumentacaoAcademicaRegistroDiplomaDigital_v1.05.xsd	Define as estruturas internas da documentação acadêmica destinada à emissão e ao registro.	Fonte complementar para documentação acadêmica, requisições documentais e relações de proveniência associadas ao registro.

Arquivo XSD / Categoria	Função no pacote MEC	Utilização na OntoDiplomaDigital
Auxiliar: leiauteHistoricoEscolar_v1.05.xsd	Define as estruturas internas do Histórico Escolar Digital.	Fonte complementar para curso, trajetória acadêmica e documentos associados.
Auxiliar: leiauteListaDiplomasAnulados_v1.05.xsd	Define as estruturas internas da lista de diplomas anulados.	Fonte complementar para situações de anulação e estados documentais não desenvolvidos integralmente na ontologia.
Auxiliar: leiauteArquivoFiscalizacao_v1.05.xsd	Define as estruturas internas do arquivo de fiscalização.	Referência complementar para processos de fiscalização mantidos fora do núcleo modelado.
Auxiliar: leiauteCurriculoEscolar_v1.05.xsd	Define as estruturas internas do Currículo Escolar Digital.	Fonte complementar para componentes curriculares e para delimitação dos dados mantidos fora do núcleo operacional.
Auxiliar: xmldsig-core-schema_v1.1.xsd	Especifica os elementos XML utilizados na assinatura digital.	Fundamenta a representação de AssinaturaDigital e sua distinção conceitual em relação aos agentes signatários.

Fonte: Elaborado pelo autor com base nos *schemas* XSD v1.05 do Ministério da Educação.

A distinção entre fontes normativas e fontes técnicas foi necessária porque cada conjunto contribuiu de modo diferente para a construção da ontologia. As normas indicaram o significado institucional e jurídico dos processos, enquanto os XSDs evidenciaram a estrutura informacional dos documentos. A modelagem da OntoDiplomaDigital resultou da articulação entre esses dois planos, evitando que a ontologia fosse construída apenas como transposição da sintaxe XML para RDF.

5.2.2 Elementos extraídos e artefatos produzidos

A análise das fontes resultou na identificação de candidatos a entidades, atributos, relações e valores controlados do subdomínio do Registro do Diploma. Entre os elementos extraídos estiveram Diplomado, Diploma Digital, Instituição de Ensino Superior, IES Emissora, IES Registradora, Curso, Registro do Diploma, Assinatura Digital, Carimbo de Tempo, Código de Validação, Documentação Acadêmica e Histórico Escolar Digital. Também foram identificados atributos relacionados à identificação, às datas, aos códigos institucionais, ao registro acadêmico e aos mecanismos de autenticidade, além de relações como “tem titular”, “emitido por”, “registrado por”, “tem curso”, “tem registro”, “tem assinatura” e “tem código de validação”. O Quadro 31 distinguiu os três glossários resultantes da Etapa 2 dos registros auxiliares utilizados para preservar evidências da extração.

Quadro 31 - Artefatos produzidos da Ontologia

Artefato produzido	Conteúdo consolidado
Glossário preliminar de conceitos	Termos candidatos a classes, entidades do domínio ou conceitos controlados, acompanhados de indicação de fonte, nota de uso e possível correspondência com elementos normativos ou técnicos.
Glossário preliminar de verbos	Verbos e expressões relacionais identificados nas fontes analisadas, utilizados como candidatos à formulação de propriedades ontológicas e relações entre recursos.
Glossário preliminar de relações	Relações candidatas entre entidades do domínio, registradas com recurso de origem, recurso de destino, leitura semântica preliminar, possível domínio, possível alcance e fonte de extração.
Lista preliminar de atributos	Elementos candidatos a propriedades de dados, incluindo identificadores, datas, códigos institucionais, nomes, números de registro, informações administrativas e demais valores literais associados aos recursos.
Lista preliminar de valores controlados	Termos candidatos a vocabulários controlados, como grau conferido, título conferido, modalidade de oferta, situação do diploma, ambiente operacional, formato de assinatura e papel da assinatura.
Registro de rastreabilidade das fontes	Associação entre elementos extraídos e suas fontes normativas ou técnicas, especialmente portarias, instruções normativas, *schemas* XSD, caminhos estruturais, elementos e tipos de dados utilizados como evidência para a modelagem.

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience (MENDONÇA, 2015)

Os três glossários constituíram os resultados formais da aquisição e extração do conhecimento e funcionaram como ligação entre as fontes analisadas e a conceitualização da ontologia. Neles, os termos, os verbos e as relações foram mantidos como candidatos, sem pressupor sua conversão automática em classes, propriedades ou conceitos SKOS. As listas auxiliares preservaram evidências sobre atributos e categorias terminológicas identificadas, enquanto o registro de rastreabilidade manteve a referência às fontes que fundamentaram as análises posteriores.

O Glossário de Conceitos, apresentado no Quadro 32, reuniu os termos candidatos identificados na aquisição do conhecimento e serviu como entrada para a Etapa 3 da OntoForInfoScience. Nessa etapa posterior, os elementos foram examinados quanto ao significado, à função no domínio, às variantes terminológicas e à pertinência para a modelagem conceitual da OntoDiplomaDigital.

Quadro 32 - Glossário de conceitos da Etapa 2 da OntoForInfoScience

Nº	Português	Inglês
1	Diploma Digital	Digital Diploma
2	Diplomado	Graduate
3	Instituição de Ensino Superior	Higher Education Institution
4	IES Emissora	Issuing HEI
5	IES Registradora	Registering HEI
6	Curso	Academic Program
7	Curso NSF	NSF Academic Program
8	Curso por Decisão Judicial	Academic Program by Court Decision
9	Habilitação	Qualification Track
10	Informações de Tramitação e-MEC	e-MEC Processing Information
11	Registro do Diploma	Diploma Registration
12	Registro do Diploma NSF	NSF Diploma Registration
13	Responsável pelo Registro	Registration Responsible Person
14	Assinante do Registro	Registration Signer
15	Assinante da Emissão	Issuance Signer
16	Signatário do Termo de Responsabilidade da Emissora	Issuance Responsibility Statement Signer
17	Representante Institucional da Emissora	Issuing HEI Institutional Representative
18	Assinatura Digital	Digital Signature
19	Carimbo de Tempo	Timestamp
20	Código de Validação	Validation Code
21	Documentação Acadêmica	Academic Documentation
22	Requisição de Documentação Acadêmica	Academic Documentation Request
23	Histórico Escolar Digital	Digital Academic Transcript

Nº	Português	Inglês
24	Processo Administrativo do Diploma	Diploma Administrative Process
25	Endereço	Address
26	Alternativa de Endereço do Curso por Decisão Judicial	Course Address Alternative by Court Decision
27	Local de Naturalidade	Birthplace
28	Agente Certificador	Certifying Agent
29	Plataforma Governamental	Government Platform
30	Valor Controlado	Controlled Value
31	Conceito de Grau Conferido	Degree Type Concept
32	Conceito de Título Conferido	Awarded Title Concept
33	Conceito de Modalidade de Oferta	Mode of Delivery Concept
34	Conceito de Situação do Diploma	Diploma Status Concept
35	Conceito de Ambiente Operacional	Operational Environment Concept
36	Conceito de Formato de Assinatura	Signature Format Concept
37	Conceito de Papel da Assinatura	Signature Role Concept

Fonte: Elaborado pelo autor (2026), com base na aquisição e extração do conhecimento a partir das fontes normativas e técnicas examinadas

O Quadro 32 demonstrou que a aquisição do conhecimento não se limitou à identificação de nomes presentes nos schemas XSD. Os termos reunidos articularam a terminologia normativa do Diploma Digital, as estruturas técnicas dos documentos e as necessidades de representação definidas para a pesquisa. Alguns corresponderam a recursos centrais do domínio, como Diploma Digital, Diplomado, Curso, Instituição de Ensino Superior e Registro do Diploma; outros indicaram agentes, processos, mecanismos de validação ou categorias terminológicas relevantes para a representação do subdomínio.

Esse glossário teve função preparatória para a Etapa 3, na qual os candidatos passaram por definição conceitual e decisão de modelagem. Portanto, seus itens foram lidos como resultado da extração do conhecimento, e não como reprodução direta da estrutura sintática dos documentos XML ou como inventário das classes posteriormente formalizadas.

A extração do conhecimento também identificou verbos e expressões relacionais utilizados para representar vínculos entre os recursos do subdomínio do Registro do Diploma destacados no Quadro 33. Esses verbos foram registrados como candidatos à formulação posterior de propriedades ontológicas, pois indicaram relações entre diploma, diplomado, curso, instituição, registro, assinatura, código de validação e plataforma governamental.

Quadro 33 - Glossário de verbos da Etapa 2 da metodologia OntoforInfoscience

Português	Inglês
acessível via	is accessible via
assina administrativamente	is registration signer of
assina institucionalmente	institutionally signs
assinado administrativamente por	has registration signer
assinado institucionalmente por	institutionally signed by
assinatura de	is signature of
carimbo de tempo de	is timestamp of
certifica tecnicamente assinatura	technically certifies signature
concluído por	completed by
concluiu curso	completed program
confere grau	awards degree
confere título	awards title
é assinante da emissão de	is issuance signer of
é processo administrativo de	is administrative process of
é requisição de documentação acadêmica de	is academic documentation request of
é signatário do termo de responsabilidade de	is issuance responsibility statement signer of
emite diploma	issues diploma
emitido por	issued by
habilitação de	is qualification of
registrado por	registered by
registra diploma	registers diploma
registro de	is registration of
responsável pelo registro de	is registration responsible of
tem ambiente operacional	has operational environment
tem assinante da emissão	has issuance signer
tem assinatura	has signature
tem carimbo de tempo	has timestamp
tem certificador técnico	has technical certifier
tem código de validação	has validation code
tem curso	has program
tem documentação acadêmica	has academic documentation
tem endereço	has address
tem formato de assinatura	has signature format
tem habilitação	has qualification track
tem histórico	has transcript
tem informações de tramitação e-MEC	has e-MEC processing information
tem local de naturalidade	has birthplace
tem modalidade de oferta	has mode of delivery
tem objetivo relacionado	has related objective
tem papel da assinatura	has signature role
tem processo administrativo	has administrative process

Português	Inglês
tem registro	has registration
tem requisição de documentação acadêmica	has academic documentation request
tem responsável pelo registro	has registration responsible person
tem signatário do termo de responsabilidade da emissora	has issuance responsibility statement signer
tem situação	has status
tem titular	has holder
tem valor controlado	has controlled value
titular de	holds diploma
valida diploma	validates diploma
vinculado ao diploma	is program of diploma

Fonte: Elaborado pelo autor (2026), com base na aquisição e extração do conhecimento e confrontado com o arquivo canônico OntoDiplomaDigital_registro_diploma.ttl.

O Quadro 33 evidenciou que os verbos extraídos não foram utilizados apenas como elementos linguísticos, mas como indícios das relações que precisariam ser representadas formalmente na ontologia. Expressões como “tem titular”, “emitido por”, “registrado por”, “tem curso” e “tem registro” indicaram relações centrais entre os principais recursos do domínio. Pares como “tem titular” e “titular de”, “tem assinatura” e “assinatura de”, bem como “tem carimbo de tempo” e “carimbo de tempo de”, registraram possibilidades de leitura em sentidos complementares.

A análise também distinguiu relações centrais, relações complementares e relações associadas a valores controlados. “Tem código de validação” e “valida diploma”, por exemplo, indicaram vínculos relativos à consulta pública do diploma; “tem certificador técnico” e “certifica tecnicamente assinatura” registraram a cadeia técnica associada à assinatura digital. Por sua vez, “tem valor controlado” foi mantida como expressão de agrupamento para relações cujos objetos corresponderam a categorias terminológicas, como grau conferido, modalidade de oferta, situação do diploma, formato de assinatura e papel da assinatura. Essas distinções subsidiaram a definição conceitual das relações na Etapa 3.

O Glossário de Relações apresentado no Quadro 34 consolidou os vínculos identificados entre os elementos do subdomínio Registro do Diploma. Na estrutura adotada, “Elemento 1” designou o recurso de origem da relação; “Relação”, a expressão verbal que representou o vínculo; e “Elemento 2”, o recurso de destino. Essa organização permitiu registrar, ainda na Etapa 2 da OntoForInfoScience, relações candidatas que orientaram a definição conceitual na etapa subsequente.

Quadro 34 - Glossário de relações

Elemento 1	Relação	Elemento 2
Diploma Digital	tem titular	Diplomado
Diplomado	titular de	Diploma Digital
Diploma Digital	emitido por	IES Emissora
IES Emissora	emite diploma	Diploma Digital
Diploma Digital	registrado por	IES Registradora
IES Registradora	registra diploma	Diploma Digital
Diploma Digital	tem curso	Curso
Curso	vinculado ao diploma	Diploma Digital
Curso	concluído por	Diplomado
Diplomado	concluiu curso	Curso
Diploma Digital	tem registro	Registro do Diploma
Registro do Diploma	registro de	Diploma Digital
Registro do Diploma	tem responsável pelo registro	Responsável pelo Registro
Responsável pelo Registro	responsável pelo registro de	Registro do Diploma
Registro do Diploma	assinado administrativamente por	Assinante do Registro
Assinante do Registro	assina administrativamente	Registro do Diploma
Diploma Digital	tem assinante da emissão	Assinante da Emissão
Assinante da Emissão	é assinante da emissão de	Diploma Digital
Documentação Acadêmica	tem signatário do termo de responsabilidade da emissora	Signatário do Termo de Responsabilidade da Emissora
Signatário do Termo de Responsabilidade da Emissora	é signatário do termo de responsabilidade de	Documentação Acadêmica
Diploma Digital	tem assinatura	Assinatura Digital
Assinatura Digital	assinatura de	Diploma Digital
Assinatura Digital	tem carimbo de tempo	Carimbo de Tempo
Carimbo de Tempo	carimbo de tempo de	Assinatura Digital
Diploma Digital	tem código de validação	Código de Validação
Código de Validação	valida diploma	Diploma Digital
Diploma Digital	tem histórico	Histórico Escolar Digital
Diploma Digital	tem documentação acadêmica	Documentação Acadêmica
Curso	confere grau	Conceito de Grau Conferido
Curso	confere título	Conceito de Título Conferido
Curso	tem habilitação	Habilitação
Habilitação	habilitação de	Curso
Curso	tem informações de tramitação e-MEC	Informações de Tramitação e-MEC
Curso	tem modalidade de oferta	Conceito de Modalidade de Oferta
Diploma Digital	tem situação	Conceito de Situação do Diploma

Elemento 1	Relação	Elemento 2
Assinatura Digital	tem formato de assinatura	Conceito de Formato de Assinatura
Assinatura Digital	tem papel da assinatura	Conceito de Papel da Assinatura
Assinatura Digital	assinado institucionalmente por	Instituição de Ensino Superior
Instituição de Ensino Superior	assina institucionalmente	Assinatura Digital
Assinatura Digital	tem certificador técnico	Agente Certificador
Agente Certificador	certifica tecnicamente assinatura	Assinatura Digital
Diploma Digital	acessível via	Plataforma Governamental
Recurso com endereço associado	tem endereço	Endereço
Recurso relacionado a valor controlado	tem valor controlado	Valor Controlado
Questão de Competência	tem objetivo relacionado	Objetivo Específico
Registro do Diploma	tem requisição de documentação acadêmica	Requisição de Documentação Acadêmica
Requisição de Documentação Acadêmica	é requisição de documentação acadêmica de	Registro do Diploma
Registro do Diploma	tem processo administrativo	Processo Administrativo
Processo Administrativo	é processo administrativo de	Registro do Diploma
Diplomado	tem local de naturalidade	Local de Naturalidade
Diploma Digital	tem ambiente operacional	Conceito de Ambiente Operacional

Fonte: Elaborado pelo autor (2026), com base na aquisição e extração do conhecimento e confrontado com o arquivo canônico OntoDiplomaDigital_registro_diploma.ttl.

As relações apresentadas no Quadro 34 mostram que a extração do conhecimento procurou representar não apenas os elementos isolados do domínio, mas também os vínculos que lhes atribuem significado. Relações como “tem titular”, “emitido por”, “registrado por”, “tem curso” e “tem registro” estruturam o núcleo do Registro do Diploma. Outras relações, como “tem assinatura”, “tem carimbo de tempo”, “tem código de validação” e “acessível via”, representam mecanismos associados à autenticidade, validação e consulta pública. A presença de relações inversas, como “tem titular” e “titular de”, também favoreceu a leitura bidirecional do modelo e a posterior formulação de consultas SPARQL.

Além dos três glossários previstos para a Etapa 2, a extração do conhecimento também produziu uma lista auxiliar de atributos candidatos à representação de valores literais. Esses atributos incluíram identificadores, datas, nomes, códigos institucionais, números de registro, informações administrativas, dados de localização e elementos técnicos relacionados à assinatura e à validação. Para preservar a rastreabilidade entre a extração e a formalização, o Quadro 35 apresentou a correspondência desses

atributos com as propriedades de dados efetivamente declaradas no TTL canônico, organizadas segundo seus domínios e alcances.

Quadro 35- Propriedades de dados identificadas na Etapa 2 da metodologia OntoForInfoScience

Domínio declarado no TTL	Propriedades de dados identificadas	Alcance declarado
odd:DiplomaDigital	odd:version; odd:environment; odd:issueDate	xsd:string; xsd:string; xsd:date
odd:Diplomado	odd:nationality; odd:cpf	xsd:string; xsd:string
odd:LocalNaturalidade	odd:ibgeMunicipalityCode; odd:birthMunicipalityName; odd:birthFederativeUnitCode; odd:foreignBirthMunicipalityName	xsd:string
odd:Curso	odd:conclusionDate; odd:courseName; odd:workloadHours; odd:emphasis; odd:otherAwardedTitle	xsd:date; xsd:string; xsd:decimal; xsd:string; xsd:string
odd:CursoPorDecisaoJudicial	odd:courseAddressUnavailable	xsd:boolean
odd:RegistroDiploma	odd:registrationDate; odd:academicDocumentationRequestId; odd:diplomaProcessNumber; odd:degreeConferralDate; odd:diplomaExpeditionDate; odd:registryBook; odd:registrationNumber; odd:registryPage; odd:registrySequence; odd:registrationAdditionalInformation	xsd:date; xsd:string
odd:InstituicaoEnsinoSuperior	odd:institutionMecCode; odd:cnpj	xsd:string
odd:ResponsavelRegistro	odd:responsibleName; odd:responsibleCpf; odd:responsibleInternalId	xsd:string
odd:AssinanteRegistro	odd:registrationSignerCpf; odd:registrationSignerRole	xsd:string
odd:AssinanteEmissao	odd:issuanceSignerCpf; odd:issuanceSignerRole	xsd:string
odd:SignatarioTermoResponsabilidadeEmissora	odd:issuanceResponsibilitySignerName; odd:issuanceResponsibilitySignerCpf; odd:issuanceResponsibilitySignerRole; odd:issuanceResponsibilitySignerDesignationAct	xsd:string
odd:AssinaturaDigital	odd:signatureDateTime	xsd:dateTime
odd:CarimboTempo	odd:timestampDateTime	xsd:dateTime
odd:CodigoValidacao	odd:validationCodeValue	xsd:string
odd:Habilitacao	odd:qualificationName; odd:qualificationDate	xsd:string; xsd:date
odd:InformacoesTramitacaoEmec	odd:emecProcessNumber; odd:emecProcessType; odd:emecRegistrationDate; odd:emecProtocolDate	xsd:string; xsd:string; xsd:date; xsd:date
odd:Endereco	odd:streetAddress; odd:addressNumber; odd:addressComplement; odd:district; odd:municipality; odd:federativeUnit; odd:postalCode; odd:countryName	xsd:string
Não declarado no TTL	odd:identifier; odd:fullName; odd:codeEmec	xsd:string

Fonte: Elaborado pelo autor (2026), com base na aquisição e extração do conhecimento e confrontado com o arquivo canônico OntoDiplomaDigital_registro_diploma.ttl, disponibilizado no repositório público indicado no Apêndice B.

O Quadro 35 evidenciou que os atributos identificados durante a aquisição foram posteriormente representados como propriedades de dados quando expressaram valores literais relevantes para o domínio. Enquanto as relações entre recursos foram formalizadas por propriedades de objeto, os atributos registrados no quadro corresponderam a valores como nomes, datas, códigos, identificadores e informações administrativas.

A apresentação conjunta da lista extraída e de sua correspondência no TTL permitiu acompanhar a passagem entre a identificação inicial dos atributos e sua formalização. Nos casos em que o arquivo canônico não declarou domínio explícito, o quadro preservou essa condição, sem atribuir restrições não formalizadas. Dessa forma, o inventário funcionou como registro complementar da Etapa 2 e como evidência da rastreabilidade entre os elementos extraídos e a estrutura posteriormente implementada.

A extração do conhecimento também identificou valores sujeitos a controle terminológico. Esses valores foram organizados em categorias relacionadas ao ambiente operacional, à modalidade de oferta, ao grau conferido, ao título conferido, à situação do diploma, ao formato da assinatura e ao papel da assinatura. O Quadro 36 apresenta os valores controlados declarados no TTL canônico e sua vinculação às respectivas categorias.

Quadro 36 - Valores controlados identificados na Etapa 2 da metodologia OntoForInfoScience

Categoria declarada no TTL	Evidência no TTL canônico
Ambiente operacional	vocab:Producao; vocab:Homologacao
Modalidade de oferta	vocab:Presencial; vocab:Ead; vocab:Semipresencial
Grau conferido	vocab:Bacharelado; vocab:Licenciatura; vocab:Tecnologia; vocab:CursoSequencial
Título conferido	vocab:Bacharel; vocab:Licenciado; vocab:Tecnologo; vocab:Medico
Situação do diploma	vocab:Ativo; vocab:Anulado; vocab:Revogado
Formato de assinatura	vocab:XAdES; vocab:CAdES; vocab:PAdES
Papel da assinatura	vocab:AssinaturaPFEmissora; vocab:AssinaturaPJEmissora; vocab:AssinaturaPFResponsavelRegistro; vocab:AssinaturaPJRegistradora; vocab:AssinaturaPJArquivamentoDocumentacaoAcademica

Fonte: Elaborado pelo autor (2026), com base na aquisição e extração do conhecimento e confrontado com o arquivo canônico OntoDiplomaDigital_registro_diploma.ttl.

O Quadro 36 demonstra que determinados elementos do domínio não foram tratados como texto livre, mas como conceitos controlados. Essa decisão favorece a padronização terminológica e reduz ambiguidades em consultas e interpretações

futuras. Por exemplo, modalidades como Presencial, Educação a Distância e Semipresencial são representadas como valores pertencentes a uma mesma categoria, o que permite recuperar e comparar informações de forma mais consistente.

A rastreabilidade das fontes foi registrada por meio de propriedades e notas destinadas a explicitar a origem técnica ou documental dos elementos representados. Esses mecanismos permitem relacionar recursos, propriedades e decisões de modelagem às fontes utilizadas na construção da ontologia, especialmente aos schemas XSD e aos documentos normativos analisados. O Quadro 37 apresenta os mecanismos de rastreabilidade declarados no TTL canônico e suas ocorrências no arquivo.

Quadro 37 - Mecanismos de rastreabilidade declarados no TTL

Mecanismo de rastreabilidade	Função no modelo
dcterms:source	Registra a fonte normativa, técnica ou documental associada ao recurso representado.
odd:xsdPath	Indica o caminho estrutural do elemento no schema XSD utilizado como evidência técnica.
odd:xsdElement	Registra o elemento XSD associado ao recurso, propriedade ou valor modelado.
odd:xsdType	Indica o tipo XSD relacionado ao elemento extraído ou formalizado.
odd:sourceNote	Explicita observações sobre a origem documental ou técnica do elemento representado.
odd:modelingRationale	Registra a justificativa da decisão de modelagem adotada na passagem da fonte para a ontologia.

Fonte: Elaborado pelo autor (2026), com base na aquisição e extração do conhecimento e confrontado com o arquivo canônico OntoDiplomaDigital_registro_diploma.ttl.

O Quadro 37 evidencia que a rastreabilidade foi incorporada como parte da documentação da ontologia. As propriedades dcterms:source, odd:xsdPath, odd:xsdElement e odd:xsdType registram a vinculação entre os elementos modelados e suas fontes técnicas, enquanto odd:sourceNote e odd:modelingRationale explicitam observações de origem e justificativas de modelagem. Essa estrutura fortalece a auditabilidade do artefato, pois permite compreender de onde determinado elemento foi extraído e por que foi representado daquela forma.

5.3 Resultado da Conceitualização - Etapa 3 da OntoForInfoScience

A Etapa 3 consolidou a representação conceitual do subdomínio do Registro do Diploma. O resultado principal foi a organização dos elementos extraídos na etapa

anterior em conjuntos conceituais capazes de descrever o diploma digital, seu titular, o curso concluído, as instituições envolvidas, o ato de registro, os agentes responsáveis e os mecanismos de autenticidade e validação.

Foram definidos conceitos centrais relacionados ao Diplomado, ao Diploma Digital, ao Curso, às Instituições de Ensino Superior e ao Registro do Diploma. Também foram organizados conceitos de apoio associados à emissão, ao registro, à assinatura digital, ao carimbo de tempo, ao código de validação e aos agentes envolvidos. Conceitos contextuais, como Histórico Escolar Digital, Documentação Acadêmica, Agente Certificador e Plataforma Governamental, foram mantidos para representar relações necessárias com o ecossistema informacional do Diploma Digital, sem ampliar o núcleo da ontologia para todo o conjunto documental do MEC.

Os resultados da conceitualização são apresentados nas subseções seguintes por meio do dicionário de conceitos, da tabela de conceitos e propriedades, do dicionário de verbos e dos modelos conceituais gráficos. Esses artefatos demonstram como os candidatos identificados na aquisição do conhecimento foram organizados em uma estrutura conceitual intermediária, anterior à formalização em OWL e SKOS.

5.3.1 Dicionário de conceitos, Tabela de Conceitos e Propriedades

Os resultados da conceitualização foram consolidados em dois quadros complementares. O Quadro 38 apresenta o dicionário de conceitos, no qual cada elemento conceitual é acompanhado da definição adotada, das variantes terminológicas identificadas e da respectiva origem documental ou decisão de modelagem. O Quadro 39 apresenta a tabela de conceitos e propriedades, evidenciando como esses conceitos foram articulados por meio de relações, atributos e notas de uso. Em conjunto, os dois quadros demonstram a passagem dos termos extraídos das fontes para uma estrutura conceitual mais estável, utilizada posteriormente na formalização da OntoDiplomaDigital.

.

Quadro 38 - Definição e origem dos conceitos do subdomínio do Registro do Diploma

Classe ou elemento conceitual	Definição adotada	Sinônimo ou variante	Origem documental ou decisão de modelagem
Diploma Digital	Recurso documental digital que formaliza a certificação acadêmica do diplomado e articula as demais entidades e relações do subdomínio.	diploma eletrônico	Entidade central do domínio. Base documental em TDiploma e TInfDiploma, no leiauteDiplomaDigital_v1.05.xsd, expostos pelo DiplomaDigital_v1.05.xsd.
Diplomado	Pessoa que concluiu um curso de graduação e figura como titular do diploma digital correspondente.	egresso	Classe consolidada a partir dos termos “aluno”, “estudante” e “diplomado”, para reduzir ambiguidades terminológicas.
Instituição de Ensino Superior	Organização educacional pertencente ao sistema de ensino superior e envolvida, no contexto do diploma digital, em atividades de emissão, registro, manutenção ou disponibilização de dados acadêmicos. Constitui a classe institucional genérica do modelo, da qual se especializam os papéis de IES Emissora e IES Registradora.	IES; instituição de educação superior	Classe consolidada para representar genericamente as organizações de ensino superior presentes no domínio, sem fixar previamente a função exercida no processo de diplomação. Foi alinhada a schema:CollegeOrUniversity para favorecer a interoperabilidade com vocabulários externos, preservando, contudo, o significado específico adotado na OntoDiplomaDigital.
IES Emissora	Instituição de Ensino Superior responsável pela geração dos dados acadêmicos e pela emissão do diploma digital, podendo estar associada a informações institucionais e de localização por meio de uma estrutura própria de endereço.	instituição emissora	Especialização da classe Instituição de Ensino Superior, criada a partir do papel de emissão identificado nos leiautes XSD e na regulamentação do diploma digital. A separação dessa função permitiu distinguir semanticamente a emissão do registro, sem exigir que ambas sejam exercidas por organizações diferentes.
IES Registradora	Instituição de Ensino Superior responsável pelo registro institucional do diploma digital, podendo estar associada a informações institucionais e de localização por meio de uma estrutura própria de endereço.	instituição registradora	Especialização da classe Instituição de Ensino Superior, criada a partir do papel de registro identificado nos leiautes XSD e na regulamentação do diploma digital. A separação dessa função permitiu distinguir semanticamente o registro da emissão, sem exigir que ambas sejam exercidas por organizações diferentes.
Curso	Estrutura formativa concluída pelo diplomado, vinculada ao diploma digital e identificada por metadados institucionais e regulatórios.	curso superior	Corresponde ao bloco DadosCurso do leiaute oficial do MEC, incluindo a especialização TDadosCursoNSF.

Classe ou elemento conceitual	Definição adotada	Sinônimo ou variante	Origem documental ou decisão de modelagem
Curso NSF	Especialização de Curso correspondente ao ramo destinado à diplomação fora do Sistema Federal de Ensino. Mantém as características semânticas gerais de Curso, diferenciando-se pela origem estrutural e regulatória dos dados representados.	curso fora do Sistema Federal de Ensino	Subclasse criada a partir do tipo TDadosCursoNSF, distinguido de TDadosCurso pelo XSD leiauteDiplomaDigital_v1.05.xsd. A especialização preserva a diferença estabelecida no documento de origem sem duplicar toda a modelagem conceitual e todas as propriedades da classe geral Curso.
Curso em diplomação por decisão judicial	Especialização da classe Curso utilizada para representar os dados da formação acadêmica associados a um diploma cuja emissão ou registro decorreu de decisão judicial. Mantém o significado e as características gerais de Curso, diferenciando-se pela origem estrutural e processual dos dados no ramo judicial do diploma digital.	Curso associado a decisão judicial; curso do ramo judicial.	Subclasse derivada do tipo TDadosCursoPorDecisaoJudicial, integrante de TDadosDiplomaPorDecisaoJudicial no leiaute XSD do diploma digital. Sua inclusão preserva a distinção estrutural estabelecida no documento de origem e mantém coerência com a representação de Registro do Diploma por decisão judicial, sem caracterizar a decisão judicial como modalidade, grau ou natureza acadêmica do curso.
Habilitação	Estrutura acadêmica destinada ao registro de habilitações específicas do curso e de suas respectivas datas.	habilitação específica	Derivada do tipo THabilitacao, com os campos NomeHabilitacao e DataHabilitacao.
Informações de tramitação e-MEC	Estrutura destinada ao registro de dados processuais de cursos sem código e-MEC informado.	dados de tramitação e-MEC	Derivada do tipo TInformacoesTramitacaoEMEC, usado como alternativa a CodigoCursoEMEC e em estruturas regulatórias relacionadas.
Registro do Diploma	Ato institucional de registro do diploma digital, com dados de livro, folha, sequência, datas e agentes envolvidos.	ato de registro do diploma	Formaliza TDadosRegistro e TLivroRegistro, incluindo especializações NSF e judicial.
Registro do Diploma NSF	Especialização da classe Registro do Diploma utilizada para representar o ato institucional de registro associado a uma diplomação realizada fora do Sistema Federal de Ensino. Mantém o significado e as características gerais do registro do diploma, diferenciando-se pela origem estrutural e regulatória dos dados no ramo NSF.	Registro do diploma fora do Sistema Federal de Ensino; registro NSF.	Subclasse derivada do tipo TDadosRegistroNSF, previsto no leiaute XSD do diploma digital. Sua inclusão preserva a distinção estabelecida na estrutura de origem e mantém coerência com a classe Curso NSF, sem exigir a duplicação integral das relações, propriedades e restrições da classe geral Registro do Diploma.

