



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ
CURSO DE GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

JHONATTAN NASCIMENTO BARBOSA

**ONTOLOGIAS E PROCESSAMENTO DE LINGUAGEM NATURAL NA
PUBLICAÇÃO E ACESSO DE DADOS PARA CARTA DE SERVIÇOS DO CEARÁ**

QUIXADÁ

2022

JHONATTAN NASCIMENTO BARBOSA

ONTOLOGIAS E PROCESSAMENTO DE LINGUAGEM NATURAL NA PUBLICAÇÃO E
ACESSO DE DADOS PARA CARTA DE SERVIÇOS DO CEARÁ

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Sistemas de Informação
do Campus Quixadá da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Sistemas de Informação.

Orientador: Prof. Dr. Davi Romero de
Vasconcelos

QUIXADÁ

2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

B198o Barbosa, Jhonattan Nascimento.
Ontologias e processamento de linguagem natural na publicação e acesso de dados para carta de serviços do Ceará / Jhonattan Nascimento Barbosa. – 2022.
64 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Quixadá, Curso de Sistemas de Informação, Quixadá, 2022.
Orientação: Prof. Dr. Davi Romero de Vasconcelos.

1. Ontologia (Ciência da computação). 2. Serviço público. 3. Chatterbot. I. Título.

CDD 005

JHONATTAN NASCIMENTO BARBOSA

ONTOLOGIAS E PROCESSAMENTO DE LINGUAGEM NATURAL NA PUBLICAÇÃO E
ACESSO DE DADOS PARA CARTA DE SERVIÇOS DO CEARÁ

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Sistemas de Informação
do Campus Quixadá da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Sistemas de Informação.

Aprovada em: ____/____/____.

BANCA EXAMINADORA

Prof. Dr. Davi Romero de Vasconcelos (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Criston Pereira de Souza
Universidade Federal do Ceará (UFC)

Prof. Dr. Paulo de Tarso Guerra Oliveira
Universidade Federal do Ceará (UFC)

AGRADECIMENTOS

Inicialmente agradeço a Deus, por possibilitar que meus objetivos fossem alcançados, durante todos os meus anos de estudos. Agradeço a minha mãe, Francisca Nileda Batista do Nascimento Barbosa, meu pai, Ednelson Nunes Barbosa, minha irmã, Wanessa Nascimento Barbosa, minha avó, Quitéria Nunes Barbosa, por todo o carinho, oportunidades e sacrifícios que proporcionaram durante toda minha vida. Desde já, deixo meu muito obrigado e um amo vocês, pois todos influenciaram no que me tornei atualmente.

Agradeço a minha namorada, Maria Mirna Lima da Silva, por todo companheirismo, as motivações, os momentos vividos tornando-me um ser humano melhor. Saiba que amo-te por tudo, mas especial por sempre ouvir minhas lamentações e incentivar-me para revertê-las. Aos meus amigos de infância, Claivert Nascimento Fernandes, Francisco Raniele Rodrigues da Cruz, Lucas Sales de Araújo, Samuel Sousa Santos, por toda motivação, amizade e força que me deram durante todos esses anos.

Além disso, agradeço aos amigos que encontrei durante estes anos de graduação, pelos momentos de alegria, dificuldade e de estudo que compartilhamos, no qual ampliou nossas experiências longe de casa e cidade natal incrível. Em especial, agradeço ao Francisco Thierry Barros Oliveira, Gerardo Magela dos Santos Filho, Paulo Ravi Feijão Leal e Vitoria Gomes Alves, por terem proporcionado tantos momentos de felicidade e conhecimentos trocados que levarei para toda minha trajetória de vida.

Ao Prof. Dr. Davi Romero de Vasconcelos pelos anos de trabalhos como bolsista e orientando neste trabalho. Por todos as orientações, ensinamentos e por sempre acreditar em mim. Esse período foi de grande importância na minha vida pessoal, acadêmica e profissional. Aos professores Criston Pereira de Sousa e Paulo de Tarso Guerra Oliveira, pela participação na banca avaliadora, pelos conhecimentos, dicas e conselhos repassados em nossas reuniões de projeto que proporcionaram o desenvolvimento deste trabalho.

Ao Inove UFC Quixadá, fica minha gratidão a toda equipe e células de trabalho que possibilitaram uma mescla de cultural com tantas pessoas diferentes que mantive contato. Em especial agradeço, a Ex-Coordenadora Ma. Roberta Dutra de Andrade, que desde meu primeiro semestre abrigou-me como bolsista de iniciação acadêmica e com seus ensinamentos e conselhos sobre inovar e fazer sempre a diferença onde estiver. Além disso fica meu agradecimento, a Coordenadora Ma. Maria Simone Mendes Nunes, que sempre me aconselhou, durante minha estadia como bolsista de extensão e voluntário que suas dicas e oportunidades me fizeram melhor

como pessoa, e me proporcionar uma amizade imensa que levarei para vida toda. No inove passei por momentos que me fizeram amar imensamente esse projeto e desejo que cresça a cada dia mais.

A toda comunidade acadêmica da UFC Quixadá que proporcionou todo conhecimento durante esse período de graduação em ambiente de altíssima qualidade, na qual obtive lições e experiências que levarei para uma vida toda.

Ao final agradeço ao Insight Lab. por inserido em seu ambiente de trabalho como bolsista, pela oportunidade de trabalhar em vertentes inovadoras, que me tornou uma pessoa totalmente diferente e ampliando a cada dia mais meu nível de aprendizagem. E à Fundação Cearense de Apoio ao Desenvolvimento (Funcap), pelo financiamento da pesquisa via bolsa de inovação tecnológica.

“O sonho é que leva a gente para frente. Se a gente for seguir a razão, fica aquietado, acomodado.”

