



Termo de Execução Descentralizada/Projeto

**Termo de Execução Descentralizada Cade e UnB**

**Pesquisa Aplicada com fundamento no estudo de necessidades, análise informacional, concepção, prototipação e experimentação com vistas à inovação tecnológica e de processos no domínio da Defesa Econômica para o Conselho Administrativo de Defesa Econômica – Cade.**

Documento:

**Meta 7  
Pesquisa e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência**

Data de Emissão:

**20/12/2022**

Elaborado por:

**Universidade de Brasília – UnB  
Laboratório de Tecnologias da Tomada de Decisão – LATITUDE.UnB**

Digital Object Identifier:

**10.5281/zenodo.7798673**

**Conselho Administrativo de Defesa Econômica – Cade.**

**UNIVERSIDADE DE BRASÍLIA**

**Alexandre Cordeiro Macedo**  
Presidente do Cade

**Márcia Abrahão Moura**  
Reitora

**Paulo Eduardo Silva de Oliveira**  
Diretor de Administração e Planejamento

**Profª. Maria Emília Machado Telles Walter**  
Decana em exercício  
Decanato de Pesquisa e Inovação – DPI

**Vinicius Eloy dos Reis**  
Coordenador-Geral de Tecnologia da Informação

**Rafael Timóteo de Sousa Júnior**  
Coordenador do Laboratório de Tecnologias  
da Tomada de Decisão – LATITUDE

**William José da Conceição Marques**  
Chefe de Serviço de Gestão e Governança

**Leonardo de Oliveira Alves Sanches Lima**  
Chefe de Serviço de Segurança e Infraestrutura

**EQUIPE TÉCNICA**

Alessandro Lustosa Seixas Pinheiro  
Cristiane de Freitas Leal  
Felipe Alberto Moreira  
Helio Ribeiro da Costa  
Jorge Luiz Hachem Vasconcelos  
Wendel Sousa de Anchieta

**EQUIPE TÉCNICA**

**Edna Dias Canedo**  
(Pesquisador Sênior)  
**Georges Daniel Amvame Nze**  
(Pesquisador Sênior)  
**Rafael Timóteo de Sousa Júnior**  
(Pesquisador Sênior)  
**Valério Aymore Martins**  
(Pesquisador Sênior)  
Adailton Amaro Nunes  
Alessandra de Vasconcelos Sales  
Aline Lima da Silva  
Ana Beatriz de Sousa Monteiro  
Ana Flavia Araújo Santana  
André Luiz Chaves da Silva  
Bruno Justino Garcia Praciano  
Caroline Emanuele da Silva Sousa  
Danilo Freire Holanda de Paiva  
Denise Mitie Taketomi  
Edle Marcio dos Santos Silva  
Eduardo Drumond Soares  
Fábio Lúcio Lopes de Mendonça  
Felipe Evangelista dos Santos  
Felipe Lobão Lima  
Guilherme Pereira Paiva  
Hugo Bartolomeu Ferreira  
Ian Porto e Melo  
Igor de Almeida Gentil  
Isabela de Lima Silva  
Isabela Carolina Martins  
Italo Eduardo Dias Frota  
Izadora Oliveira Linhares  
Jeanne Louize Emygdio  
Julia Jamile de Oliveira Gonçalves  
Karen Ferreira Machado  
Lincoln Abreu Barbosa  
Lohana Salustiano Cardoso  
Luiz Fernando Santos de Araujo Caduda  
Marcelo Dos Santos Vianna  
Marcio Junior Vieira  
Marcos Vinicius da Silva Souza  
Mateus Alves Dantas  
Noely Gomes de Lima  
Paulo Henrique Batista Rodrigues  
Rafael de Arruda Faria  
Renato Carauta Ribeiro  
Rogerio de Oliveira Costa  
Thays de Oliveira Coelho

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.2/95
---------------	---------------------	--	----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.3/95
---------------	---------------------	--	----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## HISTÓRICO DE REVISÕES

Data	Versão	Descrição
30/11/2022	1.0	Criação do documento e sua formatação inicial



Universidade de Brasília – UnB  
 Campus Universitário Darcy Ribeiro - FT – ENE – Latitude  
 CEP 70.910-900 – Brasília-DF  
 Tel.: +55 61 3107-5598 – Fax: +55 61 3107-5590

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.4/95
---------------	---------------------	---	----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
 É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

1. INTRODUÇÃO .....	7
2. OBJETIVO.....	10
2.1 PERCURSO METODOLÓGICO ESTABELECIDO PARA A PESQUISA: EVOLUÇÃO .....	10
2.2 ATIVIDADES DA DIMENSÃO ARQUITETURA TECNOLÓGICA .....	11
2.2.1 Término da seleção de TICs .....	12
2.2.2 Instalação, configuração e testes das TICs selecionadas .....	12
2.2.3 Replicação da arquitetura no ambiente de produção no Cade .....	12
2.2.4 Disseminação da documentação produzida e representação gráfica .....	12
2.3 ATIVIDADES DA DIMENSÃO RECURSOS HUMANOS .....	12
2.3.1 Treinamento tecnológico.....	12
2.3.2 Treinamento metodológico.....	13
2.4 ATIVIDADES DA DIMENSÃO VOCABULÁRIO CONTROLADO.....	13
2.4.1 Seleção de termos .....	13
2.4.2 Análise de termos .....	13
2.4.3 Consolidação de termos entre os especialistas em informação .....	14
2.4.4 Avaliação dos termos por especialistas de domínio .....	14
2.4.5 Homologação de termos .....	15
2.4.6 Publicação de termos .....	15
2.4.7 Conversão de vocabulários.....	15
3. RESULTADOS OBTIDOS NESTA ETAPA DA PESQUISA .....	15
3.1 DIMENSÃO ARQUITETURA TECNOLÓGICA .....	16
3.1.1 Término da seleção de TICs.....	16
3.1.2 Instalação, configuração e testes das TICs selecionadas.....	16
3.1.2.1 Web Protégé e Protégé.....	16
3.1.2.2 Skosmos.....	16
3.1.2.3 Apache Jena Fuseki: ambiente de armazenamento e consulta .....	17
3.1.2.4 Apache Shiro (LDAP) .....	17
3.1.3 Replicação da arquitetura no ambiente do Cade .....	19
3.1.3.1 Web Protégé .....	19
3.1.3.2 SKOSMOS .....	19
3.1.3.3 Apache Jena Fuseki, Apache Shiro e integração das ferramentas .....	19
3.1.4 Visão gráfica da arquitetura tecnológica da frente de ontologia.....	19
3.2 DIMENSÃO RECURSOS HUMANOS.....	22
3.2.1 Treinamento metodológico.....	22

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.5/95
---------------	---------------------	---	----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

3.2.1.1	Para pesquisadores da frente de ontologia .....	22
<b>3.2.1.2</b>	<b>Para especialistas de domínio .....</b>	<b>23</b>
3.2.2	Treinamento tecnológico.....	23
3.2.2.1	Para integrantes da frente de ontologia.....	23
3.2.2.2	Para especialistas de domínio.....	24
3.3	<b>DIMENSÃO VOCABULÁRIO CONTROLADO .....</b>	<b>24</b>
3.3.1	Seleção de termos .....	24
3.3.2	Análise de termos .....	25
3.3.3	Consolidação .....	26
3.3.4	Avaliação .....	26
3.3.5	Homologação.....	27
3.3.6	Publicação de termos e conversão do vocabulário .....	27
4.	<b>ATIVIDADES EM ANDAMENTO .....</b>	<b>28</b>
5.	<b>REFERÊNCIAS.....</b>	<b>29</b>
	<b>APÊNDICES.....</b>	<b>30</b>
	APÊNDICE A: Método para construção de Tesouros .....	30
	APÊNDICE B: Instalação e configuração do Web Protégé .....	34
	APÊNDICE C: Instalação e configuração do SKOSMOS .....	36
	APÊNDICE D: Instalação e Configuração do Apache Jena Fuseki .....	37
	APÊNDICE E: Instalação e configuração do Apache Shiro .....	38
	APÊNDICE F: Procedimentos de instalação do Web Protégé no Cade .....	40
	APÊNDICE G: Procedimentos de instalação do SKOSMOS no Cade .....	44
	APÊNDICE H: Arquitetura Apache Jena Fuseki, Web Protégé e SKOSMOS no Cade .....	47
	APÊNDICE I: Relação de fontes informacionais atualizadas .....	50
	APÊNDICE J: Mapeamento de constructos de Tesouros para notação SKOS .....	57
	APÊNDICE K: Diretrizes para representação do conhecimento no padrão SKOS.....	59
	APÊNDICE L: Material instrucional de treinamento no Web Protégé.....	64
	APÊNDICE M: Diretrizes para execução das atividades de validação de termos.....	70
	APÊNDICE N: Evolução das análises sobre conceitos centrais do domínio Cade. ....	73
	APÊNDICE O: Listas de termos por tipo de classificação .....	78
	APÊNDICE P: Conversão de planilhas excel para RDF via Cellfie .....	88
	APÊNDICE Q: Percurso para publicação de vocabulários no Skosmos .....	91
6.	<b>CONCLUSÃO.....</b>	<b>94</b>

## 1. INTRODUÇÃO

O presente documento trata do Relatório Técnico referente ao Acordo de Cooperação firmado entre a Unb e o Conselho Administrativo de Defesa Econômica - Cade, que visa atender atividades pertinentes a execução do Plano de Ação. O Conselho Administrativo de Defesa Econômica – Cade – é uma Autarquia Federal, vinculada ao Ministério da Justiça, com sede e foro no Distrito Federal, que exerce, em todo o território nacional, as atribuições dadas pela Lei nº 12.529/2011.

O Cade tem como missão zelar pela livre concorrência no mercado, sendo a entidade responsável, no âmbito do poder executivo, não só por investigar e decidir, em última instância, sobre a matéria concorrencial, como também fomentar e disseminar a cultura da livre concorrência. As atribuições do Cade são definidas pela Lei nº 12.529, de 30 de novembro de 2011, e complementadas pelo regimento interno do Cade - RiCade, aprovado pela Resolução nº 20, de 07 de junho de 2017.

Essa Autarquia exerce três funções primárias:

- preventiva: analisar e posteriormente decidir sobre as fusões, aquisições de controle, incorporações e outros atos de concentração econômica entre grandes empresas que possam colocar em risco a livre concorrência.
- repressiva: investigar, em todo o território nacional, e posteriormente julgar cartéis e outras condutas nocivas à livre concorrência.
- educativa: instruir o público em geral sobre as diversas condutas que possam prejudicar a livre concorrência; incentivar e estimular estudos e pesquisas acadêmicas sobre o tema, firmando parcerias com universidades, institutos de pesquisa, associações e órgãos do governo; realizar ou apoiar cursos, palestras, seminários e eventos relacionados ao assunto; editar publicações, como a Revista de Direito da Concorrência e cartilhas.

Diante da complexidade estrutural e funcional do mercado e das interrelações entre os diversos atores, e considerando a evolução tecnológica como ferramenta no suporte as atividades de defesa econômica nacional - em particular notando a evolução tecnológica do próprio modo operante das condutas nocivas à livre concorrência. Assim, foi entendimento do Cade que se faz necessário que essa Autarquia esteja cada vez mais preparada e especializada, através de processos de P&D, para que essa possa trazer respostas efetivas

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.7/95
---------------	---------------------	---	----------

Confidencial.

aos problemas confrontados usando-se do que existe de melhor suporte tecnológico e sistêmico.

Com os avanços da ciência e das tecnologias correlatas no tratamento de dados de interesse nas atividades preventiva, repressiva e educativa, tais como as realizadas pelo Cade, é interesse deste Conselho buscar inovações tecnológicas e de processos que contribuam ao permanente aumento da qualidade dos seus serviços.

Portanto, diante da necessidade, foi estabelecida a parceria do CADE com instituição de pesquisa UnB para execução de objeto comum, a ser concretizado mediante a mútua cooperação, em defesa da concorrência. Especificamente, a parceria se fez por Termo de Execução Descentralizada (TED) para a realização de projeto de Pesquisa Desenvolvimento e Inovação, tendo como objetivo geral realizar estudos com fundamento na análise de requisitos em campo, concepção, desenvolvimento, prototipação, avaliação experimental, visando a criação e integração de inovações tecnológicas no domínio da Defesa Econômica, por meio de tecnologias aplicadas, bem como da análise de processos e da garantia da qualidade, buscando assim resolver problemas de interesse público a partir da aplicação de análise crítica, da ciência de dados e da inteligência artificial, da segurança da informação e forense computacional. O presente relatório técnico é um dos produtos resultantes do TED CADE-UnB.

Logo, a proposta do projeto ora declarado, objeto do Acordo de Cooperação em destaque é a realização de projeto de pesquisa, desenvolvimento de ideias e de inovação, que tem como objetivo geral realizar estudos tendo como fundamento a análise de requisitos dos processos do Cade junto com a concepção, desenvolvimento de um modelo - experimental e empírico por avaliação experimental - visando a criação e integração de inovações tecnológicas no domínio da defesa econômica, por meio do uso de tecnologias aplicadas, bem como da análise de processos e da garantia da qualidade, buscando assim resolver problemas de interesse público a partir da aplicação de análise crítica, da ciência de dados e da inteligência artificial, da segurança da informação e forense computacional.

Essa Cooperação, objeto da parceria com a Universidade de Brasília (UnB), busca gerar a expectativa de alcance dos seguintes benefícios institucionais para o Conselho Administrativo de Defesa Econômica:

- Melhoria dos processos internos de investigação de condutas anticompetitivas ou de análise de atos de concentração;

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.8/95
---------------	---------------------	---	----------

Confidencial.

- Produção de tecnologias que subsidiem o progresso da Defesa Econômica, mensurável por indicadores;
- Inovação das ferramentas de análises de concentração e de condutas anticompetitivas pelo uso de técnicas de mineração dos dados, composição, agrupamento, classificação e outras maneiras de raciocínio artificial, de modo a permitir a identificação automática destes eventos de interesse;
- Geração de boas práticas de dados abertos e de governança de dados, bem como de produção de relatórios gerenciais relacionados aos processos de atos de concentração e de apuração de infrações à ordem econômica;
- Possibilidade de efetuar, no espaço cibernético, ações ofensivas, defensivas e exploratórias;
- Maior efetividade no manuseio de dados e geração de análise preditivas e de inteligência artificial em estrutura multidimensional dos dados que permita elevar a eficiência da defesa econômica no País;
- Otimização dos recursos econômicos e humanos alocados nos processos de trabalho do Conselho;
- Melhoria no controle dos processos;
- Melhoria de análises gerenciais a partir de geração de relatórios estatísticos;
- Geração de estudos e pesquisas para o contexto da defesa econômica no Brasil;
- Melhoria da oferta de serviços digitais para os cidadãos; e
- Aquisição de novos conhecimento e habilidades dos servidores do Cade.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.9/95
---------------	---------------------	---	----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## 2. OBJETIVO

Este documento tem por objetivo apresentar as atividades realizadas pela frente de Ontologia, no período de junho a novembro de 2022 como requisito às entregas previstas em Projeto de Pesquisa de parceria entre UnB e Cade.

O presente relatório técnico encontra-se estruturado da seguinte forma: na Seção 2 apresenta-se a evolução das etapas previstas para o percurso metodológico, originalmente estabelecido e apresentado no RT2 e aprimorado nesta etapa. Na Seção 3 apresentam-se os resultados obtidos nesta etapa de execução da pesquisa e na Seção 4 apresentam-se as considerações finais abrangendo atividades em andamento e as próximas atividades previstas.

### 2.1 PERCURSO METODOLÓGICO ESTABELECIDO PARA A PESQUISA: EVOLUÇÃO

Nesta seção serão apresentadas as etapas do percurso metodológico estabelecido para a execução da pesquisa, compreendendo três dimensões paralelas de trabalho: i) Dimensão Arquitetura Tecnológica; ii) Dimensão Vocabulário Controlado e, iii) Dimensão Recursos Humanos. A apresentação do percurso metodológico estruturado em dimensões caracteriza um avanço na metodologia de trabalho desta frente, cujas atividades se desdobram em ações distintas e paralelas para atendimento a demandas nestes três enfoques complementares.

Na Dimensão Arquitetura Tecnológica foram previstas quatro atividades a saber: i) término da seleção das TICs aptas a suportar as atividades de pesquisa desta frente; ii) instalação, configuração e testes das TICs selecionadas em ambiente de desenvolvimento; iii) replicação da arquitetura completa para o ambiente de produção no Cade e; iv) Disseminação da documentação, incluindo uma representação gráfica da arquitetura estabelecida.

Na Dimensão Recursos Humanos foram previstas duas atividades recorrentes a serem realizadas para acolhimento e reciclagem de (novos) pesquisadores da frente de ontologia, a saber: i) treinamento metodológico e, ii) treinamento tecnológico.

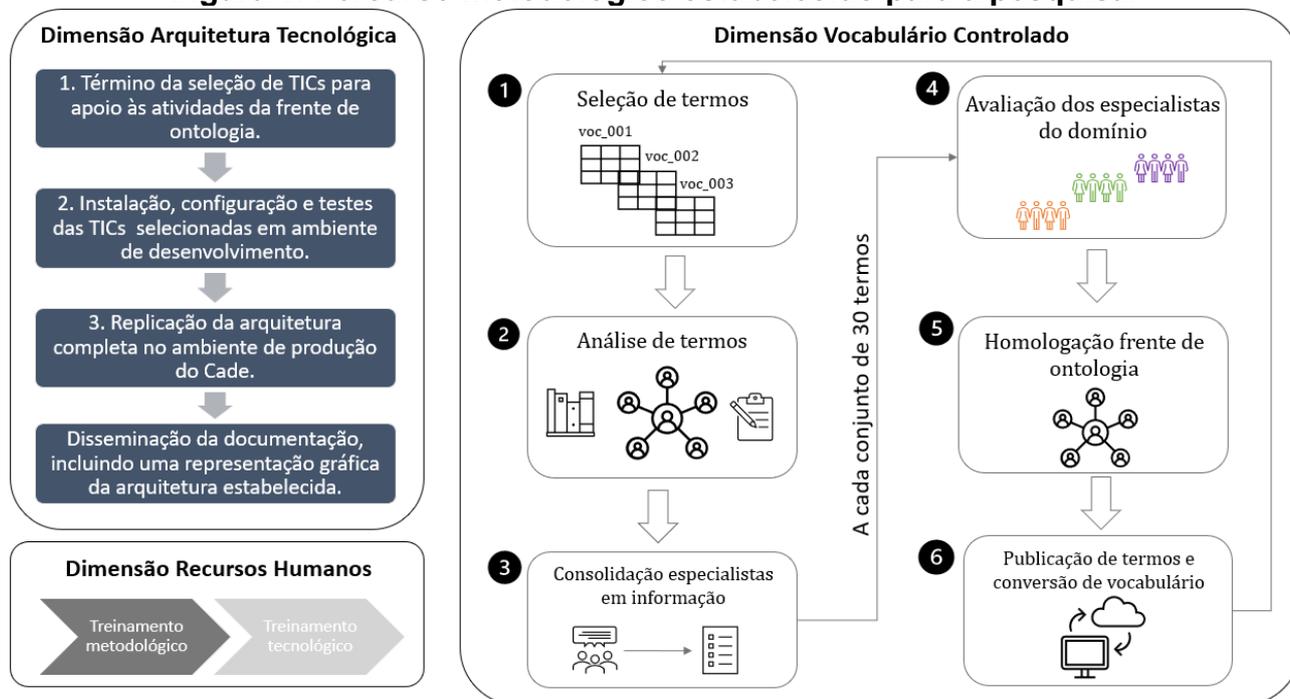
Na Dimensão Vocabulário Controlado foram previstas seis atividades a saber: i) seleção de termos; ii) análise de termos; iii) consolidação entre os especialistas em informação; iv) avaliação dos especialistas de domínio; v) homologação da frente de ontologia; vi) publicação de termos e conversão do vocabulário. Na Figura 1 ilustra-se o percurso

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.10/95
---------------	---------------------	---	-----------

Confidencial.

metodológico evoluído, em seguida apresenta-se um detalhamento de todas as atividades a serem realizadas nas dimensões.

**Figura 1. Percurso metodológico estabelecido para a pesquisa.**



Fonte: Elaborado pelos autores (2022).

O detalhamento das atividades previstas em cada dimensão será apresentado a seguir.

## 2.2 ATIVIDADES DA DIMENSÃO ARQUITETURA TECNOLÓGICA

Para esta dimensão foram previstas quatro atividades a saber: i) término da seleção das TICs aptas a suportar as atividades de pesquisa desta frente; ii) instalação, configuração e testes das TICs selecionadas em ambiente de desenvolvimento; iii) replicação da arquitetura completa para o ambiente de produção no Cade e; iv) Disseminação da documentação, incluindo uma representação gráfica da arquitetura estabelecida.

As atividades visam estabelecer a arquitetura completa de tecnologias capazes de cobrir o ciclo de representação do conhecimento em tesouros e ontologias e a conversão destes artefatos em insumos para a frente de Machine Learning (ML).

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.11/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade. É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

### 2.2.1 Término da seleção de TICs

Para esta etapa previu-se a conclusão dos testes iniciados na etapa anterior (e relatados no RT2) sobre as ferramentas candidatas a sustentar as atividades desta frente. Na etapa anterior realizou-se pesquisa teórica de papers das ferramentas mais utilizadas para a construção de Tesouros e ontologias, como forma de aquisição de conhecimento sobre as características gerais de cada ferramenta, suas funcionalidades, aspectos técnicos e limitações.

Ao final da etapa previu-se a apresentação das ferramentas selecionadas e a justificativa das escolhas.

### 2.2.2 Instalação, configuração e testes das TICs selecionadas

Nesta etapa previu-se: i) a instalação das TICs selecionadas em ambiente de desenvolvimento; ii) a elaboração de uma documentação detalhada das configurações necessárias à utilização de cada ferramenta, para fins de compartilhamento de conhecimento e posterior replicação no ambiente de produção do Cade; iii) a execução de testes funcionais sobre as ferramentas selecionadas com forma de apoiar as atividades de treinamento tecnológico e metodológico dos recursos humanos.

### 2.2.3 Replicação da arquitetura no ambiente de produção no Cade

Nesta etapa previu-se a replicação da arquitetura completa e funcional, montada no ambiente de desenvolvimento, para o ambiente de produção do Cade.

### 2.2.4 Disseminação da documentação produzida e representação gráfica

Nesta etapa previu-se a disseminação de toda a documentação produzida durante a realização das atividades desta dimensão em ambiente de acesso aos pesquisadores da frente de ontologia e aos especialistas de domínio do Cade.

Previu-se também a elaboração de uma visão gráfica da arquitetura a ser incluída na documentação.

## 2.3 ATIVIDADES DA DIMENSÃO RECURSOS HUMANOS

### 2.3.1 Treinamento tecnológico

O objetivo previsto para esta etapa foi o de ministrar treinamento tecnológico sobre as ferramentas de apoio à execução da pesquisa para cada novo integrante da equipe e reciclar os pesquisadores em atividade, quando necessário. Para tanto, estabeleceu-se as

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág. 12/95
---------------	---------------------	---	------------

Confidencial.

seguintes técnicas a serem adotadas: treinamento síncrono, treinamento assíncrono a partir de vídeos gravados durante os pontos de controle, apresentações, elaboração e compartilhamento de tutoriais disponibilizados.

### 2.3.2 Treinamento metodológico

O objetivo previsto para esta etapa foi o de ministrar treinamento metodológico sobre as atividades previstas para a Dimensão Vocabulário Controlado, para cada novo integrante da equipe e reciclar os pesquisadores em atividade, quando necessário. Para tanto estabeleceu-se as seguintes técnicas a serem adotadas: treinamento síncrono, treinamento assíncrono a partir de vídeos gravados durante os pontos de controle, apresentações, tutoriais disponibilizados e acompanhamento de lideranças durante a apropriação do processo.

## 2.4 ATIVIDADES DA DIMENSÃO VOCABULÁRIO CONTROLADO

### 2.4.1 Seleção de termos

Nesta etapa foram previstas atividades de extração de termos de três formas: i) a partir dos vocabulários indicados pelos especialistas de domínio em setembro de 2021; ii) a partir do estudo de documentos do domínio; e iii) a partir da observância dos resultados obtidos pela frente de ML que indicam as frequências dos termos identificados naquela frente de projeto. As extrações são realizadas tanto por humanos quanto por suporte computacional.

### 2.4.2 Análise de termos

Nesta etapa foram previstas atividades de revisão dos termos também sob dois aspectos concomitantes:

- partindo dos termos específicos para os termos gerais: uma vez que os vocabulários preliminares disponibilizados são do domínio da jurisprudência, há uma demanda de identificação dos termos comuns ao domínio CADE, provavelmente, classificados naqueles vocabulários, como termos relacionados;
- partindo dos termos mais genéricos identificados nos documentos do domínio para auxiliar na identificação de termos específicos e relacionados.