Classe ou elemento conceitual	Definição adotada	Sinônimo ou variante	Origem documental ou decisão de modelagem
Responsável pelo Registro	Agente humano indicado como responsável administrativo pelo registro do diploma na IES registradora.	responsável administrativo pelo registro	Classe criada na versão 1.17 para representar explicitamente o agente indicado em ResponsavelRegistro.
Registro do Diploma por decisão judicial	Especialização de Registro do Diploma utilizada quando a emissão ou o registro decorre de determinação judicial, preservando os dados institucionais do registro e acrescentando os elementos necessários à identificação do processo judicial correspondente.	Registro judicial do diploma; registro por determinação judicial.	Especialização derivada do ramo judicial previsto no leiaute XSD do diploma digital. Sua inclusão permite distinguir o fundamento judicial do registro ordinário sem tratar a decisão judicial como diploma, instituição ou agente do processo acadêmico.
Assinante do Registro	Agente humano indicado no bloco Assinantes relativo ao registro do diploma.	assinante administrativo	Classe derivada do bloco Assinantes do registro, distinta de Assinatura Digital.
Assinante da Emissão	Agente humano indicado no bloco Assinantes relativo à emissão do diploma.	assinante administrativo da emissão	Classe derivada do bloco Assinantes da emissão, distinta de Assinatura Digital.
Assinatura Digital	Artefato criptográfico associado ao diploma digital, utilizado para assegurar autenticidade e integridade.	—	Abstração semântica de ds:Signature, referenciada nos leiautes MEC e definida no xmldsig-core-schema_v1.1.xsd.
Carimbo de Tempo	Evidência temporal vinculada à assinatura digital, destinada a registrar o momento da assinatura ou da validação.	marca temporal	Classe técnica complementar derivada dos metadados temporais associados à assinatura digital no contexto da ICP-Brasil.
Código de Validação	Identificador textual público associado ao diploma digital para consulta e verificação de autenticidade.	código verificador; código de consulta	Modelado como classe OWL por representar um recurso individual verificável associado ao diploma, e não um conceito controlado.
Histórico Escolar Digital	Documento digital que registra a trajetória acadêmica do estudante e serve de apoio informacional à emissão ou validação do diploma.	histórico acadêmico digital	Derivada do leiaute do Histórico Escolar Digital e mantida como documento de apoio informacional.
Documentação Acadêmica	Conjunto documental acadêmico que reúne evidências para emissão e registro do diploma e serve de ponto de acesso ao termo de responsabilidade da emissora.	documentação comprobatória	Classe agregadora derivada dos XSDs da documentação acadêmica e do documento-raiz DocumentacaoAcademicaRegistro.
Signatário do Termo de Responsabilidade da Emissora	Agente humano que assina o Termo de Responsabilidade da Emissora e formaliza a responsabilidade da IES emissora pela documentação acadêmica submetida.	signatário da responsabilidade da emissora	Classe derivada semanticamente dos campos Nome, CPF, Cargo e AtoDesignacao de TTermoResponsabilidade e da leitura normativa sobre a responsabilidade da emissora.
Endereço	Estrutura postal associada a instituições, cursos ou outras entidades do modelo.	endereço postal	Derivada do tipo TEndereco, reutilizado nos leiautes XSD para diferentes entidades.

Classe ou elemento conceitual	Definição adotada	Sinônimo ou variante	Origem documental ou decisão de modelagem
Agente Certificador	Organização pertencente à cadeia técnica de confiança da assinatura digital, representada no modelo para identificar, quando necessário, o agente relacionado à emissão, à gestão ou à validação dos certificados utilizados no artefato criptográfico. Atua exclusivamente na dimensão técnica da assinatura e não desempenha os papéis acadêmicos de emissão ou registro do diploma.	entidade certificadora; certificador técnico da assinatura	Classe complementar introduzida a partir dos elementos de certificação e infraestrutura X.509 associados ao padrão XMLDSig. Não corresponde a uma classe explicitamente nomeada no fluxo acadêmico-regulatório dos *schemas* do MEC, tendo sido incorporada como extensão semântica para representar a cadeia técnica de confiança da assinatura.
Plataforma Governamental	Plataforma digital governamental capaz de disponibilizar, consultar ou consumir dados interoperáveis do diploma digital.	—	Classe introduzida em atendimento aos cenários de aplicação e às questões de competência; Gov.br aparece como exemplo de plataforma, não como integração comprovada.

Fonte: Elaborado pelo autor com base nos resultados da formalização da OntoDiplomaDigital (2026).

Quadro 39 - Propriedades, atributos e regras de aplicação dos conceitos

Classe / elemento conceitual	Relações centrais	Atributos centrais	Nota de uso
Diploma Digital	tem titular; emitido por; registrado por; tem curso; tem registro; tem assinante da emissão; tem assinatura; tem código de validação; tem histórico; tem documentação acadêmica; tem situação; acessível via	versão; ambiente; data de emissão; identificador	Recurso central do modelo. Possui cardinalidades específicas para titular, IES emissora, IES registradora, curso, registro, assinatura, código de validação e situação.
Diplomado	titular de; concluiu curso	CPF; nome completo	Representa a pessoa que concluiu o curso e recebeu o diploma.
Instituição de Ensino Superior	assina institucionalmente; tem endereço	nome completo; código MEC; CNPJ	Classe geral para representar a IES independentemente do papel exercido. Pode ser especializada como IES Emissora, IES Registradora ou ambas.
IES Emissora	emite diploma; tem endereço	nome completo; código MEC; CNPJ	Representa o papel institucional de emissão. A mesma instituição pode também exercer o papel de registradora.
IES Registradora	registra diploma; tem endereço	nome completo; código MEC; CNPJ	Representa o papel institucional de registro. A mesma instituição pode também exercer o papel de emissora.
Curso	vinculado ao diploma; concluído por; confere grau; confere título; tem habilitação; tem	data de conclusão; nome; carga horária; ênfase; outro título conferido; código e-MEC	Reúne os elementos acadêmicos relacionados à formação certificada.

Classe / elemento conceitual	Relações centrais	Atributos centrais	Nota de uso
	informações de tramitação e-MEC; tem modalidade de oferta		
Curso NSF	herda as relações de Curso	herda os atributos de Curso	Utilizado quando os dados de origem correspondem ao ramo TDadosCursoNSF.
Curso em diplomação por decisão judicial	herda as relações de Curso	herda os atributos de Curso	Utilizado quando os dados têm origem no ramo TDadosCursoPorDecisaoJudicial.
Habilitação	habilitação de	nome; data	Representa habilitação específica vinculada ao curso.
Informações de tramitação e-MEC	—	número do processo; tipo; data de cadastro; data de protocolo	Utilizada quando não há código e-MEC, mas existem informações processuais.
Registro do Diploma	registro de; tem responsável pelo registro; assinado administrativamente por	data de registro; identificador da documentação acadêmica; número do processo; data de colação; data de expedição; livro; número; folha; sequência; informações adicionais	Representa o ato formal de registro do diploma.
Registro do Diploma NSF	herda as relações de Registro do Diploma	herda os atributos de Registro do Diploma	Utilizado quando os dados de origem correspondem ao ramo TDadosRegistroNSF.
Registro do Diploma por decisão judicial	herda as relações de Registro do Diploma	atributos gerais do registro; número do processo judicial; autoridade judicial	Utilizado quando o registro decorre do ramo de decisão judicial.
Responsável pelo Registro	responsável pelo registro de	nome; CPF; ID ou matrícula	Representa o agente indicado como responsável pelo registro.
Assinante do Registro	assina administrativamente	CPF; cargo	Representa o papel administrativo declarado no bloco de assinantes.
Assinante da Emissão	é assinante da emissão de	CPF; cargo	Representa o agente administrativo associado à emissão.
Assinatura Digital	assinatura de; tem carimbo de tempo; tem formato de assinatura; tem papel da assinatura; assinado institucionalmente por; tem certificador técnico	data e hora da assinatura	Representa o artefato técnico e criptográfico da assinatura.
Carimbo de Tempo	carimbo de tempo de	data e hora	Representa a evidência temporal vinculada à assinatura digital.

Classe / elemento conceitual	Relações centrais	Atributos centrais	Nota de uso
Código de Validação	valida diploma	valor do código	Identificador utilizado na consulta e verificação pública do diploma.
Histórico Escolar Digital	—	—	Classe contextual relacionada ao ecossistema do diploma.
Documentação Acadêmica	tem signatário do termo de responsabilidade da emissora	—	Representação controlada da documentação acadêmica associada ao registro.
Signatário do Termo de Responsabilidade da Emissora	é signatário do termo de responsabilidade de	nome; CPF; cargo; ato de designação	Representa a pessoa identificada no termo de responsabilidade da IES emissora.
Endereço	endereço de	logradouro; município; UF; CEP; país	Estrutura postal reutilizável associada a recursos institucionais ou acadêmicos.
Agente Certificador	certifica tecnicamente assinatura	denominação; identificadores técnicos	Classe técnica opcional para representação complementar da cadeia de certificação.
Plataforma Governamental	—	—	Classe contextual para cenários de acesso externo e interoperabilidade.

Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma digital (MEC)

Os Quadro 38 e Quadro 39 demonstram como os elementos extraídos na Etapa 2 foram reorganizados em uma estrutura conceitual mais estável. O primeiro explicita o significado atribuído aos conceitos e sua relação com as fontes ou decisões de modelagem. O segundo mostra como esses conceitos foram articulados por meio de relações e atributos, preparando sua posterior formalização como classes, propriedades de objeto, propriedades de dados e conceitos controlados.

Essa organização diferencia a terminologia identificada nas fontes, a interpretação conceitual realizada na pesquisa e a estrutura formal posteriormente registrada no TTL. A conceitualização, portanto, não consistiu em simples reprodução dos XSDs, mas em uma etapa analítica na qual os elementos do Registro do Diploma foram selecionados, definidos, relacionados e delimitados segundo sua função no domínio.

Além da definição dos conceitos e de sua articulação por relações e atributos, a conceitualização registrou restrições e decisões de modelagem necessárias para delimitar o uso de determinados elementos da OntoDiplomaDigital. Essas restrições indicaram cardinalidades mínimas, máximas ou exatas para relações centrais do modelo, enquanto as decisões de modelagem explicitaram especializações e classes auxiliares criadas para preservar distinções presentes nas fontes técnicas. O Quadro 40 reúne as principais restrições e decisões associadas às classes da ontologia.

Quadro 40 - Restrições e decisões de modelagem das classes da ontologia

Classe	Restrição ou decisão de modelagem
Diploma Digital	Pelo menos 1 titular; exatamente 1 IES emissora; exatamente 1 IES registradora; pelo menos 1 curso; no máximo 1 registro; pelo menos 1 assinatura digital; no máximo 1 código de validação; no máximo 1 situação; exatamente 1 ambiente operacional.
Curso	Exatamente 1 grau conferido; no máximo 1 título conferido; exatamente 1 modalidade de oferta; no máximo 1 informação de tramitação e-MEC; no máximo 1 código e-MEC; no máximo 1 outro título conferido; exatamente 1 endereço.
Curso por Decisão Judicial	Subclasse de Curso; representa o ramo de decisão judicial previsto no XSD sem constituir categoria acadêmica própria.
Alternativa de Endereço do Curso por Decisão Judicial	Classe lógica auxiliar que formaliza a alternativa entre possuir exatamente 1 endereço ou declarar endereço indisponível como verdadeiro. A exclusividade é tratada na validação operacional, não inferida apenas pela união OWL.
Habilitação	Exatamente 1 nome e exatamente 1 data.
Registro do Diploma	Exatamente 1 data de registro; exatamente 1 requisição de documentação acadêmica; exatamente 1 processo administrativo; exatamente 1 data de colação de grau; exatamente 1 data de expedição do diploma; exatamente 1 responsável pelo registro.

Classe	Restrição ou decisão de modelagem
Registro do Diploma NSF	Subclasse de Registro do Diploma; identifica o ramo estrutural NSF previsto no XSD.
Assinatura Digital	Pelo menos 1 carimbo de tempo; no máximo 1 formato de assinatura; no máximo 1 papel da assinatura.
Agente Certificador	Classe complementar para representar a cadeia técnica de confiança da assinatura digital, sem corresponder a ator acadêmico-regulatório do processo de emissão ou registro.

Fonte: Elaborado pelo autor (2026), com base na conceitualização do subdomínio e confrontado com o arquivo canônico OntoDiplomaDigital_registro_diploma.ttl.

O Quadro 40 complementa a descrição dos conceitos ao evidenciar que a conceitualização também envolveu a delimitação de regras de uso. As restrições associadas a Diploma Digital, Curso, Registro do Diploma, Assinatura Digital e Habilitação indicam quais relações ou atributos deveriam ser obrigatórios, únicos ou limitados no modelo. Já as especializações e classes auxiliares, como Curso por Decisão Judicial, Registro do Diploma NSF e Alternativa de Endereço do Curso por Decisão Judicial, registram decisões tomadas para preservar diferenças estruturais identificadas nas fontes técnicas sem transformar cada variação documental em uma ontologia separada.

Essas restrições e decisões prepararam a passagem da conceitualização para a formalização. Na Etapa 3 da metodologia OntoForInfoScience, elas funcionaram como regras conceituais de organização do domínio; na etapa de formalização, foram representadas por recursos OWL, como owl:Restriction, cardinalidades qualificadas, subclasses e classe auxiliar baseada em união.

5.3.2 Dicionário de Verbos

O dicionário de verbos consolidou as relações identificadas durante a conceitualização do subdomínio do Registro do Diploma e estabeleceu correspondência entre os rótulos em português e as propriedades declaradas no TTL canônico da OntoDiplomaDigital. Para cada relação, foram registrados o domínio, o alcance, a propriedade OWL correspondente e a definição adotada. Esse resultado explicitou como os vínculos entre diplomas, pessoas, instituições, cursos, registros, documentos, mecanismos de validação e valores controlados foram organizados na estrutura conceitual da ontologia.

O Quadro 41 apresenta a primeira parte desse dicionário, reunindo as relações modeladas na ontologia com seus respectivos rótulos em português, propriedades OWL, domínios, alcances e definições.

Quadro 41 – Dicionário de verbos

Relação	Propriedade OWL	Domínio → Alcance	Definição
tem titular	hasHolder	Diploma Digital → Diplomado	Relaciona o diploma digital ao diplomado que dele é titular.
titular de	holdsDiploma	Diplomado → Diploma Digital	Relaciona o diplomado ao diploma digital do qual é titular.
tem local de naturalidade	hasBirthPlace	Diplomado → Local de Naturalidade	Relaciona o diplomado ao local de nascimento informado no elemento Naturalidade do XSD.
emitido por	issuedBy	Diploma Digital → IES Emissora	Relaciona o diploma digital à instituição que exerce o papel de emissora.
emite diploma	issuesDiploma	IES Emissora → Diploma Digital	Relaciona a IES emissora ao diploma digital emitido.
registrado por	registeredBy	Diploma Digital → IES Registradora	Relaciona o diploma digital à instituição que exerce o papel de registradora.
registra diploma	registersDiploma	IES Registradora → Diploma Digital	Relaciona a IES registradora ao diploma digital que ela registra.
tem curso	hasProgram	Diploma Digital → Curso	Relaciona o diploma digital ao curso concluído pelo diplomado.
vinculado ao diploma	isProgramOfDiploma	Curso → Diploma Digital	Relaciona o curso ao diploma digital ao qual está vinculado.
concluído por	completedBy	Curso → Diplomado	Relaciona o curso ao diplomado que o concluiu.
concluiu curso	completedProgram	Diplomado → Curso	Relaciona o diplomado ao curso concluído.
tem registro	hasRegistration	Diploma Digital → Registro do Diploma	Relaciona o diploma digital ao seu ato formal de registro.
registro de	isRegistrationOf	Registro do Diploma → Diploma Digital	Relaciona o registro formal ao diploma digital correspondente.
tem requisição de documentação acadêmica	hasAcademicDocumentationRequest	Registro do Diploma → Requisição de Documentação Acadêmica	Relaciona o registro do diploma à requisição de documentação acadêmica identificada no XSD.
é requisição de documentação acadêmica de	isAcademicDocumentationRequestOf	Requisição de Documentação Acadêmica → Registro do Diploma	Relação inversa de hasAcademicDocumentationRequest.
tem processo administrativo	hasAdministrativeProcess	Registro do Diploma → Processo Administrativo	Relaciona o registro do diploma ao processo administrativo associado ao registro ou à expedição.
é processo administrativo de	isAdministrativeProcessOf	Processo Administrativo → Registro do Diploma	Relação inversa de hasAdministrativeProcess.

Relação	Propriedade OWL	Domínio → Alcance	Definição
tem responsável pelo registro	hasRegistrationResponsible	Registro do Diploma → Responsável pelo Registro	Relaciona o registro do diploma à pessoa responsável indicada no bloco ResponsavelRegistro do XSD.
responsável pelo registro de	isRegistrationResponsibleOf	Responsável pelo Registro → Registro do Diploma	Relaciona a pessoa responsável ao registro do diploma correspondente.
assinado administrativamente por	hasRegistrationSigner	Registro do Diploma → Assinante do Registro	Relaciona o registro do diploma ao assinante administrativo indicado no bloco Assinantes.
assina administrativamente	isRegistrationSignerOf	Assinante do Registro → Registro do Diploma	Relaciona o assinante administrativo ao registro do diploma correspondente.
tem assinante da emissão	hasIssuanceSigner	Diploma Digital → Assinante da Emissão	Relaciona o diploma digital ao assinante administrativo indicado no bloco Assinantes da emissão.
é assinante da emissão de	isIssuanceSignerOf	Assinante da Emissão → Diploma Digital	Relaciona o assinante administrativo da emissão ao diploma digital correspondente.
tem signatário do termo de responsabilidade da emissora	hasIssuanceResponsibilitySigner	Documentação Acadêmica → Signatário do Termo de Responsabilidade da Emissora	Relaciona a documentação acadêmica ao agente humano identificado no termo de responsabilidade da IES emissora.
é signatário do termo de responsabilidade de	isIssuanceResponsibilitySignerOf	Signatário do Termo de Responsabilidade da Emissora → Documentação Acadêmica	Relaciona o signatário do termo de responsabilidade da emissora à documentação acadêmica correspondente.
tem assinatura	hasSignature	Diploma Digital → Assinatura Digital	Relaciona o diploma digital à assinatura digital associada.
assinatura de	isSignatureOf	Assinatura Digital → Diploma Digital	Relaciona a assinatura digital ao diploma digital correspondente.
tem carimbo de tempo	hasTimestamp	Assinatura Digital → Carimbo de Tempo	Relaciona a assinatura digital ao carimbo de tempo associado.
carimbo de tempo de	isTimestampOf	Carimbo de Tempo → Assinatura Digital	Relaciona o carimbo de tempo à assinatura digital correspondente.

Fonte: Elaborado pelo autor (2026), com base na conceitualização do subdomínio e confrontado com o arquivo canônico OntoDiplomaDigital_registro_diploma.ttl.

As relações apresentadas no Quadro 42 evidenciam que o núcleo relacional da OntoDiplomaDigital foi organizado por vínculos diretos e, em muitos casos, por suas leituras inversas. Essa organização permite consultar o modelo tanto a partir do diploma quanto a partir das entidades associadas a ele, como diplomado, curso, instituição emissora, instituição registradora, registro, assinatura e documentação acadêmica. A distinção entre papéis institucionais e agentes humanos também contribui para separar emissão, registro, assinatura administrativa e responsabilidade documental.

Quadro 42 - Relações semânticas complementares

Relação	Propriedade OWL	Domínio → Alcance	Definição
tem código de validação	hasValidationCode	Diploma Digital → Código de Validação	Relaciona o diploma digital ao código de validação utilizado na consulta pública.
valida diploma	validatesDiploma	Código de Validação → Diploma Digital	Relaciona o código de validação ao diploma digital que ele permite verificar.
tem histórico	hasTranscript	Diploma Digital → Histórico Escolar Digital	Relaciona o diploma digital ao histórico escolar digital do diplomado.
tem documentação acadêmica	hasAcademicDocumentation	Diploma Digital → Documentação Acadêmica	Relaciona o diploma digital ao conjunto de documentos acadêmicos que o fundamentam.
confere grau	awardsDegree	Curso → Conceito de Grau Conferido	Relaciona o curso ao conceito SKOS que representa o grau conferido ao diplomado.
confere título	awardsTitle	Curso → Conceito de Título Conferido	Relaciona o curso ao conceito SKOS que representa o título conferido ao diplomado.
tem habilitação	hasQualification	Curso → Habilitação	Relaciona o curso às habilitações específicas registradas no leiaute do diploma.
habilitação de	isQualificationOf	Habilitação → Curso	Relaciona a habilitação ao curso ao qual está associada.
tem informações de tramitação e-MEC	hasEmecProcessingInfo	Curso → Informações de Tramitação e-MEC	Relaciona o curso à estrutura usada quando não há código e-MEC disponível, mas há informação de tramitação processual.
tem modalidade de oferta	hasModeOfDelivery	Curso → Conceito de Modalidade de Oferta	Relaciona o curso ao conceito SKOS que representa sua modalidade de oferta.
tem situação	hasStatus	Diploma Digital → Conceito de Situação do Diploma	Relaciona o diploma digital ao conceito SKOS que representa sua situação administrativa ou jurídica.
tem ambiente operacional	hasEnvironment	Diploma Digital → Conceito de Ambiente Operacional	Relaciona o diploma digital ao conceito controlado que representa seu ambiente operacional.
tem formato de assinatura	hasSignatureFormat	Assinatura Digital → Conceito de Formato de Assinatura	Relaciona a assinatura digital ao conceito SKOS que representa seu formato técnico.

Relação	Propriedade OWL	Domínio → Alcance	Definição
tem papel da assinatura	hasSignatureRole	Assinatura Digital → Conceito de Papel da Assinatura	Relaciona a assinatura digital ao conceito SKOS que representa seu papel semântico-administrativo no fluxo do diploma.
assinado institucionalmente por	institutionallySignedBy	Assinatura Digital → Instituição de Ensino Superior	Relaciona a assinatura digital à instituição de ensino superior que a chancela institucionalmente como pessoa jurídica.
assina institucionalmente	institutionallySigns	Instituição de Ensino Superior → Assinatura Digital	Relaciona a instituição de ensino superior à assinatura digital que ela chancela institucionalmente.
tem certificador técnico	certifiedBy	Assinatura Digital → Agente Certificador	Relaciona a assinatura digital a um agente certificador quando se deseja explicitar a cadeia técnica de confiança ou certificação da assinatura.
certifica tecnicamente assinatura	certifiesSignature	Agente Certificador → Assinatura Digital	Relaciona o agente certificador à assinatura digital somente no plano técnico complementar da cadeia de confiança.
acessível via	isAccessibleVia	Diploma Digital → Plataforma Governamental	Relaciona o diploma digital à plataforma governamental por meio da qual pode ser consultado interoperavelmente.
tem endereço	hasAddress	Não declarado no TTL → Endereço	Relaciona instituição de ensino superior ou curso ao endereço postal correspondente.
tem valor controlado	hasControlledValue	Não declarado no TTL → Valor Controlado	Propriedade abstrata que agrupa relações cujo objeto deve ser um conceito SKOS pertencente a vocabulário controlado.
tem objetivo relacionado	hasRelatedObjective	Questão de Competência → Objetivo Específico	Relaciona uma questão de competência ao objetivo específico da pesquisa que ela operacionaliza ou valida.

Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma digital (MEC)

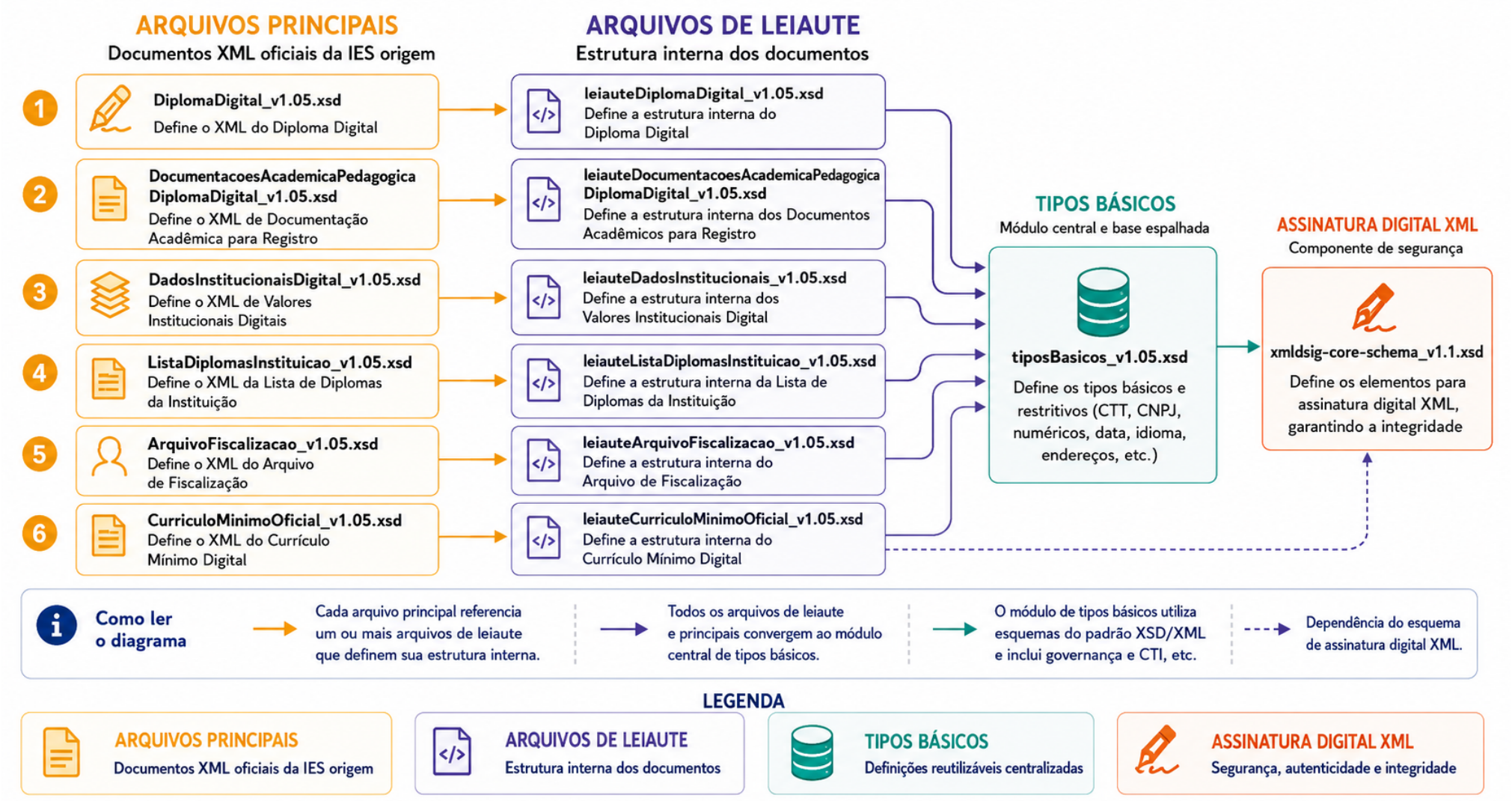
O segundo conjunto demonstra que a conceitualização não se restringiu aos vínculos estruturais diretamente extraídos dos XSDs. A OntoDiplomaDigital incorporou relações destinadas à caracterização semântica dos cursos, diplomas e assinaturas, utilizando conceitos controlados como objetos de propriedades especializadas. Relações como awardsDegree, awardsTitle, hasModeOfDelivery, hasStatus, hasSignatureFormat, hasSignatureRole e hasEnvironment foram associadas a valores controlados, evitando sua representação apenas como literais livres.

Também foram registradas relações destinadas a distinguir as dimensões administrativa, institucional e técnica da assinatura digital. As propriedades institutionallySignedBy e institutionallySigns representam a chancela institucional da Instituição de Ensino Superior, enquanto certifiedBy e certifiesSignature permanecem restritas à cadeia técnica de certificação. Outras relações, como isAccessibleVia, hasAddress, hasBirthPlace e hasRelatedObjective, ampliam o dicionário para contemplar, respectivamente, acesso por plataforma governamental, associação a endereço postal, naturalidade do diplomado e vínculo entre questões de competência e objetivos específicos da pesquisa.

5.3.3 Modelos conceituais gráficos

Os modelos conceituais gráficos consolidaram visualmente os principais resultados da Etapa 3, permitindo representar a organização das classes, das relações e das variações estruturais identificadas no subdomínio do Registro do Diploma. Esses modelos não tiveram a finalidade de reproduzir integralmente os *schemas* XSD, mas de explicitar os elementos considerados relevantes para a conceitualização da OntoDiplomaDigital.

Figura 10 - Arquitetura dos arquivos XSD/XML - MEC



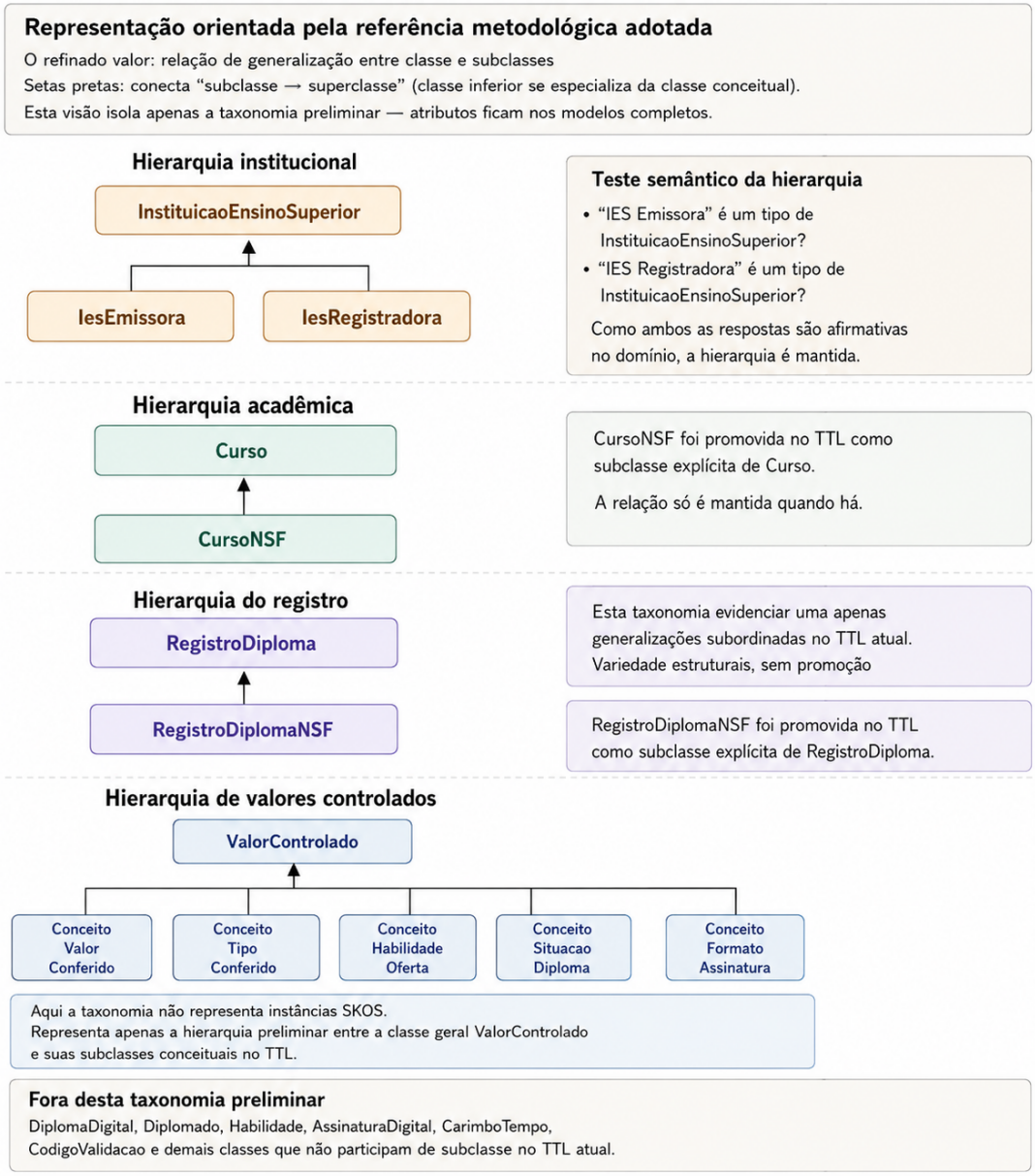
Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma digital (MEC).