(Ariano Suassuna)

RESUMO

Este projeto busca desenvolver uma base de dados sobre as cartas de serviços, implementar uma ontologia de dados para publicação de dados abertos para a carta de serviço do Ceará integrando informações de cartas de estados adversos ligada a um *chatbot* para facilitar o acesso às informações para as cidades. O trabalho implementou um comparativo de dados sobre as cartas de serviços do estados, almejando uma generalização dos termos e com apoio de tecnologias como o Protegé e a linguagem python, o framework rasa e suas bibliotecas para a criação da ontologia e *chatbot*. Definiu-se os conceitos T-Box e A-Box para preencher a ontologia. Ao preencher a ontologia, realizou-se um pequeno comparativo sobre dois serviços equivalente de cartas de serviços diferentes apontando suas diferenças em determinados campos. Aplicou-se 1400 perguntas testes e obteve-se dos resultado uma acurácia de 0,340 a nível de conversação, 0,477 no nível de entidades e 0.849 a nível de intenções para as associações realizadas pelo *chatbot* sobre sua base de dados treinada.

Palavras-chave: Ontologia (Ciência da computação). Serviço Público. Chatterbot.

ABSTRACT

This project seeks to develop a database on service letters, implement a data ontology for publishing open data for the Ceará service letter, integrating information from adverse status letters linked to a *chatbot* to facilitate access information for cities. The work implemented a comparison of data on state service charters, aiming at a generalization of the terms and with the support of technologies such as Protegé and the python language, the rasa framework and its libraries for the creation of the ontology and *chatbot* . The T-Box and A-Box concepts were defined to fill the ontology. When filling in the ontology, a small comparison was made on two equivalent services from different service letters, pointing out their differences in certain fields. 1400 test questions were applied and the results obtained an accuracy of 0.340 at the conversation level, 0.477 at the entity level and 0.849 at the intention level for the associations made by the chatbot on its trained database.

Keywords: Ontology (Computer Science). Public service. Chatbot.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de uma ontologia de dados	22
Figura 2 – Exemplo da identificação de intenções e entidades	24
Figura 3 – Etapas de procedimentos metodológicos	31
Figura 4 – Aplicação da Ontologia	40
Figura 5 – Serviços de estados diferentes	40
Figura 6 – Funcionamento da sistema proposto	41
Figura 7 – Consulta Sparql automatizada: Console da aplicação	43
Figura 8 – Consulta Sparql automatizada: Tela do Jena Fuseki	43
Figura 9 – Resposta da API do Telegran	44
Figura 10 – Matriz de Confusão de Intenções	46

LISTA DE TABELAS

Tabela 1 – Comparativo entre os trabalhos relacionados e o proposto	30
---	----

LISTA DE QUADROS

Quadro 1 – Especificação de Requisitos	35
Quadro 2 – Tabela de relacionamento entre cartas	36
Quadro 3 – Descrição do T-BOX	38
Quadro 4 – Descrição das Intentions e Actions	41
Quadro 5 – Resultado dos testes automatizados	45

LISTA DE ABREVIATURAS E SIGLAS

TTL	<i>Terse RDF Triple Language</i>
XML	<i>eXtensible Markup Language</i>
RDF	<i>Resource Description Framework</i>
OWL	<i>Web Ontology Language</i>
CSV	<i>Comma-Separated Values</i>
JSON	<i>JavaScript Object Notation</i>
SKOS	<i>Simple Knowledge Organization System</i>
RDFS	<i>Resource Description Framework Schema</i>

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivos	16
<i>1.1.1</i>	<i>Objetivo geral</i>	<i>16</i>
<i>1.1.2</i>	<i>Objetivos específicos</i>	<i>17</i>
1.2	Organização	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	Serviço público	18
<i>2.1.1</i>	<i>Carta de serviços do estado do Ceará</i>	<i>18</i>
2.2	Técnicas de processamento de linguagem natural	19
<i>2.2.1</i>	<i>Mecanismos básicos de limpeza de dados</i>	<i>19</i>
<i>2.2.2</i>	<i>Python</i>	<i>20</i>
<i>2.2.3</i>	<i>Web Scraping</i>	<i>20</i>
2.3	Ontologia	21
<i>2.3.1</i>	<i>O conceito de RDF</i>	<i>21</i>
<i>2.3.2</i>	<i>Modelo de implementação de ontologia</i>	<i>22</i>
2.4	Chatbot	23
<i>2.4.1</i>	<i>Framework Rasa</i>	<i>24</i>
3	TRABALHOS RELACIONADOS	25
3.1	Uso de técnicas de processamento de linguagem natural para identificação de similaridade de serviços públicos	25
3.2	OntoTrans: An Ontology on Transparency	26
3.3	OGDPub: Uma ontologia para publicação de dados abertos governamentais	27
3.4	Public Administration Domain Ontology for a Semantic Web Services EGovernment Framework	28
3.5	Um agente conversacional pergunta-resposta com recomendações baseadas em uma ontologia de domínio	29
3.6	Comparativo entre os trabalhos	30
4	PROCEDIMENTOS METODOLÓGICOS	31
4.1	Coleta de dados: planilha de serviços	31