Estabeleceu-se também um conjunto de oito guias para nortear as atividades de classificação de acordo com as normas e padrões de construção de tesouros já mencionados

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.13/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.

É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

neste documento. Os guias podem ser consultados no Apêndice A e compreendem o controle de vocabulário cíclico empregado na classificação dos termos. A estrutura hierárquica dos termos vem sendo definida ao longo da evolução do trabalho, assim como a definição das relações entre os termos.

Estabeleceu-se ainda que cada especialista em informação faria as análises de termos em particular e posteriormente os resultados deveriam ser discutidos em equipe.

Em função da dimensão do corpus de termos a ser analisado, são previstas atividades paralelas de análise de termos por cada pesquisador envolvido nesta frente e que tenha conhecimento dos processos de modelagem conceitual para construção de vocabulários controlados.

Inicialmente, não se prevê a necessidade de estabelecer uma ordem de escolha dos termos que serão distribuídos entre os pesquisadores. Prevê-se a organização de uma planilha de classificação que deverá ser utilizada até a seleção da ferramenta tecnológica adequada. A planilha deverá manter as seguintes abas para facilitar o acompanhamento da evolução das análises e classificações:

- Aba 1: Descritores
- Aba 2: Não-Descritores
- Aba 3: Modificadores
- Aba 4: Especificadores
- Aba 5: Termos eliminados

### 2.4.3 Consolidação de termos entre os especialistas em informação

Nesta etapa serão realizadas as discussões sobre as análises de termos realizadas individualmente na etapa anterior para se obter um consenso sobre as classificações e o entendimento do domínio.

Cada especialista deverá apresentar suas análises justificando-as de acordo com as etapas previstas nos guias de orientação.

Estabeleceu-se ainda que a cada 30 termos consolidados seria gerada uma remessa destes para validação dos especialistas de domínio. Este conjunto de termos deverá apresentar qualidade suficiente para ser submetido à avaliação de especialistas de domínio.

### 2.4.4 Avaliação dos termos por especialistas de domínio

Nesta etapa serão realizadas duas atividades: i) análise individual de especialistas do domínio CADE sobre o conjunto de termos recebidos; ii) reunião entre as equipes de

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág. 14/95
---------------	---------------------	---	------------

Confidencial.

construção do tesouro e de especialistas de domínio para discussão e validação das análises realizadas.

#### 2.4.5 Homologação de termos

Nesta etapa são previstas novas sessões de discussões entre os especialistas da frente de ontologia acerca das avaliações realizadas pelos especialistas e a confrontação dos resultados com estudos previamente realizados ou percepções advindas da reunião de avaliação.

Ao término de uma sessão de homologação de termos, todas as modificações consideradas necessárias nos termos devem ser realizadas. Cada termo deverá ocupar uma e apenas uma aba na estrutura da planilha previamente estabelecida. Todos os termos relacionados (ao termo em revisão) devem ocupar as colunas e abas específicas de acordo com a sua classificação.

#### 2.4.6 Publicação de termos

Os termos homologados devem ser disponibilizados para consulta em plataforma acessível aos pesquisadores do Cade para acompanhamento das evoluções, sendo viável ainda a disponibilização de apresentações visuais que facilitem a compreensão da evolução.

#### 2.4.7 Conversão de vocabulários

Nesta etapa prevê-se a conversão da planilha de termos para as ferramentas tecnológicas selecionadas, bem como a documentação de todos os procedimentos de conversão realizados.

Como saídas esperadas desta etapa estão: i) a disponibilização do tesouro para consulta no ambiente do CADE; ii) o fornecimento de novas classificações para apoiar a frente de ML; e iii) uma visão do Tesouro no padrão SKOS.

### 3. RESULTADOS OBTIDOS NESTA ETAPA DA PESQUISA

Nesta seção serão apresentados os resultados obtidos durante a execução das atividades previstas na seção 2.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág. 15/95
---------------	---------------------	---	------------

Confidencial.

### 3.1 DIMENSÃO ARQUITETURA TECNOLÓGICA

#### 3.1.1 Término da seleção de TICs.

As etapas de seleção de TICs candidatas, Instalação, Configuração e Testes das ferramentas foram apresentadas no RT2. Nesta etapa finalizamos a seleção das ferramentas WEB Protégé e Protégé em detrimento do VocBench em função da estabilidade das ferramentas selecionadas e da presença de pesquisadora com conhecimento nestas ferramentas na frente de ontologia, o que tende a reduzir a curva de aprendizado da equipe.

#### 3.1.2 Instalação, configuração e testes das TICs selecionadas.

##### 3.1.2.1 Web Protégé e Protégé

O Web Protégé e o Protégé são ambientes de desenvolvimento de ontologias, oriundos de dois projetos distintos, desenvolvidos pelo Stanford Center for Biomedical Informatics Research (BMIR) da Universidade de Stanford, na Califórnia (EUA). Ambos possuem comunidades ativas de usuários e desenvolvedores, sendo o Projeto Protégé mais maduro, contando com uma comunidade de 366.084 usuários e uma base de mais de 250 plugins que vêm sendo gradualmente migrados para o projeto Web Protégé. Este, por sua vez, tem por objetivo disponibilizar os recursos do Protégé em uma plataforma Web para possibilitar a realização de atividades colaborativas.

Ambas as ferramentas são desenvolvidas com a linguagem Java e oferecem suporte aos padrões W3C: Ontology web Language versão 2 (OWL2) - e suas sublinguagens, Resource Description Framework (RDF), Extensible Markup Language (XML), JavaScript Object Notation (JSON) e a Linguagem de consulta SPARQL.

Os procedimentos de instalação e configuração do Web Protegé são apresentados no Apêndice B.

##### 3.1.2.2 Skosmos

Skosmos é um navegador e ferramenta de publicação de vocabulários controlados baseado no padrão W3C Simple Knowledge Organization System (SKOS), desenvolvido com a linguagem de programação PHP (“PHP”) e JavaScript, utiliza bibliotecas de código aberto Composer, EasyRDF, Twig, jQuery, jsTree, Bootstrap, typeahead.js, URI.js, Punic e EasyRdf para acessar dados em SPARQL e RDF, além de disponibilizar a API REST para acesso a vocabulários em outras aplicações. Na ferramenta, os vocabulários são acessados por meio de endpoints SPARQLs contendo vocabulários no padrão SKOS.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.16/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

Os principais recursos oferecidos pela ferramenta são: i) Pesquisa e navegação em vocabulários; ii) Índice alfabético e Índice temático; iii) Exibição de conceito estruturado; iv) Hierarquia de conceito visualizado e v) Interface de usuário multilíngue.

Os principais requisitos para sua utilização são: i) um ou mais vocabulários baseados no padrão SKOS; ii) Servidor web compatível com PHP; iii) Armazenamento de triplas no formato RDF; iv) endpoint SPARQL para consultas, recomenda-se Apache Jena Fuseki; v) Cache HTTP (por exemplo, Varnish) - opcional, mas recomendado.

Os procedimentos de instalação e configuração do Skosmos encontram-se descritos no Apêndice C.

### 3.1.2.3 Apache Jena Fuseki: ambiente de armazenamento e consulta

Apache Jena é um framework gratuito e de código aberto para a construção de aplicativos de Web semântica e Linked Data, assim como o WebProtégé, ele é desenvolvido na com a linguagem de programação Java. O Apache Jena é composto por diferentes APIs que interagem entre si para processar dados em RDF.

O Apache Jena Fuseki é um servidor SPARQL que utiliza o framework Apache Jena para consultar dados em formato RDF. Ele pode ser executado como um serviço do sistema operacional, como um aplicativo da Web Java e como um servidor autônomo (“Apache Jena Fuseki”). Ele combina um sistema “webapp”, com uma interface do usuário para administração e consulta com servidor adequado para ser executado como parte de uma implantação maior, inclusive com Docker.

A Apache Jena Fuseki fornece os protocolos SPARQL 1.1 para consulta e atualização, bem como o protocolo SPARQL Graph Store (“SPARQL”), além de ser totalmente integrado ao TDB (Beckett) para fornecer uma camada de armazenamento persistente transacional robusta e incorpora consulta de texto Jena, conforme apresentado no Apêndice D.

### 3.1.2.4 Apache Shiro (LDAP)

Para atendimento à solicitação de controle de acesso às aplicações via LDAP, foram adicionadas configurações de acesso adequadas.

A interface gráfica do Apache Jena Fuseki é um ponto de extensão conveniente para conexão com o LDAP, a partir da escrita de um adaptador que permita o funcionamento das chamadas da GraphAPI. A versão 2 do Fuseki fornece segurança utilizando o Apache Shiro.

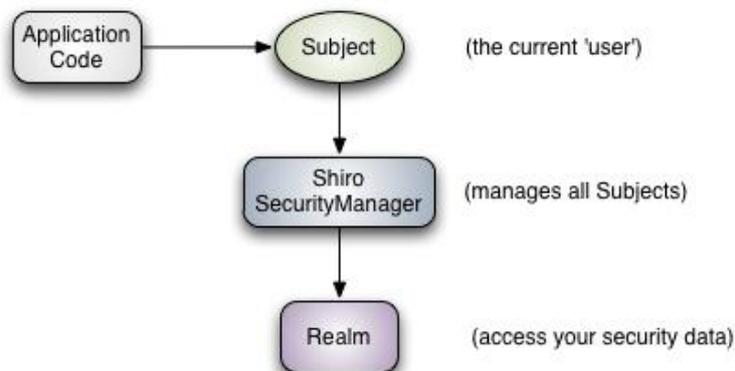
O Apache Shiro é um framework de segurança em Java para autenticação, autorização criptografia e gerenciamento de sessão. Numa visão de alto nível os conceitos principais

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág. 17/95
---------------	---------------------	---	------------

Confidencial.

do Apache Shiro são: Assunto, Gerente de Segurança e Reinos (Subject, SecurityManager e Realms). A Figura 12 ilustra a arquitetura do framework.

**Figura 12: Arquitetura de alto nível do Apache Shiro**



Fonte ASF (2022).

Na arquitetura, encontram-se os seguintes componentes:

- Assunto (Subject): representa um usuário ou um sistema ou um serviço;
- Gerente de Segurança (Security Manager): principal componente da arquitetura, ele coordena os componentes internos de segurança. Gere os reinos e os usuários (assuntos);
- Reinos (realm): atua como um conector entre os dados de segurança e o aplicativo. Ele é essencialmente um DAO (Data Access Object) que encapsula os detalhes de conexão para fontes de dados.

As principais características do Apache Shiro são:

- Baseado em assunto: quase todas as atividades realizadas no Shiro baseiam-se no usuário em execução no momento, chamado de assunto. É possível recuperar facilmente o assunto em qualquer ponto de um código fonte, o que facilita a compreensão e o trabalho com o Shiro junto aos aplicativos desejados.
- Chamada de método único: o processo de autenticação é uma chamada de método único. A necessidade de apenas uma chamada de método mantém a API simples e o código do aplicativo limpo, economizando tempo e esforço. O Shiro oferece uma hierarquia de exceção que apresenta explicações detalhadas do motivo de falha de um login, o que permite diagnosticar mais facilmente erros de código ou problemas de atendimento ao cliente relacionados à autenticação. A hierarquia também permite a criação de uma funcionalidade de autenticação mais complexa, se necessário.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.18/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade. É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

- Lembrar-me integrado: por padrão, a API Shiro possui a capacidade de lembrar os usuários caso eles retornem ao aplicativo, oferecendo melhor experiência aos usuário no uso da API com esforço mínimo de desenvolvimento.
- Fontes de dados conectáveis: através dos Realms e da disponibilização de domínios prontos, possibilita conexão segura às fontes de dados a partir de LDAP, Active Directory e JDBC. Permite criação de domínios customizados.
- Login com um ou mais reinos: permite autenticação de um usuário em um ou mais reinos, retornando uma visão unificada de sua identidade. Permite personalização do processo de autenticação por meio de estratégias que podem ser estabelecidas em arquivos de configuração para que as alterações não exijam modificações no código-fonte, o que reduz a complexidade e o esforço de manutenção.

Os procedimentos de configuração do arquivo shiro.ini encontram-se disponíveis no Apêndice E.

### 3.1.3 Replicação da arquitetura no ambiente do Cade

#### 3.1.3.1 Web Protégé

Os procedimentos de replicação da implantação do Web Protégé no Cade encontram-se descritos no Apêndice F.

#### 3.1.3.2 SKOSMOS

Os procedimentos de replicação da implantação do Web Protégé no Cade encontram-se descritos no Apêndice G.

#### 3.1.3.3 Apache Jena Fuseki, Apache Shiro e integração das ferramentas

O ambiente consolidado do WebProtégé, Apache Jena/Fuseki e Skosmos, está descrito pelo arquivo do docker compose abaixo e depende dos arquivos versionados no GIT do cade no endereço <https://gitlab.cade.gov.br/cade/busca-juridica/ontologia/-/tree/develop/01-Fontes>, conforme descrição apresentada no Apêndice H.

### 3.1.4 Visão gráfica da arquitetura tecnológica da frente de ontologia

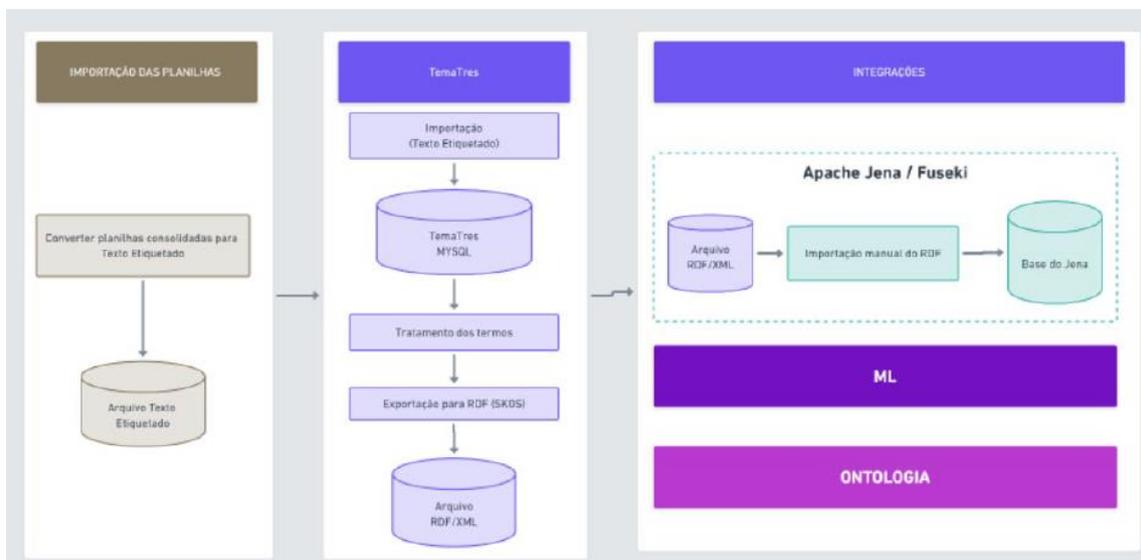
Ilustra-se na Figura 16 uma visão da arquitetura de suporte às atividades desta frente, mantida até o RT2.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág. 19/95
---------------	---------------------	---	------------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

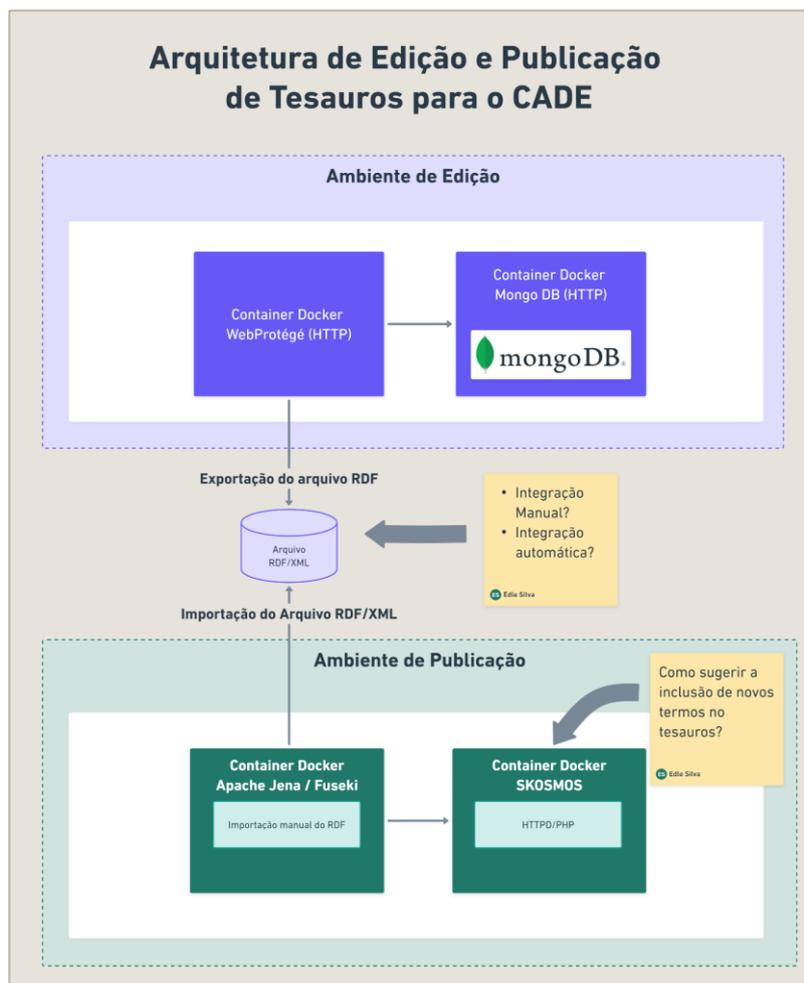
**Figura 16: Arquitetura de suporte das atividades da frente de ontologia mantida até o RT2.**



Fonte: Elaborado pelos autores (2022).

Atualmente a arquitetura é composta também por três camadas, abrangendo todas as tecnologias descritas até o momento: i) Edição; ii) Publicação e iii) Integração, conforme ilustrado na Figura 17, a seguir.

**Figura 17: Arquitetura de edição, integração e publicação de Tesouros e ontologias**



Fonte: Elaborado pelos autores (2022).

Na camada de edição encontra-se o Web Protégé, para a construção e manutenção de tesouros e ontologias.

Na camada de publicação encontram-se duas aplicações, o Apache Jena Fuseki, atuando como servidor de endpoint SPARQL, e o SKOSMOS, atuando como consumidor dos dados em formato RDF, além de disponibilizar uma interface de consulta amigável sobre os vocabulários controlados para o usuário final.

Na camada de integração, ocorrem os procedimentos necessários para publicar o material construído pelos especialistas para público específico. O que remete às atividades de exportação dos vocabulários construídos no Web Protégé para o Apache Jena Fuseki.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.21/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade. É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

### 3.2 DIMENSÃO RECURSOS HUMANOS

Nesta seção serão descritos os procedimentos de capacitação metodológica e tecnológica realizados nesta etapa da pesquisa para os pesquisadores da frente de ontologia (interessantes) e especialistas de domínio.

#### 3.2.1 Treinamento metodológico

##### 3.2.1.1 Para pesquisadores da frente de ontologia

Neste período ocorreu a seguinte movimentação de recursos humanos na frente de ontologia apresentada no Quadro 1:

**Quadro 1: Movimentação de recursos humanos na frente de ontologia**

Pesquisador	Nível	Data ingresso	Data de desligamento
Layze Lúcia Dutra Avanci	Estagiário	Abril/2022	maio/2022
Thais Ketlyn da Silva Medeiros	Estagiário	junho/2022	agosto/2022
Zaqueu Isaque Alves Cabral	Estagiário	junho/2022	agosto/2022
Noély Gomes de Lima	Estagiário	outubro/2022	—
Ana Beatriz de Sousa Monteiro	Estagiário	agosto/2022	—

Fonte: Elaborado pelos autores (2022).

Tendo em vista a complexidade das tarefas em andamento e a movimentação de recursos humanos nesta frente, neste período, tornou-se necessária a execução de um processo de capacitação metodológica abrangendo as seguintes atividades:

- Indicação de cursos abertos da Escola Nacional de Administração Pública (ENAP) relacionados à Defesa da concorrência;
- Disponibilização das fontes de informação em repositório compartilhado para a equipe do projeto. Explicação do propósito de cada fonte e da forma de se localizar conteúdos relevantes. A versão mais atualizada da lista de fontes pode ser consultada no Apêndice I;
- Apresentação da metodologia de análise, explicação dos guias e, posteriormente, análise conjunta de termos (Seção 2.3.3 deste relatório);
- Apresentação das regras de mapeamento de construtos de tesouros para padrão SKOS, disponível no Apêndice J;

- Apresentação, explicação e treinamento das Diretrizes para Representação do Conhecimento no Padrão SKOS, disponível no Apêndice K;
- Apresentação de material instrucional referente à representação do conhecimento em Tesouros utilizando o Web Protégé, disponível no Apêndice L.

### 3.2.1.2 Para especialistas de domínio

Nesta fase deu-se início às atividades de validação de termos com as especialistas do Cade:

- Iara do Espírito Santo: Chefe da Divisão de Acompanhamento Processual da Coordenação-Geral Processual do Conselho Administrativo de Defesa Econômica - Cade;
- Déborah Lins e Nóbrega: Bibliotecária do Serviço de Informação e Documentação da da Coordenação-Geral do Conselho Administrativo de Defesa Econômica - Cade.

Para nortear as atividades de validação dos termos por parte destas especialistas de domínio, definiu-se um conjunto de orientações metodológicas que pode ser consultado no Apêndice M.

## 3.2.2 Treinamento tecnológico

### 3.2.2.1 Para integrantes da frente de ontologia

Para a capacitação de cada novo integrante da equipe de ontologia, em ferramentas tecnológicas utilizadas, especificamente os estagiários, foi disponibilizada uma relação de "links úteis" para acesso aos sistemas do Cade: i) Intranet; ii) Redmine - Sistema de Gerenciamento de Projetos, utilizado para documentação das atividades gerais de pesquisa; iii) Sistema de Busca de Jurisprudência (BJ), aplicado em consultas para verificação de ocorrências de termos e suas variantes; e iv) Editor de tesouros e ontologias Web Protégé.

Todos os membros do projeto têm acesso a uma pasta no Google Drive onde ficam disponíveis os documentos previamente elaborados para o projeto, assim como as apresentações realizadas nos pontos de controle. Na Figura 18 ilustra-se a organização deste ambiente.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.23/95
---------------	---------------------	---	-----------

Confidencial.

**Figura 25: Repositório do projeto no Google Drive**

Name ↑	
1 - Cronograma	6 - Construção TCADE
2 - Estudos sobre Tesouros	7 - Relatórios técnicos
3 - Estudos sobre o CADE	8 - Materiais para estagiários
4 - Estudos sobre ferramentas	9 - Materiais para especialistas
5 - Dicionário de dados e planilhas de vocabulários	10 - Backup
	11 - Conversões

Fonte: Organizado pelos autores (2022).

Os estagiários que se destacaram na execução de suas atividades de classificação de termos na planilha do excel se tornaram aptos a iniciar as atividades no Web Protégé. Para tanto foi criado um projeto reservado às atividades de treinamento, configurado com base no ambiente de construção do tesouro oficial Projeto “Tesouro CADE”. Foi elaborado um vídeo tutorial para demonstração dos procedimentos para elaboração da descrição de termos no Web Protégé.

### 3.2.2.2 Para especialistas de domínio

Para capacitação das especialistas de domínio, Iara e Déborah Lins foram realizadas as seguintes atividades:

- Capacitação em Web Protégé a partir de sessão virtual de treinamento ministrado pela pesquisadora Jeanne;
- Gravação e disponibilização de vídeo instrucional da capacitação para acesso posterior;
- Apresentação didática em ppt dos tópicos abordados no vídeo instrucional e disponibilização para as especialistas. Este material pode ser consultado no Apêndice L.

Durante as sessões de validação de termos as especialistas sanaram diversas dúvidas de utilização e outras foram resolvidas via email.

## 3.3 DIMENSÃO VOCABULÁRIO CONTROLADO

### 3.3.1 Seleção de termos

Conforme mencionado na seção 2.3.1, não foi prevista uma ordem de escolha dos termos para as atividades de análise tendo em vista a demanda inicial de estudo e análise

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.24/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

dos termos dos 3 vocabulários base: voc\_001 (VCE do TCU), voc\_002 (Gerado pelo pesquisador Prof. Valério) e voc\_003 (Gerado pelo Especialista do Cade, Vinícius Eloy dos Reis).

No entanto, o tempo demandado para o estudo de cada termo em particular suscitou a definição de uma estratégia baseada na descoberta dos termos mais relevantes do domínio para agregarmos no entorno de cada um apenas os que tivessem algum tipo de relação com os mesmos. O resultado deste esforço pode ser consultado no Apêndice N.