A Figura 10 apresenta a organização geral dos arquivos XSD/XML do Diploma Digital Brasileiro. Esse modelo funcionou como apoio técnico para compreender a relação entre os documentos principais, os leiautes internos, os tipos básicos reutilizáveis e os componentes de assinatura digital. Sua função na conceitualização foi evidenciar que os elementos do domínio não estavam isolados, mas distribuídos em uma arquitetura documental composta por arquivos principais e auxiliares.

A Figura 11 apresenta a taxonomia da OntoDiplomaDigital, organizada a partir das principais classes identificadas no subdomínio. O modelo evidencia as relações de generalização e especialização entre os elementos, como a especialização de Instituição de Ensino Superior em IES Emissora e IES Registradora, de Curso em Curso NSF e Curso por Decisão Judicial, e de Registro do Diploma em Registro do Diploma NSF. Também diferencia as classes centrais do domínio das categorias associadas a valores controlados.

.

Figura 11 - Taxonomia da ontologia



Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma digital (MEC)

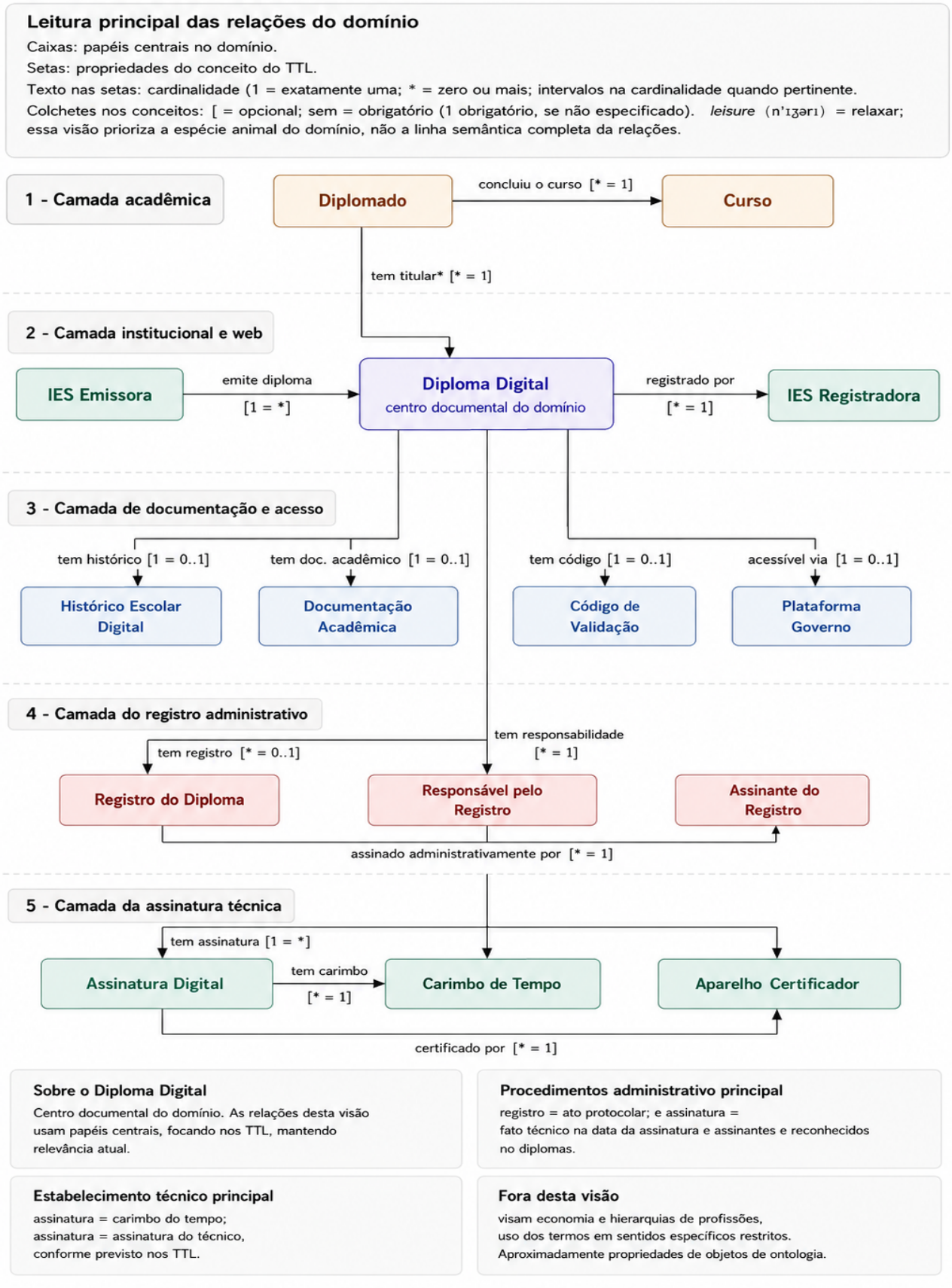
A taxonomia contribuiu para estabilizar a leitura conceitual do domínio antes da formalização. A separação entre entidades do domínio e valores controlados foi especialmente relevante, pois permitiu distinguir recursos como Diploma Digital, Curso, Registro do Diploma e Assinatura Digital de categorias terminológicas usadas para qualificar esses recursos, como grau conferido, título conferido, modalidade de oferta, situação do diploma, ambiente operacional, formato de assinatura e papel da assinatura.

A Figura 12 apresenta o encadeamento principal das relações do domínio do Registro do Diploma, organizado em camadas para facilitar a leitura conceitual do

modelo. A camada acadêmica reúne o Diplomado e o Curso, evidenciando a formação concluída que fundamenta a emissão do diploma. A camada institucional explicita os papéis da IES Emissora e da IES Registradora. A camada documental e de acesso relaciona o Diploma Digital ao Histórico Escolar Digital, à Documentação Acadêmica, ao Código de Validação e à Plataforma Governamental. A camada do registro administrativo representa o Registro do Diploma, o Responsável pelo Registro e o Assinante do Registro. Por fim, a camada técnica de autenticação reúne a Assinatura Digital, o Carimbo de Tempo e o Agente Certificador.

No centro dessas camadas, o Diploma Digital funciona como recurso articulador do domínio, pois conecta o titular, o curso concluído, as instituições envolvidas, o registro, os documentos de apoio e os mecanismos de validação. As setas indicam as relações entre os elementos, enquanto as cardinalidades representam a quantidade esperada de vínculos entre os recursos.

Figura 12 - Encadeamento da relação do domínio Registro do Diploma



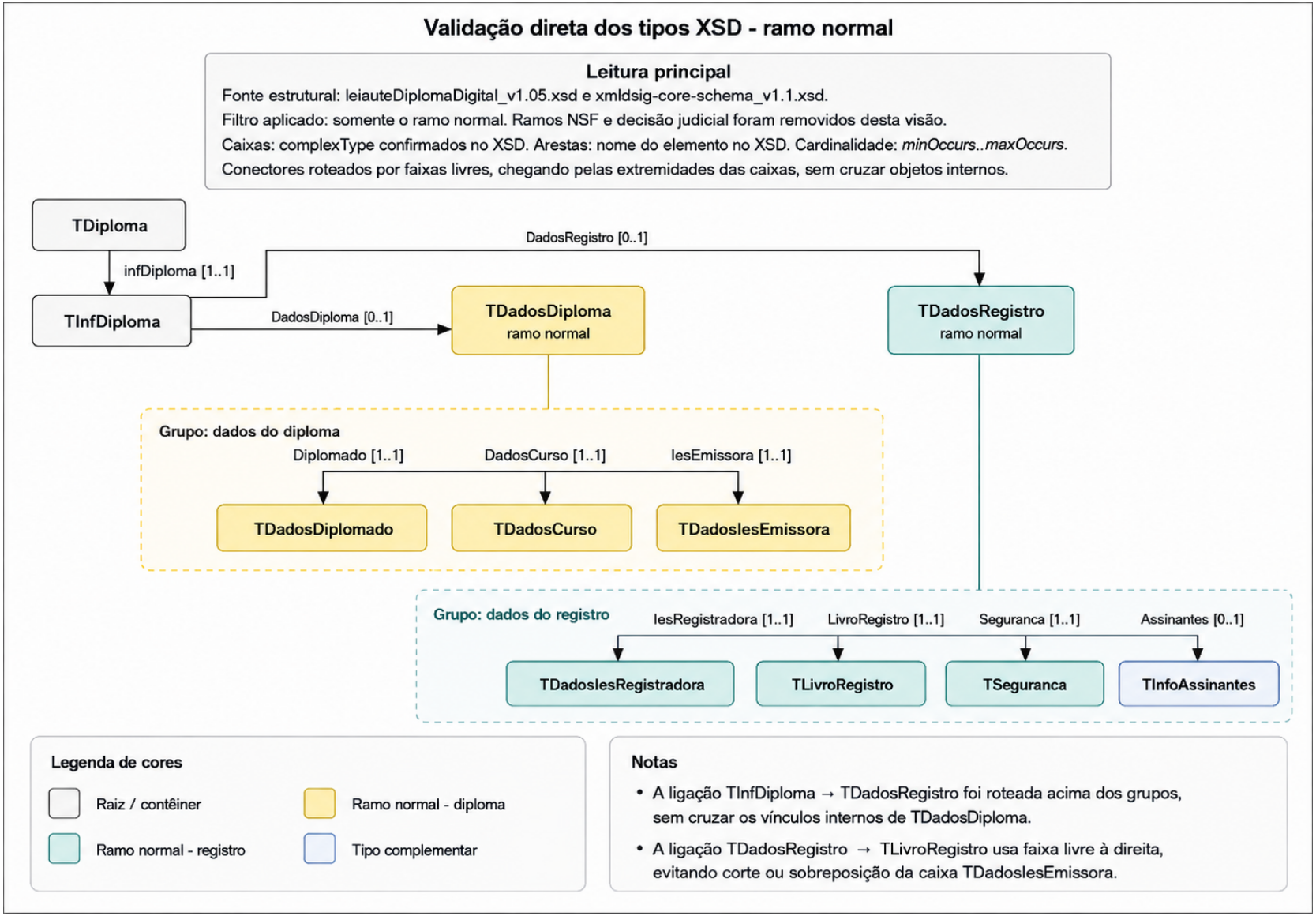
Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma digital (MEC)

O encadeamento das relações demonstrou que o diploma digital não foi tratado como documento isolado, mas como recurso informacional situado em uma rede de agentes, documentos, registros e mecanismos de validação. Essa representação também ajudou a distinguir dimensões diferentes do domínio: a formação acadêmica do diplomado, a responsabilidade institucional das IES, o ato administrativo de registro, os documentos de apoio e os recursos técnicos de autenticidade e validação.

As Figura 13, Figura 14 e Figura 15 apresentam modelos auxiliares dos ramos estruturais previstos nos *schemas* XSD do Diploma Digital. O ramo normal representa o fluxo ordinário de diplomação e registro; o ramo por decisão judicial contempla situações em que a emissão ou o registro decorre de determinação judicial; e o ramo NSF representa o fluxo associado a instituições fora do Sistema Federal de Ensino. Esses modelos foram utilizados para compreender as variações estruturais que precisavam ser consideradas na conceitualização.

.

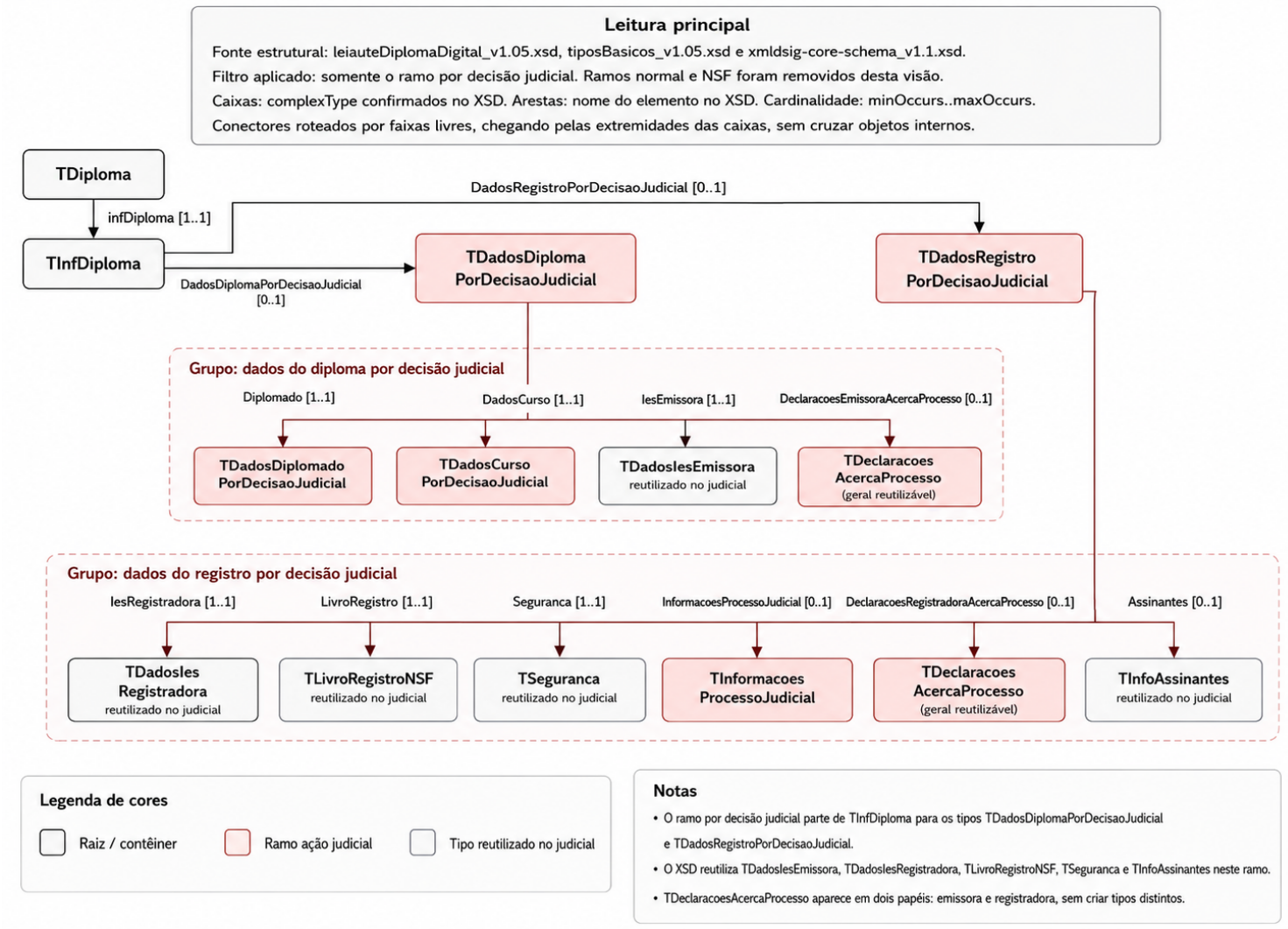
Figura 13 - Validação dos tipos XSD/XML - ramo normal



Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma digital (MEC)

O segundo corresponde ao ramo de ação judicial, destinado aos casos em que o registro do diploma decorre de determinação judicial.

Figura 14 - Validação dos tipos XSD/XML - ramo judicial

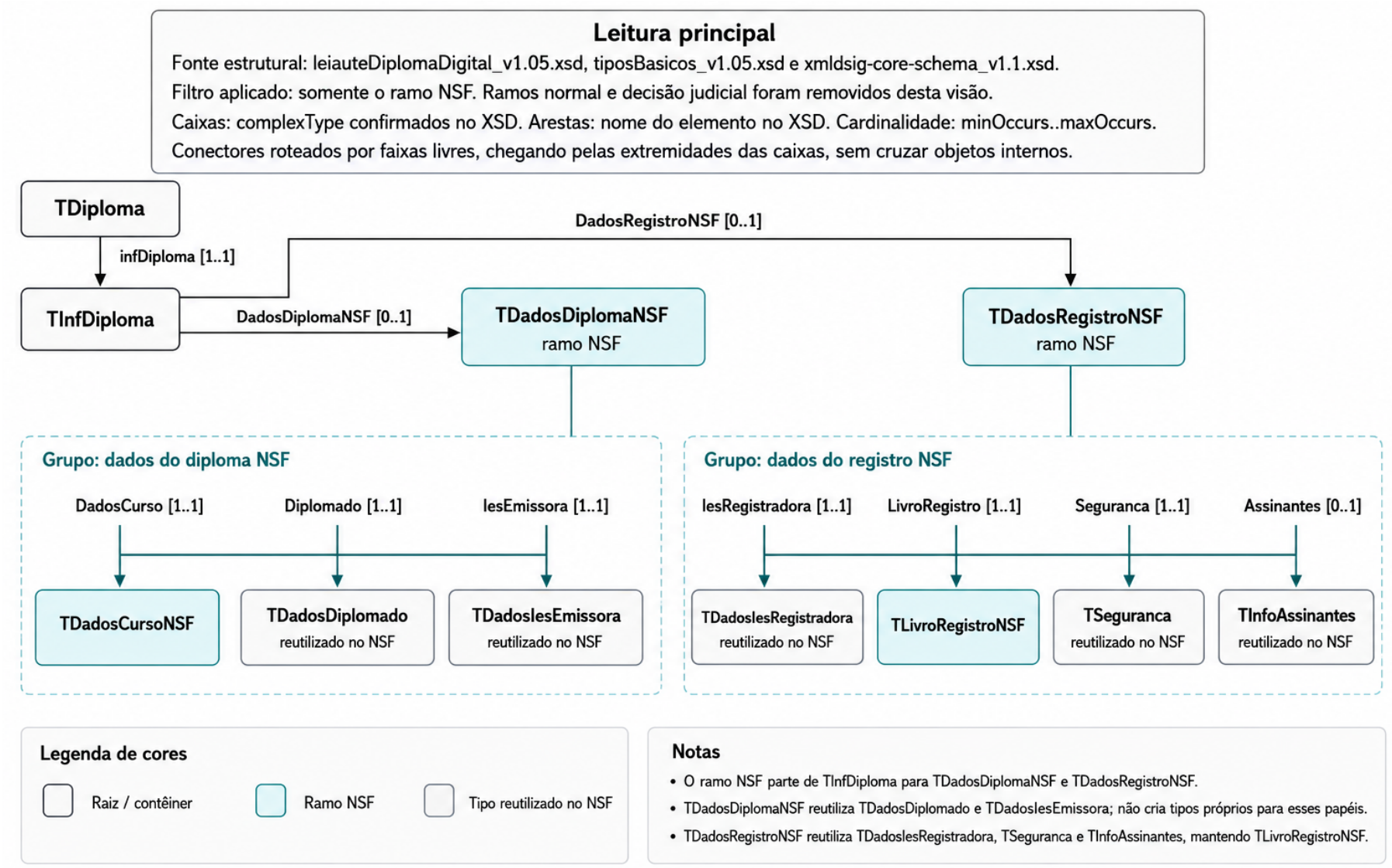


Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma Digital (MEC).

O terceiro corresponde ao ramo NSF, associado aos tipos TDadosDiplomaNSF, TDadosCursoNSF, TDadosRegistroNSF e TLivroRegistroNSF, utilizado para representar o fluxo fora do sistema federal.

Essa organização demonstra que o XSD não define apenas campos isolados, mas também alternativas documentais conforme a natureza administrativa ou jurídica do registro. No interior de TInfDiploma, o leiaute permite escolhas condicionais entre dados do diploma e dados do registro, contemplando o fluxo regular, o fluxo NSF e o fluxo por decisão judicial. Por reunir obrigatoriamente o elemento infDiploma e a assinatura digital, a classe TDiploma permanece como estrutura central do documento, pois inclui obrigatoriamente o elemento infDiploma, do tipo TInfDiploma, e o elemento de assinatura digital ds:Signature, conectando os metadados do diploma à sua dimensão jurídica e criptográfica.

Figura 15 - Validação do XSD/XML - ramo NSF



Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma Digital (MEC).

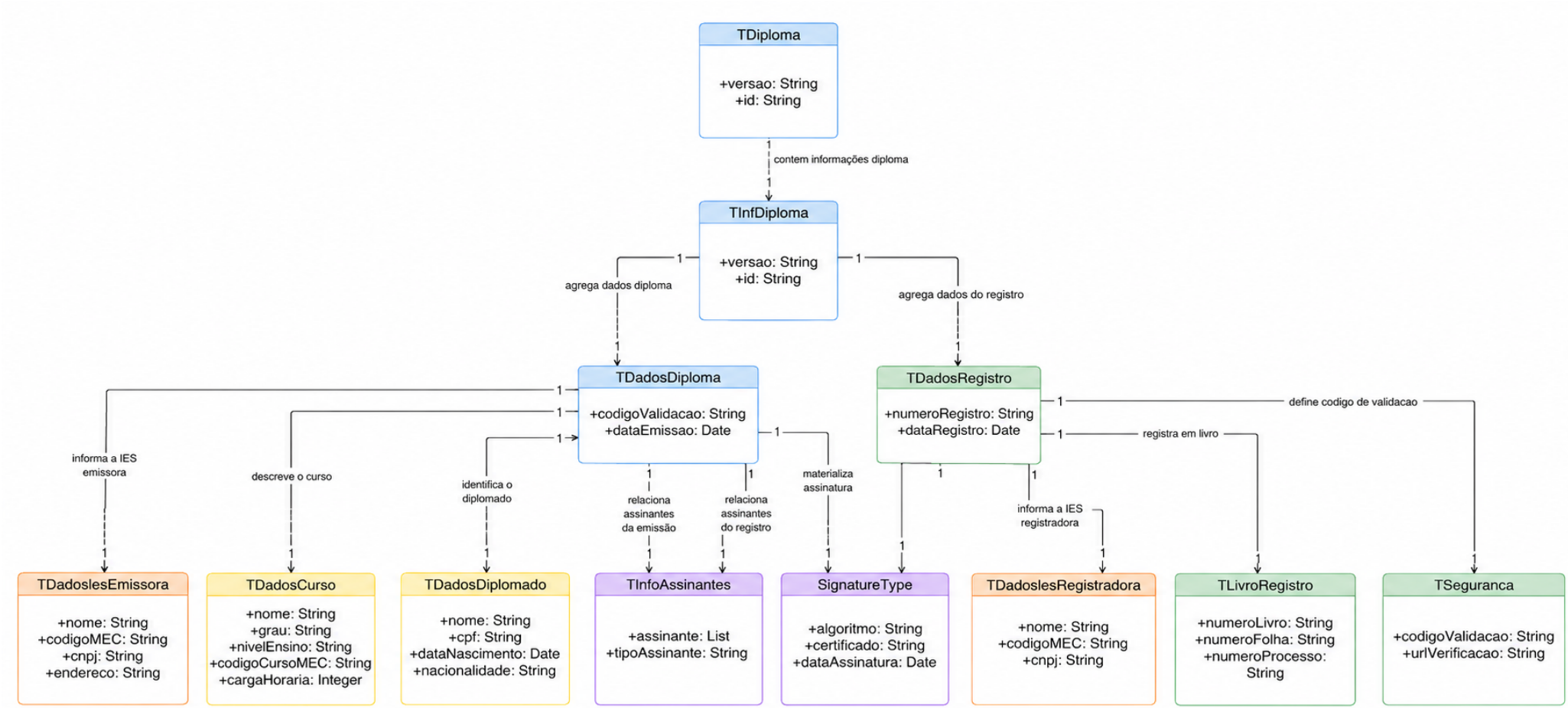
A representação dos três ramos evidenciou que o XSD define alternativas documentais conforme a natureza administrativa ou jurídica do registro. Essa distinção foi importante para evitar que variações estruturais fossem tratadas como conceitos equivalentes. Na OntoDiplomaDigital, o ramo normal, o ramo NSF e o ramo por decisão judicial foram interpretados como variações relevantes do domínio, mas sem transformar cada estrutura documental em uma ontologia separada.

Em conjunto, os modelos gráficos demonstram a passagem da leitura técnico-sintática dos XSDs para a organização conceitual da ontologia. As figuras permitiram identificar classes, especializações, relações e restrições que posteriormente orientaram a formalização em OWL, RDF e SKOS. A conceitualização gráfica, portanto, funcionou como etapa intermediária entre a análise das fontes e a implementação do TTL canônico.

A Figura 16 apresenta o modelo conceitual de encadeamento das relações do Diploma Digital. Diferentemente da Figura , que organiza o domínio em camadas de leitura, esse modelo sintetizou o encadeamento conceitual das entidades e relações que estruturam o Registro do Diploma. Sua função foi demonstrar, de forma integrada, como os elementos identificados nos XSDs e nos modelos anteriores foram reorganizados como classes, relações e pontos de articulação da OntoDiplomaDigital.

.

Figura 16 - Modelo conceitual de encadeamento das relações do diploma digital



Fonte: Elaborado pelo autor com base nos resultados da formalização do Diploma Digital (MEC)

O modelo representado na Figura 11 evidenciou a passagem da leitura técnico-sintática dos XSDs para uma estrutura conceitual orientada à ontologia. Ao explicitar as conexões entre diploma, diplomado, curso, instituições, registro, documentação, assinatura e mecanismos de validação, o gráfico ofereceu uma visão de síntese para a Etapa 3. Essa representação apoiou a identificação das classes, propriedades e restrições posteriormente formalizadas no TTL canônico, sem reduzir a ontologia a uma conversão automática dos arquivos XML.

5.4 Resultado da fundamentação ontológica – Etapa 4 da OntoForInfoScience

A Etapa 4 da OntoForInfoScience prevê a explicitação da fundamentação ontológica do modelo e considera a possibilidade de utilização de ontologias de alto nível como suporte à definição das categorias empregadas na representação. Entretanto, Mendonça (2015, p. 209–214) sugere que a aplicação dessa etapa seja adaptada às características do artefato desenvolvido, especialmente em ontologias de menor porte, ontologias leves de domínio ou subontologias. Essa possibilidade mostrou-se pertinente à OntoDiplomaDigital, cujo escopo foi delimitado ao subdomínio do Registro do Diploma e cuja finalidade consistiu na construção de uma ontologia operacional voltada à representação e à recuperação computacional dos dados desse recorte.

A decisão de não vincular a OntoDiplomaDigital, nesta versão, a uma ontologia de fundamentação decorreu, portanto, da delimitação do escopo e dos objetivos estabelecidos para a pesquisa, e não de uma incompatibilidade entre ontologias operacionais e ontologias de alto nível. Conforme Falbo (2014), ontologias operacionais correspondem a representações implementadas considerando requisitos computacionais e destinadas ao uso por aplicações. No caso desta pesquisa, o esforço de modelagem concentrou-se na passagem dos conceitos, relações e restrições identificados nas normas e nos schemas XSD do Ministério da Educação para uma representação formal que pudesse ser materializada com dados, processada e consultada computacionalmente.

A incorporação de uma ontologia de fundamentação exigiria um nível adicional de análise das categorias ontológicas presentes no modelo, estabelecendo compromissos mais gerais sobre a natureza das entidades representadas e suas

relações. Embora esse aprofundamento pudesse ampliar o rigor conceitual da ontologia, não foi considerado necessário para responder às questões de competência nem para realizar a prova de conceito definida no escopo desta dissertação. A opção adotada preservou, assim, a proximidade entre o modelo conceitual construído, as fontes normativas e técnicas examinadas e a representação operacional implementada.

Essa decisão, entretanto, não elimina a necessidade de explicitar e controlar os compromissos conceituais assumidos pela ontologia. Guizzardi (2020) destaca que a contribuição das ontologias para a interoperabilidade semântica depende da explicitação dos significados atribuídos às entidades e às relações representadas. Da mesma forma, a formalização computacional, isoladamente, não assegura uniformidade de interpretação, pois os compromissos conceituais expressos pelo modelo precisam ser adequadamente definidos, documentados e compartilhados (Guizzardi, 2020; Noy; McGuinness, 2001). Por essa razão, a não adoção de uma ontologia de fundamentação foi acompanhada pela definição explícita dos conceitos, relações, restrições, fontes e justificativas de modelagem desenvolvidas nas etapas anteriores e posteriormente formalizadas na OntoDiplomaDigital.

Como estratégia complementar de interoperabilidade, foram analisados vocabulários e modelos consolidados da Web Semântica, entre eles DCMI Metadata Terms, PROV-O, FOAF, DCAT, Schema.org, LOCN e vCard RDF. O reúso foi realizado somente quando se verificou compatibilidade entre o significado atribuído ao elemento da OntoDiplomaDigital e o significado, o domínio, o alcance e a finalidade do termo externo. Quando não foi identificada correspondência semanticamente adequada, foram mantidos termos próprios da ontologia, acompanhados de definição, fonte e justificativa de modelagem.

O resultado da Etapa 4 consistiu, portanto, na decisão fundamentada de não adotar, no escopo desta versão da OntoDiplomaDigital, uma ontologia de alto nível como estrutura superior, preservando-se os compromissos conceituais necessários ao domínio por meio da conceitualização desenvolvida nas etapas anteriores e do reúso semântico controlado de vocabulários externos. A adoção futura de uma ontologia de fundamentação permanece como possibilidade de evolução da OntoDiplomaDigital, especialmente em cenários que exijam ampliação do domínio, integração com outras ontologias ou maior explicitação das categorias ontológicas empregadas.

5.5 Resultado da Formalização da ontologia – Etapa 5 da OntoForInfoScience

A Etapa 5 resultou na formalização dos conceitos, relações, atributos, valores controlados e restrições definidos nas etapas anteriores em uma representação computacional da OntoDiplomaDigital. O resultado principal dessa etapa foi o arquivo Turtle canônico da ontologia, no qual foram registradas as classes do domínio, as propriedades de objeto, as propriedades de dados, as hierarquias, as restrições OWL, os conceitos controlados em SKOS e os elementos de documentação e rastreabilidade.

O arquivo canônico reuniu 3.009 triplas, com 39 classes locais, 51 propriedades de objeto, 58 propriedades de dados, 13 propriedades de anotação, 28 restrições OWL nomeadas e 31 conceitos SKOS. Essas contagens foram obtidas diretamente da versão canônica do arquivo Turtle e correspondem ao conjunto efetivamente formalizado na TBox da OntoDiplomaDigital, não apenas aos termos apresentados nos quadros de síntese da dissertação.

Como forma de acesso ao artefato produzido, a ontologia foi disponibilizada em repositório público no GitHub, indicado no Apêndice B, juntamente com os demais arquivos da prova de conceito. Essa disponibilização permite ao leitor consultar o TTL canônico, examinar sua estrutura formal e relacionar os elementos descritos na dissertação aos artefatos computacionais efetivamente produzidos. A disponibilização no GitHub, entretanto, não caracteriza a implantação da ontologia como serviço institucional em rede ou endpoint SPARQL permanente, mas sim como publicação acadêmica dos artefatos desenvolvidos na pesquisa.

5.5.1 Arquitetura formal implementada

A formalização da OntoDiplomaDigital resultou na combinação entre um núcleo ontológico em OWL, uma camada terminológica em SKOS, metadados descritivos e mecanismos de rastreabilidade das fontes técnicas. Essa organização permitiu separar os recursos centrais do subdomínio do Registro do Diploma, os valores controlados usados para qualificá-los, as informações de documentação da ontologia e as referências aos schemas XSD que subsidiaram a modelagem.