4.1.1	<i>Limpeza de dados</i>	31
4.2	Ontologia de dados da carta de serviços	32
4.3	Chatbot para recuperação de dados	32
4.3.1	<i>Integração: Chatbot e Ontologia</i>	33
4.3.2	<i>Integração: Chatbot e Telegram</i>	33
4.4	Testes automatizados	33
5	RESULTADOS	34
5.1	Ontologia de dados: Metodologia Neon	34
5.1.1	<i>Fase inicial: definição de requisitos</i>	34
5.1.2	<i>Fase de reuso</i>	34
5.1.3	<i>Fase de implementação</i>	35
5.2	Coleta de dados: planilha de serviços	35
5.2.1	<i>Limpeza de dados</i>	37
5.3	Ontologia de dados: Implementação	37
5.3.1	<i>Conjunto T-BOX</i>	37
5.3.2	<i>Conjunto A-BOX</i>	38
5.3.3	<i>Aplicações de uma ontologia</i>	39
5.3.4	<i>Comparação de serviços de cartas diferentes</i>	40
5.4	Chatbot para recuperação de dados	41
5.4.1	<i>Integração: Chatbot e Ontologia</i>	43
5.4.2	<i>Integração: Chatbot e Telegran</i>	44
5.5	Testes automatizados	45
6	CONCLUSÕES E TRABALHOS FUTUROS	47
7	RECONHECIMENTO	48
	REFERÊNCIAS	49
	APÊNDICE A – PLANILHA DE ANÁLISE DAS CARTAS DE SERVIÇO BRASILEIRAS	51

1 INTRODUÇÃO

A ampliação e a propagação dos sistemas governamentais brasileiros estão se tornando cada vez maiores, devido ao uso de sistemas entre os seus usuários. Cada sistema possui suas próprias peculiaridades, fazendo com que a padronização dos vocabulários de descrições dos serviços tornem-se uma prática desafiadora. Existe uma variedade de vocabulários para serviços públicos encontrados (nacionais ou internacionais) e faz-se necessário selecionar alguns modelos como base com intuito de aproveitar a utilização de conceitos preexistentes que facilitam a comunicação e a troca de informação entre os sistemas.

No estado do Ceará, uma das principais maneiras de publicação de dados abertos de serviços públicos é por intermédio da carta de serviços do estado do Ceará (PLANEJAMENTO *et al.*, 2017). Conforme a Controladoria e Ouvidoria Geral do Estado (CGE) PLANEJAMENTO *et al.* (2017), esse documento público auxilia o governo na melhoria de práticas em suas respectivas organizações. A carta de serviços do estado do Ceará é constituída de 3666 serviços oferecidos pelo governo. Além disso, na carta de serviços estão descritos os canais e formas de acesso, prazos, horários de atendimento, compromissos assumidos e padrão de atendimento sobre as atividades desenvolvidas pelo governo.

Determinadas páginas na internet possuem a implementação de uma API de acesso para solicitar dados que podem ser utilizados para projetos, estudos entre outros modelos (GLEZ-PEÑA *et al.*, 2013). Mas ainda assim, quando não se fizer disponível, pode-se utilizar a *web scraping* para extração de informações de páginas para arquivos estruturados, desde que seja efetuada em concordância com os termos de utilização que cada plataforma disponibiliza. De acordo com GLEZ-PEÑA *et al.* (2013), a técnica de *web scraping* pode ser definida como o método de extrair e combinar essas informações capturadas de maneira sistêmica para criação de bases de dados.

Segundo BERNERS-LEE *et al.* (2001), o conceito de Web Semântica consiste no modo de expressar e organizar padrões de dados colaborativos na rede. Inicialmente, o termo ontologia originou-se do grego *ontos* o “Ser” e *logia* o “Estudo”. Para GRUBER (2002), na inteligência artificial as ontologias são conjuntos de especificações de dados que visam auxiliar no compartilhamento e divulgação de informações na Web Semântica. Sendo assim, é possível expressar essas informações em vários formatos: *Terse RDF Triple Language* (TTL), *eXtensible Markup Language* (XML), *Resource Description Framework* (RDF) e *Web Ontology Language* (OWL). Além disso, a partir dessas bases de conhecimento, faz-se necessária uma linguagem de

consulta nomeada por *SPARQL*, que atua na recuperação de dados da maneira desejada.

Impulsionado pela necessidade do acesso rápido as informações, os *chatbots* surgiram na década de 60 visando simular a interação humana e proporcionar rapidez na solução de dúvidas do dia-a-dia. Por exemplo, o *chatbot* Eliza, criado pelo Massachusetts Institute of Technology (MIT), surgiu com propósito inicial de emular uma psicóloga e efetuar perguntas conforme o que seus usuários respondiam durante a conversa. De acordo com a pesquisa feita pela Salesforce GROUP *et al.* (2018), 69% dos consumidores preferem o atendimento de *chatbots* para pesquisas simples e complexas em relação aos aplicativos, devido a sua agilidade de resposta ao usuário. Nota-se que a demanda de acesso à informação em qualquer lugar é essencial na adoção de práticas como esta.

No desenvolvimento de *chatbots*, utilizam-se vários *frameworks* que auxiliam na implementação desses assistentes, como o *framework* Rasa. Consoante a sua documentação, o Rasa é uma estrutura de aprendizado de máquina de código aberto para criação de assistentes e *chatbots* (RASA, 2020). Sendo assim, este *framework* é facilmente implementável, pois não necessita de experiência em linguagem de programação para desenvolver os assistentes.

Por esses motivos, o trabalho apresenta uma ontologia de dados que descreva a carta de serviços. Além disso, a criação de um *chatbot* que facilite a recuperação de informações sobre a ontologia criada. Posto isto, a principal contribuição deste trabalho é o modo de descrever as informações dos serviços públicos, propondo-se a facilitar o acesso aos dados da carta de serviços de maneira prática para o cidadão comum.

1.1 Objetivos

Nesta seção são apresentados os objetivos gerais e específicos para este trabalho.

1.1.1 Objetivo geral

Desenvolver uma base de informações do site oficial da carta de serviços do Ceará e com a mesma propor uma ontologia que visa descrevê-lá com propósito de auxiliar na busca e recuperação de dados por intermédio de um *chatbot* treinado e desenvolvido utilizando-se da ontologia criada.

1.1.2 *Objetivos específicos*

- Criar uma base de dados em arquivo *Comma-Separated Values* (CSV) da carta de serviços com a técnica de *web scraping*;
- Realizar um comparativo de campos com as cartas de serviços de todos os estados visando uma generalização de termos;
- Propor uma ontologia para a carta de serviços do estado do Ceará utilizando a base de dados criada;
- Integrar a ontologia proposta com os dados da carta de serviços de outros estados;
- Desenvolver um *chatbot* com *framework* *rasa*, alimentado pela ontologia projetada.