Posteriormente, o Cade solicitou às especialistas de domínio que fizessem uma triagem nos termos do voc\_001 para encontrar apenas aqueles que estabelecessem algum tipo de relação com o domínio Cade.

Nova demanda foi apresentada no decorrer do projeto para que os termos do voc\_001 fossem integrados ao escopo do vocabulário em construção. Esse processo suscitou durante a análise na descoberta de novos termos que agregaram valor e compuseram a lista de termos do Tesouro Cade.

### 3.3.2 Análise de termos

As atividades de análise de termos vêm sendo realizadas por todos os especialistas, conforme já mencionado previamente.

Atualmente todos os termos estão cadastrados no Web Protégé, classificados de acordo com a situação em que se encontram:

**Quadro 5. Classificação dos termos de acordo com a situação**

Classificação	Descrição	Total de termos
Conceitos	Já identificados como termos centrais no domínio Cade.	28
Em análise	Termos que estão sendo trabalhados pelos especialistas da frente de ontologia.	18
Candidatos	Termos com grande possibilidade de agregar o Tesouro Cade	114
Total de termos em tratamento, específicos do domínio Cade		160

A lista de termos classificados encontra-se disponível no Apêndice O.

### 3.3.3 Consolidação

Durante a etapa de consolidação, realizada semanalmente, 13 termos dentre os 28 classificados como conceitos foram consolidados em conjunto entre os especialistas da frente de ontologia, após análise e estudo individual do pesquisador responsável por cada termo. Soma-se a esse conjunto outros 57 termos identificados como termos equivalentes.

### 3.3.4 Avaliação

Durante esta etapa foram realizadas 17 sessões de validação de termos consolidados, junto às especialistas de domínio.

As sessões têm caracterizado oportunidades importantes para troca de conhecimentos e aprofundamento de aprendizado em aspectos complexos do domínio Cade que, por vezes, não são perceptíveis através do estudo das fontes de informação selecionadas. Algumas descobertas são apresentadas a seguir:

- Há na literatura a atribuição de determinadas siglas a conceitos de forma arbitrária ao que ocorre no domínio, como é o caso do “Compromisso de cessação” que para Marrara (2020) pode ser identificado por “TCC”, enquanto que no domínio, TCC é sigla para o “Termo de compromisso de cessação” (CADE, 2021);
- Há sutilidades nos jargões dos especialistas que conseguimos perceber apenas no confronto dos resultados das buscas de conceitos no sistema de busca de jurisprudência e aqueles obtidos no estudo das fontes de informação. Um exemplo ocorreu em relação ao termo “Leniência” que é utilizado como sinônimo de “Acordo de leniência” em vários documentos, quando, por definição, leniência remete à tolerância, condescendência, lenidade;
- Há criação de conceitos por parte dos especialistas do domínio de maneira arbitrária ao que está estabelecido nas referências básicas do Cade. Um exemplo remete ao conceito “Aditivo ao marker”, criado para distinguir a solicitação de alteração da especificação da conduta do agente infrator no momento do pedido de marker, da “Adesão ao acordo de leniência”,

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.26/95
---------------	---------------------	---	-----------

Confidencial.

conceito próprio do domínio, aplicado para alteração do rol de pessoas físicas lenientes após assinatura do acordo de leniência original.

O envolvimento das especialistas nesta etapa do projeto agregou valor às práticas em andamento. Ambas sugeriram, na última sessão de validação, a inclusão de especialistas de outras áreas de negócio do Cade para nos permitir maior profundidade no entendimento dos termos e evitar retrabalhos futuros, uma vez que previu-se o envolvimento destes especialistas um pouco mais à frente.

### 3.3.5 Homologação

Do total de 160 termos cadastrados no Web Protégé, 13 foram homologados. Esta tarefa tem se mostrado mais lenta do que o esperado em função do impacto que a análise de novos termos acaba gerando em termos previamente analisados.

Muitas vezes, as dúvidas levantadas pela frente de ontologia sobre alguns termos tornam-se, também, dúvidas que as especialistas de domínio precisam sanar com especialistas de outras áreas no Cade para que seja possível alcançarmos uma visão final sobre os conceitos.

### 3.3.6 Publicação de termos e conversão do vocabulário

Atualmente o Tesouro do Cade tem sido construído diretamente no Web Protégé, por todos os pesquisadores da frente de ontologia. Neste ambiente também ocorrem as atividades de consolidação, validação e homologação de termos.

As diretrizes estabelecidas para representação dos termos no Web Protégé podem ser consultadas no Apêndice K.

Os procedimentos para conversão das planilhas de termos em excel para o Web Protégé utilizando o plugin Cellfie, podem ser consultados no Apêndice P.

Os procedimentos de conversão dos vocabulários para o Apache Jena Fuseki e consulta no Skosmos, podem ser consultados no Apêndice Q.

Até o momento 3 vocabulários encontram-se disponíveis para consulta: o Tesouro Cade, Entidades do VCE e Localidade do VCE.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.27/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.

É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

#### 4. ATIVIDADES EM ANDAMENTO

As atividades em andamento compreendem as descritas na Seção 3 que serão cíclicas até o término de construção dos vocabulários.

Há previsão de revisão das entidades inseridas no Tesouro Cade em função da transição de governo que pode impactar a estrutura já validada.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.28/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## 5. REFERÊNCIAS

APACHE SOFTWARE FOUNDATION (ASF). Apache Shiro (Site). 2022. **Section** Apache Shiro Architecture. Disponível em: <https://shiro.apache.org/architecture.html>. Acesso em: 10 dez. 2022.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro: ABNT, 1992. 4 p.

CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. **Regimento interno do Cade**. 5. ed. Brasília: Cade, 2021. 66 p.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-1**: information and documentation - thesauri and interoperability with other vocabularies - part 1: thesauri for information retrieval. Geneva, Switzerland: ISO, 2011. 152 p.

MARRARA, Thiago. Acordos no Direito da Concorrência. **Revista de Direito da Concorrência**, Brasília, v. 8, n. 2, p. 78-103, dez. 2020. Disponível em: <https://revista.cade.gov.br/index.php/revistadedefesadaconcorrenca/article/view/451/352>.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **ANSI/NISO Z39.19-2005 (R2010)**: guidelines for the construction, format, and management of monolingual controlled vocabularies. Baltimore, USA: NISO, 2010. 172 p. Disponível em: [https://groups.niso.org/apps/group\\_public/download.php/12591/z39-19-2005r2010.pdf](https://groups.niso.org/apps/group_public/download.php/12591/z39-19-2005r2010.pdf).

SHINTAKU, Milton *et al.* **Guia sobre a construção de tesouros**. Brasília: Ibict, 2021. 43 p. Disponível em: <https://ridi.ibict.br/handle/123456789/1167>.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.29/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## APÊNDICES

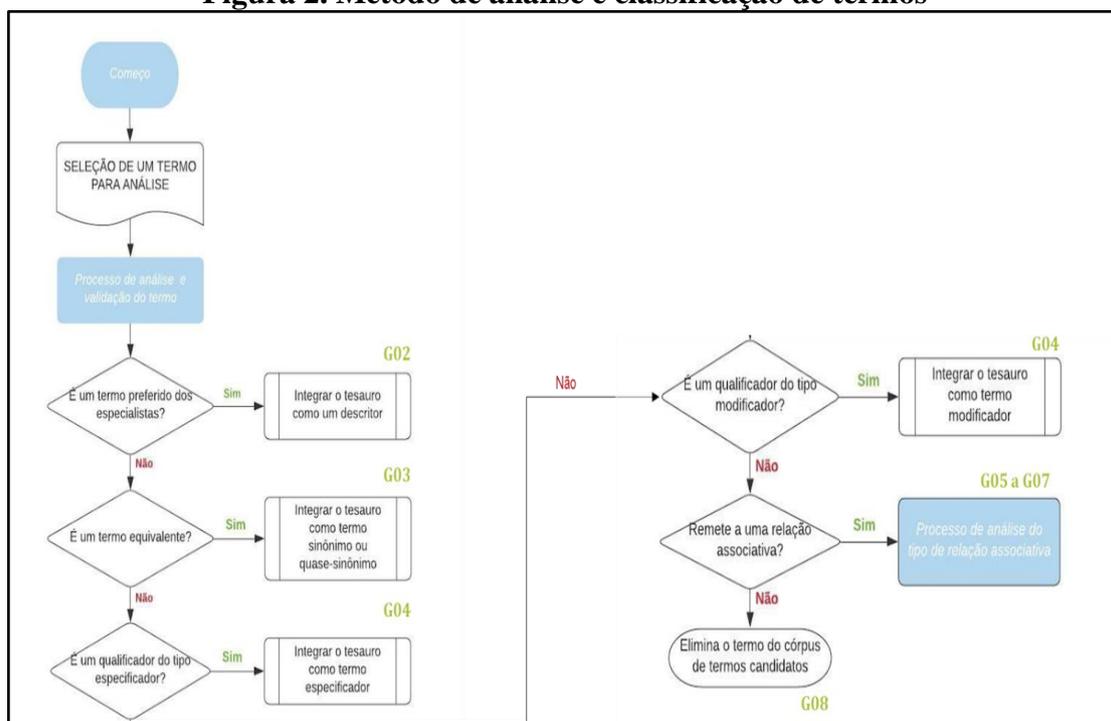
### APÊNDICE A: Método para construção de Tesouros

Conforme apresentado no RT2, para a realização da análise de termos elaborou-se um método que sintetiza as técnicas previstas para construção de tesouros nas referências: ABNT (1992), NISO (2010), ISO (2011) e Shintaku *et al.* (2021).

O método é composto por um conjunto de questões que devem ser realizadas sobre cada termo em análise para se identificar a sua natureza. Em função das respostas obtidas, parte-se para guias específicos que irão nortear a classificação de cada termo a compor o tesouro.

O método completo é composto de 8 Guias destinados a apoiar as análises para as seguintes classificações: i) classificação de descritores; ii) classificação de termos equivalentes; iii) classificação de qualificadores (especificadores e modificadores); iv) classificação de relações associativas (equivalência, hierárquicas e relacional) e; v) identificação de termos a serem eliminados do corpus. A Figura 2 ilustra o método estabelecido.

**Figura 2. Método de análise e classificação de termos**



Fonte: Elaborado pelos autores (2022).

Os guias específicos (sinalizados por G1 a G8 na Figura 1), e as questões de análise relacionadas a cada nível de classificação, são apresentadas nas Figuras 3 a 10, a seguir.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.30/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.

É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

Recordamos que o método e os guias já foram descritos no RT2, sendo, portanto, apenas indicados nesta etapa para auxiliar na compreensão do fluxo de atividades atual.

**Figura 3. Guia 1: Conjunto de questões norteadoras**

G01	SELEÇÃO E ANÁLISE DE TERMOS CANDIDATOS							
Termo candidato	CRITÉRIOS DE ANÁLISE							
	É preferido?	É equivalente?	É qualificador?		É relação associativa?			Eliminar do cópuz
			Especificador?	Modificador?	Equivalência	Hierárquica	Associativa	
Termo do BD	VER G02	VER G03	VER G04	---	VER G03	VER G05	VER G06	VER G08

Fonte: Elaborado pelos autores (2022).

**Figura 4. Guia 2: Classificação de descritores**

G02	CLASSIFICAÇÃO DE DESCRITORES
Checklist	REGRAS A SEREM APLICADAS
	1 - Caracteriza um dos tipos permitidos?  Tipos permitidos: - substantivo (ALIMENTO); - adjetivo substantivado (CONIVÊNCIA); - verbo substantivado (ARTICULAÇÃO); - estrangeirismo incorporado ao vocabulário e práticas do domínio (OUTDOOR); - estrangeirismo já aportuguesado (BALÉ, XAMPU); - formato por extenso de uma abreviação; - nomes de lugares que possuam variações por idiomas.
	2 - Não caracteriza tipos não permitidos?  Tipos não permitidos: a) adjetivos puros; b) verbos puros.
	3 - Foram aplicadas as regras de normalização?  - Padrão: masculino singular - Exceção 1: termos unicamente femininos (AGRAVANTE); - Exceção 2: conjuntos de elementos (DADOS PESSOAIS, RELAÇÕES INTERNACIONAIS);
	4 - Foi possível encontrar uma definição para o termo?  Materiais a serem utilizados por ordem (disponíveis na pasta 6- Construção TCADE, no drive): - F11 - Cade Mecum - F12 - Regimento Interno do CADE - F13 - Cartilha do CADE - F18 - Os fundamentos do antitruste
	5 - Há termos não-preferenciais relacionados? (VER G03)
	6 - Expressa um conceito de forma única? (VER G07)
	7 - Representar na ferramenta de edição  - Rótulo do termo - Definição - Nota de escopo - Designação do termo como um conceito
	8 - Preparar relatório de validação de especialistas de domínio

Fonte: Elaborado pelos autores (2022).

**Figura 5. Guia 3: Classificação de equivalências**

<b>G03</b>	<b>CLASSIFICAÇÃO DE EQUIVALÊNCIAS</b>
<b>Checklist</b>	<b>REGRAS A SEREM APLICADAS</b>
	1 - O termo representa um conceito para o qual já se estabeleceu um descritor?
	2 - Foi identificado o tipo de equivalência? a) Entre termos sinônimos Ex. Pedra preciosa, Esmeralda, Diamante, Turmalina b) Estrangeirismo e suas traduções (estrangeirismo já incorporado ao jargão da área e suas práticas deve ser designado descritor) Ex. Milkshake, Outdoor c) Abreviações Ex. CADE d) Nomes de lugares e suas variações por idiomas Ex. Nova Iorque e New York
	3 - Foi apresentada a lista de termos para que o especialista do domínio indique qual dentre eles é o termo preferido?
	4 - Foram estabelecidas as relações de equivalências entre o termo preferido e TODOS os termos não preferidos? Ex. PEDRA PRECIOSA UP Esmeralda UP Diamante UP Turmalina
	5 - Foram estabelecidas as relações de equivalência entre cada termo não-preferido e o termo preferido? Ex. Esmeralda USE PEDRA PRECIOSA Diamante USE PEDRA PRECIOSA Turmalina USE PEDRA PRECIOSA

Fonte: Elaborado pelos autores (2022).

**Figura 6. Guia 4: Classificação de qualificadores**

<b>G04</b>	<b>CLASSIFICAÇÃO DE QUALIFICADORES</b>
<b>Checklist</b>	<b>REGRAS A SEREM APLICADAS</b>
	1 - O termo elimina ambiguidade entre homônimos (termos polissêmicos: termos similares que possuem sentidos distintos).  Ocorrem em dois tipos: - Homógrafos: mesma escrita. Ex.: MERCÚRIO (PLANETA), MERCÚRIO (METAL) e MERCÚRIO (MITOLOGIA); - Homófonos: mesma fonética. Ex. CESSÃO (DE BENS), SEÇÃO (DE REFERÊNCIA), SESSÃO (DE CINEMA).
	2 - O termo foi classificado em lista específica para qualificadores?
	3 - A lista de qualificadores foi validada pelos especialistas do domínio?

Fonte: Elaborado pelos autores (2022).

**Figura 7. Guia 5: Classificação de relações hierárquicas**

<b>G05</b>	<b>CLASSIFICAÇÃO DE RELAÇÕES HIERÁRQUICAS</b>
<b>Checklist</b>	<b>REGRAS A SEREM APLICADAS</b>
	1 - Foi identificado o tipo de relação hierárquica entre os termos?  - Genérica (entre categorias e seus membros) - Ex. Pássaros (TG) e Papagaios (TE)  - Todo-parte (parte de uma entidade ou sistema que pertença a um todo particular) 4 classes principais de aplicação: i) sistemas e órgãos do corpo; ii) localizações geográficas; iii) disciplinas ou campos de discurso; iv) estruturas sociais e hierárquicas Ex. TGP - CIÊNCIAS TEP - BIOLOGIA TEP - BOTÂNICA  - Instância (entre uma categoria e uma instância desta categoria - geralmente um nome próprio) Ex.: TGI - REGIÕES MONTANHOSAS TEI - ALPES HIMALAIAS
	2 - Foram estabelecidas as relações hierárquicas inversas para as relações hierárquicas identificadas?
	3 - Foram identificados relacionamentos poli-hierárquicos?  - Um mesmo termo estabelece relação com mais de uma categoria Ex1. TG: INSTRUMENTO DE CORDA TE: PIANO Ex2. TG: INSTRUMENTO DE PERCUSSÃO TE: PIANO
	4 - Foram estabelecidas as relações poli-hierárquicas em ambos os sentidos?
	5 - Foram identificados rótulos de nó? (temas=facetar) - termos que agregam assuntos - mostram princípios de divisão entre termos irmãos - possuem a mesma função dos termos gerais mas não devem ser usados como descritores - devem ser apresentados entre colchetes, itálico, etc. para distinção dos descritores
	6 - As relações hierárquicas foram apresentadas para validação do especialista do domínio?

Fonte: Elaborado pelos autores (2022).

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.32/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

**Figura 8. Guia 6: Classificação de relações associativas**

G06	CLASSIFICAÇÃO DE RELAÇÕES ASSOCIATIVAS
Checklist	REGRAS A SEREM APLICADAS
	1 - Foi identificado que a relação não é do tipo equivalente, nem hierárquica?
	2 - A relação estabelece vínculo semântico entre termos de mesma hierarquia (compartilham o mesmo termo geral) ou hierarquias diferentes? Ex. CÉLULAS e CITOLOGIA
	3 - Foram estabelecidas as relações associativas entre todos os termos relacionados, em ambos os sentidos? Ex. CÉLULAS TR: CITOLOGIA
	4 - Foram estabelecidas as relações associativas em ambos os sentidos? Ex. CITOLOGIA TR: CÉLULAS
	5 - As relações identificadas foram apresentadas para validação do especialista de domínio?

Fonte: Elaborado pelos autores (2022).

**Figura 9. Guia 7: Identificação de categorias**

G07	CATEGORIAS
Checklist	REGRAS A SEREM APLICADAS
	A partir da identificação de um conceito, foi possível identificar uma categoria à qual ele remeta?
	* Para Aristóteles, aquilo que pode ser dito sobre algo, ou a respeito de algo estabelece um método para distinguir as categorias. * Unidades que agrupam características comuns a vários conceitos
	Tipos de categorias - Coisas e as suas partes físicas (Ex. Membros do corpo) - Materiais (Ex. Borracha, adesivo) - Atividades e processos (Ex. Fertilização, Acondicionamento) - Eventos e ocorrências (Ex. Pandemias, Guerras Cívicas) - Entidades abstratas, propriedades de pessoas, coisas, materiais ou ações (Ex. Personalidade, Elasticidade) - Disciplinas ou campos temáticos (Defesa da concorrência) - Unidades de medida (Ex. Quilômetros, Hertz) - Tipos de pessoas e organizações (Ex. Organizações governamentais, pessoa com deficiência) - Entidades únicas expressas por nomes próprios (pessoas, documentos e objetos) podem ser inseridas no tesouro para efeito de registro. Quando não forem assunto principal de um documento devem constar em um vocabulário auxiliar, mas não na estrutura do tesouro. Caracterizam-se em três tipos: i) lugares; ii) objetos específicos, recursos topográficos e outras entidades; iii) pessoas físicas, cargos organizacionais e entidades coletivas (Ex. Aristóteles, Diretor geral da Organização Mundial da Saúde, Organização Mundial da Saúde)

Fonte: Elaborado pelos autores (2022).

**Figura 10. Guia 8: Identificação de termos a serem eliminados do corpúsculo**

G08	ELIMINAÇÃO DO TERMO DO CÓRPUS
Checklist	REGRAS A SEREM APLICADAS
	1 - Registrar o termo na aba de termos eliminados
	2 - Registrar o motivo da exclusão a) G02-2a: adjetivos puros b) G02-2b: verbos puros c) G02-5a: o termo remete a mais de um conceito c) G08-1: Não pertence ao domínio CADE (Categorias adjacentes) d) G08-2: Não encontrado na literatura da Defesa Econômica e) G08-3: Grafia incorreta
	3 - Analisar termos equivalentes 3.1 - Excluir os termos equivalentes e suas relações

Fonte: Elaborado pelos autores (2022).

## APÊNDICE B: Instalação e configuração do Web Protégé

### Procedimentos de instalação e configuração do Web Protégé

O *Web Protégé* é uma aplicação web desenvolvida em Java, seus pré-requisitos são basicamente o Java, na versão 11 e um servidor de aplicação com suporte a Servlet (“Servlets”), como por exemplo o Apache Tomcat, que é o servidor recomendado, e o Jetty (“Eclipse Jetty”).

#### Pré-instalação:

1. Configurações dos diretórios com permissões de leitura e escrita (padrão recomendado)
  - a. **Criar o diretório de dados**
    - i. **Windows:** c:\srv\webprotege
    - ii. **Linux:** /srv/webprotege
  - b. **Criar o diretório de configurações**
    - i. Existem dois arquivos de configuração que por padrão ficam em WEB-INF/classes:
      1. **webprotege.properties** - possui as configurações do diretório de dados e da conexão com o banco de dados;
      2. **mail.properties** - possui as configurações do servidor de e-mail;
    - ii. Criar o diretório de configuração
      1. Windows: C:\ProgramData\WebProtege;
      2. Linux: /etc/webprotege;
    - iii. Criar uma cópia do arquivo **webprotege.properties** e **mail.properties** no diretório de configuração e tenha certeza que ele pode ser lido pelo Tomcat ou Jetty;
  - c. **Criar o diretório de logs**
    - i. **Windows:** C:\ProgramData\WebProtege\Logos;
    - ii. **Linux:** /var/log/webprotege;

#### Instalação e configuração:

- d. Com o Tomcat ou Jetty parado;
- e. Baixa o arquivo .war (“Releases WebProtege”);
- f. Altere o nome do arquivo para **webprotege.war**;

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.34/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

- g. Copie o **webprotege.war** para o diretório webapps do Tomcat ou para o similar no Jetty;
- h. Reinicie o Tomcat ou Jetty;
- i. Acesse o endereço <https://localhost:8080/webprotege>

Iniciando o *Web Protégé* e configurando a conta do Administrador

- j. Baixar a ferramenta WebProtégé CLI (Command Line Tool) (“WebProtégé CLI”);
- k. Abrir o terminal no diretório onde baixou a ferramenta CLI;
- l. Executar o comando:
  - i. **java -jar webprotege-cli.jar create-admin-account**
  - ii. Siga as instruções da ferramenta e crie um usuário e senha
- m. Acessar o endereço <https://localhost:8080/webprotege> e entrar com as credenciais criadas no passo anterior
- n. Acessar o endereço <https://my.server.address:8080/webprotege#application/settings> para finalizar as configurações da ferramenta.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.35/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## APÊNDICE C: Instalação e configuração do SKOSMOS

### Procedimentos de instalação e configuração do SKOSMOS

#### Passos para instalação:

Pré Requisitos: Apache Jena Fuseki, Docker, Docker Compose (“Docker”) e Git

1. Baixar o código fonte: <https://github.com/NatLibFi/Skosmos>
2. Incluir as configurações do vocabulário no arquivo  
Skosmos/dockerfiles/config/config-docker-compose.ttl
3. Executar o comando do docker compose para subir o ambiente  
Entre no diretório Skosmos/dockerfiles  
O arquivo docker-compose.yml contém as configuração do ambiente  
Execute o comando para subir o ambiente  
`docker-compose up -d`
4. Para acessar o Skosmos digite no navegador o endereço  
<http://localhost:9090/>

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.36/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## APÊNDICE D: Instalação e Configuração do Apache Jena Fuseki

### Procedimentos de instalação e configuração do Apache Jena Fuseki

#### Passos para instalação:

- Baixar e descompacte da distribuição: <https://jena.apache.org/download/>.
- Copiar o arquivo WAR para o diretório “webapp” do Apache Tomcat (“Apache Tomcat”), sob o nome 'fuseki';
- Se o usuário sob o qual o Apache Tomcat está sendo executado não tiver acesso de gravação a /etc, certifique-se de definir a variável de ambiente FUSEKI\_BASE, enquanto o valor deve ser um diretório no qual o usuário que executa o Apache Tomcat possa gravar;
- Em um navegador, acesse <http://localhost:8080/fuseki>, detalhes como número da porta dependem da configuração do Apache Tomcat;
- A tela inicial do Apache Jena Fuseki será exibida na Figura 11.

**Figura 11: Página inicial do Apache Jena Fuseki**



Fonte: Capturado pelos autores (2022).

Outras informações sobre configurações do Apache Jena Fuseki podem ser encontradas em: (<https://jena.apache.org/documentation/fuseki2/>).