Quadro 43 - Arquitetura formal implementada na ontologia

Eixo	Resultado consolidado	Evidências no TTL
Núcleo formal em OWL	As entidades e relações centrais do Registro do Diploma foram representadas como classes, propriedades de objeto e propriedades de dados, acompanhadas de hierarquias, restrições e características lógicas.	owl:Class, owl:ObjectProperty, owl:DatatypeProperty, rdfs:subClassOf, owl:Restriction, restrições de cardinalidade, propriedades inversas e propriedades funcionais.
Vocabulários controlados em SKOS	Foi definido o uso de esquemas de conceitos para representar categorias que exigiam padronização terminológica e distinção entre o recurso descrito e o valor classificatório associado.	skos:ConceptScheme, skos:Concept, skos:Collection, skos:prefLabel, skos:altLabel, skos:broader e skos:narrower.
Relações com valores controlados	As propriedades destinadas a valores classificatórios foram reunidas sob uma relação abstrata comum, indicando que seus objetos devem corresponder a conceitos SKOS, e não a literais livres.	odd:hasControlledValue e suas subpropriedades, entre as quais odd:awardsDegree, odd:awardsTitle, odd:hasModeOfDelivery, odd:hasStatus, odd:hasSignatureFormat e odd:hasSignatureRole.
Definições, comentários e notas de escopo	Os recursos da ontologia receberam rótulos, definições, comentários e notas de escopo para explicitar seu significado, sua função e seus limites de uso.	rdfs:label, rdfs:comment, skos:definition, skos:scopeNote, skos:editorialNote e skos:historyNote.
Justificativas de modelagem	Foram registradas justificativas para decisões de modelagem, especialmente nos casos em que a estrutura dos XSDs exigiu interpretação conceitual antes da formalização ontológica.	odd:modelingRationale, odd:sourceNote e comentários associados às classes, propriedades e valores controlados.
Formatos de assinatura representados	O vocabulário contemplou conceitos correspondentes a XAdES, CAdES e PAdES, permitindo distinguir o formato técnico da assinatura de seu papel administrativo ou institucional no fluxo do diploma.	vocab:XAdES, vocab:CAdES, vocab:PAdES, vocab:CategoriaFormatoAssinatura e vocab:CategoriaPapelAssinatura.
Rastreabilidade das fontes técnicas	Foram registradas referências aos schemas XSD utilizados como fontes estruturais e terminológicas da modelagem, preservando a ligação entre os elementos ontológicos e sua origem documental.	dcterms:source, odd:xsdPath, odd:xsdElement e odd:xsdType.
Metadados e identificação da ontologia	A ontologia recebeu metadados de identificação, descrição, autoria, namespace, versão, licença, fontes e conformidade, favorecendo sua documentação e futura disponibilização.	dcterms:title, dcterms:description, dcterms:creator, dcterms:license, vann:preferredNamespacePrefix, vann:preferredNamespaceUri, owl:versionInfo e owl:versionIRI.

Fonte: Elaborado pelo autor com base na formalização da OntoDiplomaDigital em TTL.

Os eixos apresentados no Quadro 43 demonstram que a arquitetura da OntoDiplomaDigital combinou diferentes padrões de representação para cumprir

funções complementares. O núcleo em OWL concentrou as classes, propriedades e restrições do subdomínio do Registro do Diploma; os vocabulários em SKOS organizaram valores sujeitos a controle terminológico; e os metadados descritivos identificaram a ontologia, sua versão, autoria, licença, fontes e conformidade.

As evidências registradas no TTL também indicam que a ontologia não resultou de uma conversão automática dos schemas XML. Por exemplo, estruturas dos XSDs associadas ao diplomado, ao curso, ao diploma e ao registro foram interpretadas como recursos do domínio e representadas por classes e propriedades em OWL. Já valores como grau conferido, modalidade de oferta e situação do diploma foram tratados como conceitos controlados em SKOS, e não como simples literais. Esses exemplos mostram que a formalização exigiu interpretação conceitual, documentação das decisões tomadas e preservação da rastreabilidade entre o artefato ontológico e suas fontes técnicas.

5.5.2 Correspondência entre os passos metodológicos e a formalização realizada

A formalização foi examinada segundo os sete passos previstos para a Etapa 5 da OntoForInfoScience. O Quadro 44 relaciona cada entrega metodológica aos componentes efetivamente formalizados na OntoDiplomaDigital, permitindo identificar a correspondência entre os passos da metodologia e os elementos presentes no arquivo canônico. Essa síntese apresenta o resultado da formalização sem retomar os procedimentos de edição ou o funcionamento das ferramentas utilizadas.

Quadro 44 - Correspondência entre os passos da Etapa 5 e a formalização da OntoDiplomaDigital

Passo da metodologia	Entrega esperada	Materialização na OntoDiplomaDigital
5.1 — Construir a taxonomia geral	Organizar classes e subclasses e pesquisar termos equivalentes para reúso.	Declarações owl:Class, relações rdfs:subClassOf, disjunções e alinhamentos com classes de PROV-O, FOAF, Schema.org, LOCN, vCard e DBpedia Ontology.
5.2 — Definir propriedades descritivas das classes	Registrar identificador, rótulo, definição, sinônimos, origem, exemplos e comentários.	Anotações rdfs:label, rdfs:comment, skos:prefLabel, skos:altLabel, skos:definition, skos:scopeNote, skos:example, dcterms:source, odd:sourceNote e odd:modelingRationale.
5.3 — Criar definições formais	Expressar condições lógicas sobre as classes.	Subclasses, disjunções, domínios, alcances, 28 restrições OWL nomeadas e uma união formal. As restrições estabelecem condições necessárias; a versão auditada não contém owl:equivalentClass.
5.4 — Definir propriedades de dados	Determinar atributos literais, tipos de dados e restrições.	58 declarações owl:DatatypeProperty, todas com alcance tipado, 55 com domínio explícito e 14 também declaradas funcionais.
5.5 — Criar instâncias	Representar ocorrências particulares coerentes com as classes.	31 conceitos SKOS e 13 indivíduos metodológicos no TTL; indivíduos do domínio mantidos em ABox operacional separada.
5.6 — Especificar relações ontológicas	Reutilizar ou criar relações entre entidades do domínio.	51 declarações owl:ObjectProperty, alinhamentos por rdfs:subPropertyOf, 18 pares de relações inversas, relações locais e propriedades especializadas para valores controlados.
5.7 — Definir propriedades das relações	Registrar propriedades descritivas e lógicas das relações.	Rótulos, comentários, domínio, alcance, inversa, superpropriedade, funcionalidade, fonte XSD e justificativa de modelagem.

Fonte: Elaborado pelo autor (2026), com base nos resultados da formalização e no arquivo canônico OntoDiplomaDigital_registro_diploma.ttl, disponibilizado no repositório público indicado no Apêndice B.

O quadro evidenciou que todos os passos metodológicos encontraram correspondência em componentes formalizados da OntoDiplomaDigital, embora com diferentes níveis de completude. A taxonomia, as propriedades de dados e as relações ontológicas foram implementadas diretamente; a documentação descritiva apresentou ampla cobertura; e a axiomatização incorporou restrições, disjunções e uma união formal. A criação de instâncias foi distribuída entre os conceitos SKOS, os indivíduos metodológicos necessários à documentação do modelo e a ABox operacional mantida separadamente da TBox.

A principal ressalva concentrou-se no passo 5.3, dedicado à criação de definições formais. As restrições foram associadas às classes por meio de rdfs:subClassOf, e não por owl:equivalentClass; por essa razão, expressaram condições necessárias para a pertença às classes, mas não definições necessárias e

suficientes. Esse resultado delimitou o nível de formalização alcançado pela ontologia nessa etapa.

5.5.3 Núcleo conceitual efetivamente implementado

O núcleo conceitual da OntoDiplomaDigital reuniu os recursos necessários à representação das principais entidades e relações do subdomínio do Registro do Diploma. O Quadro 45 apresentou um recorte funcional das classes implementadas, destinado a evidenciar os componentes centrais do modelo e as informações passíveis de associação e recuperação, sem constituir inventário integral do arquivo Turtle.

Das 39 classes locais declaradas, 28 foram utilizadas na representação do domínio ou em funções de apoio, oito organizaram categorias de valores controlados, duas registraram elementos metodológicos e uma desempenhou função lógica auxiliar. As relações taxonômicas concentraram-se nas especializações efetivamente formalizadas, como IesEmissora e IesRegistradora em relação a InstituicaoEnsinoSuperior, CursoNSF e CursoPorDecisaoJudicial em relação a Curso, e RegistroDiplomaNSF em relação a RegistroDiploma. Por essa razão, o recorte geral foi apresentado como núcleo conceitual, enquanto o termo taxonomia foi reservado às relações hierárquicas explicitamente declaradas.

Quadro 45 - Recorte do núcleo conceitual implementado na OntoDiplomaDigital

Grupo	Componentes representativos	Resultado no modelo
Classes centrais do domínio	DiplomaDigital, Diplomado, InstituicaoEnsinoSuperior, Curso, RegistroDiploma	Representação do documento, de seu titular, da formação, das instituições e do ato administrativo de registro.
Classes institucionais e especializações	IesEmissora, IesRegistradora, CursoNSF, CursoPorDecisaoJudicial, RegistroDiplomaNSF	Distinção dos papéis institucionais e das variantes estruturais previstas nos documentos de origem.
Classes documentais, espaciais e de apoio	DocumentacaoAcademica, HistoricoEscolarDigital, Endereco, LocalNaturalidade, ProcessoAdministrativo, RequisicaoDocumentacaoAcademica, Habilitacao, InformacoesTramitacaoEmec	Representação dos documentos, locais e estruturas auxiliares necessárias ao fluxo do diploma.
Agentes e recursos técnicos	ResponsavelRegistro, AssinanteRegistro, AssinanteEmissao, SignatarioTermoResponsabilidadeEmissora, AssinaturaDigital, CarimboTempo, CodigoValidacao, AgenteCertificador, PlataformaGoverno	Identificação dos agentes administrativos e dos recursos de autenticação, validação e acesso.
Categorias e conceitos SKOS	ValorControlado e sete classes Conceito*; 31 conceitos concretos	Padronização de grau, título, modalidade, situação, ambiente

Grupo	Componentes representativos	Resultado no modelo
		operacional, formato e papel de assinatura.
Indivíduos metodológicos e anotações	Seis objetivos específicos, sete questões de competência e propriedades de anotação	Rastreabilidade entre requisitos da pesquisa, decisões de modelagem e recursos da ontologia.
Propriedades de objeto	51 propriedades, entre elas hasHolder, issuedBy, registeredBy, hasProgram, hasRegistration, hasSignature, has Address e hasBirthPlace	Ligação entre recursos identificados por IRI.
Propriedades de objeto para valores controlados	Sete subpropriedades de hasControlledValue, entre elas awardsDegree, hasEnvironment, hasModeOfDelivery, hasStatus e hasSignatureFormat	Ligação de recursos do domínio a conceitos SKOS padronizados.
Propriedades de dados	58 propriedades, entre elas cpf, cnpj, courseName, registrationNumber, issueDate, nationality, postalCode e ibgeMunicipalityCode	Registro de textos, códigos, datas, valores numéricos e outros literais tipados.

Fonte: elaborado pelo autor a partir do TTL canônico e dos inventários da Etapa 5.

As propriedades associadas a valores controlados foram formalizadas como propriedades de objeto especializadas, e não como uma terceira categoria técnica paralela às propriedades de objeto e de dados. Essas propriedades conectaram recursos do domínio, como diploma, curso ou assinatura, a conceitos SKOS identificados por IRI. A superpropriedade abstrata hasControlledValue agrupou essas relações e não foi utilizada para preenchimento direto dos dados. Essa organização manteve separados o núcleo ontológico e a camada terminológica, substituindo valores livres por conceitos controlados e reutilizáveis.

5.5.4 Termos das classes principais

A documentação das classes locais permitiu registrar os principais conceitos formalizados na OntoDiplomaDigital, preservando suas denominações no arquivo Turtle e explicitando rótulos, definições, variantes terminológicas e notas de uso. Na versão auditada, as 39 classes locais apresentaram rdfs:comment e skos:definition. Esse resultado evidencia que a formalização não se limitou à criação de identificadores técnicos, mas incorporou uma camada descritiva destinada a tornar os conceitos compreensíveis, rastreáveis e reutilizáveis.

As classes centrais da ontologia representaram os elementos estruturantes do Registro do Diploma. DiplomaDigital, também registrado como diploma eletrônico, representou o documento digital oficial associado ao diplomado, ao curso, às

instituições, ao registro e aos mecanismos de autenticidade. Diplomado, com a variante egresso, representou a pessoa titular do diploma. Curso, correspondente a curso superior, representou a formação concluída e reuniu informações como nome, carga horária, modalidade, grau, título, habilitação e endereço. RegistroDiploma, também denominado ato de registro do diploma, representou o ato administrativo associado a datas, livro, folha, sequência, processo, requisição documental e agentes responsáveis.

No eixo institucional, InstituicaoEnsinoSuperior, também registrada como IES ou instituição de educação superior, representou a organização participante do ciclo institucional do diploma. As classes IesEmissora e IesRegistradora especializaram essa classe para distinguir os papéis de emissão e de registro, sem pressupor que tais funções fossem necessariamente exercidas por organizações distintas em todos os casos.

A ontologia também registrou especializações documentais derivadas das estruturas analisadas nos schemas. CursoNSF especializou Curso para representar o ramo NSF do leiaute, enquanto CursoPorDecisaoJudicial registrou a variante associada às alternativas previstas para situações decorrentes de decisão judicial. AlternativaEnderecoCursoJudicial representou a alternativa estrutural relativa ao endereço do curso no contexto de decisão judicial, e RegistroDiplomaNSF especializou RegistroDiploma para o ramo documental correspondente. Essas classes preservaram diferenças estruturais consideradas relevantes na fonte, sem transformar automaticamente todo elemento técnico dos XSDs em classe ontológica.

As classes documentais e espaciais complementaram a representação do ecossistema informacional do diploma. HistoricoEscolarDigital representou o documento relativo à trajetória acadêmica do diplomado, e DocumentacaoAcademica reuniu os documentos vinculados aos processos de emissão e registro. Endereco foi utilizado como estrutura postal reutilizável, enquanto LocalNaturalidade registrou o lugar de nascimento declarado para o diplomado, com informações de município, código IBGE e unidade federativa quando aplicáveis.

Os agentes do fluxo administrativo foram representados por classes distintas, de acordo com as funções desempenhadas nos processos de emissão, registro e assinatura. ResponsavelRegistro, AssinanteRegistro, AssinanteEmissao e SignatarioTermoResponsabilidadeEmissora distinguiram responsabilidades humanas associadas ao ciclo documental. RepresentanteInstitucionalEmissora foi utilizado

como generalização local para os papéis vinculados à representação da instituição emissora.

Os recursos técnicos de autenticação e validação foram formalizados por meio de classes específicas. AssinaturaDigital representou o artefato técnico associado à autenticidade e à integridade do diploma; CarimboTempo registrou o componente temporal vinculado à assinatura; e CodigoValidacao representou o recurso utilizado na verificação pública do documento. AgenteCertificador foi mantido como classe complementar associada à cadeia técnica de confiança.

Também foram formalizados recursos administrativos complementares. ProcessoAdministrativo representou o processo ao qual o registro foi vinculado; RequisicaoDocumentacaoAcademica, a requisição documental associada ao registro; Habilitacao, uma qualificação vinculada ao curso; e InformacoesTramitacaoEmec, os dados processuais utilizados quando o código e-MEC não estava disponível.

Por fim, a dimensão de interoperabilidade externa foi representada de forma controlada. PlataformaGoverno representou o agente de software governamental por meio do qual o diploma foi relacionado a serviços de consulta. LocalNaturalidade foi alinhada de forma ampla a `dbo:Place` e `schema:Place`, enquanto hasBirthPlace especializou `dbo:birthPlace`. Esses alinhamentos foram registrados sem declarar equivalência integral entre os recursos locais e os termos externos, preservando a especificidade semântica da OntoDiplomaDigital.

5.5.5 Propriedades e relações formalizadas

A versão auditada da OntoDiplomaDigital definiu 51 propriedades de objeto, 58 propriedades de dados e um conjunto de propriedades de objeto especializadas para valores controlados em SKOS, conforme sintetizado no Quadro 46 . As propriedades de objeto representaram conexões entre recursos, enquanto as propriedades de dados registraram valores literais. As relações com conceitos controlados foram formalizadas como propriedades de objeto, preservando a distinção entre os recursos do domínio e os valores terminológicos utilizados para qualificá-los.

Entre as propriedades de objeto, todas apresentaram alcance, 49 apresentaram domínio explícito, 18 pares declararam relações inversas e 14 foram definidas como funcionais. hasAddress e hasControlledValue permaneceram sem domínio explícito: a primeira por ser aplicável a diferentes entidades endereçáveis, e a segunda por atuar como superpropriedade de agrupamento das relações

associadas a valores controlados. No total, foram identificadas 79 declarações `rdfs:subPropertyOf`, das quais 71 corresponderam a alinhamentos externos: sete incidentes sobre propriedades de objeto e 64 sobre propriedades de dados. As demais declarações registraram especializações internas, especialmente no agrupamento das propriedades associadas a conceitos controlados.

As 58 propriedades de dados apresentaram alcance explícito, 55 registraram domínio e 14 foram definidas como funcionais. identifier, fullName e codeEmec permaneceram com domínio aberto para aplicação a diferentes tipos de recursos. Os alcances utilizaram tipos como `xsd:string`, `xsd:date`, `xsd:dateTime`, `xsd:decimal` e xsd:boolean. A propriedade legada environment foi mantida como depreciada, enquanto hasEnvironment foi formalizada para relacionar o diploma aos conceitos SKOS correspondentes aos ambientes de produção e homologação.

Quadro 46 - Grupos funcionais de propriedades formalizadas

Grupo funcional	Exemplos representativos	Resultado semântico
Identificação	identifier, cpf, cnpj, institutionMecCode, registrationNumber, validationCodeValue	Identificação de pessoas, instituições, registros e recursos de validação, com alinhamentos externos quando o valor identifica o sujeito descrito.
Pessoas e instituições	hasHolder, issuedBy, registeredBy, fullName, responsibleName	Ligação do diploma ao titular e aos papéis institucionais, além da descrição nominal dos agentes.
Curso e formação acadêmica	hasProgram, completedBy, hasQualification, courseName, workloadHours, conclusionDate	Representação do curso concluído, da habilitação, da carga horária e das datas acadêmicas.
Emissão e registro	hasRegistration, hasRegistrationResponsible, hasAdministrativeProcess, hasAcademicDocumentationRequest, registrationDate, issueDate	Representação do ato administrativo, de seus agentes, processos, requisições e marcos temporais.
Assinatura e validação	hasSignature, hasTimestamp, hasValidationCode, institutionallySignedBy, signatureDateTime, timestampDateTime	Organização dos recursos técnicos de assinatura, chancela institucional, tempo e consulta pública.
Documentação acadêmica	hasTranscript, hasAcademicDocumentation, hasIssuanceResponsibilitySigner	Ligação do diploma e do registro aos documentos comprobatórios e ao termo de responsabilidade.
Endereço e localização	hasAddress, hasBirthPlace, streetAddress, municipality, postalCode, nationality, ibgeMunicipalityCode	Separação entre endereço, local de naturalidade e nacionalidade, preservando a granularidade dos dados de origem.
Interoperabilidade externa	rdfs:subPropertyOf para DCTERMS, Schema.org, FOAF, PROV-O, LOCN, vCard, DCAT e DBpedia Ontology	Recuperação dos dados por vocabulários consolidados sem apagar as propriedades locais específicas.

Grupo funcional	Exemplos representativos	Resultado semântico
Valores controlados	awardsDegree, awardsTitle, hasModeOfDelivery, hasStatus, hasEnvironment, hasSignatureFormat, hasSignatureRole	Substituição de categorias textuais por conceitos SKOS identificados e reutilizáveis.

Fonte: Elaborado pelo autor (2026), com base nos inventários de propriedades de objeto e de dados da OntoDiplomaDigital e no arquivo canônico OntoDiplomaDigital_registro_diploma.ttl, disponibilizado no repositório público indicado no Apêndice B.

O uso de propriedades locais preservou relações específicas do domínio, como titularidade, emissão, registro e assinatura, nos casos em que propriedades externas mais genéricas não apresentaram o mesmo nível de precisão semântica. Os alinhamentos externos foram registrados quando a propriedade local pôde ser interpretada como especialização de uma superpropriedade existente. Nesse conjunto, courseName especializou dcterms:title e schema:name; hasRegistrationResponsible e hasRegistrationSigner especializaram prov:wasAssociatedWith; e as propriedades postais especializaram termos de LOCN e vCard. Não foram declaradas equivalências de propriedades com base apenas na semelhança entre rótulos.

5.5.6 Axiomas e restrições lógicas

Os axiomas, concebidos como declarações formais que representaram relações e condições válidas no modelo, explicitaram as regras utilizadas para qualificar classes e propriedades da OntoDiplomaDigital. Em termos simplificados, registraram afirmações como “uma classe é subclasse de outra”, “duas classes não podem representar simultaneamente o mesmo tipo de recurso” ou “determinada relação deve ocorrer uma quantidade específica de vezes”. A ontologia registrou 57 relações nomeadas de subclasse, das quais 15 organizaram a hierarquia local e 42 alinharam classes locais a classes externas. Foram também identificadas dez declarações owl:disjointWith, uma construção owl:unionOf, 28 propriedades funcionais, sendo 14 de objeto e 14 de dados, e 28 restrições OWL nomeadas. Essas restrições se distribuíram em 14 cardinalidades exatas, nove máximas, quatro mínimas e uma condição owl:hasValue.

As restrições, conforme ilustrado no Quadro 47 , impactaram diretamente as classes DiplomaDigital, Curso, CursoNSF, RegistroDiploma, AssinaturaDigital e Habilitacao. Adicionalmente, uma classe auxiliar foi criada para consolidar, por meio da propriedade owl:unionOf, as alternativas pertinentes à classe

CursoPorDecisaoJudicial. Os domínios e alcances complementaram essa axiomatização ao especificar os tipos esperados em cada extremidade das relações. Para o leitor não especializado, essas declarações podem ser interpretadas como regras formais do modelo, capazes de expressar condições como “esta relação faz parte de um diploma”, “o valor desta relação deve ser uma instituição emissora”, “esta informação deve ser única” ou “estas duas classes representam categorias distintas”.

Quadro 47 - Exemplos de restrições formalizadas e sua leitura em linguagem natural

Restrição no TTL	Leitura em linguagem natural	Função no modelo
DiplomaDigital com hasHolder e cardinalidade mínima qualificada 1 para Diplomado	Todo diploma deve estar ligado a pelo menos um diplomado.	Exigir a presença de titular compatível com a classe esperada.
DiplomaDigital com issuedBy e cardinalidade qualificada exatamente 1 para IesEmissora	Todo diploma deve estar ligado a exatamente uma instituição emissora.	Estabilizar o papel institucional principal da emissão.
RegistroDiploma com hasRegistrationResponsible e cardinalidade qualificada exatamente 1 para ResponsavelRegistro	Todo registro deve possuir exatamente um responsável pelo registro.	Qualificar o agente administrativo requerido para o ato de registro.
AssinaturaDigital com hasTimestamp e cardinalidade mínima qualificada 1 para CarimboTempo	Toda assinatura deve possuir pelo menos um carimbo de tempo.	Assegurar a associação da assinatura a uma evidência temporal.
Curso com awardsDegree e cardinalidade qualificada exatamente 1 para ConceitoGrauConferido	Todo curso deve conferir exatamente um grau representado por conceito controlado.	Restringir a categoria e a quantidade do valor de grau.

Fonte: elaborado pelo autor a partir do inventário de restrições OWL da OntoDiplomaDigital.

Esses axiomas produziram uma axiomatização parcial e verificável, mas não devem ser descritos como definições lógicas completas de todas as classes. Como não foram encontrados axiomas owl:equivalentClass, as restrições associadas por rdfs:subClassOf expressam condições necessárias para os membros das classes, sem afirmar que qualquer recurso que satisfaça essas condições pertença automaticamente a elas. A definição textual registrada em skos:definition e rdfs:comment, a axiomatização por subclasses e restrições e uma eventual definição necessária e suficiente constituem camadas distintas; somente as duas primeiras estavam presentes na versão analisada.

5.5.7 Instâncias, conceitos controlados e separação da ABox

A instanciação foi organizada em três conjuntos com finalidades distintas. O primeiro reuniu 31 conceitos SKOS utilizados como valores controlados para grau, título, modalidade de oferta, situação do diploma, ambiente operacional, formato de assinatura e papel da assinatura. Esses conceitos integraram a formalização como recursos identificáveis e foram associados a propriedades como awardsDegree, hasStatus e hasEnvironment. O segundo conjunto reuniu 13 indivíduos metodológicos, correspondentes a seis objetivos específicos e sete questões de competência, empregados para registrar a relação entre o escopo da pesquisa e os requisitos informacionais da ontologia.

O terceiro conjunto correspondeu aos indivíduos operacionais do domínio e foi mantido fora do TTL canônico, em arquivo próprio de ABox. Essa separação preservou a distinção entre a definição do vocabulário ontológico, concentrada na TBox, e as ocorrências particulares representadas na ABox. Na versão pública disponibilizada no repositório indicado no Apêndice B, a ABox reuniu 25.553 triplas com 3.861 indivíduos fictícios relativos a diplomas, diplomados, cursos, instituições, registros, endereços, assinaturas, carimbos de tempo, códigos de validação, processos administrativos, documentação acadêmica e demais recursos utilizados na prova de conceito.

Os conceitos SKOS foram representados como indivíduos de skos:Concept, vinculados aos respectivos esquemas e categorias locais, e não como subclasses adicionais da ontologia. Da mesma forma, os indivíduos metodológicos não representaram entidades do domínio do Diploma Digital, mas elementos de documentação relacionados aos objetivos específicos e às questões de competência. Essa organização distinguiu três níveis da implementação: as categorias ontológicas definidas na TBox, os valores terminológicos controlados em SKOS e as ocorrências fictícias representadas na ABox.

A formalização concluída nesta etapa estabeleceu a estrutura utilizada posteriormente na demonstração operacional. A TBox definiu classes, propriedades e restrições; os conceitos SKOS registraram os valores categóricos controlados; e a ABox permaneceu como camada independente destinada às instâncias do domínio. Os resultados referentes ao mapeamento RDF no Ontotext Refine, à carga no

GraphDB e às consultas SPARQL foram apresentados na etapa seguinte, dedicada à verificação operacional da ontologia..

5.6 Resultado da avaliação da ontologia – Etapa 6 da OntoForInfoScience

A Etapa 6 reuniu evidências sobre a adequação da OntoDiplomaDigital aos requisitos da pesquisa, sua correspondência semântica com o subdomínio do Registro do Diploma e a correção formal e operacional dos artefatos implementados. Esses resultados foram organizados em três dimensões complementares, avaliação, validação e verificação, sintetizadas no Quadro 48. Essa organização permitiu distinguir o atendimento aos objetivos da pesquisa, a adequação dos conceitos e das relações ao domínio representado e o processamento computacional da ontologia.

Quadro 48 - Síntese dos resultados da avaliação, validação e verificação da OntoDiplomaDigital

Dimensão	Foco da análise	Resultado observado
Avaliação	Atendimento aos objetivos e às questões de competência da pesquisa.	A ontologia representou o recorte definido para o Registro do Diploma e permitiu recuperar, no conjunto fictício, as informações previstas nas questões de competência.
Validação	Correspondência semântica entre as fontes consultadas e os conceitos, as relações e os valores controlados formalizados.	Os elementos centrais do subdomínio foram representados de forma coerente com as fontes normativas e técnicas e com as decisões conceituais registradas na pesquisa.
Verificação	Correção sintática, estrutural e operacional dos artefatos computacionais.	Os arquivos Turtle foram processados no ambiente utilizado, e a integração entre TBox e ABox permitiu a execução das consultas previstas na prova de conceito.

Fonte: Elaborado pelo autor (2026), com base em Mendonça (2015) e nos resultados da avaliação da OntoDiplomaDigital.

A síntese evidenciou que as três dimensões cumpriram funções diferentes no exame da ontologia. A avaliação considerou o atendimento aos objetivos e às necessidades informacionais da pesquisa; a validação examinou a correspondência entre os significados formalizados e o domínio representado; e a verificação incidiu sobre a correção e o processamento dos artefatos computacionais. Essa distinção evitou que questões conceituais, informacionais e técnicas fossem tratadas como um único resultado.

Sob a perspectiva da Ciência da Informação, os resultados evidenciaram como a delimitação do domínio, a definição dos conceitos, o controle terminológico e a explicitação das relações orientaram a construção e a análise da ontologia. A distinção entre instituição emissora e instituição registradora, por exemplo, preservou o

significado de dois papéis institucionais; a representação de grau e modalidade por conceitos controlados organizou os valores utilizados para caracterizar a formação acadêmica; e o registro das fontes permitiu relacionar os elementos formalizados à documentação que fundamentou sua representação. Esses componentes tornaram examináveis tanto o significado das informações quanto sua recuperação por meio de consultas.

5.6.1 Resultados da validação semântica interna da OntoDiplomaDigital

A validação semântica interna, conforme apresentado no Quadro 49, examinou a congruência entre a OntoDiplomaDigital e o subdomínio representado nas fontes normativas, técnicas e conceituais utilizadas na pesquisa. A análise concentrou-se no significado atribuído às classes, propriedades, relações e restrições formalizadas, confrontando-as com as fontes que fundamentaram a modelagem e com as questões de competência estabelecidas na especificação. Esse procedimento teve caráter interno e documental e não correspondeu a uma validação externa da ontologia por especialistas do domínio ou por instituições usuárias do Diploma Digital.

Quadro 49 – Resultados da validação semântica interna da OntoDiplomaDigital

Critério	Evidência esperada	Resultado alcançado
Aderência semântica	As classes e as relações representaram de forma coerente os conceitos centrais do domínio.	O modelo preservou a leitura do diploma digital, do registro, da assinatura, da documentação e dos agentes envolvidos.
Coerência entre fontes e modelagem	As decisões ontológicas puderam ser justificadas pelas fontes e pela análise conceitual.	O reúso de vocabulários externos e as escolhas locais ficaram documentados.
Separação entre níveis semânticos	Classes, relações, propriedades de dados e conceitos controlados não foram confundidos.	SKOS, RDFS e OWL foram usados em camadas distintas e complementares.

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience de Mendonça (2015).

Os resultados indicaram correspondência entre os conceitos e as relações formalizados e o conjunto de fontes e requisitos adotados na pesquisa. A prova de conceito complementou esse exame ao permitir observar a utilização da representação em situações de recuperação previstas nas questões de competência. A consulta que distinguiu as instituições emissora e registradora, por exemplo, mostrou que os dois papéis institucionais puderam ser recuperados separadamente.

As consultas sobre valores controlados e proveniência também permitiram examinar como essas categorias e relações foram utilizadas no conjunto sintético. Essas respostas apoiaram o exame da correspondência entre a interpretação adotada na modelagem e as informações recuperadas, sem substituir sua confrontação com as fontes do domínio.

A contribuição da prova de conceito à validação semântica interna deve ser compreendida nesses limites. A execução das consultas demonstrou a utilização das relações modeladas nas condições testadas, mas não comprovou sua adequação às práticas ou às interpretações de instituições reais. A validação permaneceu interna e documental, sem participação de especialistas externos ou instituições usuárias. Os resultados detalhados das consultas são apresentados na Seção 5.6.3, enquanto os aspectos sintáticos, estruturais e operacionais da implementação são tratados na Seção 5.6.2 a seguir.

5.6.2 Resultados da verificação formal e operacional da ontologia

A verificação representada no Quadro 50, analisou a correção técnica e lógica da OntoDiplomaDigital após sua formalização. Foram examinados a sintaxe dos arquivos Turtle, os prefixos e namespaces, as declarações de classes e propriedades, os domínios e alcances, as restrições, as relações inversas e a correspondência entre a TBox e a ABox utilizadas na prova de conceito.

Quadro 50 – Resultados da verificação formal da Ontologia

Critério	O que foi conferido	Resultado alcançado
Sintaxe do Turtle	A escrita RDF/Turtle foi processada sem erro de parsing.	Arquivo computacionalmente legível.
Consistência das declarações	Prefixos, classes, propriedades, indivíduos e anotações foram declarados de forma compatível.	Estrutura formal estável.
Uso correto de RDFS e OWL	Domínios, alcances, restrições e hierarquias foram expressos em conformidade com o vocabulário adotado.	Declarações processadas e estruturalmente compatíveis no ambiente utilizado
Processabilidade por ferramentas	O TTL pôde ser lido por ferramentas de análise, consulta e validação.	Modelo operacionalmente utilizável.
Coerência entre classes, propriedades e instâncias	As relações entre TBox, ABox e anotações não foram identificadas incompatibilidades nos	Estrutura processada sem erros nos procedimentos de verificação executados

Critério	O que foi conferido	Resultado alcançado
	controles estruturais e operacionais realizados.	