1.2 **Organização**

Este trabalho está dividido da seguinte forma: na Seção 2, são apresentados conceitos necessários para o entendimento do trabalho. Na Seção 3, são retratados alguns trabalhos relacionados com o método proposto. Na Seção 4, são apresentados os procedimentos metodológicos que foram seguidos neste trabalho. Na Seção 5, são mostrados os resultados obtidos durante a elaboração do projeto. Na Seção 6, são representadas algumas considerações finais sobre o projeto atrelado de possíveis trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, são apresentados os conceitos fundamentais para o entendimento e desenvolvimento do trabalho.

2.1 Serviço público

Entende-se o conceito de serviço público como uma prestação de serviço a população exercida por responsabilidade do Estado ou por uma instituição privada, por intermédio de uma delegação estatal, com base no que a lei determina pelo artigo 216 da Constituição Federal Brasil (1988). No Brasil, é uma obrigação do Estado a publicação dos dados dos serviços públicos fornecidos pelo mesmo, instituído pela Cartilha de Acesso à Informação pela lei de nº 12.527, no decreto nº 7.724 UNIÃO (2011).

Conforme o trabalho de PEREIRA *et al.* (2021), foi identificado que nem todos os estados possuem uma forma de acesso aos dados abertos simplificado ao cidadão leigo. De acordo com UNIÃO (2017), pela lei de nº 13.460, o art. 7 descreve o conceito de carta de serviços ao cidadão que visa informar ao usuário as informações sobre os serviços prestados por cada respectivo órgão, secretária ou entidade sob controle do estado. Logo, é importante disponibilizar a carta de forma online, simplificando o acesso aos dados para o cidadão comum.

2.1.1 Carta de serviços do estado do Ceará

No estado do Ceará, uma das principais maneiras de publicação de dados abertos de serviços públicos é por intermédio da carta de serviços do estado do Ceará (PLANEJAMENTO *et al.*, 2017). Conforme a Controladoria e Ouvidoria Geral do Estado (CGE) PLANEJAMENTO *et al.* (2017), esse documento público auxilia o governo na melhoria de práticas em suas respectivas organizações.

A carta de serviços do estado do Ceará¹ é constituída de 3666 serviços oferecidos e fornecidos pelo governo. Além disso, ela serve como ponte entre os dados públicos e os cidadãos, onde estão descritos os canais e formas de acesso, prazos, horários de atendimento, compromissos assumidos e padrão de atendimento sobre as atividades desenvolvidas no governo.

¹ Site oficial: cartadeservicos.ce.gov.br

2.2 Técnicas de processamento de linguagem natural

Conforme BEYSOLOW (2018), o Processamento de Linguagem Natural (PLN) é uma área de estudo da inteligência artificial que auxilia na manipulação e interpretação de dados da linguagem humana. Desse modo, um dos grandes desafios previstos em PLN é a viabilidade de extrair o sentido e contexto da linguagem natural e torná-la compreensível para a máquina (BEYSOLOW, 2018). Além disso, existem outros mecanismos que visam ajudar o computador, como a limpeza de dados, objetivando a eliminação de informações desnecessárias e redução do vocabulário, mantendo a essencialidade do contexto textual.

Segundo CAMPOS *et al.* (2005), a máquina trabalha com generalizações e hierarquias entre as entidades identificadas, considerando suas respectivas classes e subclasses. O conceito de aprendizado de máquina, definido por Mitchell (1997), explora a pesquisa e a construção de algoritmos que podem aprender com os erros, mas também realizar previsões sobre os dados. Esses algoritmos atuam construindo modelos através de entradas predefinidas para fazer previsões ou decisões com base em dados, em vez de simplesmente seguir instruções de programação imperativas e estáticas.

2.2.1 Mecanismos básicos de limpeza de dados

De fato, a variedade linguística do português é uma das principais causas que dificultam o processamento de dados. Com intuito de preservar a essência do texto, a utilização das técnicas de limpeza de dados são de grande importância para impulsionar o crescimento do aprendizado da máquina. Conforme HAN *et al.* (2006), a redução de preposições, palavras e outros símbolos desnecessários aumentam a compreensão e processamento de informações da máquina. Em busca da melhoria terminológica dos dados neste presente trabalho, as técnicas de limpeza a serem aplicadas são:

- Transformação do texto em minúscula (Sensibilidade ao caso);
- Remoção de palavras de parada (*Stop Words*);
- Radicalização das palavras (*Lematização*);
- Remoção das palavras curtas (palavras até dois caracteres);
- Remoção de pontuações.

2.2.2 Python

A linguagem *python* foi criada por Guido van Rossum, com auxílio de dois colegas, Jack Jansen e Sjoerd Mullender, como um simples passatempo entre os amigos. O objetivo dos criadores era desenvolver uma linguagem orientada a objetos, com alta portabilidade e menos complexa do que Java ou C++ (SONGINI, 2005). O autor ainda incrementa que os desenvolvedores normalmente atraem-se pelo *python* por causa de sua rapidez em processamento, visto que isso aumenta sua produtividade e amplitude no desenvolvimento de aplicações.

Por outro lado, *python* vem tornando-se muito conhecida por adeptos de aplicações de código aberto e suas comunidades. Além disso, recentemente obteve apoio em empresas, como Google, Amazon, Facebook e entre outras que utilizam a linguagem em áreas da ciência de dados, com apoio de diversos *frameworks* e bibliotecas que apoiam este estilo de desenvolvimento.