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.37/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## APÊNDICE E: Instalação e configuração do Apache Shiro

### Instalação e configuração do Apache Shiro

A configuração do gerenciador de segurança do Apache Shiro baseada em texto é através de um arquivo no formato INI. O INI é fácil de ler, fácil de configurar e simples de configurar e se adapta bem à maioria dos aplicativos. A seguir temos dois exemplos do arquivo INI.

#### Configuração padrão para LDAP

```
[main]
# Development
ssl.enabled = false

ldapRealm = org.apache.shiro.realm.ldap.DefaultLdapRealm
ldapRealm.userDnTemplate = uid={0},ou=users,dc=example,dc=org
ldapRealm.contextFactory.url = ldap://ldapHost:389
ldapRealm.contextFactory.authenticationMechanism = DIGEST-MD5
ldapRealm.contextFactory.environment[some.obscure.jndi.key] = ssl

securityManager.realms = $ldapRealm

[urls]
/login.html = authc
/logout = logout
/assets/** = anon
/fonts/** = anon

/**=authc
```

#### Configuração para Active Directory

```
[main]

shiro.loginUrl = /login.html

#Para redirecionar uma página diferente de "/"
## authc.successUrl = /whatever

cacheManager = org.apache.shiro.cache.MemoryConstrainedCacheManager
securityManager.cacheManager = $cacheManager

#ACTIVE
contextFactory = org.apache.shiro.realm.ldap.JndiLdapContextFactory
contextFactory.url = ldap://cade.gov.br:389
```

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.38/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

```
contextFactory.systemUsername = USUARIO@cade.gov.br
contextFactory.systemPassword = SENHA

ldapRealm = org.apache.shiro.realm.activedirectory.ActiveDirectoryRealm
ldapRealm.ldapContextFactory = $contextFactory
ldapRealm.searchBase = "CN=CADE,dc=cade,dc=gov,dc=br"

securityManager.realms = $ldapRealm

[urls]
/login.html = authc
/logout = logout
/assets/** = anon
/fonts/** = anon

/**=authc
```

## APÊNDICE F: Procedimentos de instalação do Web Protégé no Cade

### Procedimentos para replicação da instalação do Web Protégé no Cade

O ambiente do *Web Protégé* no Cade foi configurado utilizando o Docker, que é um sistema de virtualização que gerencia containers e seus recursos de forma isolada.

#### Passos para instalação:

- Pré Requisitos: Docker, Docker Compose (“Docker”), Git
- Baixar o código fonte do repositório git do Cade
  - <https://gitlab.cade.gov.br/cade/busca-juridica/ontologia/-/tree/develop/01-Fontes>
- Entrar no diretório **ontologia/01-Fontes**
- Executar o comando para subir o ambiente configurado no arquivo docker-compose.yml
  - docker-compose -p ontologia up -d
- Observação: se necessário altere o caminho dos volumes no arquivo docker-compose.yml

```

version: "3.7"

services:
  wp_mongo:
    container_name: wp_mongo
    hostname: wp_mongo
    image: mongo:4.1-bionic
    restart: unless-stopped
    volumes:
      - /var/docker_data/mongo/db:/data/db

  wp_mongo-express:
    container_name: wp_mongo-express
    hostname: wp_mongo-express
    image: mongo-express:1.0.0-alpha.4
    restart: unless-stopped
    ports:
      - 8091:8081
    environment:
      ME_CONFIG_MONGODB_ADMINUSERNAME: root
      ME_CONFIG_MONGODB_ADMINPASSWORD: senhaaqui
      ME_CONFIG_MONGODB_URL: mongodbs://wp_mongo:27017/

  C:
    container_name: wp_webprotege
    hostname: wp_webprotege
    
```

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.40/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

```

build: ./ontologia/01-Fontes/webprotege
depends_on:
  - wp_mongo
restart: unless-stopped
environment:
  - webprotege.mongodb.host=wp_mongo
  - CATALINA_OPTS=-Xms512M -Xmx2g

volumes:
  - /var/docker_data/protege:/srv/webprotege
ports:
  - 8090:8080
    
```

- Entrar no container **wp\_webprotege**, do WebProtégé para criar o usuário ADMIN e executar o comando:  
 docker exec -it web\_webprotege java -jar /webprotege-cli.jar create-admin-account
- Para finalizar as configurações, depois de fazer o primeiro login
  - [http://NOME\\_MAQUINA:PORTA/#application/settings](http://NOME_MAQUINA:PORTA/#application/settings)
- Configurações obrigatórias
  - Nome da aplicação;
  - E-mail para notificações do sistema;
  - Esquema HTTP ou HTTPS;
  - Host - Cadastrar o DNS criado para o sistema;
  - Port - Setar a porta, caso não seja a padrão. Exemplo: 80;
  - Habilitar criação de projetos;
  - Habilitar upload de projetos.

O Web Protégé demanda a instalação, configuração e disponibilização do MongoDB, um banco de dados que utiliza o modelo de documentos, no padrão JSON para armazenamento de dados.

No ambiente de edição foi incluído o software mongo-express de administração com interface gráfica baseada na web para visualização e manipulação do banco de dados se necessário em ambiente de Testes/Homologação. Documentações sobre o MongoDB podem ser encontradas em: <https://github.com/mongo-express/mongo-express> e <https://www.mongodb.com/pt-br/what-is-mongodb>.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.41/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
 É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## Configurações necessárias à utilização do Web Protégé pelos pesquisadores:

Acessar “Project Preferences” e configurar:

1. Box Preferências para entidades novas: IRI Prefix, IRI Suffix e Digit count, conforme ilustrado na Figura 13;
2. Box Configurações de linguagem para novas entidades: Annotation property e Language Tag, conforme ilustrado na Figura 14;
3. Box Configurações de exibição de nomes: Display name property and language priority, conforme ilustrado na Figura 15.

**Figura 13: Preferências para entidades novas**



**Preferências para Entidades Novas**

**IRI Prefix**  
http://cadethes/cade\_

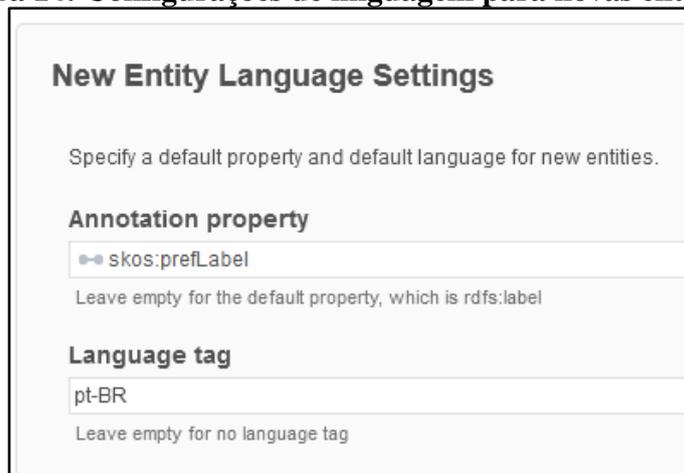
**IRI Suffix**  
Auto-generated OBO Style Id

**Digit count**  
7

**User specific ranges**  
Type user name

Fonte: Capturada pelos autores (2022).

**Figura 14: Configurações de linguagem para novas entidades**



**New Entity Language Settings**

Specify a default property and default language for new entities.

**Annotation property**  
skos:prefLabel  
Leave empty for the default property, which is rdfs:label

**Language tag**  
pt-BR  
Leave empty for no language tag

Fonte: Capturada pelos autores (2022).

**Figura 15: Configurações de exibição de nomes**

### Display Name Settings

Specify a list of annotation property and language pairs that should be used for entity display n  
properties and languages that appear closer to the bottom of the list.

**Display name property and language priority**

↑
↓

→ rdfs:label	en	✕
→ skos:prefLabel	pt-BR	✕
→ http://www.w3.org/2004/02/skos/core#altLabel	pt-BR	✕
→ rdfs:label	pt-BR	✕

+

**Display Name Fallback**

Fall back to IRI local name  
If an entity does not have any annotations that match the above property language pairs then checkin

Reset to referenced languages

Resets the list of display property language pairs to the list of pairs that are used in the project ordered by

Fonte: Capturada pelos autores (2022).

## APÊNDICE G: Procedimentos de instalação do SKOSMOS no Cade

### Procedimentos para replicação da instalação do SKOSMOS no Cade.

Optou-se por executar o Skosmos em um contêiner do Docker. Os procedimentos necessários para a replicação são descritos a seguir.

- Executar o comando abaixo para baixar o ambiente:

docker-compose down

```
:tculocthes a skosmos:Vocabulary, void:Dataset ;
dc:title "VCE - TCU Localidades"@pt-BR ;
skosmos:shortName "VCELOC";
dc:subject :cat_general ;
void:uriSpace "http://tculocthes/";
skosmos:language "pt", "pt-BR", "en";
skosmos:defaultLanguage "pt-BR";
skosmos:showTopConcepts true ;
skosmos:defaultSidebarView 'alphabetical';
skosmos:fullAlphabeticalIndex true ;
void:sparqlEndpoint <http://fuseki-cache:80/skosmos/sparql> ;
skosmos:sparqlGraph <http://tculocthes/> .
```

- Executar o comando para subir o ambiente configurado no arquivo docker-compose.yml
  - docker-compose -p ontologia up -d
- Observação: se necessário altere o caminho dos volumes no arquivo docker-compose.yml

```
version: "3.7"

services:
  skosmos_fuseki:
    container_name: skosmos_fuseki
    hostname: fuseki
    build: ./ontologia/01-Fontes/ApacheJena/dockerfiles
    restart: unless-stopped
    environment:
      - ADMIN_PASSWORD=cade#ontologia#2022
      - JVM_ARGS=-Xmx2g
    ports:
      - 8092:3030
    volumes:
      - /var/docker_data/fuseki:/fuseki
      - type: bind
        source: /var/docker_data/ontologia/01-Fontes/Skosmos/dockerfiles/config/skosmos.ttl
        target: /fuseki/configuration/skosmos.ttl
      - type: bind
        source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/login.html
```

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.44/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

```

target: /jena-fuseki/webapp/login.html
- type: bind
source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/assets/layout-light.css
target: /jena-fuseki/webapp/assets/layout-light.css
- type: bind
source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/assets/logo-vcde-ontologia-transparente.png
target: /jena-fuseki/webapp/assets/logo-vcde-ontologia-transparente.png
- type: bind
source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/assets/theme-light.css
target: /jena-fuseki/webapp/assets/theme-light.css
- type: bind
source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/fonts/open-sans-v18-latin-ext_latin-600.woff2
target: /jena-fuseki/webapp/fonts/open-sans-v18-latin-ext_latin-600.woff2
- type: bind
source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/fonst/open-sans-v18-latin-ext_latin-regular.woff2
target: /jena-fuseki/webapp/fonts/open-sans-v18-latin-ext_latin-regular.woff2
- type: bind
source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/fonst/open-sans-v18-latin-ext_latin-regular.woff
target: /jena-fuseki/webapp/fonts/open-sans-v18-latin-ext_latin-regular.woff
- type: bind
source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/fonst/open-sans-v18-latin-ext_latin-regular.ttf
target: /jena-fuseki/webapp/fonts/open-sans-v18-latin-ext_latin-regular.ttf
skosmos_fuseki-cache:
container_name: skosmos_fuseki-cache
hostname: skosmos_fuseki-cache
image: varnish
ports:
- 8093:80
restart: unless-stopped
volumes:
- type: bind
source: /var/docker_data/ontologia/01-Fontes/Skosmos/dockerfiles/config/varnish-default.vcl
target: /etc/varnish/default.vcl
skosmos:
container_name: skosmos
hostname: skosmos
build:
context: /var/docker_data/ontologia/01-Fontes/Skosmos
dockerfile: dockerfiles/Dockerfile.ubuntu
args:
SMTP_HOST: smtp.cade.gov.br
SMTP_PORT: 25
ports:
- 8094:443
restart: unless-stopped
environment:

```

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.45/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

```
- FEEDBACK_MAIL_TO=vcde@cade.gov.br
volumes:
- type: bind
  source: /var/docker_data/ontologia/01-Fontes/Skosmos/dockerfiles/config/config-docker-compose.ttl
  target: /var/www/html/config.ttl
```

Maiores informações sobre instalação do Skosmos com Fuseki no Docker podem ser encontradas em: (<https://github.com/NatLibFi/Skosmos/wiki/Install-Skosmos-with-Fuseki-in-Docker>).

## APÊNDICE H: Arquitetura Apache Jena Fuseki, Web Protégé e SKOSMOS no Cade

version: "3.7"

services:

**wp\_mongo:**

container\_name: wp\_mongo  
hostname: wp\_mongo  
image: mongo:4.1-bionic  
restart: unless-stopped  
volumes:  
- /var/docker\_data/mongo/db:/data/db

**wp\_mongo-express:**

container\_name: wp\_mongo-express  
hostname: wp\_mongo-express  
image: mongo-express:1.0.0-alpha.4  
restart: unless-stopped  
ports:  
- 8091:8081  
environment:  
ME\_CONFIG\_MONGODB\_ADMINUSERNAME: root  
ME\_CONFIG\_MONGODB\_ADMINPASSWORD: cade#ontologia#2022  
ME\_CONFIG\_MONGODB\_URL: mongodb://wp\_mongo:27017/

**wp\_webprotege:**

container\_name: wp\_webprotege  
hostname: wp\_webprotege  
build: ./ontologia/01-Fontes/webprotege  
depends\_on:  
- wp\_mongo  
restart: unless-stopped  
environment:  
- webprotege.mongodb.host=wp\_mongo  
- CATALINA\_OPTS=-Xms512M -Xmx2g

volumes:

- /var/docker\_data/protege:/srv/webprotege

ports:

- 8090:8080

**skosmos\_fuseki:**

container\_name: skosmos\_fuseki  
hostname: fuseki  
build: ./ontologia/01-Fontes/ApacheJena/dockerfiles  
restart: unless-stopped  
environment:

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.47/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

```

- ADMIN_PASSWORD=cade#ontologia#2022
- JVM_ARGS=-Xmx2g
ports:
- 8092:3030
volumes:
- /var/docker_data/fuseki:/fuseki
- type: bind
  source: /var/docker_data/ontologia/01-Fontes/Skosmos/dockerfiles/config/skosmos.ttl
  target: /fuseki/configuration/skosmos.ttl
- type: bind
  source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/login.html
  target: /jena-fuseki/webapp/login.html
- type: bind
  source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/assets/layout-light.css
  target: /jena-fuseki/webapp/assets/layout-light.css
- type: bind
  source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/assets/logo-vcde-ontologia-transparente.png
  target: /jena-fuseki/webapp/assets/logo-vcde-ontologia-transparente.png
- type: bind
  source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/assets/theme-light.css
  target: /jena-fuseki/webapp/assets/theme-light.css
- type: bind
  source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/fonts/open-sans-v18-latin-ext_latin-600.woff2
  target: /jena-fuseki/webapp/fonts/open-sans-v18-latin-ext_latin-600.woff2
- type: bind
  source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/fonst/open-sans-v18-latin-ext_latin-regular.woff2
  target: /jena-fuseki/webapp/fonts/open-sans-v18-latin-ext_latin-regular.woff2
- type: bind
  source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/fonst/open-sans-v18-latin-ext_latin-regular.woff
  target: /jena-fuseki/webapp/fonts/open-sans-v18-latin-ext_latin-regular.woff
- type: bind
  source: ./ontologia/01-Fontes/ApacheJena/dockerfiles/web/fonst/open-sans-v18-latin-ext_latin-regular.ttf
  target: /jena-fuseki/webapp/fonts/open-sans-v18-latin-ext_latin-regular.ttf

```

**skosmos\_fuseki-cache:**

```

container_name: skosmos_fuseki-cache
hostname: skosmos_fuseki-cache
image: varnish
ports:
- 8093:80
restart: unless-stopped
volumes:
- type: bind

```

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.48/95
---------------	---------------------	---	-----------

Confidencial.

```

source: /var/docker_data/ontologia/01-Fontes/Skosmos/dockerfiles/config/varnish-default.vcl
target: /etc/varnish/default.vcl

```

**skosmos:**

container\_name: skosmos

hostname: skosmos

build:

context: /var/docker\_data/ontologia/01-Fontes/Skosmos

dockerfile: dockerfiles/Dockerfile.ubuntu

args:

SMTP\_HOST: smtp.cade.gov.br

SMTP\_PORT: 25

ports:

- 8094:443

restart: unless-stopped

environment:

- FEEDBACK\_MAIL\_TO=vcde@cade.gov.br

volumes:

- type: bind

source: /var/docker\_data/ontologia/01-Fontes/Skosmos/dockerfiles/config/config-docker-compose.ttl

target: /var/www/html/config.ttl

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.49/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.

É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

**APÊNDICE I: Relação de fontes informacionais atualizadas**
**Fontes de informação para estudo do domínio CADE,  
seleção de termos e definições.**
**Pesquisadores da frente de ontologia**
**Especialistas do CADE: Débora Lins e Nóbrega e Iara do Espírito Santo**

<b>Controle de versão deste documento</b>			
<b>Versão</b>	<b>Data</b>	<b>Motivo</b>	<b>Responsável</b>
1.0.5	04/08/2022	Inclusão PAN12.7	Caroline
1.0.4	27/07/2022	Inclusão PAN13	Jeanne
1.0.3	22/07/2022	Inclusão PAN12 a PAN12.6	Izadora
1.0.2	21/07/2022	Inclusão PAN11	Jeanne
1.0.1	21/06/2022	Substituição de Forgioni (2015) por Forgioni (2018). Inclusão de fontes: PAN5.1; PAN5.2	Jeanne e Izadora
1.0	12/05/2022	Inclusão: PAN6; PI6; PI7; VE5; VE6; VE7; VE8. Exclusão: VOC_004; VOC_005 e VOC_006.	Jeanne, atendendo à revisão das especialistas do CADE.
Beta_0.3	28/04/2022	Acréscimo de periódicos e bases de dados	Caroline
Beta_0.2	25/03/2022	Acréscimo de glossários	Caroline
Beta_0.1	17/02/2022	Para validação de especialistas de domínio	Jeanne e Caroline

Apresenta-se no Quadro 1 as fontes informacionais identificadas para estudo do domínio, seleção de termos e definições. O Quadro encontra-se organizado em seis categorias de acordo com o tipo de publicação encontrada (em parêntesis o total das publicações por tipo) a saber: 1) Publicações nacionais (10); ii) Publicações internacionais (7); iii) Vocabulários especializados (13); iv) Periódicos especializados (2); v) Bases de dados especializadas

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.50/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.

É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

(10); e vi) Dicionários linguísticos (6). As numerações destacadas em verde e roxo sinalizam as referências preferenciais. Em roxo sinalizam-se indicações de especialistas do domínio.

**Quadro 2. Fontes informacionais para estudo do domínio**

<b>PUBLICAÇÕES DE AUTORES NACIONAIS (PAN)</b>		
<b>No.</b>	<b>Fonte informacional</b>	<b>Citação</b>
1	CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Cade Mecum</b> : coletânea de normativos brasileiros de defesa da concorrência. Brasília: Cade, 2021a. 450 p.	Cade (2021a)
2	CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Regimento interno do Cade</b> . 5. ed. Brasília: Cade, 2021b. 66 p.	Cade (2021b)
3	CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Cartilha do CADE</b> . Brasília: Cade, 2016.	Cade (2016)
4	FORGIONI, Paula. A. <b>Os fundamentos do antitruste</b> . 10. ed. rev., atual. e ampl. – São Paulo: Thomson Reuters Brasil, 2018.	Forgioni (2018)
5	ESCOLA NACIONAL DE ADMINISTRAÇÃO PÚBLICA (ENAP). <b>Introdução à defesa da concorrência</b> . Brasília: ENAP, 2021. + <a href="#">Apostilas de outros temas relacionados, organizada pelo ENAP</a> + 5.1 ESCOLA NACIONAL DE ADMINISTRAÇÃO PÚBLICA (ENAP). <b>Análise de condutas unilaterais restritivas à concorrência</b> . Brasília: ENAP, 2021. + 5.2 ESCOLA NACIONAL DE ADMINISTRAÇÃO PÚBLICA (ENAP). <b>Prevenção e detecção de cartéis em licitações</b> . Brasília: ENAP, 2021.	ENAP (2021)
6	SILVEIRA, Paulo Burnier. <b>Direito da Concorrência</b> . Rio de Janeiro: Forense, 2021. ISBN: 9788530991159. <a href="https://bit.ly/3P0Gj8W">https://bit.ly/3P0Gj8W</a>	Silveira (2021)
7	MATIAS-PEREIRA, José. Defesa da concorrência e regulação econômica no Brasil. <b>Revista de Administração Mackenzie</b> , v. 5, n. 1, p. 35–55, 2004.	Matias-Pereira (2004)
8	MATIAS-PEREIRA, José. Políticas de defesa da concorrência e de regulação econômica: as deficiências do sistema brasileiro de defesa da concorrência. <b>Revista de Administração Contemporânea</b> , v. 10, n. 2, p. 51-73, abr./jun. 2006.	Matias-Pereira (2006)

9	DRAGO, Bruno De Luca; MORSELLI, Fabianna Vieira Barbosa. Clarifying gun jumping through guidelines: the Brazilian experience. <b>Journal of European Competition Law &amp; Practice</b> , v. 7, n. 2, p. 130–134, 1 fev. 2016.	Drago e Mor-selli (2016)
10	MOTTA, Massimo; SALGADO, Lucia Helena. <b>Política de concorrência</b> : teoria e prática e sua aplicação no Brasil. Rio de Janeiro: Elsevier, 2015.	Motta e Salgado (2015)
11	OLIVEIRA, Gesner; RODAS, João Grandino. <b>Direito e economia da concorrência</b> . 2. ed rev. e atual. São Paulo: Editora Revista dos Tribunais, 2013. 495 p. (Biblioteca de Direito e Economia, v. 3).	Oliveira e Rodas (2013)
12	<p>Guias do Cade</p> <ul style="list-style-type: none"> <li>+ 12 CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Guia para programas de compliance</b>. Brasília: Cade, 2016. 43 p.</li> <li>+ 12.1 CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Guia Programa de Leniência Antitruste do Cade</b>. Brasília: Cade, 2021. 73 p.</li> <li>+ 12.2 CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Guia para análise de atos de concentração horizontal</b>. Brasília: Cade, 2016. 59 p.</li> <li>+ 12.3 CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Guia Recomendações probatórias para propostas de acordo de leniência com o Cade</b>. Brasília: Cade, 2021. 347 p.</li> <li>+ 12.4 CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Guia de Remédios Antitruste</b>. Brasília: Cade, [201-]. 67 p.</li> <li>+ 12.5 CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Guia</b>: Termo de Compromisso de Cessação para casos de cartel. Brasília: Cade, 2017. 132 p.</li> <li>+ 12.6 CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Guia para análise da consumação prévia de atos de concentração econômica</b>. Brasília: Cade, 2015. 15 p.</li> <li>+ 12.7 CONSELHO ADMINISTRATIVO DE DEFESA ECONÔMICA. <b>Guia de combate a cartéis em licitação</b>.</li> </ul>	

	Brasília: Cade, [2020?]. 58 p.	
13	SALOMÃO FILHO, Calixto. <b>Direito concorrencial:</b> as condutas. São Paulo: Malheiros, 2003. 334 p.	Salomão Filho (2003)
<b>PUBLICAÇÕES DE AUTORES INTERNACIONAIS (PI)</b>		
No.	Fonte informacional	Citação
1	ESTADOS UNIDOS. U.S Department of Justice. Federal Trade Commission. <b>Horizontal Merger Guidelines.</b> Washington, DC: DOJ, 2010. Disponível em: <a href="https://www.justice.gov/atr/horizontal-merger-guidelines-08192010">https://www.justice.gov/atr/horizontal-merger-guidelines-08192010</a> . Acesso em: 25 mar. 2022.	Estados Unidos (2010)
2	ESTADOS UNIDOS. U.S Department of Justice. Antitrust Division. <b>Antitrust Division Manual.</b> 5th ed. Washington, DC: DOJ, 2018.	Estados Unidos (2018)
3	ESTADOS UNIDOS. U.S Department of Justice. Federal Trade Commission. <b>Vertical Merger Guidelines.</b> [S. l]: FTC, 2020. Disponível em: <a href="https://www.ftc.gov/system/files/documents/reports/us-department-justice-federal-trade-commission-vertical-merger-guidelines/vertical-merger-guidelines-6-30-20.pdf">https://www.ftc.gov/system/files/documents/reports/us-department-justice-federal-trade-commission-vertical-merger-guidelines/vertical-merger-guidelines-6-30-20.pdf</a> . Acesso em: 25 mar.	Estados Unidos (2020)
4	LIEBESKIND, Richard. <b>Gun-jumping:</b> antitrust issues before closing the merger. Palestra proferida no ABA Section of Business Law, Antitrust Committee, ABA Annual Meeting, 2003. San Francisco, Califórnia, 08 ago. 2003. Disponível em: <a href="https://www.pillsburylaw.com/images/content/1/8/v2/1842/16ADC9E2C53CF6E9F97E3F0A3F6F3242.pdf">https://www.pillsburylaw.com/images/content/1/8/v2/1842/16ADC9E2C53CF6E9F97E3F0A3F6F3242.pdf</a> . Acesso em: 25 mar. 2022.	Liebeskind (2003)
5	ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). Competition Division. <b>Suspensory Effects of Merger Notifications and Gun Jumping.</b> [S. l]: OECD, 2018. Disponível em: <a href="https://one.oecd.org/document/DAF/COMP(2018)11/en/pdf">https://one.oecd.org/document/DAF/COMP(2018)11/en/pdf</a> . Acesso em: 25 mar. 2022.	OECD (2018)
6	Antitrust and Competition Policy - Cheltenham, Inglaterra: Edward Elgar, 2005. 641 p. ISBN: 1843763192 Link acessível: <a href="#">Koha online Catálogo &gt; Detalhes para: Antitrust and Competition Policy/ (cade.gov.br)</a>  Link inacessível: <a href="https://bit.ly/3stmCH2">https://bit.ly/3stmCH2</a>	
7	The Future of Antitrust/: <a href="https://bit.ly/3stmCH2">https://bit.ly/3stmCH2</a>	IBRAC (2020)
<b>VOCABULÁRIOS ESPECIALIZADOS (VE)</b>		