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience de Mendonça (2015).

Os resultados da verificação registraram o comportamento formal da ontologia nos testes executados, distinguindo essa análise da validação semântica apresentada na subseção anterior. Enquanto a validação incidiu sobre a adequação do significado representado, a verificação concentrou-se na consistência da implementação e no processamento dos artefatos pelas ferramentas empregadas na prova de conceito.

A verificação realizada não incluiu o uso de um *reasoner* OWL para teste automatizado de consistência lógica; portanto, as evidências apresentadas referem-se à correção sintática, estrutural e operacional examinada pelos procedimentos descritos nesta pesquisa.

A demonstração operacional evidenciou a transformação das 252 unidades da fonte tabular em recursos e relações RDF compatíveis com a TBox da OntoDiplomaDigital. O mapeamento manteve a correspondência entre os registros de entrada e os 252 diplomas e diplomados representados, além de produzir recursos para cursos, instituições, registros, assinaturas, códigos de validação, processos administrativos, endereços e outros componentes do subdomínio. A representação distinguiu entidades identificáveis, relações entre recursos e valores literais, conforme a estrutura semântica definida na ontologia.

A ABox produzida reuniu 25.553 triplas com 3.961 indivíduos operacionais. Sua integração com as 3.009 triplas da TBox resultou em um grafo com 28.562 declarações explícitas nos arquivos auditados. Esse conjunto permitiu processar conjuntamente o vocabulário formal da OntoDiplomaDigital e as instâncias fictícias da prova de conceito. Os alinhamentos registrados na TBox também estabeleceram correspondências com vocabulários externos, preservando as propriedades locais quando estas apresentaram maior especificidade para o domínio do Registro do Diploma.

As consultas SPARQL recuperaram os recursos e as relações previstos nas questões de competência e nas demonstrações complementares. Foram obtidas informações sobre titulares, cursos, instituições emissoras e registradoras, registros, assinaturas, valores controlados, mecanismos de validação e proveniência. As duas consultas federadas acrescentaram evidências de recuperação distribuída ao

combinarem dados locais fictícios com informações externas da DBpedia e do Wikidata por meio de identificadores compartilhados.

A Figura 12 - Encadeamento da relação do domínio Registro do Diploma sintetizou uma dessas demonstrações. A consulta utilizou um CPF fictício como parâmetro de entrada para localizar o respectivo diplomado e recuperar o grafo associado ao seu diploma. A visualização reuniu o curso, o grau e o título conferidos, a modalidade de oferta, as instituições emissora e registradora, o registro, o processo administrativo, os agentes envolvidos, a assinatura digital, o carimbo de tempo, o formato da assinatura, o código de validação, a situação do documento e a plataforma de acesso.

5.6.3 Resultados da prova de conceito e da verificação operacional

A prova de conceito examinou se a OntoDiplomaDigital permitia representar e recuperar as informações previstas em sua especificação. O procedimento foi realizado em ambiente controlado, com dados integralmente sintéticos. A TBox forneceu os conceitos e as relações do modelo, enquanto a ABox reuniu os recursos e os dados utilizados nos testes. Sua integração no GraphDB permitiu executar consultas SPARQL e examinar as respostas obtidas.

A fonte tabular compreendeu 252 registros sintéticos. O mapeamento desses registros no Ontotext Refine produziu recursos RDF relacionados a diplomas, diplomados, cursos, instituições, registros, assinaturas e demais componentes do subdomínio. Uma linha da tabela pôde originar diferentes recursos e relações; portanto, a quantidade de registros de entrada não corresponde diretamente à quantidade de indivíduos da ABox.

O conjunto completo de dados sintéticos utilizado na prova de conceito está disponível no repositório público da pesquisa, tanto em arquivos tabulares quanto na ABox em formato Turtle. O repositório também reúne a TBox, os códigos completos das consultas SPARQL e os relatórios de verificação. Os quadros apresentados nesta seção e no Apêndice B contêm exemplos ou recortes dos resultados, enquanto os arquivos disponibilizados permitem examinar a base de teste e executar as consultas. Os endereços de acesso e as orientações de reprodução encontram-se no Apêndice B.

A verificação foi organizada em onze consultas: sete vinculadas às questões de competência e quatro demonstrações complementares. As questões de

competência expressaram as informações que a ontologia deveria permitir recuperar e forneceram uma referência para examinar seu funcionamento. As demonstrações complementares exploraram a recuperação integrada, a visualização em grafo e a combinação de informações locais com fontes externas. Os arquivos das consultas e os relatórios estão disponíveis no repositório indicado no Apêndice B.

O Quadro 51 relaciona cada questão de competência aos elementos consultados, à resposta esperada e ao resultado registrado. As conclusões referem-se exclusivamente às consultas e aos dados sintéticos empregados na prova de conceito.

Quadro 51 - Questões de competência e evidências da verificação operacional da OntoDiplomaDigital

QC	Consulta SPARQL correspondente	Entidades mobilizadas	Elementos da ontologia mobilizados	Resultado observado	Conclusão limitada ao teste
QC1	Consulta 01 — identificação das instituições e de seus papéis	Diploma, instituições emissora e registradora e relações de emissão e registro	Classe: DiplomaDigital; propriedades: identifier, issuedBy, registeredBy e fullName. Identificam o diploma, distinguem os papéis institucionais e recuperam os nomes das instituições.	Duas linhas, distinguindo emissão e registro	Os papéis institucionais puderam ser recuperados separadamente
QC2	Consulta 02 — recuperação dos elementos documentais	Diploma, registro, assinatura, carimbo de tempo e código de validação	Classe: DiplomaDigital; propriedades: identifier, hasRegistration, registrationDate, hasSignature, signatureDateTime, hasTimestamp, timestampDateTime, hasSignatureFormat, hasSignatureRole, hasValidationCode e validationCodeValue; propriedade skos:prefLabel. Relacionam o diploma aos elementos documentais e recuperam seus valores e rótulos.	Uma linha com os elementos previstos	Os metadados puderam ser recuperados, sem validação criptográfica
QC3	Consulta 03 — seleção por instituição, período e modalidade	Diploma, instituição emissora, curso, data de emissão e modalidade	Classe: DiplomaDigital; propriedades: issuedBy, issueDate, hasProgram, fullName, courseName e hasModeOfDelivery; propriedade skos:prefLabel. Relacionam o diploma à instituição, à data de emissão, ao curso e à modalidade utilizados na seleção.	Um diploma associado ao curso fictício de Bacharelado em Gestão do Conhecimento Aplicado, na modalidade presencial	Os critérios permitiram selecionar um diploma no conjunto sintético
QC4	Consulta 04 — recuperação do titular e da formação acadêmica	Diploma, diplomado, curso, grau e título	Classe: DiplomaDigital; propriedades: identifier, hasHolder, hasProgram, fullName, cpf, courseName, awardsDegree e awardsTitle; propriedade skos: prefLabel. As propriedades completedProgram e conclusionDate aparecem em trechos opcionais. A ligação entre titular e curso é estabelecida pelo diploma.	Uma linha com titular, curso, data de conclusão, grau Bacharelado e título Bacharel	As informações acadêmicas puderam ser recuperadas de forma relacionada
QC5	Consulta 05 — identificação e contagem de uso dos valores controlados	Conceitos SKOS e relações que qualificam cursos,	Classe skos:Concept; propriedades awardsDegree, awardsTitle, hasModeOfDelivery, hasStatus, hasSignatureFormat e hasSignatureRole; propriedade skos:prefLabel. Associam os recursos	Vinte linhas relativas a grau, título, modalidade, situação,	Os valores controlados puderam ser identificados e

QC	Consulta SPARQL correspondente	Entidades mobilizadas	Elementos da ontologia mobilizados	Resultado observado	Conclusão limitada ao teste
		diplomas e assinaturas	aos conceitos controlados e recuperam seus rótulos para apresentação e contagem.	formato e papel da assinatura	seus usos contabilizados
QC6	Consulta 06 — recuperação da proveniência do registro	Diploma, instituição registradora, registro, processo administrativo, responsável e assinantes	Classe: DiplomaDigital; propriedades registeredBy, hasRegistration, fullName, registrationDate, hasAdministrativeProcess, identifier, hasRegistrationResponsible, responsibleName e responsibleCpf. As propriedades opcionais hasRegistrationSigner, registrationSignerCpf e registrationSignerRole acrescentam informações sobre os assinantes.	303 linhas com combinações dessas informações	A proveniência administrativa pôde ser recuperada; as linhas não representam diplomas distintos
QC7	Consulta 07 — recuperação da plataforma associada ao diploma por CPF fictício	Diplomado, diploma, plataforma e código de validação	Classes: Diplomado, DiplomaDigital e PlataformaGoverno; propriedades cpf, hasHolder, isAccessibleVia e fullName. As propriedades opcionais hasValidationCode e validationCodeValue acrescentam o código de validação, quando disponível.	Uma linha com diploma, plataforma fictícia e código d	A relação de acesso pôde ser recuperada no grafo, sem integração com o Gov.br

Fonte: Elaborado pelo autor (2026), com base nas consultas SPARQL e nos resultados da prova de conceito.

Para exemplificar como uma questão de competência foi traduzida em consulta, o Quadro 52 apresenta uma versão abreviada da consulta QC1, que busca identificar as instituições emissora e registradora de um diploma digital. O comando `VALUES` informa o identificador sintético “urn:synthetic:diploma:001”, utilizado para localizar o diploma no grafo. As propriedades `issuedBy` e `registeredBy` permitem encontrar as instituições responsáveis, respectivamente, pela emissão e pelo registro. O comando `UNION` reúne os resultados dessas duas buscas, enquanto `BIND` indica o papel de cada instituição na resposta. Por fim, `SELECT` define as informações a apresentar: o papel institucional e o nome da instituição.

Quadro 52 - Exemplo abreviado de consulta SPARQL para a QC1

```
PREFIX odd: <https://w3id.org/ufes/ontodiplomadigital#>

SELECT ?papel ?nomeIes
WHERE {
  VALUES ?identificador { "urn:synthetic:diploma:001" }

  ?diploma a odd:DiplomaDigital ;
           odd:identifier ?identificador .

  {
    ?diploma odd:issuedBy ?ies .
    BIND("emissão" AS ?papel)
  }
  UNION
  {
    ?diploma odd:registeredBy ?ies .
    BIND("registro" AS ?papel)
  }

  ?ies odd:fullName ?nomeIes .
}
ORDER BY ?papel
```

Fonte: Elaborado pelo autor (2026), a partir da consulta QC1 da prova de conceito.

A consulta respondeu à questão “Quais são a instituição emissora e a instituição registradora de determinado diploma digital?”. Para o diploma sintético identificado por “`urn:synthetic:diploma:001`”, foram recuperadas duas linhas, uma para cada papel institucional. A resposta do Quadro 53 apresenta o papel desempenhado e o nome da respectiva instituição.

Quadro 53 - Dados retornados pela consulta QC1 para o diploma sintético selecionado

Papel institucional (`?papel`)	Nome da instituição (`?nomeIes`)
emissão	Universidade do Vale do Conhecimento
registro	Universidade Nacional da Curadoria

Fonte: Elaborado pelo autor, a partir dos resultados da consulta QC1 sobre a base sintética da prova de conceito.

O exemplo relaciona a questão de competência às relações modeladas na ontologia e aos dados apresentados na resposta. As instituições são fictícias, e o resultado se limita à recuperação dessas relações no conjunto de teste, sem comprovar a emissão ou o registro de um diploma real.

As demais consultas examinaram diferentes partes dessa representação. A QC2 recuperou os elementos documentais vinculados ao diploma, sem verificar criptograficamente sua autenticidade. As QCs 3 e 4 permitiram selecionar diplomas por critérios combinados e relacionar o titular à formação acadêmica. A QC5 examinou o uso dos conceitos controlados, enquanto a QC6 recuperou informações sobre a proveniência administrativa do registro. Nesta última, a presença de diferentes assinantes produziu combinações de respostas, razão pela qual as 303 linhas não devem ser interpretadas como 303 diplomas. A QC7 recuperou uma plataforma fictícia associada ao diploma, sem acessar serviços governamentais.

As consultas complementares 08 e 09 ampliaram a recuperação por CPF fictício. A consulta 08 reuniu, em uma visão tabular, informações acadêmicas, institucionais, administrativas e de assinatura. A consulta 09 apresentou as conexões entre esses recursos em grafo, conforme ilustrado na Figura 17.

Figura 17 — Grafo de informações recuperadas para um diplomado na prova de conceito



Fonte: Elaborado pelo autor (2026), a partir da consulta SPARQL executada no GraphDB com dados sintéticos da OntoDiplomaDigital.

A visualização representa o diploma como um recurso relacionado ao titular, ao curso, às instituições, ao registro e aos demais elementos. Para a Ciência da Informação, evidencia a recuperação contextualizada de informações distribuídas e contribui para a verificação operacional das relações modeladas. Contudo, não comprova, isoladamente, a adequação semântica ao domínio, examinada na Seção 5.6.1.

As consultas 10 e 11 integraram dados locais a fontes externas, como DBpedia e Wikidata. Essas consultas são federadas porque combinam, em uma única operação, dados locais e de ao menos uma base externa. Na consulta 11, dados sintéticos da OntoDiplomaDigital foram articulados a informações do Wikidata, acessadas pelo QLever[13]. O código IBGE do município, comum às bases, permitiu contabilizar localmente os diplomados fictícios e recuperar externamente o nome do município e o país. O Quadro 54 apresenta uma versão abreviada dessa integração.

Quadro 54 - Consulta federada para relacionar a naturalidade dos diplomados fictícios ao município no Wikidata

```
PREFIX odd:  <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX wdt:  <http://www.wikidata.org/prop/direct/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?codigoIbge ?municipioLabel ?paisLabel ?qtdDiplomados
WHERE {
  {
    SELECT ?codigoIbge
           (COUNT(DISTINCT ?diplomado) AS ?qtdDiplomados)
    WHERE {
      ?diplomado a odd:Diplomado ;
                 odd:hasBirthPlace ?localNaturalidade .
      ?localNaturalidade odd:ibgeMunicipalityCode ?codigoIbge .
      FILTER(STR(?codigoIbge) = "3550308")
    }
    GROUP BY ?codigoIbge
  }

  SERVICE <https://qlever.dev/api/wikidata> {
    ?municipio wdt:P1585 ?codigoIbge ;
               rdfs:label ?municipioLabel ;
               wdt:P17 ?pais .
    ?pais rdfs:label ?paisLabel .

    FILTER(LANG(?municipioLabel) = "pt")
    FILTER(LANG(?paisLabel) = "pt")
  }}
```

Fonte: Elaborado pelo autor, a partir da consulta federada da prova de conceito.

[13] QLever é um sistema de processamento de consultas SPARQL que permite recuperar informações de bases RDF, como o Wikidata. Nesta pesquisa, seu serviço público foi utilizado para consultar os dados geográficos externos e combiná-los com os dados sintéticos da OntoDiplomaDigital. Disponível em: https://qlever.dev/wikidata. Acesso em: 3 out. 2026.

A primeira parte da consulta busca, na base local, os diplomados associados a um local de naturalidade com o código IBGE “3550308”. O comando `FILTER` restringe a busca a esse código, enquanto COUNT(DISTINCT ?diplomado) conta cada diplomado uma única vez.

O bloco `SERVICE` encaminha a busca externa ao QLever, que disponibiliza os dados do Wikidata. A ligação entre as fontes ocorre pela variável `?codigoIbge`: seu valor é utilizado em `wdt:P1585`, propriedade que representa o código IBGE, para localizar o município correspondente. A partir desse recurso, `rdfs:label` recupera seu nome e `wdt:P17`, que representa a relação com o país, identifica o Brasil. Os filtros de idioma selecionam os nomes em português, e `SELECT` define as informações apresentadas na resposta.

A execução registrada retornou uma linha, reunindo a contagem realizada na base local e as informações geográficas recuperadas da fonte externa. O Quadro 55 apresenta esses dados. O Quadro 55 apresenta os dados retornados pela consulta.

Quadro 55 — Dados retornados pela consulta federada ao Wikidata

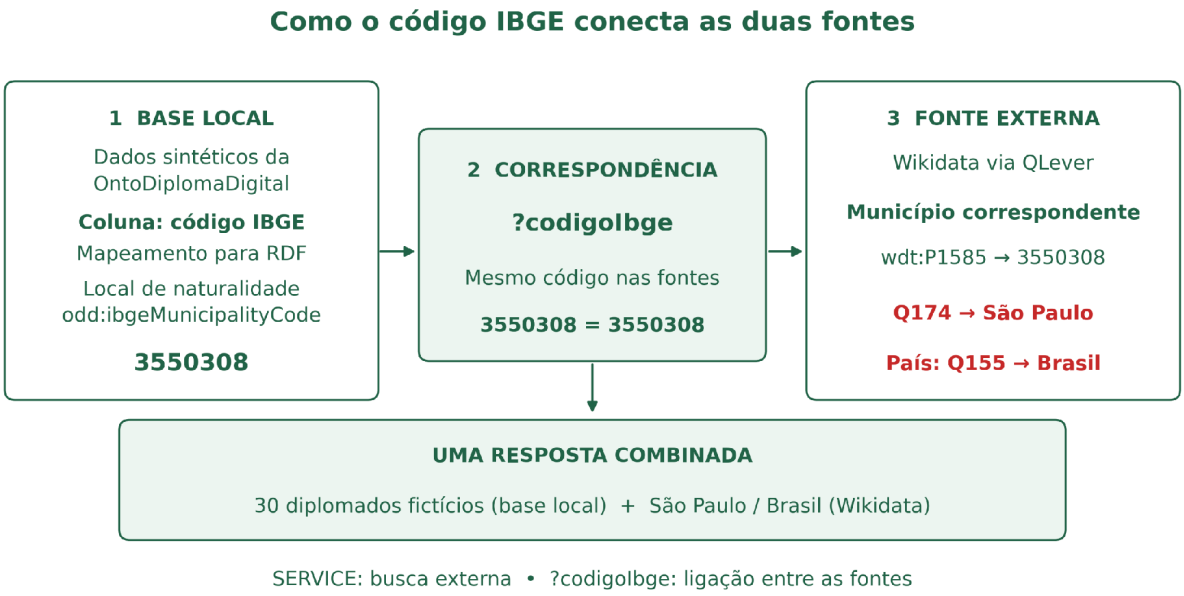
Código IBGE	Município	País	Quantidade de diplomados fictícios
3550308	São Paulo	Brasil	30

Fonte: Elaborado pelo autor, a partir do resultado registrado da consulta federada executada no GraphDB em 2 de setembro de 2026.

A resposta reuniu informações de duas fontes: a contagem de trinta diplomados fictícios foi calculada na ABox local, enquanto a identificação e os nomes do município e do país foram recuperados do Wikidata. Essa combinação utilizou propriedades de vocabulários diferentes com uma correspondência explícita de significado: ambas representam o código IBGE do município. O identificador compartilhado permitiu conectar o local de naturalidade representado na ontologia à descrição geográfica externa. O resultado demonstra uma recuperação federada delimitada, sem incorporação permanente dos dados externos à base local. A contagem refere-se exclusivamente aos indivíduos sintéticos da prova de conceito; a consulta não acessou o sistema do IBGE nem verificou registros acadêmicos reais.

A Figura 18 sintetiza graficamente a ligação entre a base local e a fonte externa. O fluxo começa na coluna com o código IBGE dos dados sintéticos, passa pelo mapeamento desse valor na OntoDiplomaDigital e mostra como a variável `?codigoIbge` estabelece a correspondência com o município no Wikidata.

Figura 18 - Ligação entre a base sintética da OntoDiplomaDigital e o Wikidata por código IBGE



Fonte: Elaborado pelo autor, com base no mapeamento RDF dos dados sintéticos, na consulta federada e no resultado da prova de conceito (2026).

O código “3550308” conecta o local de naturalidade representado na base sintética ao recurso `Q174` do Wikidata, que identifica “São Paulo” e o `Q155` que identifica o País. Essa correspondência permite reunir, na mesma resposta, a contagem dos diplomados fictícios calculada localmente e as informações sobre o município e seu país recuperadas da fonte externa.

Os resultados evidenciaram que, no recorte testado, a OntoDiplomaDigital permitiu representar, relacionar e recuperar as informações previstas nas questões de competência. As demonstrações complementares também mostraram possibilidades de recuperação integrada e de combinação com fontes externas. Essas evidências permanecem limitadas aos dados sintéticos, às consultas e às configurações utilizadas. Não constituem validação institucional, teste com dados acadêmicos reais, validação criptográfica de assinaturas, comprovação de integração com plataformas governamentais ou avaliação de desempenho em escala.

5.7 Resultado da documentação da ontologia – Etapa 7 da OntoForInfoScience

A Etapa 7 resultou na consolidação da documentação acadêmica e técnica da OntoDiplomaDigital. A dissertação registrou a construção da ontologia, suas fontes, decisões conceituais, resultados de formalização e evidências de avaliação. Como

complemento, os artefatos computacionais da prova de conceito foram organizados em repositório público no GitHub, indicado no Apêndice B.

A documentação reuniu elementos que permitem relacionar as fontes normativas e técnicas, as decisões de modelagem, a representação formal da ontologia e as evidências da demonstração operacional. Esse conjunto favoreceu a rastreabilidade das escolhas realizadas, a transparência dos resultados e a consulta dos artefatos computacionais associados à pesquisa.

5.7.1 Justificativa das escolhas semânticas

A documentação registrou as justificativas para a representação dos elementos como classes, propriedades, conceitos SKOS ou anotações, bem como as decisões de reúso de termos provenientes de vocabulários externos. A adoção de correspondências considerou significado, categoria ontológica, domínio, alcance e restrições, sem estabelecer equivalência apenas pela semelhança entre rótulos. Quando houve compatibilidade semântica, foram utilizados rdfs:subClassOf, rdfs:subPropertyOf ou propriedades de mapeamento SKOS; nos casos de especificidade normativa do Diploma Digital Brasileiro, os elementos foram mantidos no namespace local da OntoDiplomaDigital.

Entre as decisões documentadas, destacou-se a distinção entre ResponsavelRegistro, AssinanteRegistro e AssinaturaDigital. Os dois primeiros representaram agentes ou papéis administrativos vinculados ao processo de registro, enquanto AssinaturaDigital representou o recurso técnico associado à autenticidade e à integridade do documento. O registro dessa decisão preservou a diferença semântica entre o agente que realiza a assinatura e o artefato criptográfico associado ao documento.

5.7.2 Consistência terminológica

Para nomear os elementos da ontologia, adotou-se a convenção UpperCamelCase para as classes e lowerCamelCase para as propriedades[14]. As classes foram grafadas em UpperCamelCase, as propriedades de objeto e de dados em lowerCamelCase, e os rótulos destinados à leitura humana foram registrados em

[14] Essas convenções unem palavras sem espaços. Em UpperCamelCase, cada palavra começa com letra maiúscula, como em `DiplomaDigital`; em lowerCamelCase, a primeira começa com minúscula e as seguintes com maiúscula, como em `hasHolder`

linguagem natural com marcação de idioma. Nos elementos utilizados para rastreabilidade dos XSDs, foram preservadas as denominações técnicas de origem, enquanto a camada semântica empregou nomes correspondentes ao significado atribuído aos recursos no domínio.

As propriedades de objeto receberam rótulos verbais, como “emitido por”, “registrado por”, “tem responsável pelo registro” e “assinado administrativamente por”. Os sinônimos incorporados ao vocabulário foram registrados por skos:altLabel, evitando a promoção automática de variantes documentais não consolidadas. Essas convenções foram preservadas na documentação e no arquivo Turtle da ontologia.

5.7.3 Memória e evolução do projeto

A memória do projeto reuniu registros de versões, alterações de modelagem, notas técnicas e documentos históricos utilizados durante a evolução da OntoDiplomaDigital. Foi mantida uma fonte canônica do arquivo Turtle, enquanto versões numeradas foram reservadas para marcos acadêmicos ou alterações relevantes. Os XSDs permaneceram como referência estrutural, a especificação semântica registrou as decisões conceituais, o TTL formalizou essas decisões e os arquivos PlantUML, tabelas e relatórios foram mantidos como artefatos derivados.

Também foram preservados registros das alterações que afetaram a relação entre TBox e ABox. Mudanças em classes, propriedades ou restrições foram acompanhadas da revisão dos dados sintéticos, da regeneração dos artefatos correspondentes e da execução dos testes de verificação. Fragilidades, lacunas e ambiguidades identificadas nas fontes oficiais também foram documentadas juntamente com as decisões adotadas para sua representação.

5.7.4 Comunicação acadêmica da ontologia

A comunicação acadêmica da OntoDiplomaDigital foi realizada pela articulação entre a dissertação, os apêndices e o repositório público indicado no Apêndice B. A dissertação apresentou o percurso de construção da ontologia, os quadros de síntese, os modelos conceituais e a interpretação dos resultados; o repositório disponibilizou os artefatos computacionais que sustentaram a demonstração operacional.

Essa organização permitiu apresentar a ontologia sob duas perspectivas complementares: como representação formal processável por máquinas e como artefato informacional explicado para pesquisadores, estudantes e profissionais do

domínio. Os modelos gráficos e os trechos exemplificativos incluídos na dissertação apoiaram a leitura dos elementos formalizados, sem substituir o acesso ao arquivo Turtle e aos demais materiais técnicos disponibilizados no repositório.

5.7.5 Artefatos documentais consolidados

Os artefatos documentais consolidados foram distribuídos entre a dissertação e o repositório público da prova de conceito, indicado no Apêndice B. Enquanto a dissertação registrou a explicação acadêmica do percurso e dos resultados, o repositório reuniu os arquivos computacionais, os dados fictícios, as consultas e os registros de verificação associados à demonstração operacional. O Quadro 56 sintetiza os principais artefatos documentais consolidados e suas respectivas funções.

Quadro 56 - Artefatos documentais consolidados

Artefato	Função documental
Dissertação	Registrou a síntese acadêmica do processo de construção da OntoDiplomaDigital, incluindo metodologia, resultados, discussão, limites da prova de conceito e interpretação científica dos artefatos produzidos.
Arquivo TTL da TBox	Registrou a representação formal da OntoDiplomaDigital, incluindo classes, propriedades, restrições, vocabulários controlados, metadados, alinhamentos e anotações de documentação.
ABox pública com instâncias fictícias	Materializou exemplos de dados fictícios do subdomínio do Registro do Diploma, permitindo demonstrar a aplicação da ontologia sem exposição de dados pessoais, acadêmicos ou institucionais reais.
Arquivos tabulares	Preservaram os dados de entrada e versões auxiliares utilizadas para inspeção da prova de conceito, favorecendo a compreensão da passagem da estrutura tabular para a representação RDF.
Glossários, dicionários e quadros de síntese	Reuniram conceitos, propriedades, relações, definições, variantes terminológicas, fontes, notas de uso e decisões de modelagem apresentadas ao longo da dissertação.
Modelos conceituais e diagramas	Representaram visualmente classes, relações, atributos, ramos estruturais e encadeamentos do subdomínio do Registro do Diploma, apoiando a leitura conceitual da ontologia.
Consultas SPARQL	Preservaram as consultas utilizadas para verificar as questões de competência, recuperar relações do grafo e demonstrar a leitura operacional dos dados representados.
Consultas federadas demonstrativas	Demonstraram a possibilidade técnica de combinar dados locais do repositório diplomaDB com fontes externas de dados ligados, sem configurar integração real com sistemas oficiais.
Relatórios de sanitização e validação	Registraram evidências de verificação dos arquivos públicos, incluindo sanitização dos dados, validação da TBox e da ABox e execução das consultas da prova de conceito.
Repositório GitHub v1.0	Disponibilizou publicamente os artefatos computacionais da prova de conceito, organizados em estrutura navegável e citável como complemento técnico da dissertação.

Fonte: Elaborado pelo autor com base na metodologia OntoForInfoScience de Mendonça (2015) e nos artefatos documentais da OntoDiplomaDigital.

Em conjunto, os artefatos documentais preservaram a relação entre a representação formal da OntoDiplomaDigital, as fontes que fundamentaram sua construção e as evidências produzidas na prova de conceito. A organização do material no repositório ampliou a possibilidade de consulta e verificação dos resultados, nos limites dos dados fictícios e das orientações de reprodução disponibilizadas.

5.8 Resultado da disponibilização da ontologia – Etapa 8 da OntoForInfoScience

A Etapa 8 tornou públicos os artefatos da OntoDiplomaDigital em um repositório no GitHub, na versão estável v1.0, conforme o Apêndice B. Essa disponibilização reuniu os principais produtos computacionais da prova de conceito, permitindo que o leitor acesse, inspecione e reproduza parcialmente os resultados da dissertação.

O repositório contém o arquivo Turtle da ontologia, a ABox com instâncias fictícias, os dados tabulares utilizados na demonstração, as consultas SPARQL, as consultas federadas demonstrativas, os registros de validação e a documentação de apoio para carga e consulta no GraphDB. Desse modo, a disponibilização não se limitou à apresentação textual da ontologia na dissertação, mas incluiu os artefatos técnicos necessários para examinar sua estrutura formal e verificar sua aplicação no conjunto sintético utilizado.

Como resultado, a OntoDiplomaDigital passou a contar com uma forma pública de acesso acadêmico, vinculada a uma versão identificável do repositório. Essa publicação favorece a transparência da pesquisa, a rastreabilidade dos artefatos produzidos e a possibilidade de reutilização ou extensão futura da ontologia. Contudo, a disponibilização realizada não caracteriza a implantação da ontologia como serviço institucional em rede, endpoint SPARQL permanente ou publicação plena em ambiente de dados ligados.

5.9 Discussão integrada dos resultados à luz da Ciência da Informação e campos interdisciplinares

A construção da OntoDiplomaDigital exigiu decidir como os elementos documentais do Diploma Digital seriam interpretados e representados para atender às

necessidades de recuperação estabelecidas na pesquisa. Os resultados permitem discutir essa passagem das fontes para o modelo e do modelo para sua utilização computacional. A contribuição observada foi a explicitação de conceitos e relações do Registro do Diploma que puderam ser recuperados no conjunto sintético, mantendo distinta essa evidência das condições necessárias para uma aplicação institucional.

A seleção dos elementos do modelo pode ser compreendida pela perspectiva de Hjørland (2016), que relaciona os sistemas de organização do conhecimento à seleção de conceitos e relações semânticas e reconhece as perspectivas que fundamentam a análise de um domínio. Na pesquisa, as normas e os schemas XML/XSD do MEC forneceram o corpus documental, enquanto o escopo e as questões de competência orientaram a identificação dos elementos pertinentes à representação. Os componentes dos schemas não foram convertidos indiscriminadamente em classes e propriedades: foram examinados segundo seu significado, sua função e sua contribuição para as informações que se pretendia recuperar. O modelo expressa, portanto, uma interpretação delimitada desse corpus, sem abranger todas as práticas e perspectivas dos profissionais envolvidos no Diploma Digital.

A metodologia OntoForInfoScience organizou esse exame ao exigir, em sua etapa preliminar, a avaliação da necessidade de uma ontologia diante de outros instrumentos de representação (Mendonça, 2015). A escolha decorreu da necessidade de representar recursos identificáveis, diferentes relações e restrições, além da padronização dos termos. As questões de competência transformaram essa necessidade em requisitos informacionais, utilizados na seleção dos conceitos e na elaboração das consultas. Glossários, dicionários e modelos conceituais apoiaram a passagem entre os termos encontrados nas fontes e sua expressão formal; a documentação preservou as definições e as justificativas das escolhas. Esses artefatos permitiram examinar como o conteúdo documental foi interpretado antes de sua utilização computacional.