2.2.3 Web Scraping

De acordo com GLEZ-PEÑA *et al.* (2013), a técnica de *web scraping* pode ser definida como o método de extrair e combinar essas informações capturadas de maneira sistêmica. Nesse processo, o *bot* visa imitar as interações de um humano durante a navegação nas páginas da web. Passo a passo, o *bot* visita uma determinada quantidade de sites e analisa o conteúdo, para encontrar e extrair dados que o desenvolver define em código. Por fim, cria uma base de dados com o conteúdo obtido em um arquivo de dados estruturados.

Além disso, as APIs e estruturas do *web scraping* na Web abordam as tarefas mais comuns que alguns rastreadores de dados utilizam, conforme descrito abaixo:

- **Acesso ao site:** O *bot* comunica-se com o site de destino por meio do protocolo HTTP, que consiste em um protocolo de internet sem estado e baseado em texto, que coordena as transações de solicitação-resposta. Uma vez conectado ao documento HTML recuperado, o *bot* pode extrair o conteúdo do seu interesse.
- **Análise de HTML e extração de conteúdo:** Nesta fase inicia-se a análise do documento HTML para possível extração de informações. Os testes nas marcações encontradas na página são realizadas para maximizar a quantidade de conteúdo coletado. Para este propósito, a combinação entre as estruturas do documento HTML e as linguagens baseadas em seletor, como (XPath) e sintaxe de seletor (CSS), visam encontrar os dados o mais fiéis possível ao que encontra-se na página da web.

- **Construção de saída:** Ao final desse percurso, o objetivo principal da técnica é transformar o conteúdo extraído em uma representação estruturada adequada para posterior análise e armazenamento. Esta última etapa é trivial para *web scraping* e algumas ferramentas cientes dos resultados de pós-processamento, fornecem estruturas de dados na memória e soluções baseadas em texto, como strings ou arquivos (geralmente arquivos XML, *JavaScript Object Notation* (JSON) ou CSV) para armazenamento dos dados.

2.3 Ontologia

Inicialmente, o termo ontologia originou-se do grego *ontos* o “Ser” e *logia* o “Estudo” formando então o “Estudo do Ser”. Sendo assim, o termo foi absorvido pela ciência da computação, especificamente na Inteligência Artificial, para especificação de bases de conhecimento onde permita-se interligar dados de diversas bases relacionadas, transformando o seu sentido original da palavra.

Para GRUBER (2002), ontologias são conjuntos de especificações de dados que visam auxiliar no compartilhamento e divulgação de informações, almejando a inferência de dados entre outras ontologias. O conceito de ontologia, na computação, surge como uma estrutura de informações que é uma das principais bases da Web Semântica. Devido às hierarquias definidas por meio dela, suas relações visam deduzir uma representação que auxilie a identificação de dados para máquina.

2.3.1 O conceito de RDF

De acordo com World Wide Web Consortium (2004), RDF é um conjunto de especificações da W3C, inicialmente desenvolvido como um modelo de dados e metadados. Com ela é possível descrever informações conceituais na internet, utilizando-se de sintaxes e formatos de serialização.

No contexto de ontologias, necessita-se do conceito de RDF sendo uma linguagem de metadados para explicitar as relações entre os recursos introduzidos. Conforme Manola *et al.* (2004), o RDF permite definir quais são os atributos e propriedades relacionados com seus respectivos domínios representados. A W3C padroniza o conceito de RDF para ser interpretado pela máquina por apresentar uma forma comum de descrições. É possível classificar suas definições como:

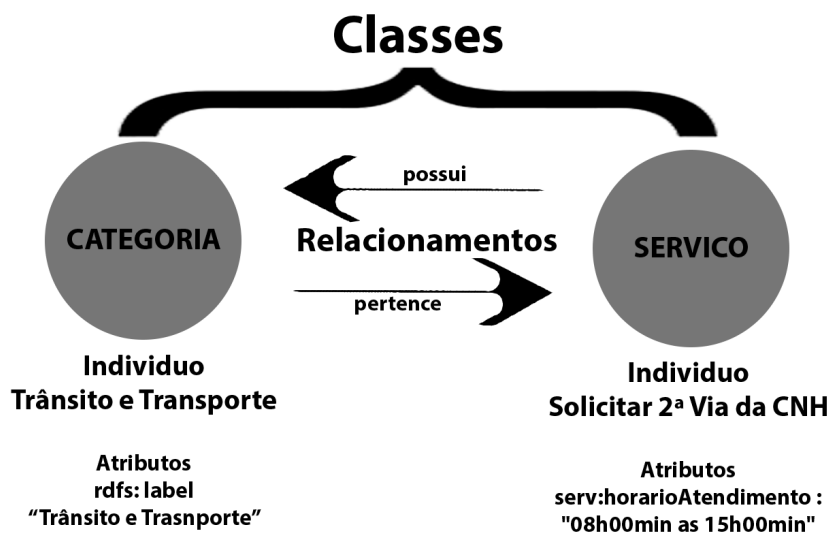
- O **recurso** pode ser qualquer definição que possa ser representada por um identificador na Web. Por exemplo: "https://dbpedia.org/data/Cristiano_Ronaldo.ttl";
- A **propriedade** indica um dado específico de um recurso. Por exemplo: "Person" ou "birthDate";
- O **valor** é expressamente o valor da propriedade de determinado recurso. Por exemplo: "Manchester United FC" ou "https://dbpedia.org/page/Manchester_United_F.C.".

2.3.2 Modelo de implementação de ontologia

Os modelos de implementações em ontologias compartilham semelhanças estruturais, independentemente do contexto e da linguagem em que são expressas. Essas semelhanças são estruturadas em indivíduos, classes, atributos e relacionamentos, podendo ser definidas como:

- Os **indivíduos** (*individuals*) são os objetos propriamente exemplificados;
- As **classes** (*class*) são os conjuntos, as coleções ou categorias de objetos desenvolvidos;
- Os **atributos** (*datatype properties*) são as propriedades ou características atribuídas aos indivíduos que podem possuir ou compartilhar entre os outros;
- Os **relacionamentos** (*object properties*) são as formas como os objetos se relacionam uns com os outros de maneira a criar hierarquias entre si.