No.	Fonte informacional	Sigla
1	<b>Vocabulário de controle externo do TCU</b> <a href="https://portal.tcu.gov.br/biblioteca-digital/vce-vocabulario-de-controle-externo-do-tribunal-de-contas-da-uniao-8A81881F7AB5B041017B1C4673783D0A.htm">https://portal.tcu.gov.br/biblioteca-digital/vce-vocabulario-de-controle-externo-do-tribunal-de-contas-da-uniao-8A81881F7AB5B041017B1C4673783D0A.htm</a>	VOC_001
2	<b>Vocabulário preliminar do Prof. Valério</b> Compartilhado no ambiente do projeto.	VOC_002
3	<b>Vocabulário preliminar do Eloy</b> Compartilhado no ambiente do projeto.	VOC_003
4	<b>Eurovoc:</b> <a href="http://eurovoc.europa.eu/">http://eurovoc.europa.eu/</a>	VOC_007
5	Dictionary of Trade Terms – Competition Policy: <a href="https://bit.ly/39NyIdU">https://bit.ly/39NyIdU</a>	
6	Glossary of terms used in EU competition policy: <a href="https://bit.ly/3wLhUCI">https://bit.ly/3wLhUCI</a>	
7	Competition Glossary da UNCTAD: <a href="https://bit.ly/39Og0Tt">https://bit.ly/39Og0Tt</a>	
8	BRASIL. Senado Federal. Vocabulário Controlado Básico:VCB. Brasília: Secretaria de Gestão de Informação e Documentação - Sgidoc, 2017. Disponível em: <a href="https://bit.ly/3P6T7uq">https://bit.ly/3P6T7uq</a>	BRASIL (2017)
9	CONCURRENCES. <b>Global dictionary of Competition Law.</b> [S.l]: Concurrences, 2022. Disponível em: <a href="https://www.concurrences.com/en/dictionary">https://www.concurrences.com/en/dictionary</a> . Acesso em 10 mar. 2022	Concurrences (2022)
10	ORGANIZATION OF AMERICAN STATES (OAS). Foreign Trade Information System. <b>Dictionary of trade terms.</b> [S.l]: SICE, 2022. Disponível em: <a href="http://www.sice.oas.org/Dictionary/Dictio_e.asp">http://www.sice.oas.org/Dictionary/Dictio_e.asp</a> . Acesso em: 10 mar.	OAS (2022)
<b>VOCABULÁRIOS ESPECIALIZADOS (VE)</b>		
No.	Fonte informacional	Sigla
11	ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). <b>Glossary of industrial organisation economics and competition law.</b> [S. l]: OECD, 1993. Disponível em: <a href="https://www.oecd.org/competition/abuse/2376087.pdf">https://www.oecd.org/competition/abuse/2376087.pdf</a> . Acesso em: 13 mar. 2022.	OECD (1993)
12	SANDRONI, Paulo. <b>Novíssimo Dicionário de Economia.</b> São Paulo: Editora Best Seller, 1999. Disponível em: <a href="http://www2.fct.unesp.br/docentes/geo/magaldi/GEO_ECONOMICA_2019/dicionario-de-economia-sandroni.pdf">http://www2.fct.unesp.br/docentes/geo/magaldi/GEO_ECONOMICA_2019/dicionario-de-economia-sandroni.pdf</a> . Acesso em: 17 fev. 2022.	Sandroni (1999)

13	SANTOS, W. DOS. <b>Dicionário jurídico brasileiro</b> . Belo Horizonte: Del Rey, 2001. Disponível em: <a href="https://roberiosoares.files.wordpress.com/2010/03/dicionario-juridico.pdf">https://roberiosoares.files.wordpress.com/2010/03/dicionario-juridico.pdf</a> . Acesso em: 17 fev. 2022.	Santos (2001)
----	--	---------------

**PERIÓDICOS ESPECIALIZADOS (PE)**

No.	Fonte informacional	Citação
1	REVISTA DE DEFESA DA CONCORRÊNCIA. Brasília, DF: CADE, 2013- . ISSN 2318-2253. Disponível em: <a href="https://revista.cade.gov.br/index.php/revistadedefesadaconcorrencia">https://revista.cade.gov.br/index.php/revistadedefesadaconcorrencia</a> .	
2	REVISTA DO IBRAC. São Paulo, SP: IBRAC, 1994- . ISSN 1517-1957 versão <i>online</i> . Disponível em: <a href="https://ibrac.org.br/revistas.htm">https://ibrac.org.br/revistas.htm</a> .	

**BASES DE DADOS ESPECIALIZADAS (BE)**

No.	Fonte informacional	Citação
1	EBSCO <a href="http://bit.ly/32UdgLY">http://bit.ly/32UdgLY</a> .	
2	JSTOR <a href="https://www.jstor.org">https://www.jstor.org</a> .	
3	Springer <a href="https://link.springer.com">https://link.springer.com</a> .	
4	Concurrences <a href="https://bit.ly/3dX6zzw">https://bit.ly/3dX6zzw</a> .	
5	Kluwer Law Online <a href="http://www.kluwerlawonline.com">http://www.kluwerlawonline.com</a> .	
6	Kluwer Competition Law <a href="https://www.wolterskluwer.com/en/solutions/kluwercompetitionlaw">https://www.wolterskluwer.com/en/solutions/kluwercompetitionlaw</a> .	
7	Passport Euromonitor International (Euromonitor) <a href="https://www.portal.euromonitor.com">https://www.portal.euromonitor.com</a> .	
8	Global Competition Review – GCR <a href="http://www.globalcompetitionreview.com">http://www.globalcompetitionreview.com</a> .	

**BASES DE DADOS ESPECIALIZADAS (BE)**

No.	Fonte informacional	Citação
9	MLex <a href="http://www.mlex.com/GlobalAntitrust/Home.aspx?siteid=190">http://www.mlex.com/GlobalAntitrust/Home.aspx?siteid=190</a> .	
10	Portal de Periódicos da CAPES <a href="https://www.periodicos.capes.gov.br">https://www.periodicos.capes.gov.br</a> .	
<b>DICIONÁRIOS LINGUÍSTICOS (DL)</b>		
No.	Fonte informacional	Citação
1	MERRIAM-WEBSTER. English Dictionary. 2022. Disponível em: <a href="https://www.merriam-webster.com/dictionary/">https://www.merriam-webster.com/dictionary/</a> . Acesso em: 17 fev. 2022.	Merriam-Webster (2022)
2	MICHAELIS. <b>Dicionário Brasileiro da Língua Portuguesa</b> . Disponível em: <a href="https://michaelis.uol.com.br/moderno-portugues/">https://michaelis.uol.com.br/moderno-portugues/</a> . Acesso em: 17 fev. 2022.	Michaelis (2022)
3	7GRAUS. <b>Dicionário de Sinônimos online de português do Brasil</b> . (2022) Disponível em: <a href="https://www.sinonimos.com.br/">https://www.sinonimos.com.br/</a> . Acesso em: 17 fev. 2022.	7GRAUS (2022)
4	REVERSO TECHNOLOGIES INC. <b>Reverso: Dicionário de sinônimos</b> . 2022. Disponível em: <a href="https://synonyms.reverso.net/dicionario-sinonimos/pt/">https://synonyms.reverso.net/dicionario-sinonimos/pt/</a> . Acesso em 17 fev. 2022.	Reverso (2022)
5	WOXIKON. <b>Dicionário de sinônimos</b> . Disponível em: <a href="https://sinonimos.woxikon.com.br/pt">https://sinonimos.woxikon.com.br/pt</a> . Acesso em: 17 fev. 2022.	Woxikon (2022)
6	POMBO, Rocha. <b>Dicionário de sinônimos da língua portuguesa</b> . 2. ed. Rio de Janeiro: Academia Brasileira de Letras, 2011. Disponível em: <a href="https://academia.org.br/sites/default/files/publicacoes/arquivos/cams-10-dicionario_de_sinonimos_da_lingua_portuguesa-para_internet.pdf">https://academia.org.br/sites/default/files/publicacoes/arquivos/cams-10-dicionario_de_sinonimos_da_lingua_portuguesa-para_internet.pdf</a> . Acesso em: 03 mar. 2022	Pombo (2011)

Fonte: Consolidado pelos autores (2022).

Duas outras listas de bibliografias foram identificadas e serão analisadas caso seja necessário utilizá-las:

- IBRAC - Instituto Brasileiro de Estudos de Concorrência, Consumo e Comércio Internacional: <https://www.ibrac.org.br/bibliografias.htm>;
- Referências indicadas por um ex-superintendente-geral do CADE: <https://www.jota.info/carreira/10-livros-fundamentais-para-atuar-na-area-do-direito-concorrencial-02022018>.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.56/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## Observação da Déborah, via email em 28/07/2022

“Destaco que também pode ser de ajuda na formulação de conceitos alguns artigos e materiais que podem ser obtidos através de pesquisa bibliográfica realizada pela Biblioteca. O formulário para solicitação está disponível na intranet em: <https://intranet.cade.gov.br/pesquisa-bibliografica/public/pesquisa>”

## APÊNDICE J: Mapeamento de constructos de Tesouros para notação SKOS

**Quadro 3: Mapeamento de constructos de Tesouros para notação SKOS**

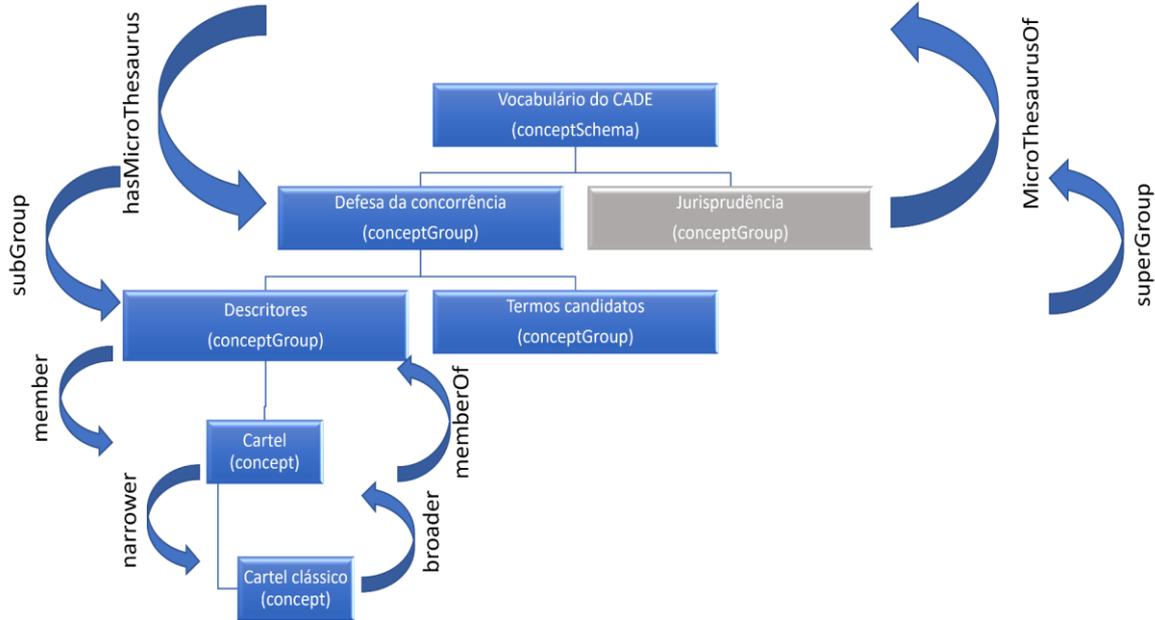
Convenção (Títulos das colunas na planilha)	UNESKOS	Objetivo (o que deve ser registrado)	Procedimentos
<b>Propriedades de anotações</b>			
Descritor	skos:prefLabel	<ul style="list-style-type: none"> <li>Registrar o termo que é preferido pelos especialistas do domínio em estudo.</li> </ul>	<ul style="list-style-type: none"> <li>Registro dos termos preferidos</li> <li>Apenas 1 (e somente 1) para cada língua</li> </ul>
Termo equivalente	skos:altLabel	<ul style="list-style-type: none"> <li>IRI do termo alternativo</li> </ul>	<ul style="list-style-type: none"> <li>Usar rdfs:label quando Type = Não descritor</li> </ul>
Definição (DEF)	skos:definition	<ul style="list-style-type: none"> <li>Definição(ões) para o conceito, precedidas de [DEF1], [DEF2]...[DEFn], para cada definição identificada na literatura.</li> </ul>	<ul style="list-style-type: none"> <li>Inicialmente: reúne todas as definições encontradas na literatura + fontes.</li> <li>Posteriormente: definição real, sem fontes. As fontes usadas para construir a definição deverão ser inseridas na nota bibliográfica.</li> </ul>
Nota de escopo (NE)	skos:scopeNote	<ul style="list-style-type: none"> <li>Situações em que um conceito é utilizado no domínio de especialidade;</li> <li>Anotações mais gerais sobre o conceito.</li> </ul>	
Nota de validação	cade:validation-Note	<ul style="list-style-type: none"> <li>Registrar a data de validação do conceito e histórico.</li> </ul> <p>Formato: [Data-Primeiro nome] &lt;Descrição da mudança&gt;. Incluir fonte bibliográfica se necessário.</p>	<ul style="list-style-type: none"> <li>Preenchido sempre que houver mudança em conceitos de acordo com validação de especialistas do domínio.</li> <li>Visível durante a construção do vocabulário, para todos os colaboradores.</li> </ul>

Nota de homologação	cade:statementNote	<ul style="list-style-type: none"> <li>Data de homologação do conceito.</li> <li>Observações necessárias.</li> </ul>	<ul style="list-style-type: none"> <li>Preenchido apenas quando for concluído.</li> <li>Visível apenas para a frente de ontologia.</li> </ul>
Nota Bibliográfica	cade:bibliographicNote	<ul style="list-style-type: none"> <li>Fontes utilizadas para construção da definição de um conceito;</li> <li>Fontes onde foram encontrados os termos alternativos.</li> </ul>	<ul style="list-style-type: none"> <li>Preenchido apenas quando for elaborada a definição real para o conceito.</li> </ul>
Nota Histórica (NH)	skos:historyNote	<ul style="list-style-type: none"> <li>Mudanças dos conceitos do domínio ao longo do tempo. Exemplo: SEAE.</li> </ul>	<ul style="list-style-type: none"> <li>Gerar um skos:altLabel para cada mudança histórica.</li> </ul>
Nota de Exemplo (NX)	skos:exampleNote	<ul style="list-style-type: none"> <li>Exemplos de uso dos conceitos do domínio de especialidade encontrados na literatura.</li> </ul>	<ul style="list-style-type: none"> <li>Incluir links dos processos administrativos relacionados (se houver).</li> </ul>
Nota Editorial	skos:editorialNote	<ul style="list-style-type: none"> <li>Manutenções realizadas no vocabulário.</li> </ul>	
Nota de mudança	skos:changeNote	<ul style="list-style-type: none"> <li>Mudanças na classificação de conceitos previamente “homologados” no tesouro. Ex.: Um descritor que se torna não-descritor.</li> </ul>	
<b>Propriedades de objetos</b>			
TG	skos:broader	<ul style="list-style-type: none"> <li>Representação de termos gerais.</li> </ul>	
TE	skos:narrower	<ul style="list-style-type: none"> <li>Representação de termos específicos.</li> </ul>	
TR	skos:related	<ul style="list-style-type: none"> <li>Representação de termos relacionados.</li> </ul>	

Estabeleceu-se a regra de não utilizar links encurtados para quaisquer referências a documentos externos disponíveis na Web, em função da incerteza sobre a sua duração e disponibilidade.

## APÊNDICE K: Diretrizes para representação do conhecimento no padrão SKOS

Figura 18: Estrutura projetada para o vocabulário



Fonte: Elaborada pelos autores (2022).

Figura 19: Visão interna das conexões no Web Protégé

Individual: C00001

IRI  
 http://skos.um.es/unescothes/C00001

**Annotations**

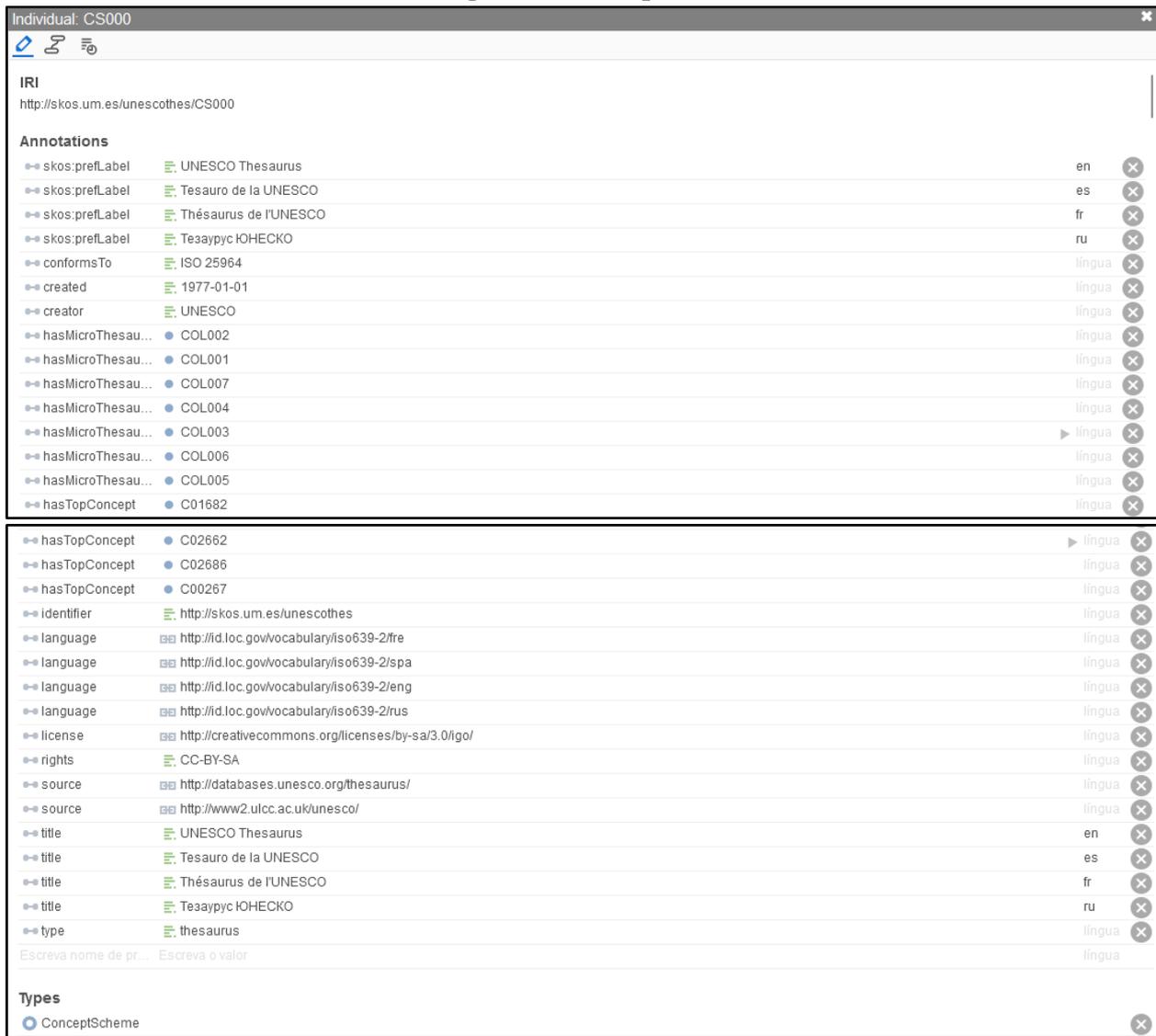
skos:prefLabel	Abandoned children	en	X
skos:altLabel	Deserted children	en	X
skos:prefLabel	Niño abandonado	es	X
skos:altLabel	Abandono de menores	es	X
skos:altLabel	Abandono infantil	es	X
skos:prefLabel	Enfant abandonné	fr	X
skos:altLabel	Enfant délaissé	fr	X
skos:prefLabel	Беспризорные дети	ru	X
skos:altLabel	Брошенные дети	ru	X
broader	C01058	língua	X
inScheme	CS000	língua	X
memberOf	COL420	língua	X
related	C00570	língua	X
related	C03850	língua	X

Escreva nome de pr... Escreva o valor

Fonte: Capturado pelos autores (2022).

Usa **TODOS** os marcadores como propriedades de anotações

Figura 20: ConceptSchema



The screenshot shows a web interface for a ConceptSchema. At the top, it displays the IRI: `http://skos.um.es/unescothes/CS000`. Below this, there are two main sections: 'Annotations' and 'Types'.

**Annotations:**

- skos:prefLabel: UNESCO Thesaurus (en)
- skos:prefLabel: Tesouro de la UNESCO (es)
- skos:prefLabel: Thésaurus de l'UNESCO (fr)
- skos:prefLabel: Тезаурыс ЮНЕСКО (ru)
- conformsTo: ISO 25964 (língua)
- created: 1977-01-01 (língua)
- creator: UNESCO (língua)
- hasMicroThesau...: COL002 (língua)
- hasMicroThesau...: COL001 (língua)
- hasMicroThesau...: COL007 (língua)
- hasMicroThesau...: COL004 (língua)
- hasMicroThesau...: COL003 (língua)
- hasMicroThesau...: COL006 (língua)
- hasMicroThesau...: COL005 (língua)
- hasTopConcept: C01682 (língua)

**Types:**

- hasTopConcept: C02662 (língua)
- hasTopConcept: C02686 (língua)
- hasTopConcept: C00267 (língua)
- identifier: http://skos.um.es/unescothes (língua)
- language: http://id.loc.gov/vocabulary/iso639-2/fre (língua)
- language: http://id.loc.gov/vocabulary/iso639-2/spa (língua)
- language: http://id.loc.gov/vocabulary/iso639-2/eng (língua)
- language: http://id.loc.gov/vocabulary/iso639-2/rus (língua)
- license: http://creativecommons.org/licenses/by-sa/3.0/igo/ (língua)
- rights: CC-BY-SA (língua)
- source: http://databases.unesco.org/thesaurus/ (língua)
- source: http://www2.ulcc.ac.uk/unesco/ (língua)
- title: UNESCO Thesaurus (en)
- title: Tesouro de la UNESCO (es)
- title: Thésaurus de l'UNESCO (fr)
- title: Тезаурыс ЮНЕСКО (ru)
- type: thesaurus (língua)

At the bottom, there is a search bar with the text 'Escreva nome de pr...' and 'Escreva o valor', and a 'Types' section with a radio button selected for 'ConceptScheme'.

Fonte: Capturado pelos autores (2022).