A distinção entre instituição emissora e instituição registradora exemplifica uma decisão cujo significado precisa ser preservado na representação. As duas funções podem ser desempenhadas pela mesma instituição, mas correspondem a relações distintas com o diploma. A modelagem explicitou essa diferença, e a consulta QC1 recuperou separadamente os papéis de emissão e registro associados ao diploma sintético selecionado. A distinção entre responsável pelo registro, assinantes e

assinatura digital também separou agentes envolvidos no processo de um recurso técnico associado ao documento. Esses resultados mostram que a seleção das relações permitiu recuperar diferenças relevantes para os requisitos da pesquisa, em vez de tratar os elementos apenas como campos isolados.

O tratamento das categorias e dos valores controlados envolveu outra decisão interpretativa. Mazzocchi (2018) destaca que a formalização dos sistemas de organização do conhecimento não elimina as escolhas relativas à finalidade e à representação do domínio. Na OntoDiplomaDigital, foi necessário distinguir categorias representadas por classes, relações expressas por propriedades, valores terminológicos organizados como conceitos SKOS e ocorrências sintéticas reunidas na ABox. Graus, títulos e modalidades receberam identificação e controle vocabular, e a QC5 permitiu examinar seus usos no conjunto de teste. O reúso de termos externos também considerou seus significados, domínios, alcances e restrições, evitando correspondências baseadas apenas na semelhança dos nomes. A formalização tornou essas escolhas explícitas e examináveis; as consultas demonstraram sua utilização, sem medir redução de ambiguidades ou superioridade sobre outras representações possíveis.

A passagem para RDF precisa ser interpretada pela função dessa representação, e não apenas pela mudança de formato. Riley (2017) distingue a organização dos metadados em árvores XML de sua representação em grafos RDF, além de separar a padronização estrutural, o controle dos valores e as orientações de conteúdo. Essa distinção ajuda a compreender a complementaridade entre os schemas oficiais e a ontologia: os XML/XSD mantêm sua função na estruturação e validação documental, enquanto o modelo ontológico explicita categorias e relações pertinentes ao Registro do Diploma. Na prova de conceito, os vínculos entre diploma, titular, curso, instituições e registro puderam ser percorridos pelas consultas. Não houve comparação experimental entre XML e RDF; a evidência obtida refere-se ao uso das relações representadas no grafo.

O caráter operacional da ontologia decorreu da associação entre conceitualização, implementação e uso. Falbo (2014) distingue a descrição conceitual de um domínio de sua implementação para processamento por máquinas, considerando o ambiente e os requisitos computacionais. Essa distinção permite interpretar a formalização em RDF, OWL e SKOS, a produção da ABox e a execução das consultas no GraphDB como partes de uma representação destinada à utilização.

A recuperação demonstrada também se aproxima da proposta de Berners-Lee, Hendler e Lassila (2001) de explicitar significados para o processamento das informações por programas: os identificadores foram associados a conceitos e relações definidos, que as consultas puderam utilizar. A aproximação com Falbo é conceitual e operacional, sem caracterizar adoção da SABiO, pois o percurso metodológico foi conduzido pela OntoForInfoScience.

As evidências da Seção 5.6.3 estabelecem a ligação entre essas decisões e os requisitos informacionais. As sete consultas vinculadas às questões de competência recuperaram informações acadêmicas, institucionais, administrativas, documentais e de proveniência no conjunto sintético. A recuperação dos elementos associados à assinatura e à validação descreveu esses recursos e seus vínculos, sem verificar criptograficamente a autenticidade dos documentos. Conforme a distinção entre validação e verificação apresentada por Mendonça (2015), responder às questões constitui um critério de avaliação, mas não equivale ao cumprimento de todos os critérios metodológicos. A pesquisa realizou confrontação interna e documental e verificações sintáticas, estruturais e operacionais; não houve validação externa com especialistas nem teste automatizado de consistência lógica por um reasoner OWL.

A interpretação compartilhada dos dados exige uma análise adicional ao funcionamento das consultas. Guizzardi (2020) argumenta que a interoperabilidade semântica depende de explicitar os compromissos conceituais da representação, pois termos coincidentes podem designar entidades ou relações diferentes. O autor defende o apoio de teorias e ontologias de fundamentação para examinar essas distinções. A OntoDiplomaDigital não adotou uma ontologia de fundamentação, decisão situada no recorte operacional e na possibilidade de adaptação admitida por Mendonça (2015, p. 210) para projetos menores, ontologias leves de domínio ou subontologias. Essa possibilidade metodológica não comprova a adequação das categorias escolhidas nem elimina o limite apontado por Guizzardi. A documentação e o reúso de vocabulários auxiliam o exame das escolhas, mas não substituem sua análise com apoio de uma fundamentação ontológica.

As consultas federadas acrescentaram evidências de recuperação distribuída, cuja interpretação encontra apoio na abordagem de Dados Ligados de Bizer, Heath e Berners-Lee (2009). Os autores relacionam a identificação dos recursos e os vínculos entre fontes à possibilidade de conectar descrições publicadas na Web. Na demonstração com a DBpedia, a consulta acrescentou classificações do recurso

Brasil como contexto previamente definido. Na consulta ao Wikidata, acessado por meio do QLever, o código IBGE foi utilizado para relacionar a localidade dos dados fictícios ao município de São Paulo. A primeira demonstração reuniu contexto externo; a segunda mostrou a função de um identificador compartilhado na correspondência entre fontes, sem incorporar permanentemente os dados externos à ABox local.

Esses exemplos não demonstram a integração semântica geral entre sistemas acadêmicos. Bizer, Heath e Berners-Lee (2009) ressaltam que integrar dados requer tratar diferenças entre vocabulários, conflitos e manutenção dos vínculos. As consultas executadas utilizaram fontes e correspondências delimitadas, sem examinar esses problemas em escala institucional. Também não equivalem à publicação plena da OntoDiplomaDigital em Dados Ligados: a disponibilização dos arquivos e a consulta a fontes externas não demonstram, para todos os recursos locais, o acesso às suas descrições na Web e sua utilização por outras aplicações.

Os princípios FAIR permitem avaliar as condições de utilização dos artefatos para além de sua recuperação nas consultas. Wilkinson et al. (2016) consideram encontrabilidade, acessibilidade, interoperabilidade e reúso por pessoas e máquinas, sem prescrever uma tecnologia específica. A identificação dos recursos, as definições, as relações e a documentação produzidas favorecem o exame da representação; a disponibilização dos artefatos permite acessar o modelo e os procedimentos empregados. A prova de conceito, entretanto, não avaliou sistematicamente a persistência dos identificadores, a indexação e a continuidade de acesso aos metadados ou as condições de licença e proveniência requeridas para o reúso. Há contribuições compatíveis com a aplicação de FAIR, mas não evidência suficiente para afirmar atendimento ao conjunto dos princípios.

A comparação com o Estado da Arte situa essa contribuição entre propostas de finalidades mais amplas. Halloluwa e Silva (2025) apresentam uma arquitetura de gestão da aprendizagem que combina blockchain e grafo de conhecimento para organizar informações da trajetória educacional; Kontzinos et al. (2020) discutem o QualiChain, que prevê recursos semânticos e outros serviços para qualificações, recrutamento e gestão de competências. Esses trabalhos mostram que a representação relacional e o reúso de vocabulários integram a agenda de projetos educacionais, mas várias funções descritas permanecem propositivas. A contribuição da OntoDiplomaDigital está no recorte do Registro do Diploma Brasileiro, vinculado às fontes normativas e técnicas do MEC, e na utilização demonstrada dessa

representação em consultas. A comparação delimita seu objeto e suas evidências, sem sustentar superioridade entre as propostas ou equivalência entre suas funcionalidades.

A relevância desse recorte pode ser discutida a partir da experiência institucional relatada por Silva e Zárate (2022). Na implantação do Diploma Digital na UFMG, as autoras identificaram registros que não se ajustavam diretamente aos requisitos da solução utilizada, incluindo avaliações sem nota numérica e identificadores repetidos de atividades. Esses casos exigiram esclarecimentos e adaptações e evidenciaram a importância da definição dos conceitos diante da diversidade de registros acadêmicos. A OntoDiplomaDigital não foi aplicada a esses problemas, vários dos quais ultrapassam seu núcleo operacional. Sua contribuição foi tornar explícitos os conceitos e as relações do Registro do Diploma, oferecendo uma base documentada para discutir as interpretações adotadas; a utilidade dessa base na implantação ainda depende de avaliação institucional.

A aplicação também exige condições técnicas e organizacionais que a representação, isoladamente, não fornece. Scheid (2022) desenvolve e avalia mecanismos de interação entre plataformas blockchain heterogêneas, enfrentando diferenças de interfaces e detalhes técnicos, sem ter as credenciais acadêmicas como objeto específico. El Chawra (2020), ao investigar experiências de portabilidade de dados estudantis, identifica valores, prioridades, liderança e compreensão compartilhada do propósito como fatores relacionados à adoção. Os estudos ajudam a distinguir a definição dos significados, os mecanismos de intercâmbio e a incorporação às práticas institucionais. A prova de conceito examinou a primeira dimensão em seu uso computacional delimitado, mas não implementou a integração de sistemas acadêmicos nem avaliou a adoção da ontologia por seus profissionais.

Para o Diploma Digital, os resultados permitem compreender a ontologia como uma representação complementar à estrutura documental: ela explicita o que os recursos representam e como se relacionam para atender a necessidades de recuperação. Os fundamentos da Organização e Representação do Conhecimento deram suporte ao exame dos conceitos, papéis e valores; a metodologia organizou sua transformação em artefatos documentados; e os recursos da Web Semântica permitiram utilizar as relações nas consultas. Essa articulação foi demonstrada nas condições da prova de conceito. Sua ampliação requer avaliar a adequação conceitual com especialistas, verificar a consistência lógica, testar dados institucionais

autorizados e estabelecer condições de integração e manutenção, preservando a diferença entre uma representação utilizável e sua aplicação efetiva entre instituições.

6 Considerações finais

Esta dissertação teve como objetivo propor uma ontologia operacional para organizar e representar o conhecimento do subdomínio do Registro do Diploma Digital Brasileiro, fundamentada na Organização e Representação do Conhecimento e implementada por meio de padrões da Web Semântica. A construção da OntoDiplomaDigital e sua utilização na prova de conceito permitiram explicitar conceitos e relações desse subdomínio e recuperar as informações previstas nas questões de competência, nos limites do conjunto sintético e do ambiente de teste.

Em resposta à questão de pesquisa, os resultados mostraram que a organização dos dados do Registro do Diploma pode ser apoiada por uma representação que articule categorias do domínio, relações entre recursos, atributos e valores controlados. Essa articulação permitiu consultar o diploma em relação ao titular, à formação acadêmica, às instituições e aos elementos administrativos e documentais associados ao registro. A contribuição demonstrada consistiu na construção e na utilização operacional dessa representação, oferecendo uma base para investigar a interoperabilidade semântica em aplicações posteriores.

Um dos principais aprendizados foi compreender que a passagem das estruturas XML/XSD para uma representação ontológica exigiu escolhas conceituais. Os esquemas oficiais forneceram os elementos documentais e as regras estruturais que fundamentaram a pesquisa. A construção da ontologia demandou examinar o significado e a função desses elementos, definir quais seriam representados como categorias, relações ou atributos e organizar os valores que exigiam controle terminológico. A Organização e Representação do Conhecimento orientou essa interpretação, relacionando as fontes utilizadas às necessidades informacionais expressas nas questões de competência.

A distinção entre instituição emissora e instituição registradora exemplificou esse processo. A representação explicitou dois papéis institucionais, inclusive quando desempenhados pela mesma organização, e permitiu recuperá-los separadamente nas consultas. A diferenciação entre agentes administrativos, assinantes e assinatura digital preservou funções distintas no processo de registro. O tratamento de graus, títulos e modalidades como conceitos controlados também tornou explícitas as categorias utilizadas para descrever a formação acadêmica. Esses resultados mostram como a análise conceitual apoiou a definição das relações necessárias à

recuperação pretendida, embora não comprovem a eliminação de ambiguidades ou uma melhoria de eficácia em comparação com outros sistemas.

A ontologia acrescentou aos esquemas XML/XSD uma representação em grafo, na qual os recursos puderam ser identificados e relacionados segundo os significados definidos para o subdomínio. Essa camada preservou a função dos esquemas oficiais na estruturação e na validação documental e permitiu consultar vínculos acadêmicos, institucionais e administrativos em uma representação comum. O resultado evidenciou que a contribuição semântica depende da explicitação dos conceitos e das relações utilizados nas consultas, além da transformação dos dados para outro formato.

A aplicação da OntoForInfoScience apoiou a construção dessa representação ao articular a delimitação do problema, a aquisição do conhecimento, a conceitualização, a formalização e a avaliação. As questões de competência estabeleceram a ligação entre os requisitos informacionais e as consultas; os glossários, os dicionários e os modelos conceituais registraram as escolhas anteriores à implementação; e a documentação permitiu relacionar os elementos formalizados às fontes que fundamentaram sua construção. O aprendizado metodológico esteve nessa continuidade entre interpretação, representação e verificação, que tornou as decisões passíveis de exame e revisão. A aplicação também evidenciou dimensões de avaliação que permaneceram pendentes, especialmente a validação externa e a verificação automatizada da consistência lógica.

O percurso contemplou os seis objetivos específicos ao analisar as fontes normativas e técnicas, identificar os vocabulários pertinentes, construir o modelo conceitual, estabelecer as correspondências para a ontologia operacional, implementar a representação e realizar a prova de conceito. A prova de conceito utilizou 252 registros sintéticos e permitiu executar as sete consultas vinculadas às questões de competência, acompanhadas de quatro demonstrações complementares. As respostas forneceram evidências de correspondência entre os requisitos definidos, as relações modeladas e as informações recuperadas. As consultas federadas também exploraram a combinação de informações locais com fontes externas, incluindo uma correspondência estabelecida por código IBGE. Essas evidências demonstraram o funcionamento da representação nas condições testadas.

O alcance da interoperabilidade examinada precisa ser delimitado. A recuperação das relações no conjunto sintético mostrou que a interpretação

formalizada pôde ser utilizada computacionalmente. A demonstração federada indicou como um identificador compartilhado pode apoiar a correspondência entre recursos de fontes distintas. Esses resultados não comprovaram, entretanto, que sistemas acadêmicos de diferentes instituições atribuam os mesmos significados aos conceitos e às relações da ontologia. Essa avaliação exige confrontar a representação com dados, práticas e interpretações institucionais que não integraram a prova de conceito.

Para a Ciência da Informação, a pesquisa contribuiu ao documentar a articulação entre organização conceitual, controle terminológico e implementação computacional no Registro do Diploma Digital. A disponibilização dos artefatos, das consultas e dos registros de verificação permitiu examinar tanto a representação produzida quanto sua utilização nos testes. A documentação no GitHub tornou examináveis as interpretações adotadas na passagem das fontes XML/XSD para os conceitos, propriedades e relações da ontologia. Esse registro permite que pesquisadores e profissionais confrontem as escolhas com suas fontes e proponham revisões ou novas relações conceituais. As propostas de ampliação deverão ser avaliadas quanto à pertinência ao domínio e à coerência com o modelo, preservando a rastreabilidade das alterações. Essa possibilidade de exame e revisão relaciona a implementação computacional às preocupações da Ciência da Informação com a organização, a representação e o uso contextualizado da informação.

No plano social, a organização semântica dos dados acadêmicos pode apoiar serviços que facilitem a compreensão e a circulação autorizada das informações sobre a formação dos cidadãos. Uma possibilidade futura seria constituir um ambiente governamental colaborativo entre instituições federais de ensino, inspirado na experiência de intercâmbio de registros acadêmicos do EMREX. Esse ambiente poderia apoiar estudantes em mobilidade nacional e internacional, profissionais que necessitam apresentar sua formação e gestores responsáveis pela conferência e pelo tratamento dessas informações. Sua concretização exigiria ampliar o recorte da pesquisa, pois a mobilidade acadêmica também envolve resultados de aprendizagem e históricos escolares, não representados integralmente pela OntoDiplomaDigital.

A distinção entre titularidade, emissão e registro pode contribuir para tornar as informações mais compreensíveis. Identificar quem recebeu o diploma, qual instituição o emitiu e qual instituição realizou seu registro permite recuperar separadamente responsabilidades que podem estar reunidas em um mesmo documento. A representação semântica complementa os mecanismos de assinatura

e validação ao explicitar o significado dessas relações. A pesquisa demonstrou sua recuperação no conjunto sintético, sem realizar validação criptográfica de assinaturas ou avaliar serviços institucionais.

A transformação dessas possibilidades em serviços dependerá de acordos institucionais sobre significados, responsabilidades, qualidade dos dados e condições de acesso. Também exigirá infraestrutura para manter os artefatos, estabelecer correspondências entre vocabulários e executar consultas sobre fontes autorizadas. O enriquecimento semântico pressupõe interpretar os recursos e suas relações para responder a necessidades informacionais, além de coletar ou conectar valores. A ontologia oferece uma base para esse desenvolvimento, mas seus benefícios para cidadãos, profissionais e gestores ainda precisam ser examinados com instituições e usuários em situações reais.

As limitações da pesquisa abrangem o recorte do subdomínio e os procedimentos de avaliação. A OntoDiplomaDigital não representou integralmente a Documentação Acadêmica, o Histórico Escolar Digital ou os demais componentes do ecossistema do Diploma Digital. A avaliação semântica teve caráter interno e documental, sem validação externa por especialistas ou instituições usuárias. Também não foi empregado um *reasoner* OWL para verificar automaticamente a consistência lógica da ontologia. As verificações realizadas concentraram-se nos aspectos sintáticos, estruturais e operacionais examinados, não permitindo afirmar a ausência de todas as possíveis inconsistências dos axiomas e das restrições formalizados. A prova de conceito não validou criptograficamente assinaturas, não integrou sistemas acadêmicos reais e não avaliou desempenho em escala.

As investigações futuras devem aprofundar a avaliação antes de ampliar as aplicações da proposta. São prioridades a análise com especialistas e usuários do domínio, a verificação automatizada da consistência lógica e os testes com dados institucionais autorizados. A ampliação modular para outros componentes do Diploma Digital poderá ser acompanhada pela avaliação das correspondências de significado entre instituições. A disponibilização de descrições acessíveis pelos identificadores dos recursos, a manutenção dos vínculos externos e a definição de condições de governança também constituem etapas necessárias para investigar sua utilização contínua em Dados Ligados e sua relação com os princípios FAIR. Um portal semântico institucional poderá ser considerado como aplicação futura, condicionado à avaliação dessas condições e das necessidades de seus usuários.

A OntoDiplomaDigital apresentou uma representação operacional do Registro do Diploma Digital Brasileiro e evidências de sua utilização em consultas com dados sintéticos. A pesquisa tornou explícita a ligação entre as fontes documentais, as escolhas conceituais e a recuperação das relações modeladas. Sua contribuição oferece uma base documentada para que a adequação da representação e suas possibilidades de aplicação sejam examinadas em contextos institucionais.

7 Referências

ALBUQUERQUE, Maria Elizabeth Baltar; MARTINS, Gracy Kelli; MOTA, Denysson Axel Ribeiro. Organização e representação da Informação e do Conhecimento: intersecções teórico-sociais. [S. l.]: Editora UFPB, 2019. Disponível em: https://www.editora.ufpb.br/sistema/press5/index.php/UFPB/catalog/book/355. Acesso em: 24 fev. 2025.

ALMEIDA, M.; SOUZA, R.; FONSECA, F. Semantics in the Semantic Web: A Critical Evaluation. Knowledge Organization, [s. l.], v. 38, n. 3, p. 187–203, 2011.

ARAKAKI, A. C. S.; ARAKAKI, F. A. Dados e metadados: conceitos e relações: concepts and relationships. Ciência da Informação, [s. l.], v. 49, n. 3, 2020. Disponível em: https://revista.ibict.br/ciinf/article/view/5504. Acesso em: 16 nov. 2025.

ARAÚJO, C. A. Á. Fundamentos da ciência da informação: correntes teóricas e o conceito de informação. Perspectivas em Gestão & Conhecimento, [s. l.], v. 4, n. 1, p. 57–79, 2014.

ATEMEZING, G. A. et al. WaSABi'13: Proceedings of the 2013th International Conference on Semantic Web Enterprise Adoption and Best Practice. [s. l.], v. 1106, n. CEUR-WS.org, 2013.

AUER, Sören; LEHMANN, Jens; NGONGA NGOMO, Axel-Cyrille; ZAVERI, Amrapali. Introduction to linked data and its lifecycle on the Web. In: RUDOLPH, Sebastian; GOTTLOB, Georg; HORROCKS, Ian; VAN HARMELEN, Frank (ed.). Reasoning Web: semantic technologies for intelligent data access. Berlin: Springer, 2013. p. 1–90. DOI: 10.1007/978-3-642-39784-4_1. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-39784-4_1. Acesso em: 11 ago. 2026.

BAADER, Franz et al. (ed.). The description logic handbook: theory, implementation, and applications. Cambridge: Cambridge University Press, 2003.

BACA, Murtha (ed.). Introduction to metadata. 3. ed. Los Angeles: Getty Research Institute, 2016.

BAHNEMANN, Greta et al. Transforming metadata into linked data to improve digital collection discoverability: a CONTENTdm pilot project. Dublin, Ohio, USA: OCLC, 2021. Disponível em: https://doi.org/10.25333/fzcv-0851. Acesso em: 25 nov. 2021.

BARITÉ, Mario. Sistemas de organización del conocimiento: una tipología actualizada. Información, Cultura y Sociedad, n. 25, p. 15-36, 2011.

BAUER, Florian; KALTENBÖCK, Martin. Linked open data: the essentials: the climate knowledge brokering edition. 2. ed. Vienna, Austria: edition mono/monochrom, 2016. Disponível em: https://www.reeep.org/sites/default/files/LOD-TheEssentials2016.pdf. Acesso em: 12 out. 2025.

BERNERS-LEE, Tim. Linked Data. Web page: W3C, 2006. Disponível em: https://www.w3.org/DesignIssues/LinkedData.html. Acesso em: 24 fev. 2016.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The semantic web. Scientific American, [s. l.], v. 284, n. 5, p. 28–37, 2001.

BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked data - the story so far. International Journal on Semantic Web and Information Systems (IJSWIS), [s. l.], v. 5, n. 3, p. 1–22, 2009.

BORST, Willem Nico. Construction of engineering ontologies for knowledge sharing and reuse. 1997. - University of Twente, [s. l.], 1997. Disponível em: https://research.utwente.nl/en/publications/construction-of-engineering-ontologies-for-knowledge-sharing-and-. Acesso em: 11 ago. 2026.

BRASIL. Decreto nº 10.332, de 28 de abril de 2020. Institui a Estratégia de Governo Digital para o período de 2020 a 2022, no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional e dá outras providências. Brasília, DF: Presidência da República,

2020a. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/decreto/d10332.htm. Acesso em: 11 ago. 2026.

BRASIL. Decreto nº 12.069, de 21 de junho de 2024. Dispõe sobre a Estratégia Nacional de Governo Digital e a Rede Nacional de Governo Digital – Rede Gov.br e institui a Estratégia Nacional de Governo Digital para o período de 2024 a 2027. Brasília, DF: Presidência da República, 2024a. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2024/Decreto/D12069.htm. Acesso em: 16 nov. 2025.

BRASIL. Decreto nº 12.198, de 24 de setembro de 2024. Institui a Estratégia Federal de Governo Digital para o período de 2024 a 2027 e a Infraestrutura Nacional de Dados, no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional. Brasília, DF: Presidência da República, 2024b. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2024/decreto/d12198.htm. Acesso em: 11 fev. 2025.

BRASIL. Decreto nº 9.756, de 11 de abril de 2019. Institui o portal único “gov.br” e dispõe sobre as regras de unificação dos canais digitais do Governo federal. Brasília, DF: Presidência da República, 2019a. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/d9756.htm. Acesso em: 16 nov. 2025.

BRASIL. Lei nº 14.129, de 29 de março de 2021. Dispõe sobre princípios, regras e instrumentos para o Governo Digital e para o aumento da eficiência pública e altera a Lei nº 7.116, de 29 de agosto de 1983, a Lei nº 12.527, de 18 de novembro de 2011 (Lei de Acesso à Informação), a Lei nº 12.682, de 9 de julho de 2012, e a Lei nº 13.460, de 26 de junho de 2017. Brasília, DF: Presidência da República, 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/l14129.htm. Acesso em: 11 ago. 2026.

BRASIL. Ministério da Educação. Instrução Normativa nº 1, de 15 de dezembro de 2020. Dispõe sobre a regulamentação técnica para a emissão e o registro de diploma de graduação, por meio digital, pelas Instituições de Ensino Superior (IES) pertencentes ao Sistema Federal de Ensino. Brasília, DF: Ministério da Educação, 2020b. Disponível em: https://portal.mec.gov.br/diplomadigital/arquivos/in_01_15122020.pdf. Acesso em: 16 nov. 2025.

BRASIL. Ministério da Educação. Instrução Normativa nº 5, de 14 de outubro de 2022. Aprova a versão 1.05 dos Anexos I, II e III da Instrução Normativa IN/SESU nº 1, de 15 de dezembro de 2020. Brasília, DF: Ministério da Educação, 2022. Disponível em: https://www.gov.br/mec/pt-br/diploma-digital/documentos/in-05-versao-completa-anexos-i-ii-e-iii-v1-05.pdf. Acesso em: 27 out. 2026.

BRASIL. Ministério da Educação. Nota Técnica nº 13/2019/DIFES/SESU/SESU: Diploma Digital. Versão 1.0. Brasília, DF: Ministério da Educação, 2019b. Disponível em: https://portal.mec.gov.br/diplomadigital/arquivos/nota_tecnica_NOTA_13_versao_1.00.pdf. Acesso em: 27 out. 2026.

BRASIL. Ministério da Educação. Pacote XSD v1.05. Brasília, DF: Ministério da Educação, [s. d.]. Disponível em: https://www.gov.br/mec/pt-br/diploma-digital/dados. Acesso em: 27 out. 2026.

BRASIL. Ministério da Educação. Portaria MEC nº 70, de 24 de janeiro de 2025. Altera a Portaria MEC nº 554, de 11 de março de 2019, que dispõe sobre a emissão e o registro de diploma de graduação, por meio digital, pelas Instituições de Ensino Superior – IES pertencentes ao Sistema Federal de Ensino. Diário Oficial da União: seção 1, Brasília, DF, n. 19, p. 28-29, 28 jan. 2025. Disponível em: https://pesquisa.in.gov.br/imprensa/servlet/INPDFViewer?captchafield=firstAccess&data=28%2F01%2F2025&jornal=515&pagina=28. Acesso em: 11 ago. 2026.

BRASIL. Ministério da Educação. Portaria nº 330, de 5 de abril de 2018. Dispõe sobre a emissão de diplomas em formato digital nas instituições de ensino superior pertencentes ao sistema federal de ensino. Brasília, DF: Ministério da Educação, 2018. Disponível em: https://www.gov.br/conarq/pt-

br/legislacao-arquivistica/portarias-federais/portaria-no-330-de-5-de-abril-de-2018. Acesso em: 27 out. 2026.

BRASIL. Ministério da Educação. Portaria nº 554, de 11 de março de 2019. Dispõe sobre a emissão e o registro de diploma de graduação, por meio digital, pelas Instituições de Ensino Superior – IES pertencentes ao Sistema Federal de Ensino. Brasília, DF: Ministério da Educação, 2019c. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/66544171/do1-2019-03-12-portaria-n-554-de-11-de-marco-de-2019-66543842. Acesso em: 6 jul. 2025.

BRASIL. Ministério da Gestão e da Inovação em Serviços Públicos. Secretaria de Governo Digital. Guia de implementação dos programas de governança de dados. Brasília, DF: Secretaria de Governo Digital, 2024c. Disponível em: https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/governancadedados/pgdados. Acesso em: 14 ago. 2026.

BRASIL. Ministério da Gestão e da Inovação em Serviços Públicos. Secretaria de Governo Digital. Portaria SGD/MGI nº 6.618, de 25 de setembro de 2024. Estabelece os princípios, os objetivos e as iniciativas para o alcance da Estratégia Federal de Governo Digital para o período de 2024 a 2027. Brasília, DF: Secretaria de Governo Digital, 2024d. Disponível em: https://www.gov.br/governodigital/pt-br/estrategias-e-governanca-digital/EFGD/portaria-sgd-mgi-no-6-618-de-25-de-setembro-de-2024. Acesso em: 1 dez. 2025.

CAPURRO, Rafael. Epistemologia e ciência da informação. 2003, Belo Horizonte. Anais [...]. Belo Horizonte: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação, 2003. Disponível em: http://www.capurro.de/enancib_p.htm. Acesso em: 23 fev. 2026.

CAPURRO, Rafael. et al. O conceito de informação. Perspectivas em Ciência da Informação, [s. l.], v. 12, n. 1, 2007. Disponível em: https://periodicos.ufmg.br/index.php/pci/article/view/22360. Acesso em: 2 mar. 2025.

CARDOSO, J. C.; LUZ, A. R. Os arquivos e os sistemas de gestão da qualidade. [s. l.], 2005.

CHUE HONG, N. P. et al. FAIR Principles for Research Software (FAIR4RS Principles). [s. l.], 2021. Disponível em: https://zenodo.org/record/6623556#.YqCJTJNBwlw. Acesso em: 25 jul. 2025.

CORREIA, MA. Plano de organização dos dossiês acadêmicos digitais da Universidade Federal de Sergipe: modelo RiC. … de graduação da Universidade Federal de Sergipe, [s. l.], 2025. Disponível em: https://ri.ufs.br/handle/riufs/24375.

CRESWELL, John W.; CRESWELL, J. David. Projeto de pesquisa: métodos qualitativo, quantitativo e misto. 5. ed. Porto Alegre: Penso, 2021. Disponível em: https://loja.grupoa.com.br/projeto-de-pesquisa-metodos-qualitativo-5ed9786581334185-p1016926. Acesso em: 27 out. 2026.

CRISTOVÃO, H. M. Dados ligados e nanopublicações, 2026. Notas de aula. Programa de Pós-graduação em Ciência da Informação (PPGCI), CCJE, UFES. [web page]. Disponível em: https://bit.ly/cristovaohm-dln.

CRISTOVÃO, H. M. Mapas conceituais na organização do conhecimento. [S. l.], 2022. webpage, mapa conceitual. Disponível em: https://cmapscloud.ihmc.us/viewer/cmap/1VK20K4CM-9JFFKS-G69. Acesso em: 25 nov. 2025.

CYGANIAK, R. et al. RDF 1.1 concepts and abstract syntax. Web page: W3C, 2014. Disponível em: http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225. Acesso em: 24 out. 2025.

DIPLOMA DIGITAL. [S. l.], 2023. Disponível em: https://ajuda.rnp.br/diplomas-digitais. Acesso em: 16 nov. 2025.

EL CHAWRA, T. Nevine. Implementing digital student data portability: experiences and lessons learned. 2020. Dissertação (Master of Arts) — University of Toronto, Toronto, 2020.

EMREX. About us. [S. l.], [s. d.]a. Disponível em: https://emrex.eu/about-us/. Acesso em: 30 ago. 2026.

EMREX. EMREX handbook: for administrators, coordinators etcetera. [S. l.]: EMREX, [s. d.]b. Disponível em: https://emrex.eu/wp-content/uploads/2019/03/EMREX-handbook.pdf. Acesso em: 30 ago. 2026.

EMREX. Technical guide to EMREX. [S. l.]: EMREX, 2020. Disponível em: https://emrex.eu/technical/. Acesso em: 13 ago. 2026.

EUROPEAN COMMISSION. Directorate-General for Research and Innovation. Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data. Luxembourg: Publications Office of the European Union, 2018. ISBN 978-92-79-96546-3. DOI: 10.2777/1524. Disponível em: https://doi.org/10.2777/1524. Acesso em: 14 ago. 2026.

EUROPEAN COMMISSION. European Learning Model for stakeholders. Europass, [s. d.]a. Disponível em: https://europass.europa.eu/en/european-learning-model-stakeholders. Acesso em: 30 ago. 2026.