Figura 1 – Exemplo de uma ontologia de dados



Fonte: Elaborado pelo autor (2022).

Na Figura 1, encontra-se um exemplo descrito de uma pequena ontologia. A mesma foi definida por duas classes (categoria e serviço), que se relacionam diretamente entre si nas

object properties (pertence e possui), onde um serviço possui uma categoria da mesma forma que uma categoria pertence a um serviço. Além disso, são definidas por indivíduos que representam a mesma (Trânsito e Transporte e Solicitar 2ª via da CNH) e, por fim, estão exemplificados alguns dos atributos daquele indivíduo (*label*: Trânsito e Transporte e o *horarioAtendimento*: 08h00min às 15h00min).

A metodologia NeOn de Suárez-Figueroa (2010) é um dos modelos para criação de ontologia no qual permite a abordagem de desenvolvimento de software, fornecendo detalhes em diversas etapas descritas. Além disso, esta metodologia trata a atividade de levantamento de requisitos como uma etapa do processo de elaboração de ontologias, fornecendo orientações explícitas neste quesito (SUÁREZ-FIGUEROA, 2010). O desenvolvimento de ontologias utiliza de cinco etapas básicas da metodologia NeOn, sendo elas: 1- Fase inicial: definição de requisitos, 2- Fase de reuso, 3- Fase de Reengenharia, 4-Fase de Design e 5- Fase de implementação.

2.4 Chatbot

Os *chatbots* surgiram de uma necessidade de facilitar e automatizar a busca e troca de informações. Um dos primeiros conceitos para construção de um *chatbot* inteligente baseia-se no artigo de Turing (1950), onde o Teste de Turing tinha o propósito de colocar uma máquina a teste que buscavam persuadir os juízes passando-se por um humano. Esse foi um dos primeiros passos onde o computador, em uma conversação com humanos, alcançou o objetivo de ampliar o nível de informações que poderiam ser fornecidas e trocadas entre eles. É possível observar que, atualmente, o Teste de Turing foi um dos principais fatores que proporcionou a evolução tecnológica de inúmeros sistemas com inteligência artificial que surgiram por meio desses estudos, como: *Siri* (2010), *Google Now* (2012), *Alexa* (2015) etc.

Atualmente, os *chatbots* são implementados em diferentes plataformas, desde aplicações web até aplicativos de mensagens, buscando viabilidade e aumento da experiência do usuário com o que ele deseja encontrar. O *chatbot* possui um diferencial para o usuário, pois consiste em uma ferramenta que pretende reproduzir uma discussão inteligente com os humanos utilizando linguagem comum, além de aprender com o passar do tempo com base no que lhe é ensinado.

2.4.1 Framework Rasa

No desenvolvimento de *chatbots*, utilizam-se vários *frameworks* que auxiliam na implementação desses assistentes, como o *framework* Rasa. Consoante a sua documentação, o Rasa é uma estrutura de aprendizado de máquina de código aberto para criação de assistentes e *chatbots* (RASA, 2020). Este *framework* é facilmente implementável por não necessitar de um vasto conhecimento em linguagem de programação para desenvolver os assistentes. Além disso, na constituição do *chatbot* Rasa, são estabelecidos padrões que auxiliam no desenvolvimento, como:

- **Intents**: É classificado como a intenção da mensagem recebida do usuário, com base no que foi anteriormente determinado;
- **Entities**: Pode ser denominado como parte específica da mensagem que auxilia na compreensão do que o usuário deseja saber, a partir do que foi definido previamente;
- **Actions**: É responsável pelo o que o *chatbot* pode fazer ou acessar, segundo o que foi identificado pela *intents* e *entities*;
- **Stories**: São os caminhos para onde a conversa pode prosseguir de acordo com o passo a passo previamente estabelecido, visando ampliar o alcance que o *chatbot* pode reagir de acordo com o usuário.

Na Figura 2, o cidadão efetua a seguinte pergunta: "Quais os quesitos para solicitar uma Cirurgia dermatológica?". O Rasa, por possuir uma base pré-treinada com exemplares de perguntas que podem ser feitas, identifica inicialmente a intenção da pergunta recebida por ele (no exemplo, reqServ). Por fim, averígua-se as entidades preestabelecidas na base treinada e compara-se com as presentes na intenção, como encontrado no exemplo (*requisito*: quesitos e *nomServico*: Cirurgia dermatológica).

Figura 2 – Exemplo da identificação de intenções e entidades

Pergunta:

**Quais os [quesitos] para solicitar
uma [Cirurgia dermatológica]?**



Intenção: reqServ
Entidades: requisitos e nomServico

Fonte: Elaborado pelo autor (2022).

3 TRABALHOS RELACIONADOS

O campo da inteligência artificial é uma das áreas mais estudadas por cientistas de dados em todo planeta. Neste trabalho focou-se no desenvolvimento de ontologias para carta de serviços do ceará e integrar cartas de outros estados, com apoio de técnicas de processamento de linguagem natural e automatização de consulta de dados por meio de um *chatbot* *rasa framework*. Utilizou-se cinco trabalhos relacionados tomados como base para este experimento.

3.1 Uso de técnicas de processamento de linguagem natural para identificação de similaridade de serviços públicos

O estudo realizado Campos e Figueiredo (2021), pretende projetar um experimento para avaliar alternativas de prestação de serviços públicos aos cidadãos, com base em serviços públicos semelhantes aos que normalmente utilizam. Além disso, utiliza-se de uma conduta de filtragem com base no conteúdo de descrição dos serviços públicos disponíveis, operando quanto ao texto com técnicas de processamento de linguagem natural (PLN). Tendo em vista as técnicas utilizadas, busca-se ainda analisar a similaridade de assuntos entre os serviços utilizados pelo cidadão.