### Cada Tesouro é:

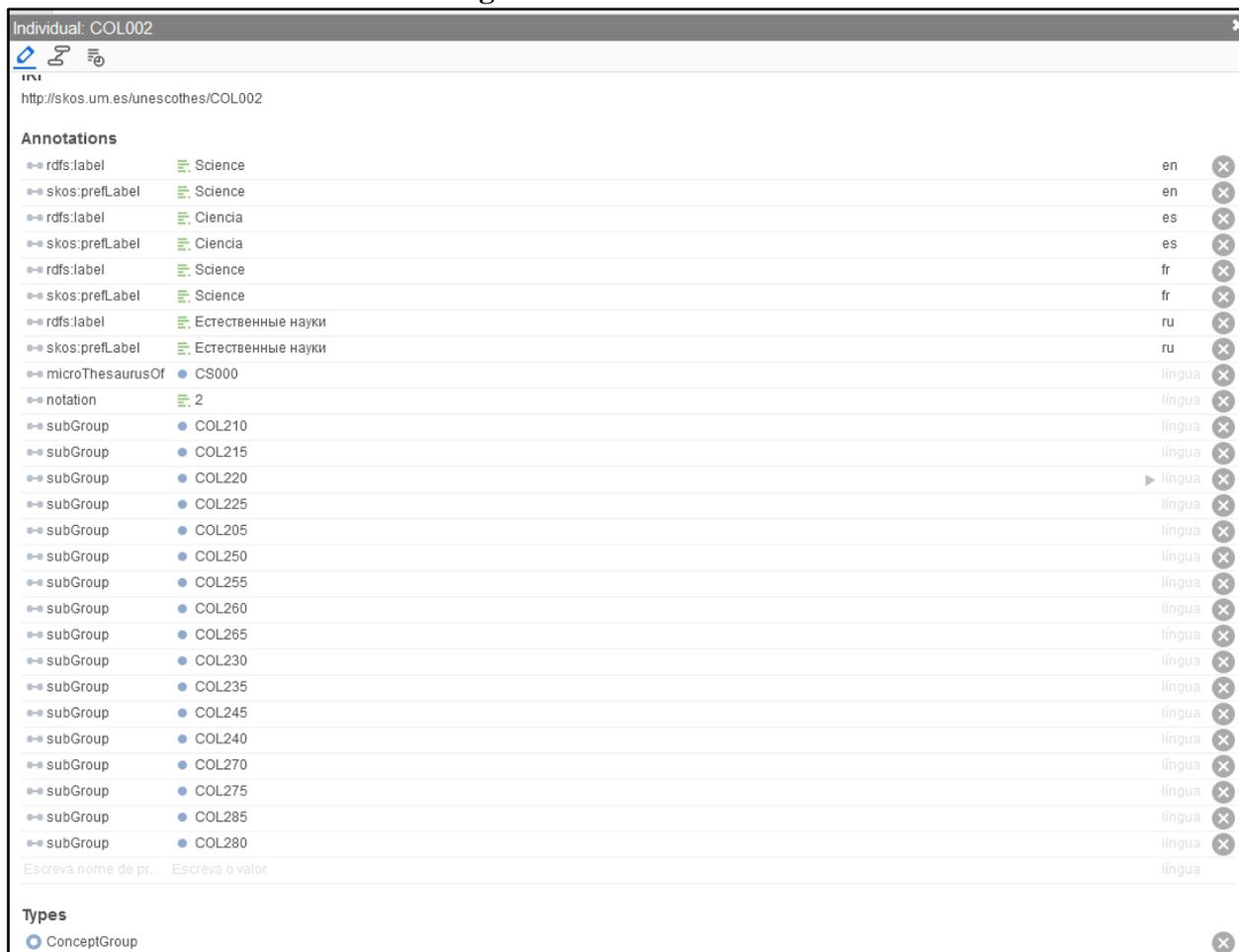
- Um ConceptScheme
- Deve estabelecer links (hasMicroThesaurus) cada grupo de conceitos que o compõe
- Deve estabelecer links (hasTopConcept) cada conceito de maior abstração dentro dos grupos de conceitos que o compõem
- Deve ser anotado com marcadores para:
  - skos:prefLabel (1 para cada língua)
  - dc:created (Data de criação do vocabulário)
  - dc:creator (Criador do vocabulário)
  - dc:language (1 para cada língua)
  - dc:identifier (IRI do vocabulário)
  - dc:title (Título do vocabulário)

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.60/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

**Figura 21: MicroThesaurusOf**

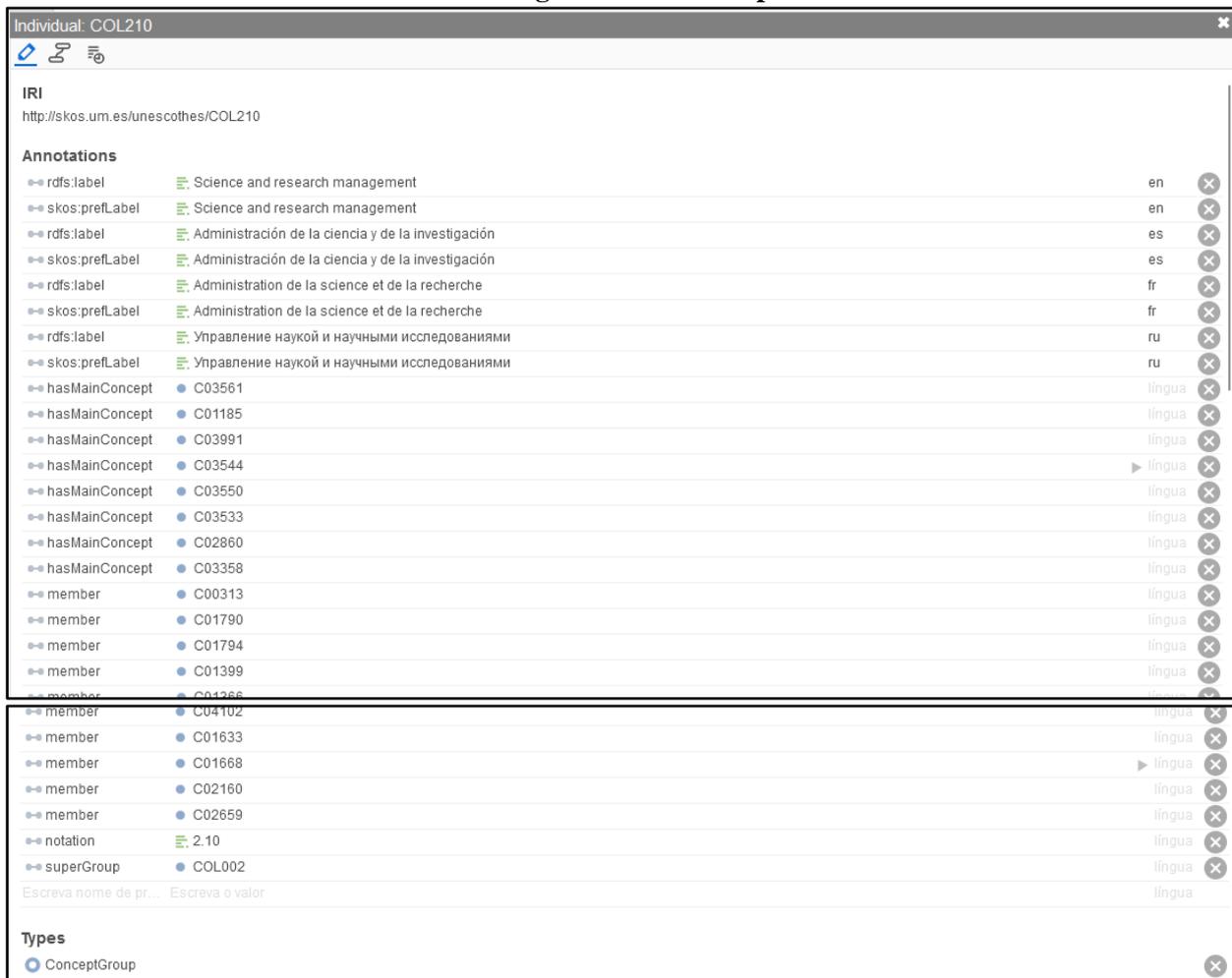


Fonte: Capturado pelos autores (2022).

### Cada Micro-Tesouro é:

- Deve estabelecer links (microThesaurusOf) com cada ConceptSchema do qual seja parte
- Deve ser tipado como ConceptGroup
- Deve estabelecer links (subGroup) com cada sub-grupo que fará parte dele
- Deve ser anotado com marcadores para:
  - skos:prefLabel (1 para cada língua)
  - rdfs:label (1 para cada língua)

Figura 22: subGroup



Individual: COL210

IRI  
http://skos.um.es/unescothes/COL210

Annotations

Property	Value	Language
rdfs:label	Science and research management	en
skos:prefLabel	Science and research management	en
rdfs:label	Administración de la ciencia y de la investigación	es
skos:prefLabel	Administración de la ciencia y de la investigación	es
rdfs:label	Administration de la science et de la recherche	fr
skos:prefLabel	Administration de la science et de la recherche	fr
rdfs:label	Управление наукой и научными исследованиями	ru
skos:prefLabel	Управление наукой и научными исследованиями	ru
hasMainConcept	C03561	língua
hasMainConcept	C01185	língua
hasMainConcept	C03991	língua
hasMainConcept	C03544	língua
hasMainConcept	C03550	língua
hasMainConcept	C03533	língua
hasMainConcept	C02860	língua
hasMainConcept	C03358	língua
member	C00313	língua
member	C01790	língua
member	C01794	língua
member	C01399	língua
member	C04266	língua
member	C04102	língua
member	C01633	língua
member	C01668	língua
member	C02160	língua
member	C02659	língua
notation	2.10	língua
superGroup	COL002	língua

Types

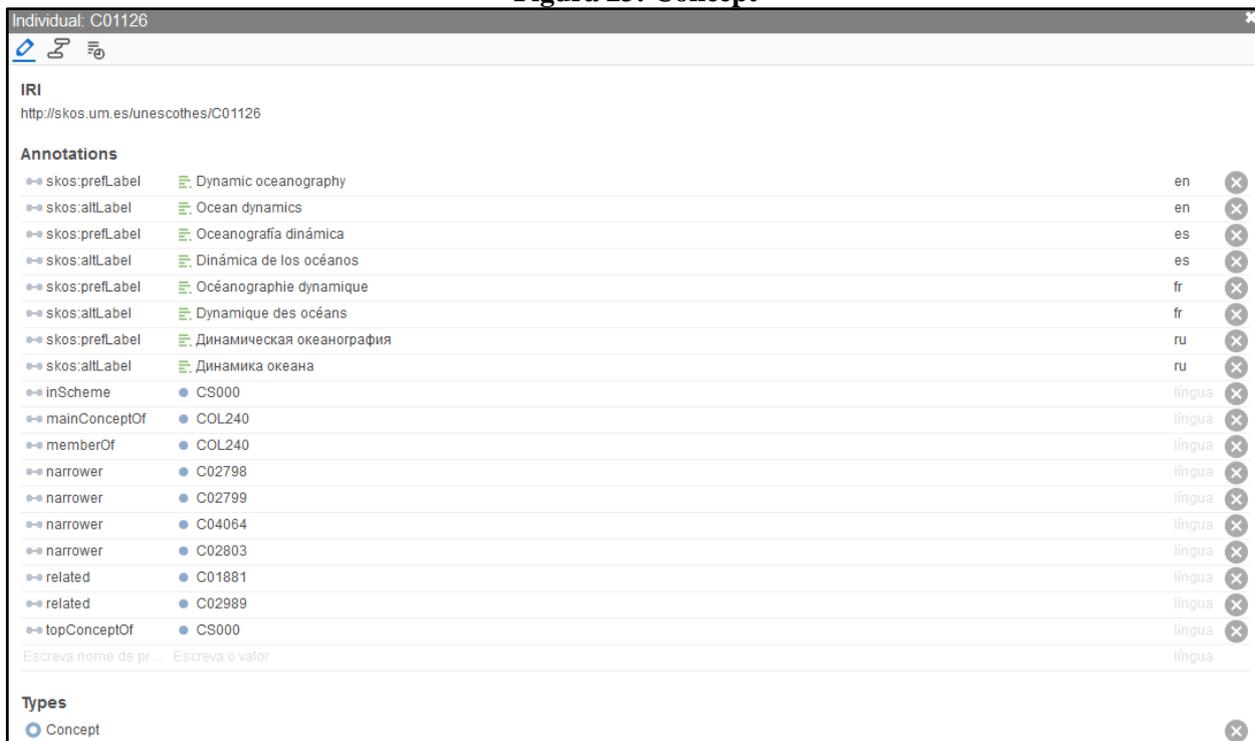
- ConceptGroup

Fonte: Capturado pelos autores (2022).

### Cada subGroup:

- Deve ser tipado como ConceptGroup
- Deve estabelecer links (superGroup) com cada grupo do qual seja sub-grupo
- Deve estabelecer links (hasMainConcept) para cada conceito principal que o compõe
- Deve estabelecer links (member) com cada termo que deverá compor o grupo
- Deve ser anotado com marcadores para:
  - skos:prefLabel (1 para cada língua)
  - rdfs:label (1 para cada língua)

Figura 23: Concept



Fonte: Capturado pelos autores (2022).

### Cada Concept:

- Deve ser tipado como Concept
- Deve estabelecer links (inScheme) com cada ConceptScheme do qual seja parte
- Deve estabelecer links (mainConceptOf) caso seja conceito principal de um grupo
- Deve estabelecer links (memberOf) com todos os grupos dos quais for parte
- Deve estabelecer links (narrower, related, broader) de acordo com as relações que mantiver com os demais termos do vocabulário
- Deve ser anotado com marcadores para:
  - skos:prefLabel (1 para cada língua)
  - skos:altlabel (1 para cada língua) — Não são links (não são termos cadastrados distintamente no tesauro)

**APÊNDICE L: Material instrucional de treinamento no Web Protégé.**

**Figura 24: Conjunto de 13 slides do material instrucional**




## Web Protégé

**Projeto CADE - Frente de Ontologia**  
**Professora: Jeanne Louize Emygdio**

Brasília  
18 de Agosto de 2022

## Agenda



- 01 Apresentação
- 02 Protégé em números
- 03 Web Protégé: área de criação/seleção de projetos
- 04 Web Protégé: página inicial de um projeto
- 05 Web Protégé: padrão semântico adotado para o vocabulário
- 06 Web Protégé: Tipos de propriedades
- 07 Web Protégé: Instâncias
- 08 Web Protégé: Consultas sobre instâncias
- 09 Web Protégé: Revisando termos



Projeto CADE - Frente Ontologia - Profa. Jeanne Louize Emygdio - Brasília - 18/08/2022
2

## Apresentação



Site: <https://protege.stanford.edu/>

O Protégé e o Web Protégé são desenvolvidos pelo *Stanford Center for Biomedical Informatics Research (BMIR)* da **Universidade de Stanford**, na Califórnia, EUA.

As pesquisas do BMIR promovem avanços em:

- Tecnologias semânticas;
- Bioestatísticas e;
- **Modelagem de sistemas biomédicos.**

Tais avanços beneficiam a pesquisa clínica e translacional assim como os cuidados com pacientes.



Projeto CADE – Frente Ontologia – Profa. Jeanne Louize Emygdio – Brasília – 18/08/2022

3

## Protégé em números



Projeto com quase 30  
anos de desenvolvimento



**366.084**  
Usuários registrados



**+ de 250**  
Plugins

Começando  
migração para  
Web Protégé



Projeto com quase 15  
anos de desenvolvimento

**18.207**  
Membros cadastrados na  
lista *Protege-announce*.

**2.376**  
Membros cadastrados na  
lista *Protégé-user*

Informações biomédicas  
*Code exemple*  
Importação/ Exportação  
Inferência  
PLN  
Navegação  
Gestão de Projetos  
Consultas  
*Reasoning*  
*Semantic Web*  
*Engenharia de software*  
Terminologias  
Validação  
Visualização



Projeto CADE – Frente Ontologia – Profa. Jeanne Louize Emygdio – Brasília – 18/08/2022

4

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.65/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## Web Protégé: área de criação/seleção de projetos

**1** Criar Projecto

**2** Nome do Projecto

**3** Dono Última vez ... Última mod...

**4** jeanne.emygdio

**5** Open

Usuário logado (dono da conta)

1. Criação de novos projetos  
 2. Visualização de projetos criados pelo usuário dono da conta  
 3. Visualização de projetos criados por outros usuários e compartilhados com o dono da conta.  
 4. Visualização de projetos excluídos (permite recuperação)

1. Exibe nome de todos os projetos que o usuário tem acesso.  
 2. Para entrar em cada projeto, basta clicar sobre o nome do mesmo.

1. Informações de propriedade dos projetos  
 2. Informações sobre acessos de entrada e modificação

1. Abrir  
 2. Abrir em nova janela  
 3. Download  
 4. Excluir

latitude  
 Laboratório de tecnologias da tomada de decisão

Projeto CADE – Frente Ontologia – Profa. Jeanne Louize Emygdio – Brasília – 18/08/2022

5

## Web Protégé: Página inicial de um projeto

**1** Tesouro CADE

**2** Classes

**3** owl:Thing

**4** IRI

**5** Annotations

**6** Comments: Concept

**7** Project Feed

**8** Terminar Sessão

Menu principal

Abas de navegação

Identificador internacional de recursos

Área de consulta e/ou criação de classes e subclasses que poderão ser utilizadas para a criação do vocabulário.

Área de anotações sobre os construtos do padrão semântico adotado.

Área de colaboração entre os membros da equipe sobre os termos em análise.

Exibição de últimas modificações no vocabulário

latitude  
 Laboratório de tecnologias da tomada de decisão

Projeto CADE – Frente Ontologia – Profa. Jeanne Louize Emygdio – Brasília – 18/08/2022

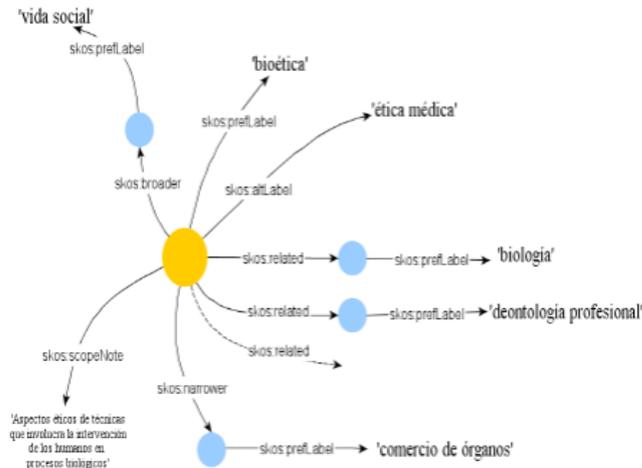
6

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.66/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
 É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

## Web Protégé: Padrão semântico adotado para o vocabulário

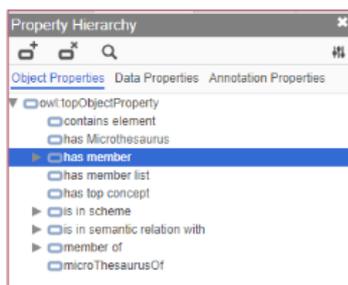


O padrão semântico utilizado se baseia no padrão SKOS, porém, mais enriquecido em termos de construtos (símbolos) para representação.

Padrão adotado é o padrão utilizado pela UNESCO.

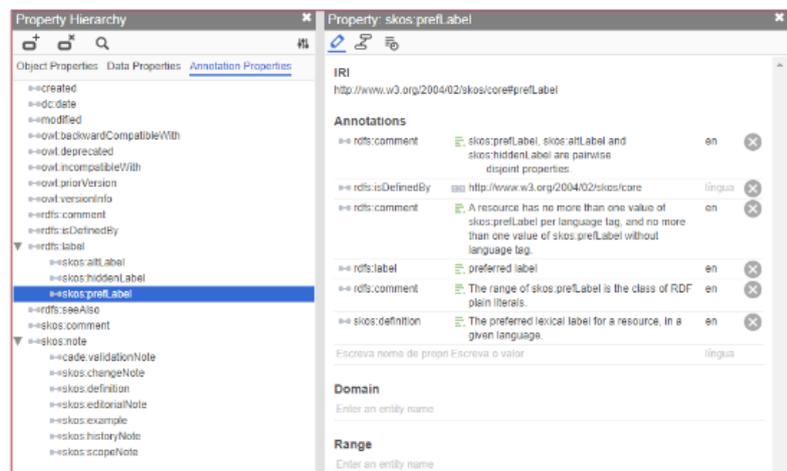
## Web Protégé: Tipos de propriedades

### Propriedades de objetos



São usadas para representação das relações entre os termos/conceitos do vocabulário.

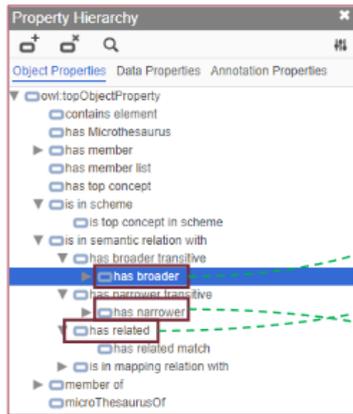
### Propriedades de anotações



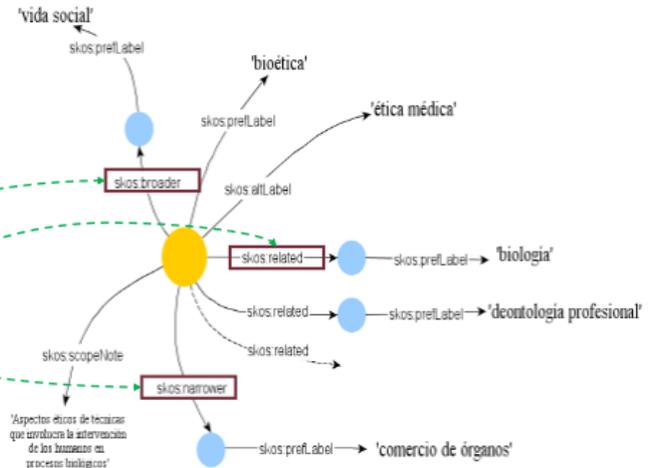
São usadas para documentar os termos/conceitos do vocabulário.

## Web Protégé: Propriedades de objetos

### Propriedades de objetos



São usadas para representação das relações entre os termos/conceitos do vocabulário.

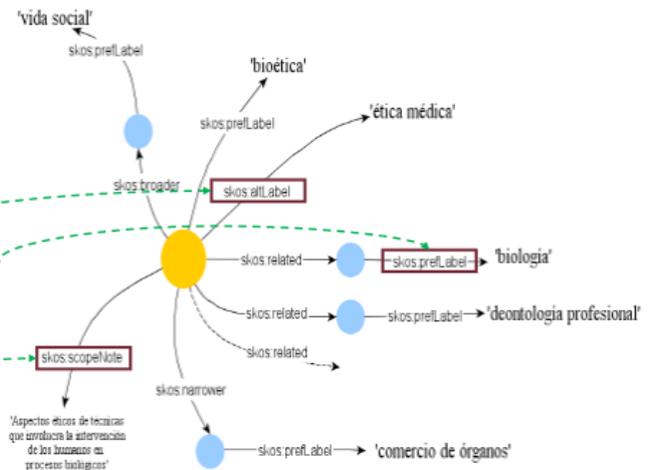


## Web Protégé: Propriedades de anotações

### Propriedades de anotações



São usadas para documentar os termos/conceitos do vocabulário.



## Web Protégé: Instâncias

**1** Termos/Conceitos, Vocabulários, Coleções, Grupos de termos e listas de termos inseridos como constituintes do vocabulário.

**2** IRI do termo

**3** IRI do termo

**4** Área de anotações sobre as instâncias (termos/conceitos) do vocabulário.

**5** Área de relacionamento entre as instâncias (termos/conceitos) do vocabulário.

Projeto CADE – Frente Ontologia – Profª. Jeanne Louize Emygdio – Brasília – 18/08/2022

## Web Protégé: Consultas sobre instâncias

**1** Selecionar entre os tipos de instâncias inseridas no vocabulário.

- Coleções (*collection*)
- Conceito (*concept*)
- Descritor
- Não descritor
- Termo candidato
- Esquema conceitual (vocabulário)

Projeto CADE – Frente Ontologia – Profª. Jeanne Louize Emygdio – Brasília – 18/08/2022

## Web Protégé: Revisando termos

**1** nova conversa  
 jeanne.emygdio  
 Less than one minute ago  
 Sobre o acordo horizontal... <inserir críticas e ajustes que julgarem necessários no termo>  
 Responder Editar

**2** Editar  
 Sobre o acordo horizontal... <inserir críticas e ajustes que julgarem necessários no termo>  
 Markdown suportada  
 Cancelar Ok

**3** jeanne.emygdio  
 Less than one minute ago  
 Sobre o acordo horizontal... <inserir críticas e ajustes que julgarem necessários no termo>  
 Responder Editar

jeanne.emygdio  
 7 days ago  
 Última modificação 7 days ago  
 Verificar se as condições para acordo horizontal ocorrerem em um mesmo mercado relevante (geográfico e material) são excludentes ou não. Recordar definição da Paula Forgiotti.  
 Responder Editar

**4** Clases | Propriedades | Individual | Comentários | History |  
 Commented entities  
 Ordenar por Emissoe  
 Acordo horizontal  
 3 Último comentário less than one minute ago por jeanne.emygdio 2 Resolvidos  
 Tribunal Administrativo de Defesa Econômica  
 1 Último comentário yesterday at 17:10 por Izadora Linares 1 Não-resolvido  
 Catê  
 1 Último comentário 7 days ago por jeanne.emygdio resolvido

Após revisão do termo pela equipe de ontologia os comentários podem ser resolvidos/fechados.  
 Exibe/Ocultar comentários resolvidos.  
 Comentários em aberto podem ser editados e respondidos.