EUROPEAN COMMISSION. Launch event: European Learning Model. Europass, 11 maio 2023. Disponível em: https://europass.europa.eu/en/events/launch-event-european-learning-model. Acesso em: 30 ago. 2026.

FALBO, R. de A. SABiO: Systematic approach for building ontologies. In: 1ST JOINT WORKSHOP ONTO.COM/ODISE ON ONTOLOGIES IN CONCEPTUAL MODELING AND INFORMATION SYSTEMS ENGINEERING CO-LOCATED WITH 8TH INTERNATIONAL CONFERENCE ON FORMAL ONTOLOGY IN INFORMATION SYSTEMS, 2014, Rio de Janeiro, RJ. Anais [...]. Rio de Janeiro, RJ: CEUR Workshop Proceedings, 2014. p. 14. Disponível em: http://ceur-ws.org/Vol-1301/ontocomodise2014_2.pdf.

GIL, Antônio Carlos. Métodos e técnicas de pesquisa social. São Paulo: Atlas, 2008.

GILLILAND, Anne J. Setting the stage. In: BACA, Murtha (ed.). Introduction to metadata. 3. ed. Los Angeles: Getty Research Institute, 2016. p. 1-20. Disponível em: https://www.getty.edu/publications/intrometadata/setting-the-stage/. Acesso em: 15 mar 2026.

GO FAIR Initiative: make your data & services FAIR. [S. l.], 2025. Disponível em: https://www.go-fair.org/. Acesso em: 25 jul. 2025.

GRUBER, Thomas R. A translation approach to portable ontology specifications. Knowledge Acquisition, [s. l.], v. 5, n. 2, p. 199–220, 1993.

GUIZZARDI, G. Ontology, Ontologies and the “I” of FAIR. Data Intelligence, [s. l.], v. 2, n. 1–2, p. 181–191, 2020.

HALLOLUWA, Thilina; DE SILVA, Chameera. Blockchain in education technology: “a tool for challenges in education technology”: a review paper. Preprints, [S. l.], 3 mar. 2025. Preprint. DOI: https://doi.org/10.20944/preprints202503.0130.v1. Disponível em: https://www.preprints.org/manuscript/202503.0130/v1. Acesso em: 19 ago. 2026.

HARZING, AW (2007) Publish ou Perish, disponível em https://harzing.com/resources/publish-or-perish

HEATH, Tom; BIZER, Christian. Linked data: evolving the Web into a global data space. 1. ed. San Rafael, CA: Morgan & Claypool Publishers, 2011. (Synthesis Lectures on the Semantic Web: Theory and Technology, v. 1, n. 1). DOI: https://doi.org/10.2200/S00334ED1V01Y201102WBE001. Disponível em: http://linkeddatabook.com/editions/1.0/.

HITZLER, P. A review of the semantic web field. Commun. ACM, [s. l.], v. 64, n. 2, p. 76–83, 2021.

HJØRLAND, Birger. Knowledge organization (KO). Knowledge Organization, [S. l.], v. 43, n. 6, p. 475–484, 2016. DOI: 10.5771/0943-7444-2016-6-475. Disponível em: https://doi.org/10.5771/0943-7444-2016-6-475. Acesso em: 1 out. 2026

HOOLAND, Seth Van; VERBORGH, Ruben. Linked data for libraries, archives and museums: how to clean, link and publish your metadata. Chicago, IL: Facet Publishing, 2014.

INC, iTeh. EN 15981:2011 - EuroLMAI Standard for Learner Achievement Information. [S. l.], 2011. Disponível em: https://standards.iteh.ai/catalog/standards/cen/33566761-c9a2-4c12-bcd6-bbba0d95b4c1/en-15981-2011. Acesso em: 20 ago. 2026.

KARUNARATNE, T.; KONTOPOULOS, E. Supporting Learning Mobility With Student Data Harmonisation: A European Perspective. In: PROCEEDINGS OF THE EUROPEAN CONFERENCE ON E-LEARNING, ECEL, 2022. (Fotaris P. & Blake A., Org.)Proc. Eur. Conf. e-Learn., ECEL. [S. l.]: Academic Conferences and Publishing International Limited, 2022. p. 156–164. Disponível em: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85172310602&doi=10.34190%2fecel.21.1.908&partnerID=40&md5=d1b2216bf4d42a13e681331a9ffedc38.

KONTZINOS, Christos; MARKAKI, Ourania; KOKKINAKOS, Panagiotis; KARAKOLIS, Vagelis; KAPSALIS, Panagiotis; PSARRAS, John. Exploring blockchain, semantics and decision support to optimise qualification certification, recruitment and competency management: an assessment of challenges, current practices and opportunities. International Journal on Advances in Intelligent Systems, [S. l.], v. 13, n. 3-4, p. 192–203, 2020. ISSN 1942-2679. Disponível em: https://personales.upv.es/thinkmind/dl/journals/intsys/intsys_v13_n34_2020/intsys_v13_n34_2020_3.pdf. Acesso em: 31 ago. 2026.

LAUFER, C. Guia de web semântica. São Paulo, SP: CeWeb.br, Embaixada Britânica, iGovSP, 2015. Disponível em: https://www.cgi.br/media/docs/publicacoes/13/Guia_Web_Semantica.pdf.

LEMOS, D. L. da S.; MARTINS, D. L.; CARMO, D. do. Padrões de qualidade para dados e metadados endereçados a aplicações em Ciência de Dados. Advanced Notes in Information Science, [s. l.], v. 2, p. 161–170, 2022.

MAZZOCCHI, F. Knowledge organization system (KOS). Knowledge Organization, [s. l.], v. 45, n. 1, p. 54–78, 2018.

MENDONÇA, F. M. OntoForInfoScience: metodologia para construção de ontologias pelos cientistas da informação. Uma aplicação prática no desenvolvimento da ontologia sobre componentes do sangue humano (HEMONTO). 2015. Tese - (Doutorado em Ciência da Informação) - Escola de Ciência da Informação da Universidade Federal de Minas Gerais, Belo Horizonte, MG, 2015. Disponível em: https://hdl.handle.net/1843/BUBD-A35H3K.

MOREIRA, M. A. Aprendizagem significativa, organizadores prévios, mapas conceituais, diagramas V e unidades de ensino potencialmente significativas. [s. l.], 2012.

NOVAK, J. D.; CANAS, A. J. A teoria subjacente aos mapas conceituais e como elabora-los e usa-los. Praxis Educativa, [s. l.], v. 5, n. 1, p. 9–29, 2010.

NOY, Natalya F.; MCGUINNESS, Deborah L. Ontology Development 101: A Guide to Creating Your First Ontology. Stanford: Stanford Knowledge Systems Laboratory, 2001. Technical Report KSL-01-05; SMI-2001-0880. Disponível em: https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html. Acesso em: 14 ago. 2026.

PADRÕES. [S. l.], [s. d.]. Disponível em: https://www.w3c.br/padroes/. Acesso em: 24 ago. 2025.

PEREIRA, C. M.; LUZ, L. P. da; SANTAREM SEGUNDO, J. E. Papéis do profissional da informação na evolução da Web semântica. 2019, Florianópolis, SC. XX ENANCIB. Florianópolis, SC: ANCIB, 2019. Disponível em: https://brapci.inf.br/index.php/res/v/123723.

Planos de transformação digital. [S. l.], 2025. Disponível em: https://www.gov.br/governodigital/pt-br/estrategias-e-governanca-digital/planos-de-transformacao-digital/planos-de-transformacao-digital. Acesso em: 1 dez. 2025.

RAMALHO, Rogério Aparecido Sá. Ontologias e Knowledge Organization System (SKOS): aproximações e diferenças. In: GUIMARÃES, José Augusto Chaves; DODEBEI, Vera (org.). Organização do conhecimento e diversidade cultural. Marília: ISKO-Brasil; FUNDEPE, 2015. p. 100–107. Disponível em: https://isko.org.br/wp-content/uploads/2021/05/Proceedings-ISKO-Brasil-2015.pdf. Acesso em: 27 out. 2026.

RIBEIRO, Renato Carauta; ALMEIDA, Murilo Góes de; CANEDO, Edna Dias. A digital signature model using XAdES standard as a REST service. Information, [s. l.], v. 12, n. 8, art. 289, 2021. DOI: 10.3390/info12080289. Disponível em: https://www.mdpi.com/2078-2489/12/8/289. Acesso em: 27 out. 2026.

RICHARDSON, Roberto Jarry; PERES, José Augusto de Souza. Pesquisa social: métodos e técnicas. 3. ed. rev. e ampl. São Paulo: Atlas, 1999. Disponível em: https://climatechangemoz.com/wp-content/uploads/2020/04/Metodologia-de-Pesquisa-Social-Richardson.pdf.

RILEY, J. Understanding metadata: what is metadata, and what is it for. Baltimore, MD: National Information Standards Organization, 2017. (NISO Primer series).

RODRIGUES DIAS, Thiago Magela (org.). Informaçao, Dados e Tecnologia. Volume 2. Tallin: Colnes Publishing, 2022.

SAA - SOCIETY OF AMERICAN ARCHIVISTS. Metadata. [S. l.]: [s. d.], 2005. Disponível em: https://dictionary.archivists.org/entry/metadata.html. Acesso em: 4 fev. 2022.

SANTANA HÜLLER, Keylha; COELHO GOMES, Jorge Henrique; MARQUES LEITE DOS SANTOS, Vivianni. Mapeamento de sistemas de gerenciamento de documentos acadêmicos e expedição de diploma digital em universidades. Ciência da Informação, [s. l.], v. 51, n. 3, 2022. Disponível em: https://revista.ibict.br/ciinf/article/view/5497. Acesso em: 30 mar. 2026.

SANT'ANA, Ricardo César Gonçalves. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. Informação & Informação, [s. l.], v. 21, n. 2, p. 116, 2016.

SAYÃO, L. F. (Org) et al. Princípios FAIR aplicados à gestão de dados de pesquisa. [S. l.]: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2021. Disponível em: https://ridi.ibict.br/handle/123456789/1182. Acesso em: 17 ago. 2025.

SCHEID, Eder John. An intent-based blockchain-agnostic interaction environment. 2022. 229 p. Tese (Doutorado) — University of Zurich, Zurich, 2022. DOI: https://doi.org/10.5167/uzh-220698. Disponível em: https://www.zora.uzh.ch/id/eprint/220698/. Acesso em: 31 ago. 2026.

SILVA, D.; SOUZA, R.; ALMEIDA, M. Ontologias e vocabulários controlados: comparação de metodologias para construção. [s. l.], v. Ci. Inf., Brasília, v. 37, n. 3, p. 60–75, 2008.

SILVA, Patrícia Nascimento; ZÁRATE, Daniele Claudia Matta Fagundes. Desafios na implementação do Diploma Digital: um relato de experiência na UFMG. In: WORKSHOP DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO DAS INSTITUIÇÕES FEDERAIS DE ENSINO SUPERIOR DO Brasil (WTICIFES), 14., 2022, Aracaju. *Anais [...]*. Aracaju: [s. n.], 2022. p. 48-52. Disponível em: https://hdl.handle.net/1843/61842. Acesso em: 8 ago. 2026.

SOUSA, J. C. da S. A implantação do Diploma Digital na Universidade Federal do Piauí. [s. l.], 2023. Disponível em: https://profiap.org.br/wp-content/uploads/2025/02/JEAN-CARLOS-DA-SILVA-SOUSA.pdf.

SOUZA, Renan Teixeira de. OntoDiplomaDigital proof of concept: versão 1.0. GitHub, 2026. Disponível em: https://github.com/renanteixeirasouza/ontodiplomadigital-proof-of-concept/releases/tag/v1.0. Acesso em: 2 set. 2026.

STRACK, H et al. Progress on Digitization of Higher Education Processes towards Standards EU &DE. ENIS, [s. l.], 2021. Disponível em: https://easychair.org/publications/paper/1lxg/open.

Trajetória da transformação digital. [S. l.], [s. d.]. Disponível em: https://www.gov.br/governodigital/pt-br/estrategias-e-governanca-digital/transformacao-digital/trajetoria-da-transformacao-digital/trajetoria-da-transformacao-digital. Acesso em: 14 ago. 2025.

TRISTÃO, Ana Maria Delazari; FACHIN, Gleisy Regina Bóries; ALARCON, Orestes Estevam. Sistema de classificação facetada e tesauros: instrumentos para organização do conhecimento. Ciência da Informação, [s. l.], v. 33, n. 2, p. 161–171, 2004.

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO (UFES). Novos diplomas de graduação da Ufes serão em formato digital. 2023. Reprodução disponível no portal da Associação das Autoridades de Registro do Brasil (AARB). Disponível em: https://aarb.org.br/novos-diplomas-de-graduacao-da-ufes-serao-em-formato-digital/. Acesso em: 11 ago. 2026.

VANDENBUSSCHE, P.-Y. et al. Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web. Semantic Web, [s. l.], v. 8, n. 3, p. 437–452, 2016.

VEIGA, V. S. D. O. et al. Plano de gestão de dados FAIR: uma proposta para a Fiocruz | Fair data management plan: a proposal for Fiocruz. Liinc em Revista, [s. l.], v. 15, n. 2, 2019. Disponível em: http://revista.ibict.br/liinc/article/view/5030. Acesso em: 26 jul. 2025.

W3C. Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C Recommendation, 26 nov. 2008. Disponível em: https://www.w3.org/TR/xml/. Acesso em: 25 nov. 2025.

W3C. SPARQL 1.1 Query Language. W3C Recommendation, 21 mar. 2013. Disponível em: https://www.w3.org/TR/sparql11-query/. Acesso em: 18 jul. 2025.

W3C. W3C XML Schema Definition Language (XSD) 1.1 Part 1: Structures. W3C Recommendation, 5 abr. 2012. Disponível em: https://www.w3.org/TR/xmlschema11-1/. Acesso em: 25 nov. 2025.

W3C. Web Semântica. [S. l.], [s. d.]. Disponível em: https://www.w3c.br/padroes/web-semantica/. Acesso em: 29 ago. 2025.

WILKINSON, Mark D. et al. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, [s. l.], v. 3, n. 1, p. 160018, 2016.

WOODLEY, M. S. Metadata matters: connecting people and information. In: BACA, M. (org.). Introduction to metadata. 3. ed. Los Angeles, CA: Getty Publications, 2016. Disponível em: https://www.getty.edu/publications/intrometadata/metadata-matters/.

ZENG, M. L. Knowledge organization systems (KOS). KO, [s. l.], v. 35, n. 2–3, p. 160–182, 2008.

ZENG, Marcia Lei; QIN, Jian. Metadata. 3. ed. Chicago: Neal-Schuman, 2022. Disponível em: https://metadataetc.org/book-website3rd/index.html.

Apêndice A – Namespaces e prefixos utilizados

Este apêndice consolida os namespaces e os prefixos empregados na dissertação e nos artefatos da OntoDiplomaDigital disponibilizados no repositório público. A relação apoia a leitura dos termos apresentados nos quadros, nos exemplos RDF/Turtle e nas consultas SPARQL, especialmente no capítulo 5, e permite identificar os vocabulários e as bases de recursos a que esses termos pertencem.

O prefixo funciona como uma abreviação do *namespace*. Em odd:DiplomaDigital, por exemplo, odd: corresponde a https://w3id.org/ufes/ontodiplomadigital#; a combinação com DiplomaDigital identifica a classe. Os prefixos odd:, ont: e vocab: apontam para essa mesma base nos arquivos examinados, enquanto rec: identifica os recursos individuais da ABox pública fictícia. Essa convenção de escrita não cria *namespaces* distintos para classes e conceitos controlados.

Quadro A-1 – Namespaces, prefixos e funções nos materiais da pesquisa

Prefixo	Namespace	Função e uso
odd:, ont:, vocab:	https://w3id.org/ufes/ontodiplomadigital#	Namespace da OntoDiplomaDigital. odd: e ont: identificam termos da ontologia; vocab: destaca os conceitos controlados. São aliases da mesma base.
rec:	https://w3id.org/ufes/ontodiplomadigital/instancias-ficticias#	Recursos individuais da ABox pública fictícia, como diplomas e diplomados.
rdf:	http://www.w3.org/1999/02/22-rdf-syntax-ns#	Vocabulário RDF; identifica, entre outros termos, rdf:type.
rdfs:	http://www.w3.org/2000/01/rdf-schema#	Classes, hierarquias, domínios, alcances, rótulos e comentários.
owl:	http://www.w3.org/2002/07/owl#	Classes e propriedades OWL, restrições e axiomas da ontologia.
xsd:	http://www.w3.org/2001/XMLSchema#	Tipos de dados de literais RDF, como xsd:string, xsd:date e xsd:boolean.
skos:	http://www.w3.org/2004/02/skos/core#	Conceitos controlados, definições, rótulos e relações de organização e mapeamento conceitual.
dcterms:	http://purl.org/dc/terms/	Metadados descritivos, autoria, fontes, datas, licença e identificação.
foaf:	http://xmlns.com/foaf/0.1/	Vocabulário externo para a representação de pessoas e agentes.
prov:	http://www.w3.org/ns/prov#	Vocabulário de proveniência para entidades, atividades, agentes e suas relações.
schema:	https://schema.org/	Vocabulário externo utilizado em alinhamentos de classes e propriedades, incluindo instituições, locais e nomes.

Prefixo	Namespace	Função e uso
vann:	http://purl.org/vocab/vann/	Metadados do vocabulário, incluindo o prefixo preferencial e a URI de seu namespace.
dcat:	http://www.w3.org/ns/dcat#	Vocabulário de descrição de catálogos e recursos de dados reutilizado nos alinhamentos da ontologia.
locn:	http://www.w3.org/ns/locn#	Vocabulário de localização e endereço reutilizado na representação espacial.
vcard:	http://www.w3.org/2006/vcard/ns#	Vocabulário de informações de contato e endereço utilizado nos alinhamentos.
dbo:	http://dbpedia.org/ontology/	Classes e propriedades da ontologia DBpedia, utilizadas em alinhamentos e na consulta federada demonstrativa.
dbr:	http://dbpedia.org/resource/	Recursos individuais da DBpedia, como dbr:Brazil, na consulta federada demonstrativa.
wdt:	http://www.wikidata.org/prop/direct/	Propriedades diretas do Wikidata empregadas na consulta demonstrativa via QLever.
xs:	https://www.w3.org/2001/XMLSchema	Declaração encontrada nos XSD publicados. Refere-se à camada de esquemas XML; observar a distinção indicada após o quadro.
ds:	https://www.w3.org/2000/09/xmldsig#	Declaração de assinatura XML encontrada nos XSD publicados; observar a distinção indicada após o quadro.
(sem prefixo)	https://portal.mec.gov.br/diplomadigital/arquivos-em-xsd	Namespace padrão e de destino nos schemas de domínio publicados no repositório.
(sem prefixo)	https://www.w3.org/2001/XMLSchema	Namespace padrão declarado no arquivo xmldsig-core-schema_v1-1.xsd publicado.

Fonte: Elaborado pelo autor (2026), a partir da dissertação, das declarações da TBox e da ABox públicas, das consultas SPARQL e dos schemas XSD disponibilizados no repositório ontodiplomadigital-proof-of-concept.

A distinção entre xsd: e xs: deve ser preservada. O primeiro identifica tipos de dados em RDF e utiliza a base terminada em #; o segundo aparece na escrita dos esquemas XML. As cópias XSD examinadas declaram xs: e ds: com https, conforme registrado no Quadro A-1. As especificações W3C identificam os namespaces XML Schema e XML Signature, respectivamente, como http://www.w3.org/2001/XMLSchema e http://www.w3.org/2000/09/xmldsig#. As formas com http e https são identificadores distintos; sua ocorrência no quadro documenta os arquivos consultados e não estabelece equivalência entre elas.

Os prefixos dbr: e wdt: são empregados nas consultas federadas demonstrativas. As orientações de acesso aos arquivos e de reprodução da prova de conceito constam do Apêndice B.

Apêndice B – Consultas SPARQL e resultados da prova de conceito

Este conjunto apresenta os códigos completos de onze consultas SPARQL. As consultas 01 a 07 correspondem às sete questões de competência definidas na especificação da OntoDiplomaDigital e verificam a recuperação das informações previstas nesses requisitos. As consultas 08 a 11 são demonstrações complementares. A consulta 08 reúne informações de um diploma em uma resposta tabular, enquanto a consulta 09 constrói um grafo de suas relações. A consulta 10 combina dados locais com informações sobre o recurso Brasil, selecionado diretamente na DBpedia. A consulta 11 utiliza o código IBGE como identificador compartilhado para relacionar a naturalidade dos diplomados ao município e ao país descritos no Wikidata, consultado por meio do QLever. Essas demonstrações ampliam a exploração das relações representadas, sem constituir novas questões de competência. Cada consulta é acompanhada de uma explicação de seu funcionamento, dos resultados disponíveis e de sua interpretação. Todos os dados acadêmicos locais são sintéticos.

Os resultados das consultas 01 a 08 foram conferidos por execução local sobre a TBox e a ABox públicas, sem inferência adicional, e confrontados com os relatórios do projeto. Os resultados das consultas 09 a 11 são apresentados a partir dos registros de verificação disponíveis. Nos quadros com mais de dez linhas de dados, são exibidas as cinco primeiras e as cinco últimas, com indicação da quantidade omitida. Nas respostas tabulares, mantém-se a ordenação estabelecida pela consulta, quando definida; na apresentação vertical de uma única resposta, mantém-se a ordem das variáveis selecionadas. O relatório da consulta 09 informa 18 entradas de resultado, mas disponibiliza somente uma tripla como amostra. O relatório da consulta 10 registra três linhas retornadas, sem apresentar seus valores completos. Essas limitações são explicitadas nos respectivos quadros. As datas futuras presentes na base são valores fictícios e não correspondem às datas de execução das consultas.

Para abreviar os identificadores nos quadros, inst: representa https://w3id.org/ufes/ontodiplomadigital/instancias-ficticias#; odd: representa https://w3id.org/ufes/ontodiplomadigital#; dbr: e dbo: identificam, respectivamente, recursos e classes da DBpedia; wd: identifica recursos do Wikidata. Os códigos mantêm os prefixos declarados nos arquivos originais.

B.1 QC1 — Instituições emissora e registradora

A consulta identifica as instituições emissora e registradora do diploma selecionado pelo identificador sintético “urn:synthetic:diploma:001”.

Arquivo da consulta: 01_qc1_instituicoes_emissao_registro.rq

Quadro B-1 – Código completo da consulta 01

```
PREFIX odd: <https://w3id.org/ufes/ontodiplomadigital#>
SELECT ?diploma ?identificadorDiploma
       ?papelInstitucional ?ies ?nomeIes
WHERE {
  VALUES ?identificadorDiploma {
    "urn:synthetic:diploma:001"
  }
  ?diploma a odd:DiplomaDigital ;
           odd:identifier ?identificadorDiploma .
  {
    ?diploma odd:issuedBy ?ies .
    BIND("emissão"@pt AS ?papelInstitucional)
  }
  UNION
  {
    ?diploma odd:registeredBy ?ies .
    BIND("registro"@pt AS ?papelInstitucional)
  }
  ?ies odd:fullName ?nomeIes .
}
ORDER BY ?papelInstitucional ?nomeIes
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

VALUES define o identificador do diploma. As propriedades issuedBy e registeredBy localizam as instituições; UNION reúne as duas buscas e BIND identifica seus papéis. SELECT define as colunas e ORDER BY organiza a resposta.

Duas linhas recuperadas; apresentam-se as colunas de papel institucional e nome da instituição.

Quadro B-2 – Resultado da consulta 01

Papel institucional	Nome da instituição
emissão	Universidade do Vale do Conhecimento
registro	Universidade Nacional da Curadoria

Fonte: Elaborado pelo autor, a partir da execução local da consulta sobre os artefatos públicos da prova de conceito, conferida com o relatório validacao_local_consultas_publicas.json.

As duas linhas distinguem os papéis de emissão e registro do mesmo diploma. Os nomes são fictícios e não comprovam atos institucionais reais.

B.2 QC2 — Registro, assinatura e validação

A consulta recupera o registro, a assinatura, o carimbo de tempo e o código de validação do diploma sintético “urn:synthetic:diploma:001”.

Arquivo da consulta: 02_qc2_autenticidade_assinatura_registro_validacao.rq

Quadro B-3 – Código completo da consulta 02

```
PREFIX odd:  <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT ?diploma ?identificadorDiploma
       ?registro ?dataRegistro
       ?assinatura ?dataHoraAssinatura
       ?carimboTempo ?dataHoraCarimbo
       ?formatoAssinaturaLabel ?papelAssinaturaLabel
       ?codigoValidacao ?valorCodigoValidacao
WHERE {
  VALUES ?identificadorDiploma {
    "urn:synthetic:diploma:001"
  }

  ?diploma a odd:DiplomaDigital ;
           odd:identifier ?identificadorDiploma ;
           odd:hasRegistration ?registro ;
           odd:hasSignature ?assinatura ;
           odd:hasValidationCode ?codigoValidacao .

  ?registro odd:registrationDate ?dataRegistro .

  ?assinatura odd:signatureDateTime ?dataHoraAssinatura ;
              odd:hasTimestamp ?carimboTempo ;
              odd:hasSignatureFormat ?formatoAssinatura ;
              odd:hasSignatureRole ?papelAssinatura .

  ?carimboTempo odd:timestampDateTime ?dataHoraCarimbo .
  ?codigoValidacao odd:validationCodeValue ?valorCodigoValidacao .

  ?formatoAssinatura skos:prefLabel ?formatoAssinaturaLabel .
  ?papelAssinatura skos:prefLabel ?papelAssinaturaLabel .

  FILTER(LANG(?formatoAssinaturaLabel) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?papelAssinaturaLabel) IN ("pt", "pt-BR", ""))
}
ORDER BY ?assinatura
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

As propriedades hasRegistration, hasSignature e hasValidationCode iniciam os caminhos de recuperação. hasTimestamp liga a assinatura ao carimbo de tempo. skos:prefLabel recupera os rótulos dos conceitos controlados, e FILTER seleciona os idiomas admitidos.

Foi recuperada uma linha. Seus valores são apresentados verticalmente, com uma variável por linha do quadro. Das 12 variáveis selecionadas, apresentam-se as cinco primeiras e as cinco últimas; 2 variáveis intermediárias foram omitidas do quadro, permanecendo explicitadas no código.

Quadro B-4 – Resultado da consulta 02

Variável selecionada	Valor recuperado
?diploma	inst:diplomaCaso1
?identificadorDiploma	urn:synthetic:diploma:001
?registro	inst:registroCaso1
?dataRegistro	2027-01-15
?assinatura	inst:assinaturaCaso1
?dataHoraCarimbo	2027-01-15T14:10:00
?formatoAssinaturaLabel	XAdES
?papelAssinaturaLabel	assinatura PJ da registradora
?codigoValidacao	inst:codigoValidacaoPublico0757
?valorCodigoValidacao	CODIGO-VALIDACAO-FICTICIO-0002

Fonte: Elaborado pelo autor, a partir da execução local da consulta sobre os artefatos públicos da prova de conceito, conferida com o relatório validacao_local_consultas_publicas.json.

A resposta recupera os elementos documentais vinculados ao diploma. Sua presença no grafo não constitui verificação criptográfica da assinatura nem consulta externa ao código de validação.

B.3 QC3 — Diplomas por instituição, período e modalidade

A consulta seleciona diplomas emitidos pela Universidade do Vale do Conhecimento em 2027, nas modalidades Presencial ou Educação a Distância.

Arquivo da consulta: 03_qc3_diplomas_ies_periodo_modalidade.rq

Quadro B-5 – Código completo da consulta 03

```
PREFIX odd:  <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX xsd:  <http://www.w3.org/2001/XMLSchema#>

SELECT ?diploma ?dataEmissao ?ies ?nomeIes ?curso ?nomeCurso
       ?modalidade ?modalidadeLabel
WHERE {
  VALUES ?nomeIes {
    "Universidade do Vale do Conhecimento"
  }
  VALUES ?modalidadeLabel {
    "Presencial"@pt
    "Educação a Distância"@pt
  }

  ?diploma a odd:DiplomaDigital ;
           odd:issuedBy ?ies ;
           odd:issueDate ?dataEmissao ;
           odd:hasProgram ?curso .

  ?ies odd:fullName ?nomeIes .
  ?curso odd:courseName ?nomeCurso ;
         odd:hasModeOfDelivery ?modalidade .
  ?modalidade skos:prefLabel ?modalidadeLabel .

  FILTER(?dataEmissao >= "2027-01-01"^^xsd:date &&
         ?dataEmissao <= "2027-12-31"^^xsd:date)
}
ORDER BY ?dataEmissao ?modalidadeLabel ?nomeCurso
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

VALUES define a instituição e as modalidades aceitas. issuedBy, issueDate e hasProgram relacionam o diploma aos critérios da seleção. FILTER delimita o período de emissão e ORDER BY organiza os resultados.

Foi recuperada uma linha. Seus valores são apresentados verticalmente, com uma variável por linha do quadro.

Quadro B-6 – Resultado da consulta 03

Variável selecionada	Valor recuperado
?diploma	inst:diplomaCaso1
?dataEmissao	2027-01-10
?ies	inst:iesEmissoraCaso1
?nomeIes	Universidade do Vale do Conhecimento
?curso	inst:cursoCaso1
?nomeCurso	Bacharelado em Gestao do Conhecimento Aplicado
?modalidade	odd:Presencial
?modalidadeLabel	Presencial

Fonte: Elaborado pelo autor, a partir da execução local da consulta sobre os artefatos públicos da prova de conceito, conferida com o relatório validacao_local_consultas_publicas.json.

O retorno de um diploma mostra a aplicação conjunta dos critérios de instituição, período e modalidade no conjunto sintético. A ausência de outras linhas não caracteriza a oferta real de cursos da instituição.

B.4 QC4 — Titular e formação acadêmica

A consulta reúne o titular, o curso, a data de conclusão, o grau e o título associados ao diploma sintético “urn:synthetic:diploma:001”.

Arquivo da consulta: 04_qc4_diplomado_curso_concluido.rq

Quadro B-7 – Código completo da consulta 04

```
PREFIX odd:  <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT ?diploma ?identificadorDiploma ?diplomado ?nomeDiplomado ?cpf
       ?curso ?nomeCurso ?dataConclusao
       ?grau ?grauLabel ?titulo ?tituloLabel
WHERE {
  VALUES ?identificadorDiploma {
    "urn:synthetic:diploma:001"
  }

  ?diploma a odd:DiplomaDigital ;
           odd:identifier ?identificadorDiploma ;
           odd:hasHolder ?diplomado ;
           odd:hasProgram ?curso .

  ?diplomado odd:fullName ?nomeDiplomado ;
             odd:cpf ?cpf .

  OPTIONAL { ?diplomado odd:completedProgram ?curso . }

  ?curso odd:courseName ?nomeCurso ;
         odd:awardsDegree ?grau ;
         odd:awardsTitle ?titulo .

  ?grau skos:prefLabel ?grauLabel .
  ?titulo skos:prefLabel ?tituloLabel .

  FILTER(LANG(?grauLabel) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?tituloLabel) IN ("pt", "pt-BR", ""))

  OPTIONAL { ?curso odd:conclusionDate ?dataConclusao . }
}
ORDER BY ?nomeDiplomado ?nomeCurso
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

`hasHolder` e `hasProgram` ligam o mesmo diploma ao titular e ao curso; `awardsDegree` e `awardsTitle` recuperam o grau e o título. `OPTIONAL` permite consultar completedProgram e conclusionDate sem exigir essas relações para obter a resposta.

Foi recuperada uma linha. Seus valores são apresentados verticalmente, com uma variável por linha do quadro. Das 12 variáveis selecionadas, apresentam-se as

cinco primeiras e as cinco últimas; 2 variáveis intermediárias foram omitidas do quadro, permanecendo explicitadas no código.

Quadro B-8 – Resultado da consulta 04

Variável selecionada	Valor recuperado
?diploma	inst:diplomaCaso1
?identificadorDiploma	urn:synthetic:diploma:001
?diplomado	inst:diplomadoCaso1
?nomeDiplomado	Pessoa Fictícia-0001
?cpf	CPF-FICTICIO-0001
?dataConclusao	2026-12-18
?grau	odd:Bacharelado
?grauLabel	Bacharelado
?titulo	odd:Bacharel
?tituloLabel	Bacharel

Fonte: Elaborado pelo autor, a partir da execução local da consulta sobre os artefatos públicos da prova de conceito, conferida com o relatório validacao_local_consultas_publicas.json.