Nesse projeto utilizou-se pré-processamento para abstrair o nível de complexidade do idioma, com propósito de tornar o conteúdo essencial e simplificado para a máquina. Outra técnica importante utilizada na análise determina a contagem com a importância das palavras no texto, *Term Frequency Inverse Document Frequency* (TF-IDF), com ela buscou-se medir estatisticamente a importância das palavras em um texto.

Além disso, a base de dados utilizada neste trabalho configura-se nos Serviços Públicos do Estado do Mato Grosso. A mesma possui a relação entre setenta e um serviços disponíveis ao cidadão e os dados foram coletados especificamente dos portais das instituições públicas responsáveis. Durante o trabalho, foram analisadas quatro características sobre os respectivos serviços, contendo um identificador do serviço (*id_servico*), o nome do serviço (*servico_nome*), uma descrição para o serviço (*servico_descricao*) e a área de negócio responsável (*servico_area*).

Por fim, os dados foram combinados com características textuais relacionadas aos seus respectivos serviços e definiu-se as oito técnicas utilizadas para higienização dos mesmos. Com objetivo de analisar como se comportam os dados após a limpeza feita. As formas de

limpeza utilizadas consistem em:

- Separação do texto em palavras (*Tokenização*)
- Transformação do texto em minúscula (Sensibilidade ao caso)
- Remoção de palavras de parada (*Stop Words*)
- Remoção de palavras específicas (Palavras excluídas)
- Radicalização das palavras (*Lematização*)
- Remoção das palavras curtas (palavras até dois caracteres)
- Remoção de pontuações
- Concatenação das palavras (*Join*) após limpeza

Portanto, o projeto buscou comprovar a similaridade entre os serviços para que informações sejam oferecidas ao cidadão com base no que ele mais acessa. Sendo assim, o cidadão pode acessar outros serviços relacionados com base na semelhança que ambos possuem.

3.2 OntoTrans: An Ontology on Transparency

No projeto de pesquisa de Pereira *et al.* (2020), apresenta-se uma ontologia de domínio sobre os dados dos portais de transparência do Estado brasileiro e sua codificação em OWL. A OntoTrans apoiará em uma análise, o compartilhamento e a reutilização do conhecimento nesse domínio, além de ser um artefato computacional capaz de contribuir para o entendimento das relações entre outros conceitos relacionados à transparência, e ser processável por aplicativos que manipulam esse conhecimento.

Ontologias fundacionais descrevem conceitos gerais independentes de um domínio e definem com precisão meta-propriedades para tornar a semântica de cada conceito na ontologia explícita. *Unified Foundational Ontology* (UFO) é um exemplo de uma ontologia fundamental desenvolvida com base em várias teorias da Ontologia Formal, Lógica Filosófica, Filosofia da Linguagem, Linguística e Psicologia Cognitiva (GUIZZARDI, 2005). É composto de três partes principais, a saber, UFO-A, UFO-B e UFO-C. UFO-A é uma ontologia. UFO-B é uma ontologia de perdurantes (eventos e processos). UFO-C é uma ontologia de entidades sociais (endurantes e perdurantes) construídas no topo de UFO-A e UFO-B. Como modelo conceitual estrutural, os conceitos da OntoTrans foram fundados nas construções do UFO-A.

Para desenvolver o trabalho de Pereira *et al.* (2020), adotou-se a metodologia de desenvolvimento de ontologias proposta por USCHOLD e KING (1995) e estendida por Uschold e Gruninger (1996). Essa metodologia foi escolhida devido à sua facilidade de uso, sua indepen-

dência de aplicação, a flexibilidade de seus modelos e o fato da sua amplitude de aplicações em diversos domínios de negócios (GÓMEZ-PÉREZ *et al.*, 2004). Utilizou-se o OntoUML para a construção usando a ferramenta OLED e gerando um código OWL.

Na avaliação do projeto, pode-se extrair conhecimento relevante sobre transparência codificado na ontologia e não facilmente acessível através de sua estrutura. Os resultados da consulta demonstram fortemente a correção da ontologia OWL em sua especificação conceitual, confirmando as afirmações da seção OntoTrans. Os resultados da avaliação também mostram que a OntoTrans pode ajudar a automatizar a aplicação da transparência. Além disso, os resultados obtidos no trabalho também mostram que a ontologia construída é um artefato computacional capaz de agregar valor nos mais diferentes domínios de cenários de aplicação de transparência, visto que é genérica e pode ser instanciada no domínio escolhido pelo usuário.

3.3 OGDPub: Uma ontologia para publicação de dados abertos governamentais

O trabalho de Pereira e Todesco (2020), procurou elaborar uma ontologia para publicação de dados abertos para municípios brasileiros. Buscando então disponibilizar aos cidadãos e as organizações, a par das atividades desenvolvidas por esses serviços. A ontologia foi desenvolvida a partir de uma metodologia NeOn e representada a linguagem OWL. A metodologia utilizada, propõe um fluxo flexível na sugestão de atividades para construção de redes ontológicas.

De fato, a ontologia proposta foi desenvolvida através do cenário de reuso e reengenharia da metodologia NeOn, onde, por definição, o reuso de recursos de outra ontologia para solucionar o problema. Já quanto ao conceito de reengenharia é baseado no processo de recuperação e transformação em um modelo conceitual desenvolvido e adaptado para o problema específico do projeto. Dessa maneira, a OGDPub concilia um fornecimento de metadados descritivos dos *datasets*, além de permitir uma representação municipal responsável pelo serviço. Com objetivo, de esquematizar os dados de forma mais compreensível ao cidadão.