Projeto CADE – Frente Ontologia – Profa. Jeanne Louize Emygdio – Brasília – 18/08/2022

Fonte: Capturado pelos autores (2022).

### APÊNDICE M: Diretrizes para execução das atividades de validação de termos

Este conjunto de diretrizes foi desenvolvido para apoiar as atividades de validação de termos por parte das especialistas de domínio. O documento original foi compartilhado no dia 08/09/2022.

#### Questões para nortear a validação dos termos

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.70/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade. É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

### Sobre **Definições**:

- O conteúdo das definições cobre o entendimento do termo no domínio?
- O conteúdo que foi identificado para compor uma definição é adequado? Está completo?
- O conteúdo identificado como definição é de fato adequado para definição ou se relaciona com uma nota de escopo?

### Sobre **Descritores e Não-descritores**:

- O termo identificado como descritor é realmente o termo mais adotado no domínio de especialidade?
- Há descritores que podem ser eliminados? Quais? Apresentar justificativa.
- Há sugestão de descritores não identificados?
- Há sugestão de não-descritores não identificados?

**Quadro 4: Sugestão de notas para documentação dos termos**

ANOTAÇÕES SOBRE OS TERMOS	
Nota	Propósito
Nota de escopo	<ul style="list-style-type: none"> <li>● Registro de situações em que um conceito é utilizado no domínio de especialidade;</li> <li>● Anotações mais gerais sobre o conceito.</li> </ul>
Nota de validação	<ul style="list-style-type: none"> <li>● Registro da data de validação do conceito e histórico.</li> <li>● Formato: [Data-Primeiro nome] &lt;Descrição da mudança&gt;. Incluir fonte bibliográfica se necessário.</li> </ul>
Nota de homologação	<ul style="list-style-type: none"> <li>● Registro da data de homologação do conceito quando da conclusão de todas as análises possíveis sobre o mesmo.</li> </ul>
Nota bibliográfica	<ul style="list-style-type: none"> <li>● Registro das fontes utilizadas para construção da definição de um conceito;</li> <li>● Fontes onde foram encontrados os termos alternativos.</li> </ul>
Nota histórica	<ul style="list-style-type: none"> <li>● Registro das mudanças dos conceitos do domínio ao longo do tempo. Exemplo: SEAE.</li> </ul>

Nota de exemplo	<ul style="list-style-type: none"> <li>Registro de exemplos de uso dos conceitos do domínio de especialidade encontrados na literatura.</li> </ul>
Nota editorial	<ul style="list-style-type: none"> <li>Registro das manutenções realizadas no vocabulário por parte de seus editores. Ex.: Versionamento, Termos em desuso, etc.</li> </ul>
Nota de mudança	<ul style="list-style-type: none"> <li>Registro de mudanças na classificação de conceitos previamente “homologados” no tesouro. Ex.: Um descritor que se torna não-descritor.</li> </ul>
<b>ANOTAÇÕES SOBRE O VOCABULÁRIO</b>	
Título	Título do vocabulário
Descrição	Descrição do vocabulário envolvendo escopo de cobertura e organização por grupos e subgrupos de termos.
Criador	Criadores do vocabulário
Licença de uso	Licença de uso adotada sobre o vocabulário. Ex: Creative Commons
Data de criação	Data de criação do vocabulário
Última modificação	Data da última modificação do vocabulário
Versão	Versão do vocabulário
Contribuidores	Lista de contribuidores
Língua	Línguas cobertas pelo vocabulário
URI	URI do vocabulário

Fonte: Elaborado pelos autores (2022).

**APÊNDICE N: Evolução das análises sobre conceitos centrais do domínio Cade.**

**Figura 26: Conjunto de 10 slides sobre a evolução das análise conceituais**

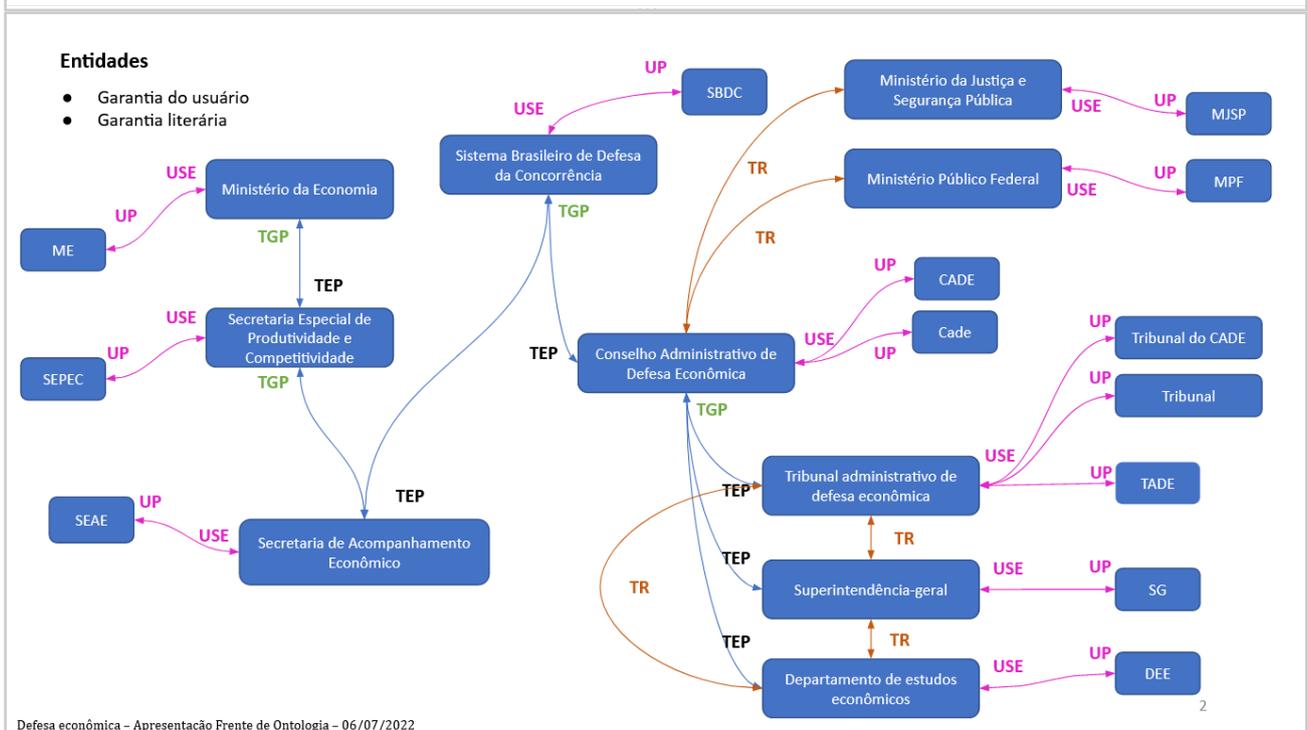
# Defesa econômica

Avanços

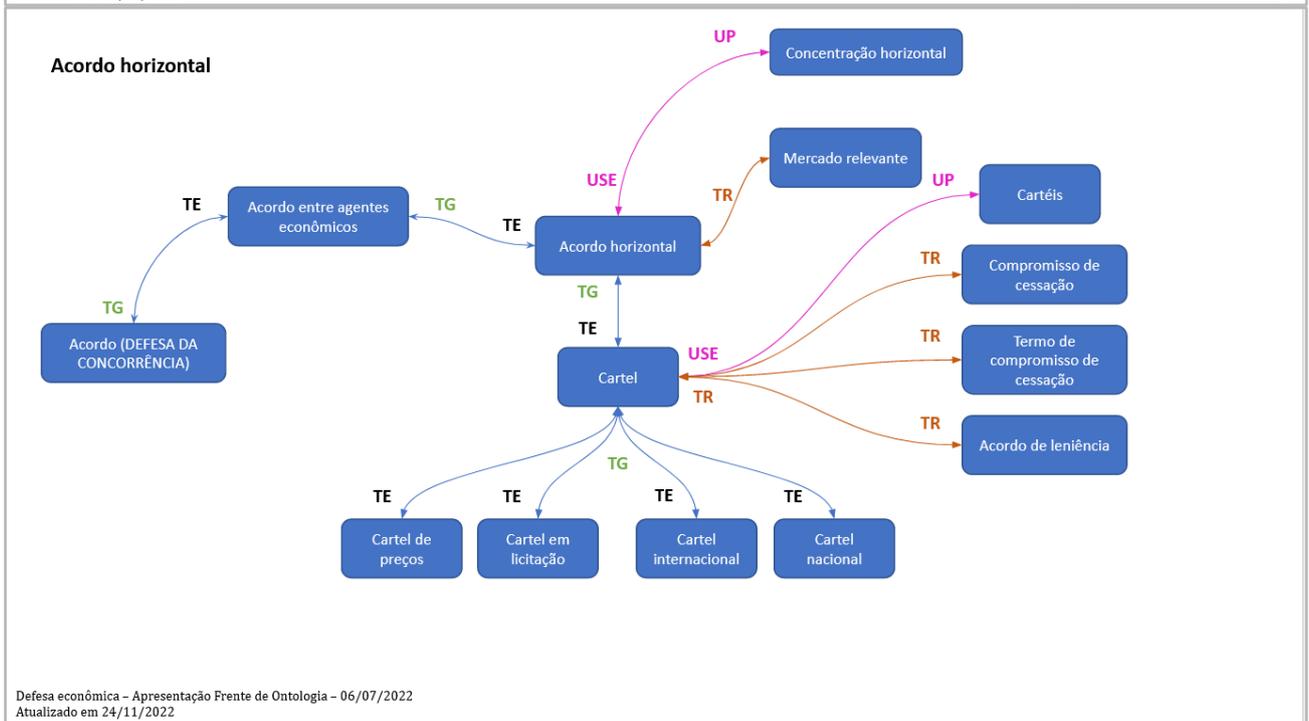
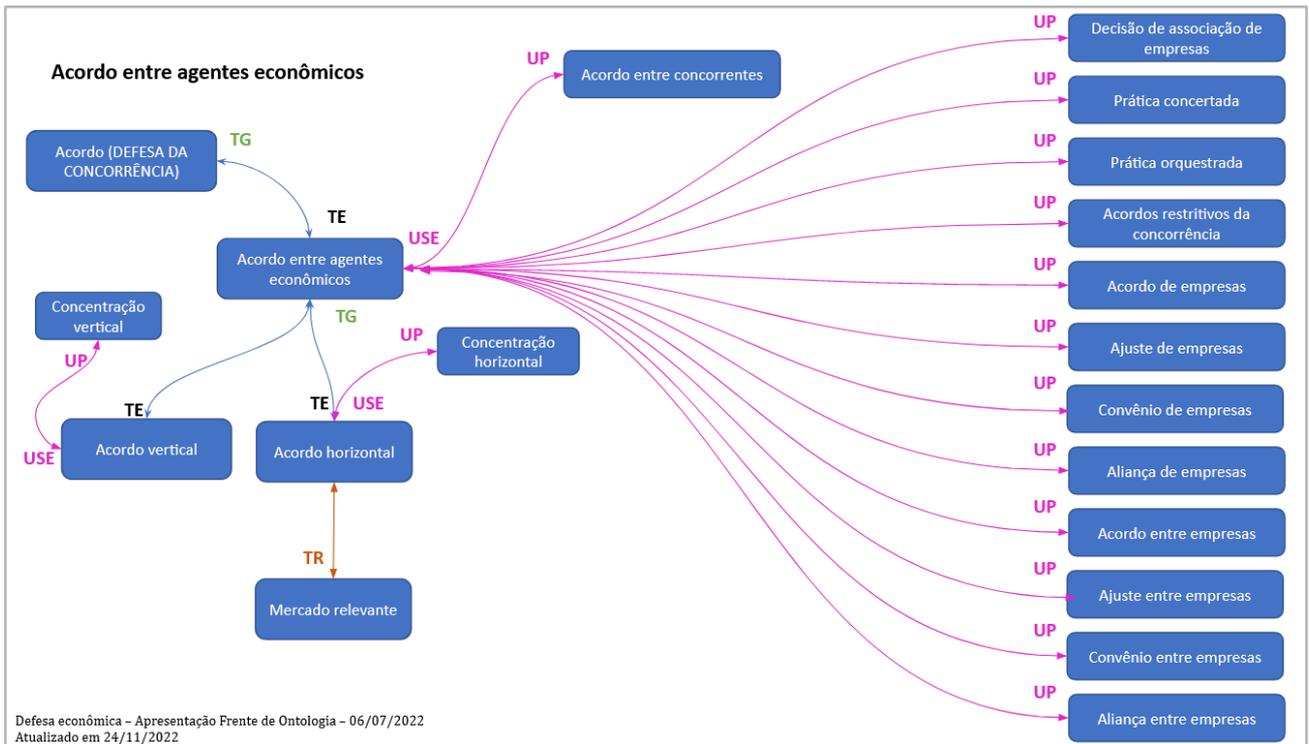
Frente de ontologia

Início: 06/07/2022

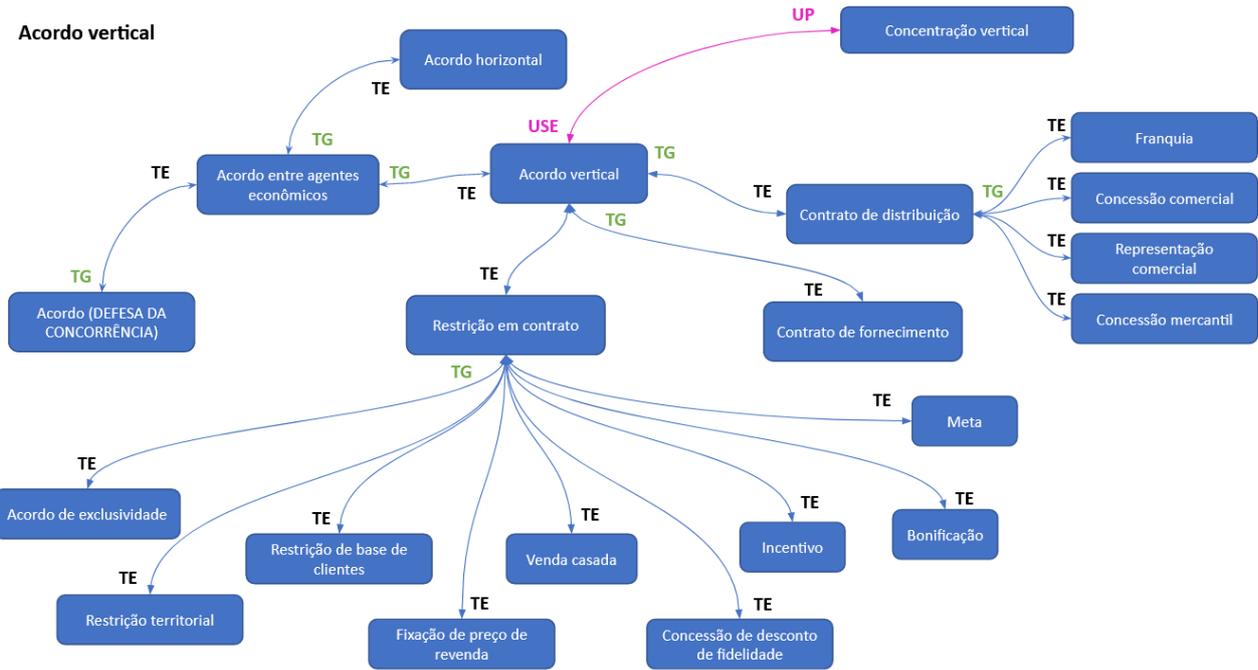
Evoluções até: Nov/2022







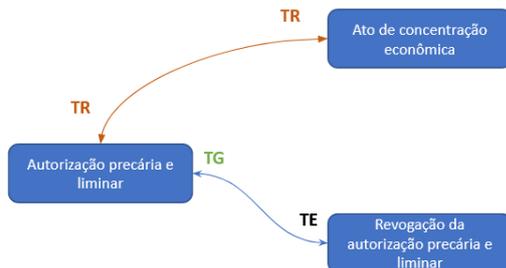
**Acordo vertical**



Defesa econômica – Apresentação Frente de Ontologia – 06/07/2022

**Autorização precária e liminar**

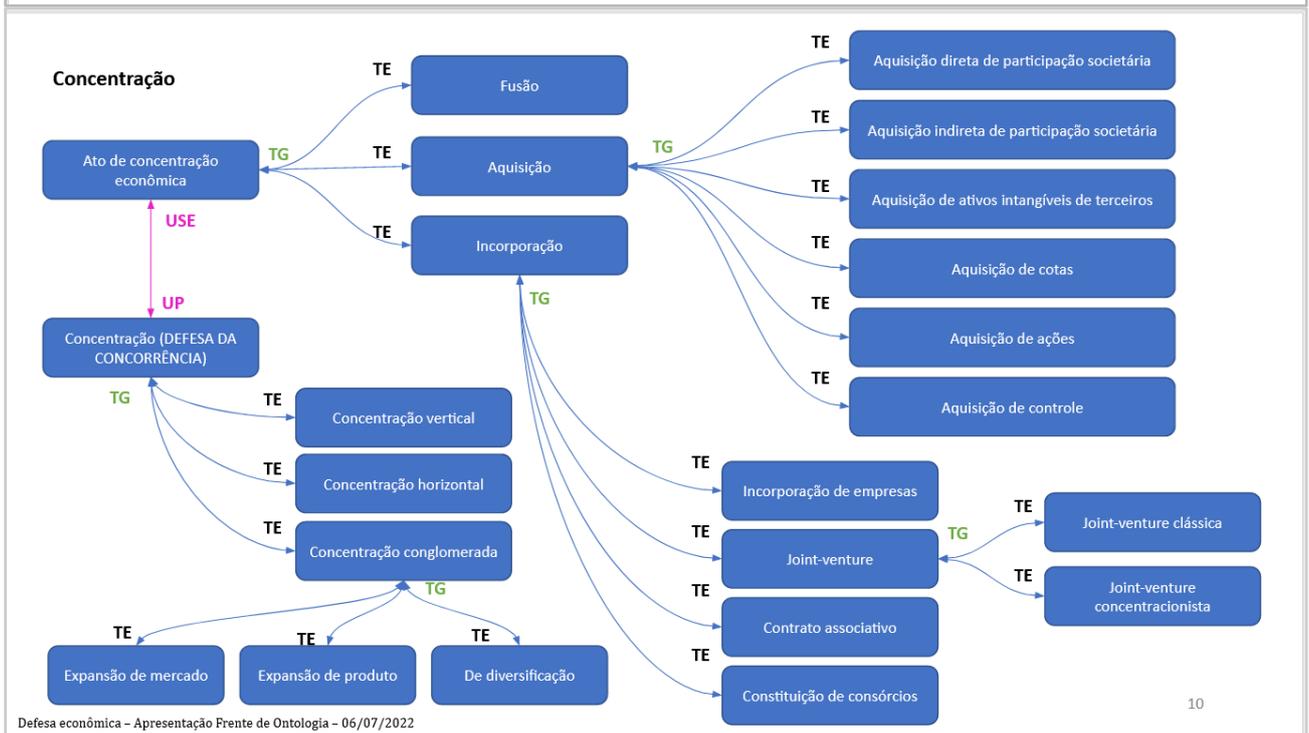
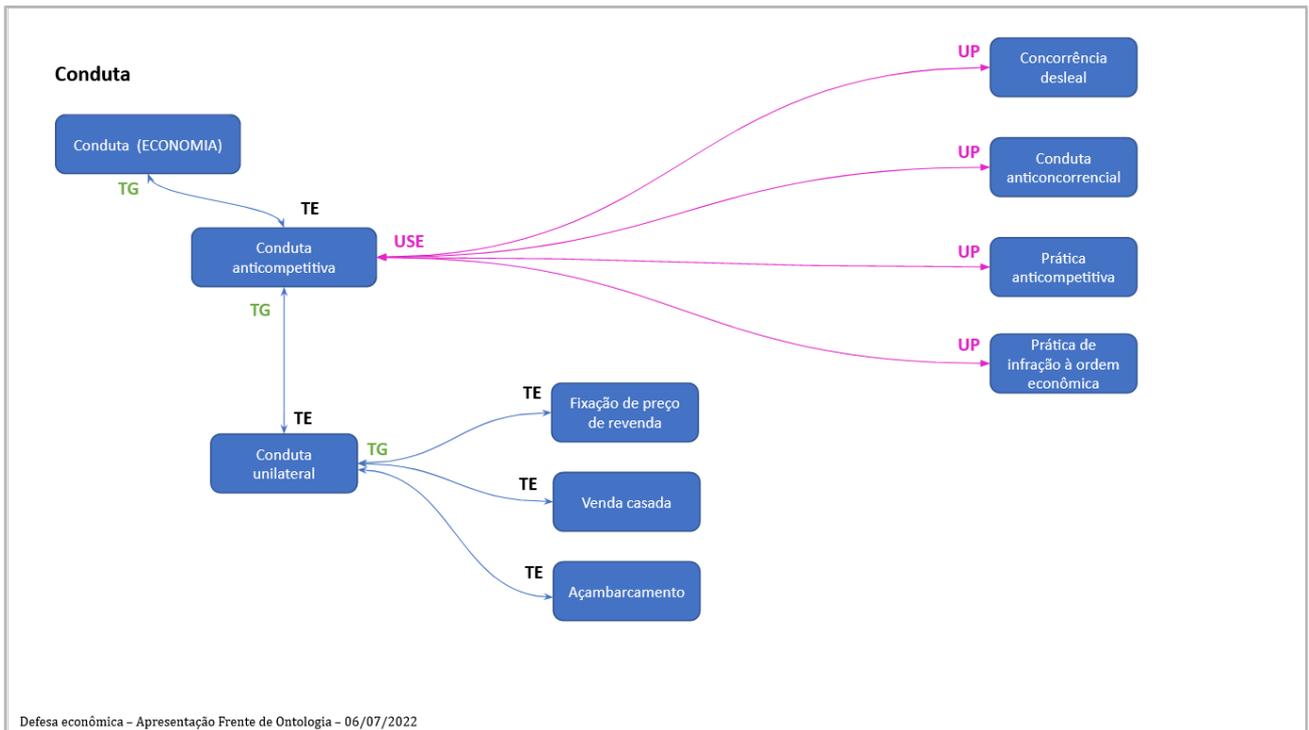
- Optamos por manter “revogação” com parte do termo “revogação da autorização precária e liminar” ao invés de modificador.



Defesa econômica – Apresentação Frente de Ontologia – 06/07/2022

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
 É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.



Fonte: Capturado pelos autores (2022).

## APÊNDICE O: Listas de termos por tipo de classificação

**Quadro 6: Lista de conceitos**

Nº	CONCEITO
1	Acordo (DEFESA DA CONCORRÊNCIA)
2	Acordo de Leniência (ADMINISTRAÇÃO PÚBLICA)
3	Acordo de Leniência (DEFESA DA CONCORRÊNCIA)
4	Acordo em controle de concentrações
5	Acordo entre agente público e empresa (DEFESA DA CONCORRÊNCIA)
6	Acordo entre agentes econômicos (DEFESA DA CONCORRÊNCIA)
7	Acordo horizontal
8	Acordo vertical
9	Cartel
10	Cartel clássico
11	Cartel de alocação de mercado
12	Cartel difuso
13	Cartel em licitação
14	Cartel explícito
15	Cartel hub-and-spoke
16	Cartel tácito

17	Compromisso de cessação
18	Conselho Administrativo de Defesa Econômica
19	Departamento de Estudos Econômicos
20	Ministério Público Federal
21	Ministério da Economia
22	Ministério da Justiça e Segurança Pública
23	Secretaria Especial de Produtividade e Competitividade
24	Secretaria de Acompanhamento Econômico
25	Sistema Brasileiro de Defesa da Concorrência
26	Superintendência-Geral
27	Termo de compromisso de cessação
28	Tribunal Administrativo de Defesa Econômica

Fonte: Organizado pelos autores (2022).

### Quadro 7: Lista de termos em análise

Nº	Termo em análise
1	Abertura de falência
2	Acordo de exclusividade
3	Acordo de preservação de reversibilidade
4	Açambarcamento

5	Cartel de compra
6	Cartel de exportação
7	Cartel de importação
8	Cartel de preço
9	Cartel de venda
10	Cartel internacional
11	Cartel nacional
12	Cláusulas restritivas à concorrência
13	Controle de concentrações
14	Controle de condutas
15	Exclusividade
16	Fixação de preços
17	Fixação de preços de revenda
18	Leniência Plus

Fonte: Organizado pelos autores (2022).