A resposta relaciona o titular à formação representada por meio do diploma, incluindo o grau e o título controlados. Os valores recuperados pertencem exclusivamente à base fictícia.

B.5 QC5 — Valores controlados e quantidades de uso

A consulta identifica os conceitos controlados efetivamente utilizados para qualificar os recursos e contabiliza a quantidade de recursos distintos associados a cada conceito.

Arquivo da consulta: 05_qc5_valores_controlados.rq

Quadro B-9 – Código completo da consulta 05

```
PREFIX odd:  <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT ?categoria ?valorControlado ?rotulo
       (COUNT(DISTINCT ?recursoQualificado) AS ?qtdUsos)
WHERE {
  {
    ?recursoQualificado odd:awardsDegree ?valorControlado .
    BIND("grau conferido"@pt AS ?categoria)
  }
  UNION
  {
    ?recursoQualificado odd:awardsTitle ?valorControlado .
    BIND("título conferido"@pt AS ?categoria)
  }
  UNION
  {
    ?recursoQualificado odd:hasModeOfDelivery ?valorControlado .
    BIND("modalidade de oferta"@pt AS ?categoria)
  }
  UNION
  {
    ?recursoQualificado odd:hasStatus ?valorControlado .
    BIND("situação do diploma"@pt AS ?categoria)
  }
  UNION
  {
    ?recursoQualificado odd:hasSignatureFormat ?valorControlado .
    BIND("formato de assinatura"@pt AS ?categoria)
  }
  UNION
  {
    ?recursoQualificado odd:hasSignatureRole ?valorControlado .
    BIND("papel da assinatura"@pt AS ?categoria)
  }

  ?valorControlado a skos:Concept ;
                   skos:prefLabel ?rotulo .
  FILTER(LANG(?rotulo) IN ("pt", "pt-BR", ""))
}
GROUP BY ?categoria ?valorControlado ?rotulo
ORDER BY ?categoria ?rotulo
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

UNION reúne seis categorias de valores controlados. BIND nomeia cada categoria e skos:prefLabel recupera o rótulo do conceito. COUNT(DISTINCT ?recursoQualificado) conta cada recurso uma única vez por grupo, definido por GROUP BY.

Foram recuperadas vinte linhas. O quadro apresenta as cinco primeiras e as cinco últimas, na ordem definida pela consulta; as dez linhas intermediárias foram omitidas.

Quadro B-10 – Resultado da consulta 05

Categoria	Conceito	Rótulo	Usos
formato de assinatura	odd:CAdES	CAdES	83
formato de assinatura	odd:PAdES	PAdES	84
formato de assinatura	odd:XAdES	XAdES	85
grau conferido	odd:Bacharelado	Bacharelado	85
grau conferido	odd:Licenciatura	Licenciatura	83
situação do diploma	odd:Ativo	Ativo	151
situação do diploma	odd:Revogado	Revogado	51
título conferido	odd:Bacharel	Bacharel	85
título conferido	odd:Licenciado	Licenciado	83
título conferido	odd:Tecnologo	Tecnólogo	84

Fonte: Elaborado pelo autor, a partir da execução local da consulta sobre os artefatos públicos da prova de conceito.

As vinte linhas mostram os conceitos usados na base e suas quantidades de associação. A contagem refere-se a recursos distintos em cada categoria, não necessariamente a diplomados ou diplomas. Um mesmo recurso pode participar de categorias diferentes.

B.6 QC6 — Proveniência administrativa do registro

A consulta recupera a proveniência administrativa dos registros, incluindo instituição registradora, data, processo, responsável e assinantes. Não restringe a busca a um único diploma.

Arquivo da consulta: 06_qc6_proveniencia_registro.rq

Quadro B-11 – Código completo da consulta 06

```
PREFIX odd: <https://w3id.org/ufes/ontodiplomadigital#>

SELECT ?diploma
       ?iesRegistradora ?nomeIesRegistradora
       ?registro ?dataRegistro
       ?processoAdministrativo ?numeroProcesso
       ?responsavel ?nomeResponsavel ?cpfResponsavel
       ?assinanteRegistro ?cpfAssinanteRegistro ?cargoAssinanteRegistro
WHERE {
  ?diploma a odd:DiplomaDigital ;
           odd:registeredBy ?iesRegistradora ;
           odd:hasRegistration ?registro .

  ?iesRegistradora odd:fullName ?nomeIesRegistradora .

  ?registro odd:registrationDate ?dataRegistro ;
            odd:hasAdministrativeProcess ?processoAdministrativo ;
            odd:hasRegistrationResponsible ?responsavel .

  ?processoAdministrativo odd:identifier ?numeroProcesso .

  ?responsavel odd:responsibleName ?nomeResponsavel ;
               odd:responsibleCpf ?cpfResponsavel .

  OPTIONAL {
    ?registro odd:hasRegistrationSigner ?assinanteRegistro .
    ?assinanteRegistro odd:registrationSignerCpf ?cpfAssinanteRegistro .
    OPTIONAL {
      ?assinanteRegistro odd:registrationSignerRole ?cargoAssinanteRegistro .
    }
  }
}
ORDER BY ?diploma ?assinanteRegistro
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

O registeredBy identifica a instituição registradora e hasRegistration conduz ao registro. As relações com o processo administrativo e o responsável recuperam sua proveniência. OPTIONAL acrescenta os assinantes e seus cargos, quando representados.

Foram recuperadas 303 linhas. O quadro apresenta as cinco primeiras e as cinco últimas, na ordem definida pela consulta; as 293 linhas intermediárias foram

omitidas. Para facilitar a leitura, apresentam-se as colunas de diploma, instituição registradora, processo, responsável, CPF do assinante e cargo. As demais variáveis permanecem identificadas no código completo.

Quadro B-12 – Resultado da consulta 06 — recorte identificado

Diploma	Instituição registradora	Processo	Responsável	CPF do assinante	Cargo do assinante
inst:diploma Caso1	Universidade Nacional da Curadoria	PROC-SYNTHETIC-2027-0001	Pessoa Fictícia-0002	CPF-FICTICIO-0003	Reitor
inst:diploma Caso1	Universidade Nacional da Curadoria	PROC-SYNTHETIC-2027-0001	Pessoa Fictícia-0002	CPF-FICTICIO-0004	Secretaria Geral Academica
inst:diploma Caso2	Centro Universitario Rede Cidada	PROC-SYNTHETIC-2026-0202	Pessoa Fictícia-0008	CPF-FICTICIO-0009	Diretoria de Registro Academico
inst:diploma Lote001	Universidade Registradora do Brasil Digital	PROC-SYNTHETIC-2026-0003	Pessoa Fictícia-0013	CPF-FICTICIO-0014	Reitoria de Registro 1
inst:diploma Lote002	Centro de Registro Interoperavel	PROC-SYNTHETIC-2026-0004	Pessoa Fictícia-0018	CPF-FICTICIO-0019	Reitoria de Registro 2
inst:diploma Lote247	Instituto de Seguranca Documental	PROC-SYNTHETIC-2028-0249	Pessoa Fictícia-0040	CPF-FICTICIO-1057	Reitoria de Registro 4
inst:diploma Lote248	Centro Universitario de Rastreabilidade Academica	PROC-SYNTHETIC-2028-0250	Pessoa Fictícia-0043	CPF-FICTICIO-1061	Reitoria de Registro 5
inst:diploma Lote249	Universidade Registradora do Brasil Digital	PROC-SYNTHETIC-2028-0251	Pessoa Fictícia-0046	CPF-FICTICIO-1065	Reitoria de Registro 6
inst:diploma Lote250	Centro de Registro Interoperavel	PROC-SYNTHETIC-2028-0252	Pessoa Fictícia-0049	CPF-FICTICIO-1069	Reitoria de Registro 7
inst:diploma Lote250	Centro de Registro Interoperavel	PROC-SYNTHETIC-2028-0252	Pessoa Fictícia-0049	CPF-FICTICIO-1070	Secretaria Academica Integrada

Fonte: Elaborado pelo autor, a partir da execução local da consulta sobre os artefatos públicos da prova de conceito, conferida com o relatório validacao_local_consultas_publicas.json.

As duas primeiras respostas correspondem ao mesmo diploma, com assinantes distintos. As 303 linhas representam combinações de informações, sem comprovar processos administrativos reais.

B.7 QC7 — Plataforma de acesso por CPF fictício

A consulta busca o diploma associado ao “CPF-FICTICIO-0012” e a plataforma representada como meio de acesso no conjunto sintético.

Arquivo da consulta: 07_qc7_acesso_por_cpf_plataforma.rq.

Quadro B-13 – Código completo da consulta 07

```
PREFIX odd: <https://w3id.org/ufes/ontodiplomadigital#>

SELECT ?cpf ?diplomado ?diploma
       ?plataforma ?nomePlataforma
       ?codigoValidacao ?valorCodigoValidacao
WHERE {
  VALUES ?cpf {
    "CPF-FICTICIO-0012"
  }

  ?diplomado a odd:Diplomado ;
             odd:cpf ?cpf .

  ?diploma a odd:DiplomaDigital ;
           odd:hasHolder ?diplomado ;
           odd:isAccessibleVia ?plataforma .

  ?plataforma a odd:PlataformaGoverno ;
              odd:fullName ?nomePlataforma .

  FILTER(CONTAINS(LCASE(STR(?nomePlataforma)), "gov"))

  OPTIONAL {
    ?diploma odd:hasValidationCode ?codigoValidacao .
    ?codigoValidacao odd:validationCodeValue ?valorCodigoValidacao .
  }
}
ORDER BY ?diploma
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

VALUES define o CPF fictício. hasHolder relaciona o diploma ao diplomado e isAccessibleVia conduz à plataforma. FILTER seleciona nomes que contenham “gov”, sem distinguir maiúsculas de minúsculas. OPTIONAL acrescenta o código de validação, quando disponível.

Foi recuperada uma linha. Seus valores são apresentados verticalmente, com uma variável por linha do quadro.

Quadro B-14 – Resultado da consulta 07

Variável selecionada	Valor recuperado
?cpf	CPF-FICTICIO-0012

Variável selecionada	Valor recuperado
?diplomado	inst:diplomadoLote001
?diploma	inst:diplomaLote001
?plataforma	inst:plataformaGovBrAcademicoFicticio
?nomePlataforma	GovBR Academico Ficticio
?codigoValidacao	inst:codigoValidacaoPublico0759
?valorCodigoValidacao	CODIGO-VALIDACAO-FICTICIO-0006

Fonte: Elaborado pelo autor, a partir da execução local da consulta sobre os artefatos públicos da prova de conceito, conferida com o relatório validacao_local_consultas_publicas.json.

A resposta evidencia a relação de acesso representada entre o diploma e uma plataforma fictícia. A consulta não acessa o Gov.br nem comprova integração operacional com serviços governamentais.

B.8 Consulta 08 — Recuperação integrada por CPF fictício

A consulta complementar reúne, em uma resposta, informações acadêmicas, institucionais, administrativas, de assinatura e de acesso do diploma associado ao “CPF-FICTICIO-0012”.

Arquivo da consulta: 08_demo_visao_360_diploma_por_cpf.rq

Quadro B-15 – Código completo da consulta 08

```
PREFIX odd:     <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX schema:  <https://schema.org/>
PREFIX skos:    <http://www.w3.org/2004/02/skos/core#>

SELECT ?cpf
       ?titular
       ?diploma
       ?identificadorDiploma
       ?cursoConcluido
       ?grau
       ?tituloConferido
       ?modalidade
       ?dataConclusao
       ?dataExpedicao
       ?dataEmissao
       ?iesEmissora
       ?iesRegistradora
       ?dataRegistro
       ?numeroProcesso
       ?responsavelRegistro
       ?situacaoDiploma
       ?formatoAssinatura
       ?dataHoraAssinatura
       ?dataHoraCarimbo
       ?codigoValidacao
       ?plataformaAcesso
       ?evidenciaDisponivel
WHERE {
  VALUES ?cpf {
    "CPF-FICTICIO-0012"
  }

  ?diplomado a odd:Diplomado ;
             odd:cpf ?cpf ;
             odd:fullName ?titular .

  ?diploma a odd:DiplomaDigital ;
           odd:identifier ?identificadorDiploma ;
           odd:hasHolder ?diplomado ;
           odd:hasProgram ?curso ;
           odd:issuedBy ?recursoIesEmissora ;
           odd:registeredBy ?recursoIesRegistradora ;
           odd:hasRegistration ?registro ;
           odd:hasStatus ?recursoSituacao ;
```

```
          odd:hasSignature ?assinatura ;
          odd:hasValidationCode ?recursoCodigo ;
          odd:isAccessibleVia ?plataforma ;
          odd:issueDate ?dataEmissao .

  ?curso odd:courseName ?cursoConcluido ;
         odd:awardsDegree ?recursoGrau ;
         odd:awardsTitle ?recursoTitulo ;
         odd:hasModeOfDelivery ?recursoModalidade ;
         odd:conclusionDate ?dataConclusao .

  ?recursoGrau skos:prefLabel ?grau .
  ?recursoTitulo skos:prefLabel ?tituloConferido .
  ?recursoModalidade skos:prefLabel ?modalidade .
  ?recursoSituacao skos:prefLabel ?situacaoDiploma .

  ?recursoIesEmissora odd:fullName ?iesEmissora .
  ?recursoIesRegistradora odd:fullName ?iesRegistradora .

  ?registro odd:diplomaExpeditionDate ?dataExpedicao ;
            odd:registrationDate ?dataRegistro ;
            odd:hasAdministrativeProcess ?processo ;
            odd:hasRegistrationResponsible ?responsavel .
  ?processo odd:identifier ?numeroProcesso .
  ?responsavel odd:responsibleName ?responsavelRegistro .

  ?assinatura odd:hasSignatureFormat ?recursoFormato ;
              odd:signatureDateTime ?dataHoraAssinatura ;
              odd:hasTimestamp ?carimbo .
  ?recursoFormato skos:prefLabel ?formatoAssinatura .
  ?carimbo odd:timestampDateTime ?dataHoraCarimbo .

  ?recursoCodigo odd:validationCodeValue ?codigoValidacao .
  ?plataforma odd:fullName ?plataformaAcesso .

  FILTER(LANG(?grau) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?tituloConferido) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?modalidade) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?situacaoDiploma) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?formatoAssinatura) IN ("pt", "pt-BR", ""))

  BIND(
    "Metadados de registro, assinatura e validação recuperados no grafo sintético"@pt
    AS ?evidenciaDisponivel
  )
}
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

O CPF definido em VALUES localiza o titular e seu diploma. As relações do diploma conduzem ao curso, às instituições, ao registro, à assinatura e à plataforma. Os rótulos SKOS tornam legíveis os valores controlados; BIND acrescenta uma descrição das evidências disponíveis.

Foi recuperada uma linha. Seus valores são apresentados verticalmente, com uma variável por linha do quadro. Das 23 variáveis selecionadas, apresentam-se as cinco primeiras e as cinco últimas; 13 variáveis intermediárias foram omitidas do quadro, permanecendo explicitadas no código.

Quadro B-16 – Resultado da consulta 08

Variável selecionada	Valor recuperado
?cpf	CPF-FICTICIO-0012
?titular	Pessoa Fictícia-0012
?diploma	inst:diplomaLote001
?identificadorDiploma	urn:synthetic:diploma:003
?cursoConcluido	Bacharelado em Ciencia de Dados Publicos
?dataHoraAssinatura	2026-02-19T09:11:00
?dataHoraCarimbo	2026-02-19T09:11:00
?codigoValidacao	CODIGO-VALIDACAO-FICTICIO-0006
?plataformaAcesso	GovBR Academico Ficticio
?evidenciaDisponivel	Metadados de registro, assinatura e validação recuperados no grafo sintético

Fonte: Elaborado pelo autor, a partir da execução local da consulta sobre os artefatos públicos da prova de conceito, conferida com o relatório validacao_local_consultas_publicas.json.

A resposta reúne informações distribuídas entre os recursos do grafo em uma visão integrada. A evidência descrita no código refere-se à recuperação de metadados sintéticos, sem validação criptográfica ou implantação institucional.

B.9 Consulta 09 — Construção de um grafo do diploma

A consulta complementar constrói um grafo de relações e rótulos do diploma associado ao “CPF-FICTICIO-0012”. Seu código é apresentado integralmente; o resultado é documentado em quadro, sem figura.

Arquivo da consulta: 09_demo_grafo_diploma_por_cpf.rq

Quadro B-17 – Código completo da consulta 09

```
PREFIX odd:  <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

CONSTRUCT {
  ?diploma rdfs:label ?rotuloDiploma ;
           odd:hasHolder ?diplomado ;
           odd:hasProgram ?curso ;
           odd:issuedBy ?iesEmissora ;
           odd:registeredBy ?iesRegistradora ;
           odd:hasRegistration ?registro ;
           odd:hasSignature ?assinatura ;
           odd:hasValidationCode ?codigoValidacao ;
           odd:isAccessibleVia ?plataforma ;
           odd:hasStatus ?situacaoDiploma .

  ?diplomado rdfs:label ?titular .

  ?curso rdfs:label ?cursoConcluido ;
         odd:awardsDegree ?grau ;
         odd:awardsTitle ?titulo ;
         odd:hasModeOfDelivery ?modalidade .

  ?iesEmissora rdfs:label ?nomeIesEmissora .
  ?iesRegistradora rdfs:label ?nomeIesRegistradora .
  ?plataforma rdfs:label ?nomePlataforma .
  ?codigoValidacao rdfs:label ?valorCodigoValidacao .

  ?situacaoDiploma rdfs:label ?labelSituacao .
  ?grau rdfs:label ?labelGrau .
  ?titulo rdfs:label ?labelTitulo .
  ?modalidade rdfs:label ?labelModalidade .

  ?registro rdfs:label ?rotuloRegistro ;
            odd:hasAdministrativeProcess ?processo ;
            odd:hasRegistrationResponsible ?responsavelRegistro ;
            odd:hasRegistrationSigner ?assinanteRegistro .

  ?processo rdfs:label ?numeroProcesso .
  ?responsavelRegistro rdfs:label ?nomeResponsavelRegistro .
  ?assinanteRegistro rdfs:label ?nomeAssinanteRegistro .

  ?assinatura rdfs:label ?rotuloAssinatura ;
              odd:hasSignatureFormat ?formatoAssinatura ;
              odd:hasSignatureRole ?papelAssinatura ;
```

```
            odd:hasTimestamp ?carimboTempo .

  ?formatoAssinatura rdfs:label ?labelFormatoAssinatura .
  ?papelAssinatura rdfs:label ?labelPapelAssinatura .
  ?carimboTempo rdfs:label ?rotuloCarimboTempo .
}
WHERE {
  VALUES ?cpf {
    "CPF-FICTICIO-0012"
  }

  ?diplomado a odd:Diplomado ;
             odd:cpf ?cpf ;
             odd:fullName ?titular .

  ?diploma a odd:DiplomaDigital ;
           odd:identifier ?identificadorDiploma ;
           odd:hasHolder ?diplomado ;
           odd:hasProgram ?curso ;
           odd:issuedBy ?iesEmissora ;
           odd:registeredBy ?iesRegistradora ;
           odd:hasRegistration ?registro ;
           odd:hasSignature ?assinatura ;
           odd:hasValidationCode ?codigoValidacao ;
           odd:isAccessibleVia ?plataforma ;
           odd:hasStatus ?situacaoDiploma .

  ?curso odd:courseName ?cursoConcluido ;
         odd:awardsDegree ?grau ;
         odd:awardsTitle ?titulo ;
         odd:hasModeOfDelivery ?modalidade .

  ?iesEmissora odd:fullName ?nomeIesEmissora .
  ?iesRegistradora odd:fullName ?nomeIesRegistradora .
  ?plataforma odd:fullName ?nomePlataforma .
  ?codigoValidacao odd:validationCodeValue ?valorCodigoValidacao .

  ?situacaoDiploma skos:prefLabel ?labelSituacao .
  ?grau skos:prefLabel ?labelGrau .
  ?titulo skos:prefLabel ?labelTitulo .
  ?modalidade skos:prefLabel ?labelModalidade .

  ?registro odd:hasAdministrativeProcess ?processo ;
            odd:hasRegistrationResponsible ?responsavelRegistro .

  OPTIONAL {
    ?registro odd:hasRegistrationSigner ?assinanteRegistro .
    ?assinanteRegistro odd:fullName ?nomeAssinanteRegistro .
  }

  ?processo odd:identifier ?numeroProcesso .
  ?responsavelRegistro odd:responsibleName ?nomeResponsavelRegistro .

  ?assinatura odd:hasSignatureFormat ?formatoAssinatura ;
              odd:hasSignatureRole ?papelAssinatura ;
```

```
              odd:hasTimestamp ?carimboTempo .

  ?formatoAssinatura skos:prefLabel ?labelFormatoAssinatura .
  ?papelAssinatura skos:prefLabel ?labelPapelAssinatura .
  ?carimboTempo odd:timestampDateTime ?dataHoraCarimbo .

  FILTER(LANG(?labelSituacao) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?labelGrau) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?labelTitulo) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?labelModalidade) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?labelFormatoAssinatura) IN ("pt", "pt-BR", ""))
  FILTER(LANG(?labelPapelAssinatura) IN ("pt", "pt-BR", ""))

  BIND("Diploma" AS ?rotuloDiploma)
  BIND("Registro" AS ?rotuloRegistro)
  BIND("Assinatura" AS ?rotuloAssinatura)
  BIND("Carimbo" AS ?rotuloCarimboTempo)
}
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

WHERE encontra os recursos relacionados ao diploma selecionado. CONSTRUCT define as triplas que formarão o grafo de resposta, incluindo relações e rótulos. BIND cria rótulos curtos para alguns recursos; OPTIONAL acrescenta informações de assinantes quando disponíveis.

O relatório de verificação registra 18 entradas para a consulta 09 e preserva uma tripla de exemplo. Apresenta-se essa tripla, sem atribuir ao trecho o caráter de resultado integral.

Quadro B-18 – Resultado da consulta 09 — recorte identificado

Sujeito	Propriedade	Objeto
inst:registroLote001	odd:hasAdministrativeProcess	inst:processoAdministrativo003

Fonte: Elaborado pelo autor, a partir da amostra registrada em validacao_local_consultas_publicas.json.

A tripla exemplificada explicita a ligação entre um registro e seu processo administrativo. O resultado da consulta é um grafo RDF, e não uma tabela de diplomas. O trecho apresentado documenta uma relação registrada, sem reproduzir todas as triplas retornadas.

B.10 Consulta 10 — Federação com a DBpedia

A consulta combina os dados locais do diploma associado ao “CPF-FICTICIO-0012” com o nome e as classificações do recurso Brasil na DBpedia. O país é escolhido diretamente no código.

Arquivo da consulta: 10_demo_consulta_federada_dbpedia_brasil.rq

Quadro B-19 – Código completo da consulta 10

```
PREFIX odd:  <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX dbr:  <http://dbpedia.org/resource/>
PREFIX dbo:  <http://dbpedia.org/ontology/>
PREFIX rdf:  <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?cpf
       ?titular
       ?diploma
       ?cursoConcluido
       ?iesEmissora
       ?paisContexto
       ?paisLabel
       ?tipoExterno
       ?fonteExterna
WHERE {
  VALUES ?cpf {
    "CPF-FICTICIO-0012"
  }

  ?diplomado a odd:Diplomado ;
             odd:cpf ?cpf ;
             odd:fullName ?titular .

  ?diploma a odd:DiplomaDigital ;
           odd:hasHolder ?diplomado ;
           odd:hasProgram ?curso ;
           odd:issuedBy ?recursoIesEmissora .

  ?curso odd:courseName ?cursoConcluido .
  ?recursoIesEmissora odd:fullName ?iesEmissora .

  SERVICE <https://dbpedia.org/sparql> {
    BIND(dbr:Brazil AS ?paisContexto)
    ?paisContexto rdfs:label ?paisLabel ;
                  rdf:type ?tipoExterno .
    FILTER(LANG(?paisLabel) = "en")
    FILTER(?tipoExterno IN (dbo:Country, dbo:Place, dbo:PopulatedPlace))
  }

  BIND("DBpedia" AS ?fonteExterna)
} LIMIT 3
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

O `SERVICE` encaminha a busca externa à DBpedia. BIND(`dbr:Brazil AS ?paisContexto`) seleciona previamente o Brasil; rdfs:label recupera seu nome e rdf:type suas classificações. Os filtros selecionam o nome em inglês e os tipos Country, Place ou PopulatedPlace. LIMIT 3 limita a resposta a três linhas. Não há uma variável de identificação compartilhada entre a busca local e a externa.

O registro da execução no GraphDB, em 2 de setembro de 2026, confirma três linhas, mas não armazena seus valores completos. O quadro documenta as informações preservadas no registro, não a transcrição da saída tabular.

Quadro B-20 – Registro do resultado da consulta 10

Elemento documentado	**Resultado registrado**
Consulta	10 — Federação com a DBpedia
Identificador local	CPF-FICTICIO-0012
Recurso externo	dbr:Brazil
Estado da execução	Executada com sucesso
Quantidade de linhas	3
Combinação observada	Dados locais do diploma fictício combinados com classes externas da DBpedia para o recurso Brazil
Data da execução registrada	02/09/2026

Fonte: Elaborado pelo autor, a partir de validacao_graphdb_consulta_federada_dbpedia.json.

O registro histórico demonstra a execução da federação, mas não preserva os valores completos das três linhas. A consulta acrescenta contexto externo predefinido; não identifica o país por um vínculo local nem comprova a nacionalidade do titular ou a localização das instituições.

B.11 Consulta 11 — Federação com o Wikidata por código IBGE

A consulta combina uma contagem local de diplomados fictícios com a descrição do município e do país no Wikidata, utilizando o código IBGE “3550308” como identificador compartilhado.

Arquivo da consulta: 11_demo_consulta_federada_qlever_wikidata_municipio_ibge.rq

Quadro B-21 – Código completo da consulta 11

```
PREFIX odd:  <https://w3id.org/ufes/ontodiplomadigital#>
PREFIX wdt:  <http://www.wikidata.org/prop/direct/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?codigoIbge
       ?municipioWikidata
       ?municipioLabel
       ?pais
       ?paisLabel
       ?qtdDiplomados
WHERE {
  {
    SELECT ?codigoIbge
           (COUNT(DISTINCT ?diplomado) AS ?qtdDiplomados)
    WHERE {
      ?diplomado a odd:Diplomado ;
                 odd:hasBirthPlace ?localNaturalidade .

      ?localNaturalidade odd:ibgeMunicipalityCode ?codigoIbge .

      FILTER(STR(?codigoIbge) = "3550308")
    }
    GROUP BY ?codigoIbge
  }

  SERVICE <https://qlever.dev/api/wikidata> {
    ?municipioWikidata
        wdt:P1585 ?codigoIbge ;
        rdfs:label ?municipioLabel ;
        wdt:P17 ?pais .

    ?pais rdfs:label ?paisLabel .

    FILTER(LANG(?municipioLabel) = "pt")
    FILTER(LANG(?paisLabel) = "pt")
  }
}
```

Fonte: Elaborado pelo autor, a partir do arquivo da consulta disponibilizado no repositório da prova de conceito.

Na preparação dos dados, o código IBGE foi representado por odd:ibgeMunicipalityCode. A subconsulta agrupa esse código e conta os diplomados

distintos. SERVICE encaminha a busca ao QLever: a mesma variável ?codigoIbge ocupa o valor de wdt:P1585, propriedade do código IBGE no Wikidata. A partir do município encontrado, rdfs:label recupera seu nome e wdt:P17 identifica o país. Os filtros selecionam os nomes em português.

A execução registrada no GraphDB em 2 de setembro de 2026 retornou uma linha. Os valores são apresentados verticalmente para preservar todos os identificadores e facilitar a leitura.

Quadro B-22 – Resultado da consulta 11

Variável selecionada	**Valor recuperado**
?codigoIbge	3550308
?municipioWikidata	wd:Q174
?municipioLabel	São Paulo
?pais	wd:Q155
?paisLabel	Brasil
?qtdDiplomados	30

Fonte: Elaborado pelo autor, a partir de validacao_graphdb_consulta_federada_qlever_wikidata_ibge.json.

A resposta reúne a contagem local de trinta diplomados fictícios com o município São Paulo (Q174) e o país Brasil (Q155), recuperados da fonte externa. O código IBGE estabelece a correspondência entre recursos descritos com vocabulários diferentes. A contagem limita-se à base sintética e a consulta não acessa o sistema do IBGE nem verifica diplomas reais.

B.12 Acesso aos artefatos e orientações de reprodução

Este Apêndice B apresenta as orientações para acesso aos artefatos computacionais e documentais da OntoDiplomaDigital, disponibilizados no repositório público *ontodiplomadigital-proof-of-concept*, no GitHub. O material complementa os resultados apresentados na dissertação e permite examinar a representação formal da ontologia, consultar os dados fictícios utilizados e executar as consultas da prova de conceito.

O acesso está disponível pelos seguintes endereços:

- Repositório público e materiais complementares.
- Versão v1.1 da prova de conceito.

A versão v1.1 reúne os arquivos da ontologia, as instâncias fictícias, os dados tabulares, as consultas SPARQL, o diagrama do fluxo da prova de conceito, as orientações de reprodução e os relatórios de verificação. A página principal também disponibiliza materiais incorporados após essa versão, incluindo os schemas XSD do Ministério da Educação, imagens extraídas da dissertação, a licença e os metadados de citação acadêmica. A organização consultada em 3 de setembro de 2026 compreende os seguintes componentes:

Localização no repositório	Conteúdo e finalidade
ontologia/	Arquivo OntoDiplomaDigital_registro_diploma.ttl, com a estrutura formal da ontologia, e arquivo separado da ABox pública, com as instâncias fictícias da prova de conceito.
dados/	Arquivos CSV que permitem examinar os dados tabulares associados à representação dos indivíduos e ao mapeamento para RDF.
consultas_sparql/	Onze consultas: sete vinculadas às questões de competência e quatro demonstrativas, destinadas à recuperação integrada de informações de um diploma, à geração de seu grafo e à consulta federada à DBpedia e ao Wikidata por meio do QLever.
validacao/	Relatórios de verificação dos arquivos, das consultas e das demonstrações federadas, além dos registros de substituição de identificadores por valores artificiais.
docs/	Guia para leitores e roteiro de carregamento da ontologia e da ABox no GraphDB, com orientações para execução das consultas.
graficos/	Diagrama do fluxo da prova de conceito.
fontes_oficiais_mec/	Cópias dos schemas XSD v1.05 utilizados como fontes técnicas, acompanhadas de informações de origem e identificadores de integridade dos arquivos.
graficos_dissertacao/	Imagens extraídas do documento, índice, relação dos arquivos e prancha de miniaturas para identificação visual.
LICENSE.md e CITATION.cff	Condições de uso dos materiais autorais e metadados para citação acadêmica do repositório.

A consulta pode ser iniciada pelo guia para leitores. Para executar a demonstração, o roteiro de reprodução no GraphDB orienta o carregamento dos dois arquivos Turtle e a utilização das consultas. As imagens podem ser examinadas diretamente no GitHub; seus nomes correspondem à ordem técnica de extração e não substituem a numeração das figuras na dissertação.

Os registros acadêmicos disponibilizados possuem finalidade demonstrativa, com nomes pessoais e identificadores substituídos por valores fictícios. As consultas permitem examinar relações entre os recursos representados, mas não realizam validação criptográfica de assinaturas nem comprovam implantação institucional. A execução das consultas federadas depende da disponibilidade dos serviços externos e das condições de acesso à internet no ambiente utilizado.

Os artefatos autorais estão disponibilizados sob a licença Creative Commons Atribuição 4.0 Internacional, conforme as condições registradas no repositório. Os schemas do MEC permanecem identificados como fontes externas, com autoria e condições próprias de disponibilização. O arquivo de citação acadêmica fornece os metadados para atribuição da autoria e referência ao material consultado.