### Quadro 8: Lista de termos candidatos

Nº	Termo candidato
1	Abuso de poder econômico
2	Abuso de posição dominante

3	Abuso no exercício de direito de propriedade intelectual
4	Acordo para geração de receita (ADMINISTRAÇÃO PÚBLICA)
5	Adesão ao acordo de leniência
6	Aditivo ao acordo de leniência
7	Agente econômico
8	Aquisição de ativos intangíveis de terceiros
9	Aquisição de ativos tangíveis de terceiros
10	Aquisição de ações
11	Aquisição de controle
12	Aquisição de cotas
13	Aquisição direta de participação societária
14	Aquisição indireta de participação societária
15	Ato de concentração
16	Ato de concentração econômico
17	Ato de concentração ordinário
18	Autoras de infração à ordem econômica
19	Autorização precária e liminar
20	Bonificação

21	Cláusulas de exclusividade
22	Cláusulas de exclusividade territorial
23	Cláusulas de nação mais favorecida
24	Cláusulas de preferência
25	Concentração conglomerada
26	Concentração conglomerada de diversificação
27	Concentrações
28	Concessão
29	Concessão comercial
30	Concessão de descontos de fidelidade
31	Concessão mercantil
32	Concorrência monopolista
33	Condenado por infração à ordem econômica
34	Conduta anticompetitiva
35	Conduta coletiva
36	Conduta colusiva
37	Conduta comercial uniforme
38	Conduta oportunista

39	Conduta unilateral
40	Conglomeração
41	Constituição de consórcios
42	Consulta- licitude de conduta
43	Contrato associativo
44	Contrato de distribuição
45	Contrato de fornecimento
46	Controle de estruturas
47	Criar dificuldade ao concorrente
48	Decisão de associação de empresas
49	Destruição de matéria-prima
50	Discriminação de preços
51	Divisão de base de clientes
52	Divisão de mercado
53	Divisão territorial
54	Excedente de consumo
55	Excedente de produtor
56	Expansão de mercado

57	Expansão de produto
58	Franquia
59	Fusão
60	Histórico da Conduta
61	Incentivo
62	Incorporação de empresas
63	Influência de conduta uniforme
64	Infração da ordem econômica
65	Inquérito administrativo para apuração de infrações à ordem econômica
66	Isenção
67	Isenção em bloco
68	Joint-venture
69	Joint-venture clássica
70	Joint-venture concentracionista
71	Legislação antitruste
72	Leniência Plus Condicional
73	Leniência parcial
74	Mercado de distribuição

75	Mercado relevante
76	Mercado relevante geográfico
77	Mercado relevante material
78	Meta
79	Monopólio
80	Monopólio natural
81	Multa
82	Novo Acordo de Leniência
83	Oligopólio
84	Ordem econômica
85	Pedido de marker
86	Penalidade
87	Poder de mercado
88	Posição dominante
89	Preparatório de inquérito administrativo para apuração de infrações à ordem econômica
90	Preços abusivos
91	Preços predatórios
92	Procedimento administrativo para apuração de ato de concentração econômica

93	Procedimento preparatório de inquérito administrativo para apuração de infrações à ordem econômica
94	Processo Administrativo
95	Processo administrativo para análise de ato de concentração econômica
96	Processo administrativo para imposição de sanções administrativas por infrações à ordem econômica
97	Processo administrativo para imposição de sanções processuais incidentais
98	Programa de leniência
99	Prática concentrada
100	Prática de infração à ordem econômica
101	Prática orquestrada
102	Recusa de contratar
103	Recusa de negociação
104	Regulação econômica
105	Restrição de base de clientes
106	Restrição de produção
107	Restrição em contratos
108	Restrição territorial
109	Sham litigation
110	Termo de ajustamento de conduta

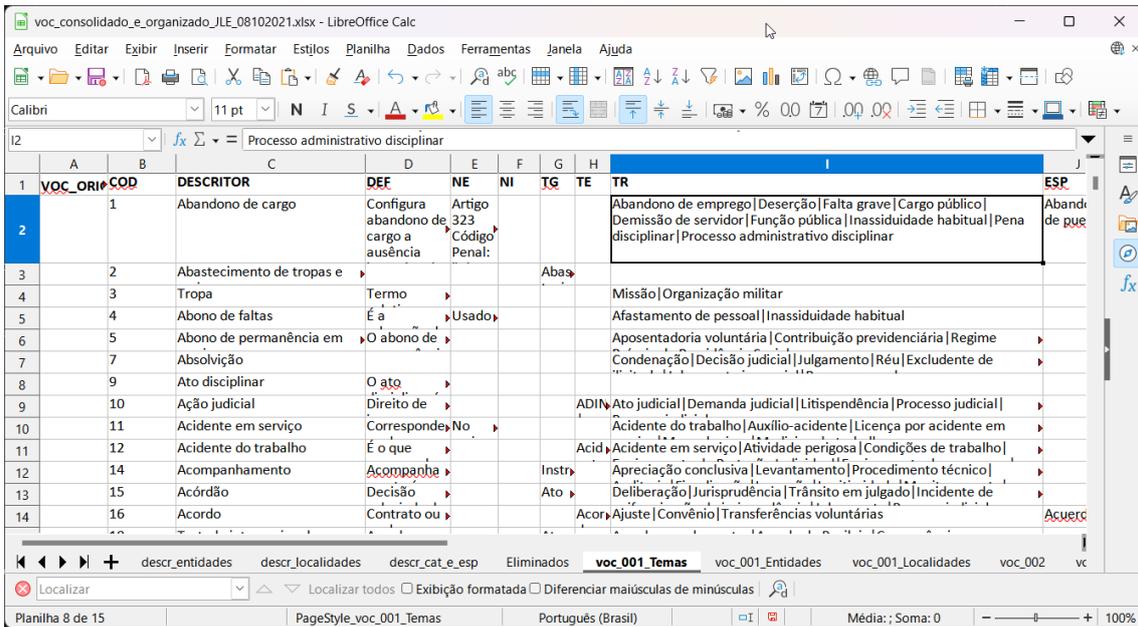
111	Termo de compromisso de desempenho
112	Unimilitância
113	Venda casada

Fonte: Organizado pelos autores (2022).

## APÊNDICE P: Conversão de planilhas excel para RDF via Cellfie

O Cellfie é um plug-in do Protégé Desktop que possibilita a importação de dados de planilhas excel para ontologias OWL. O Cellfie utiliza o MappingMaster que é uma linguagem específica de domínio (DSL) baseada na Manchester OWL Syntax, uma DSL para descrever ontologias OWL (“MappingMasterDSL”). A planilha com o vocabulário do TCU possui colunas com valores simples, exemplo coluna DESCRITOR, e colunas com múltiplos valores separados por um pipe “|”, como a coluna TR, conforme ilustrado na Figura 27.

**Figura 27: Visão da planilha do VCE do TCU**

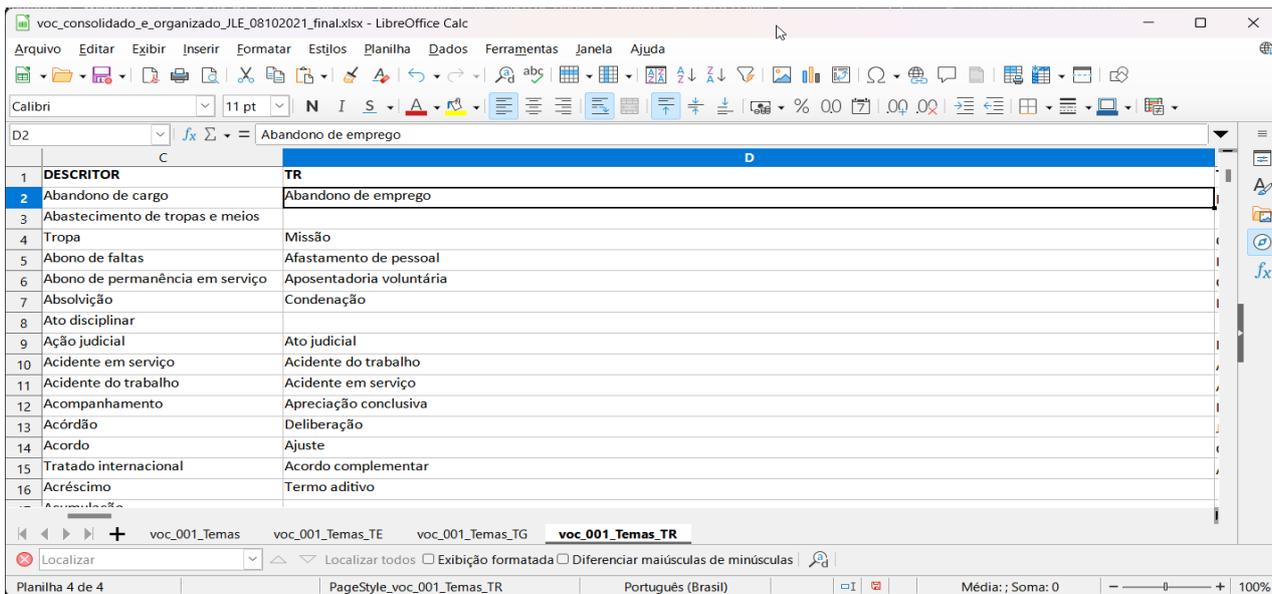


VOC_ORI	COD	DESCRITOR	DEF	NE	NI	TG	TE	TR	ESP
1	1	Abandono de cargo	Configura abandono de cargo a ausência	Artigo 323 Código Penal:				Abandono de emprego Deserção Falta grave Cargo público Demissão de servidor Função pública Inassiduidade habitual Pena disciplinar Processo administrativo disciplinar	Abandi de que
2	2	Abastecimento de tropas e				Abas			
3	3	Tropa	Termo					Missão Organização militar	
4	4	Abono de faltas	É a	Usado				Afastamento de pessoal Inassiduidade habitual	
5	5	Abono de permanência em	O abono de					Aposentadoria voluntária Contribuição previdenciária Regime	
6	7	Absolvição						Condenação Decisão judicial Julgamento Réu Excludente de	
7	9	Ato disciplinar	O ato						
8	10	Ação judicial	Direito de				ADIN	Ato judicial Demanda judicial Litispêndência Processo judicial	
9	11	Acidente em serviço	Corresponde	No				Acidente do trabalho Auxílio-acidente Licença por acidente em	
10	12	Acidente do trabalho	É o que				Acid	Acidente em serviço Atividade perigosa Condições de trabalho	
11	14	Acompanhamento	Acompanha			Instr		Apreciação conclusiva Levantamento Procedimento técnico	
12	15	Acórdão	Decisão			Ato		Deliberação Jurisprudência Trânsito em julgado Incidente de	
13	16	Acordo	Contrato ou					Acor Ajuste Convênio Transferências voluntárias	Acuer

Fonte: Capturado pelos autores (2022).

Passos para a conversão:

Transformar colunas com delimitadores em novas planilhas, exemplo coluna TR, utilizando a função como “Texto em colunas”:



DESCRITOR	TR
Abandono de cargo	Abandono de emprego
Abastecimento de tropas e meios	
Tropa	Missão
Abono de faltas	Afastamento de pessoal
Abono de permanência em serviço	Aposentadoria voluntária
Absolvição	Condenação
Ato disciplinar	
Ação judicial	Ato judicial
Acidente em serviço	Acidente do trabalho
Acidente do trabalho	Acidente em serviço
Acompanhamento	Apreciação conclusiva
Acórdão	Deliberação
Acordo	Ajuste
Tratado internacional	Acordo complementar
Acréscimo	Termo aditivo

**Figura 28: Transformando células de múltiplos valores em colunas**

Fonte: capturado pelos autores (2022).

- Criar regras de importação
  - Importando a coluna DESCRITOR

```
Individual: @C*(mm:camelCaseEncode)
Types: skos:Concept

Annotations: skos:prefLabel@C*(xml:lang="pt-BR"),skos:altLabel@J*(xml:lang="es"),skos:altLabel@K*(xml:lang="en"), skos:definition@D*(mm:SkipIfEmptyLocation xml:lang="pt-BR"), skos:scopeNote@E*(mm:SkipIfEmptyLocation xml:lang="pt-BR")
```

- Importando Termos Gerais (TG)

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.89/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

```
Individual: @**(mm:SkipIfEmptyLocation)
Types: skos:Concept
Annotations: skos:prefLabel@**(xml:lang="pt-BR")
Facts:
    skos:narrower@C*
```

○ Importando Termo específico do DESCRITOR

```
Individual: @**(mm:SkipIfEmptyLocation)
Types: skos:Concept
Annotations: skos:prefLabel@**(xml:lang="pt-BR")
Facts:
    skos:broader@C*(mm:camelCaseEncode)
```

○ Termos relacionados ao DESCRITOR

```
Individual: @**(mm:SkipIfEmptyLocation)
Types: skos:Concept
Annotations: skos:prefLabel@**(xml:lang="pt-BR")
Facts:
    skos:related@C*
```

## APÊNDICE Q: Percurso para publicação de vocabulários no Skosmos

A integração entre o WebProtégé e o Apache Jena Fuseki, para publicação do vocabulário no Skosmos é executada de forma manual, conforme será apresentado a seguir.

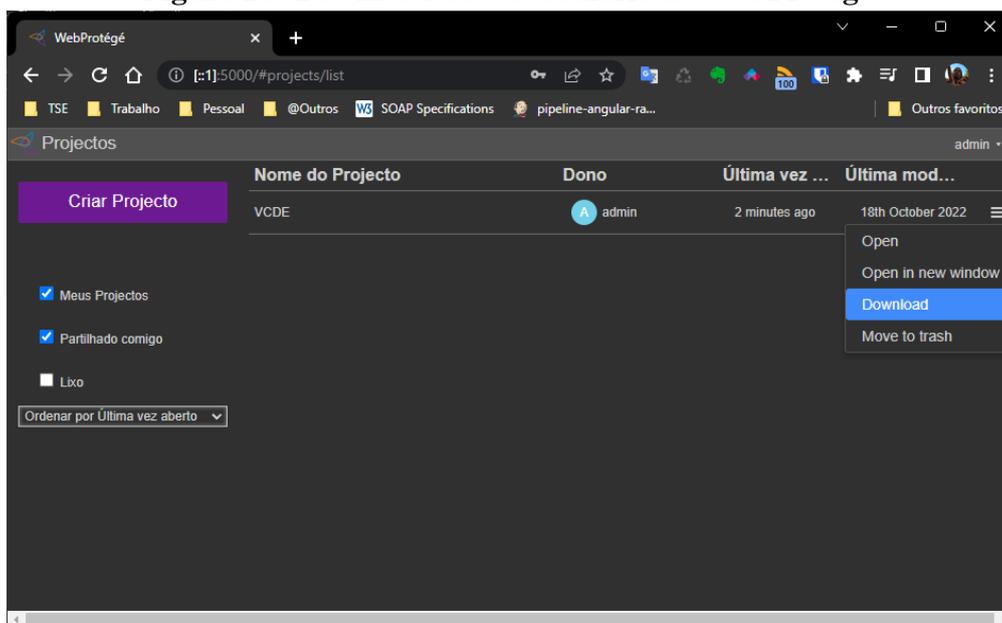
Passos para integração:

- Configurar o endpoint no Skosmos para a publicação do vocabulário, ver Apêndice G.
  - Incluir nova entrada no arquivo Skosmos\dockerfiles\config\config-docker-compose.ttl

```

:IDENTIFICADO_VACABULARIO a skosmos:Vocabulary, void:Dataset ;
  dc:title "Nome Novo Vocabulário"@pt-BR ;
  skosmos:shortName "Nome Curto";
  dc:subject :cat_general ;
  void:uriSpace "http://nome_endpoint/";
  skosmos:language "pt", "pt-BR", "en";
  skosmos:defaultLanguage "pt-BR";
  skosmos:showTopConcepts true ;
  skosmos:defaultSidebarView 'alphabetical';
  skosmos:fullAlphabeticalIndex true ;
  void:sparqlEndpoint <http://fuseki-cache:80/skosmos/sparql> ;
  skosmos:sparqlGraph <http://nome_endpoint/> .
    
```

**Figura 27: Download do Vocabulário do WebProtégé**



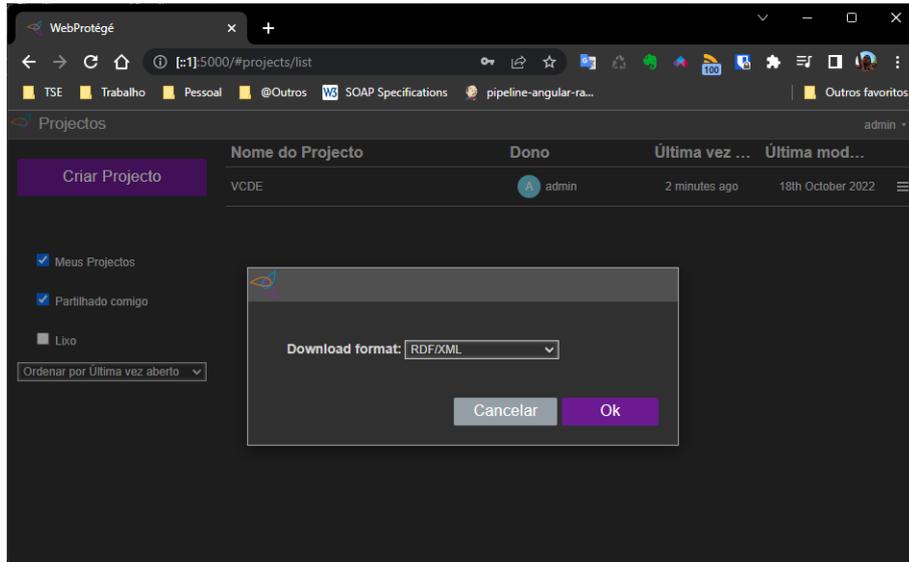
Fonte: Capturado pelos autores(2022).

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.91/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade. É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

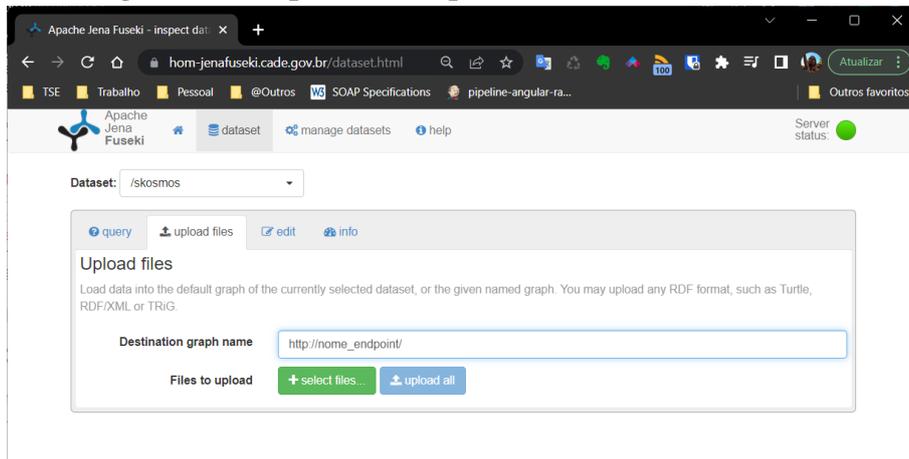
**Figura 28: Salvar o arquivo no formato XML/RDF**



Fonte: Capturado pelos autores(2022).

- Descompactar o arquivo gerado
- Importar o arquivo RDF no Apache Jena/Fuseki no endpoint configurado no Skosmos “[http://nome\\_endpoint/](http://nome_endpoint/)“
  - Acessar o menu Dataset, selecionar a aba **Upload Files** e informar o nome do grafo [http://nome\\_endpoint/](http://nome_endpoint/) no campo **Destination graph name**

**Figura 29 - Importar o arquivo no formato XML/RDF**



Fonte: Capturado pelos autores(2022).

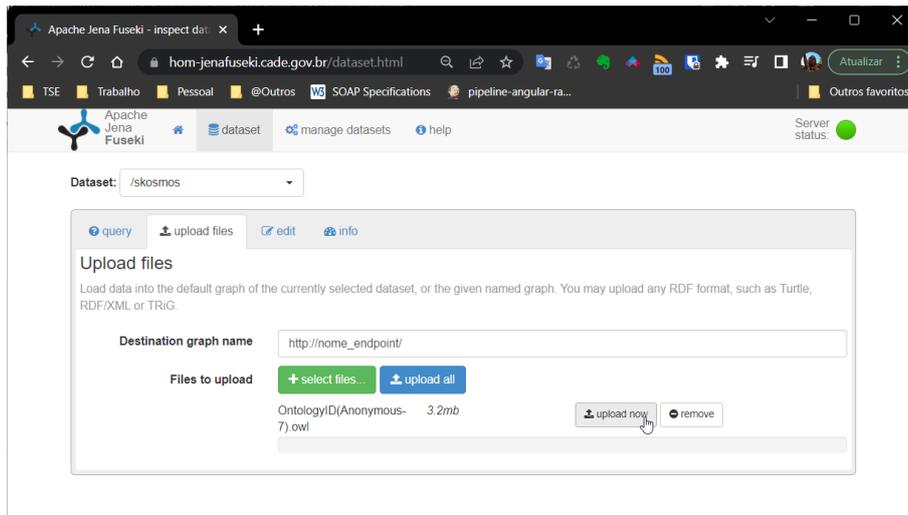
- Selecionar o arquivo descompactado e clique no botão **Upload now**

**Figura 30: Uploading do arquivo no formato XML/RDF**

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.92/95
---------------	---------------------	---	-----------

Confidencial.

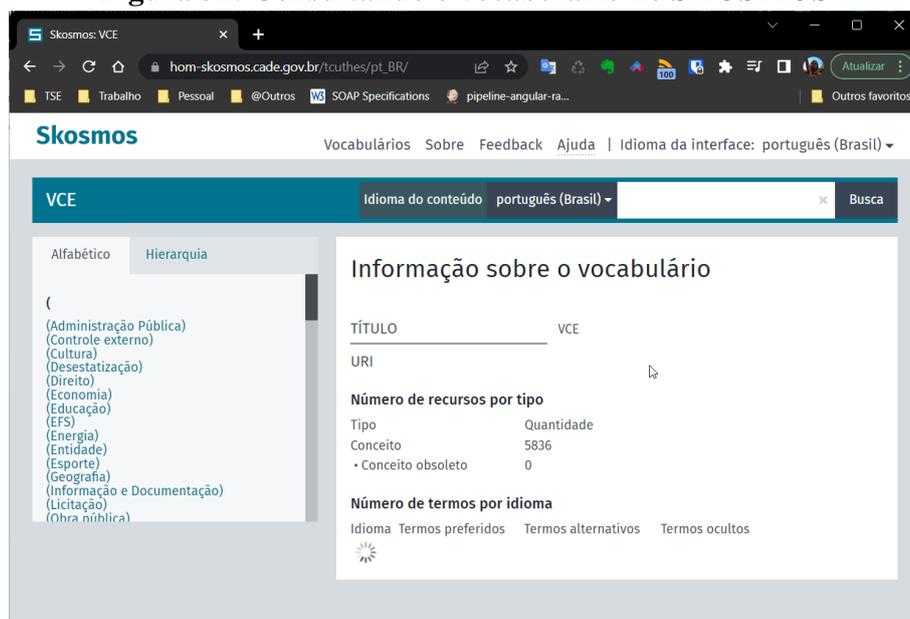
Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade. É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.



Fonte: Capturado pelos autores(2022).

- O vocabulário será exibido para consumo no Skosmos

**Figura 31: Consultando o vocabulário no SKOSMOS**



Fonte: Capturado pelos autores(2022).

## 6. CONCLUSÃO

Através de um trabalho coordenado e interdependente entre as equipes da Coordenação Geral de Tecnologia da Informação CGTI/Cade e da Universidade de Brasília, as atividades de elaboração deste RT foram planejadas, discutidas, executadas e documentadas.

As atividades envolvidas nesta etapa observaram formalmente à execução dos passos da metodologia elencada para gestão do projeto, conforme definido pelo PMI em seu guia PMBOK.

A equipe da UnB considera que teve acesso as informações necessárias à boa condução dos trabalhos e que a disponibilização dessas informações pela equipe da CGTI/Cade, assim como as atividades conjuntas de análise e discussão levaram à etapa do projeto a bom termo.

Projeto: Cade	Emissão: 20/12/2022	Arquivo: 20221220_RT_Pesquisar e desenvolvimento de uma arquitetura de Ciência da Informação relacionada à Defesa da Concorrência	Pág.94/95
---------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para o Cade.  
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, do Cade.

**Universidade de Brasília – UnB**

**Laboratório de Tecnologias da Tomada de Decisão – LATITUDE**

**Fundação de Apoio à Pesquisa - FUNAPE**

[www.unb.br](http://www.unb.br) [www.latitude.unb.br](http://www.latitude.unb.br) [www.funape.org.br](http://www.funape.org.br)