Ao final do processo de busca de dados em outros Portais Abertos para auxiliar na classificação proposta foram definidos os seguintes temas: (1) Administração Pública, (2) Agricultura, (3) Ciência e Tecnologia, (4) Cultura, Esporte e Lazer, (5) Demografia, (6) Economia, (7) Educação, (8) Finanças, (9) Indústria, Comércio e Serviços, (10) Infraestrutura e Planejamento Urbano, (11) Legislação, (12) Recursos Naturais e Meio Ambiente, (13) Saúde, (14) Segurança Pública, (15) Sociedade e Bem-estar, (16) Transporte e Mobilidade e (17) Turismo. A composição

dessa classificação é baseada no tema global usado pelos portais de análise, ou seja, um tema comum encontrado em todos ou na maioria dos portais de análise (embora a nomenclatura utilizada possa diferir em alguns casos). É importante ressaltar que o OGD_{Pub} é escalável, ou seja, novos temas podem ser implementados de acordo com as necessidades específicas de cada cidade.

Por fim, observa-se que a ontologia foi implementada com a ferramenta *Protegé* com dezesseis classes, vinte e seis *object property* e dezesseis *data property*. Além disso, utiliza conceitos oriundos de ontologias pré-existentes como: *The Organization Ontology* (Ontologia das Organizações), *Data Catalog Vocabulary* (DCAT), *Simple Knowledge Organization System* (SKOS) e *Friend of a Friend* (FOAF).

3.4 Public Administration Domain Ontology for a Semantic Web Services EGovernment Framework

O projeto de GOUDOS *et al.* (2007) propôs uma ontologia da domínio geral administração pública. Define um modelo formal para serviços de administração pública baseados em *Web Service Modeling Ontology* (WSMO). Para isso utilizou-se o modelo objeto de serviço comum da *Generic Governance Enterprise Architecture* (GEA) que fornece semântica de governança comum específica do domínio. Descrito por meio de uma ontologia usando *Web Service Modeling Language* (WSML), implementando uma ontologia de domínio arquitetura para serviços da Web Semântica de Governo Eletrônico.

A estrutura WSMO é utilizada no projeto para modelar os serviços de administração pública caracterizados em cinco tipos diferentes como: serviços de declaração, certificação, controle, autorização e produção. Construindo uma ontologia de domínio *Generic Governance Enterprise Architecture* (GEA) em cima de especificações neutras de domínio *Web Service Modeling Ontology* (WSMO) estabelecidas por essas características. Todos os elementos do modelo de serviço *Web Service Modeling Ontology* (WSMO) podem compartilhar uma ontologia *Generic Governance Enterprise Architecture* (GEA) comum. Cada objeto entidade do modelo de serviço administração pública é fornecido por um conceito de ontologia correspondente e seus relacionamentos, funções, instâncias e axiomas.

O trabalho é baseado no modelo *Generic Governance Enterprise Architecture* (GEA) implementado recentemente em OWL. Para aproveitar ao máximo os recursos da estrutura *Web Service Modeling Ontology* (WSMO) e construir serviços Web semânticos para PA, optamos

por expressar o modelo de objeto *Generic Governance Enterprise Architecture* (GEA) em *Web Service Modeling Language* (WSML).

3.5 Um agente conversacional pergunta-resposta com recomendações baseadas em uma ontologia de domínio

No projeto desenvolvido por BARBOSA (2020), uma construção de agente conversacional de recomendação baseado em ontologia. Iniciou-se o trabalho aplicando o módulo de *Interactive Question Answering* (IQA) onde realizamos o processamento da pergunta do usuário, a consulta da resposta no banco e o envio dela para o *chatbot*, bem como o envio de informações para a árvore de recomendação para realizar sua sugestão. Detalhando os processos utilizados na identificação do tipo da pergunta e o processo de recuperação de informações na base de conhecimento.

Na etapa de pré processamento do projeto, encarregou-se de realizar a *tokenização* do texto, e a retirada de caracteres indesejáveis para as tarefas posteriores. A pergunta do usuário passa pelo pré-processamento e é alimentada pelo modelo de rede neural *transformer* pré-treinado, que pode ser opcionalmente o BERT, ConveRT e GloVe. A saída será um *embedding*, que é uma representação em forma de vetor de valores reais da questão, onde sentenças similares terão representações vetoriais mais próximas. Utilizou-se o *framework* Rasa para classificação e extração conjunta de intenções e entidades.

Em seguida, inclui-se o funcionamento do sistema de recomendação baseado em ontologia. Inicialmente, tornou-se necessário converter os arquivos de ontologia em objetos *OntologyTrees*. Para coletar dados de interesse, você precisa percorrer a árvore camada por camada, uma varredura também conhecida como busca em largura. No entanto, nesta visita, poucos elementos relevantes podem ser retornados, para resolver este problema, definimos algumas restrições, como: número de sugestões; ordem de exibição; posição do nó ancestral e profundidade da consulta; com base no *domínio* e/ou *intervalo* limite.

A principal contribuição neste trabalho foi o desenvolvimento de um *chatbot* que utiliza recomendações com base na hierarquia das propriedades da ontologia. Onde primeiramente extraiu-se as relações entre as entidades identificadas na pergunta. Em seguida, faz-se uso delas com intuito de extrair propriedades ou classes relacionadas que, ao serem substituídas, mantêm seu sentido semântico.

3.6 Comparativo entre os trabalhos

Na Tabela 1, apresenta-se a tabela comparativa entre os trabalhos anteriormente citados com relação ao proposto neste projeto, onde propõe-se utilizar as técnicas de processamento de linguagem natural para higienização dos dados. Além disso, criar uma ontologia de serviços públicos para carta de serviços do estado do Ceará. Por fim, implementar um *chatbot* que alimente-se da ontologia desenvolvida auxiliando no acesso aos dados da carta.

Tabela 1 – Comparativo entre os trabalhos relacionados e o proposto

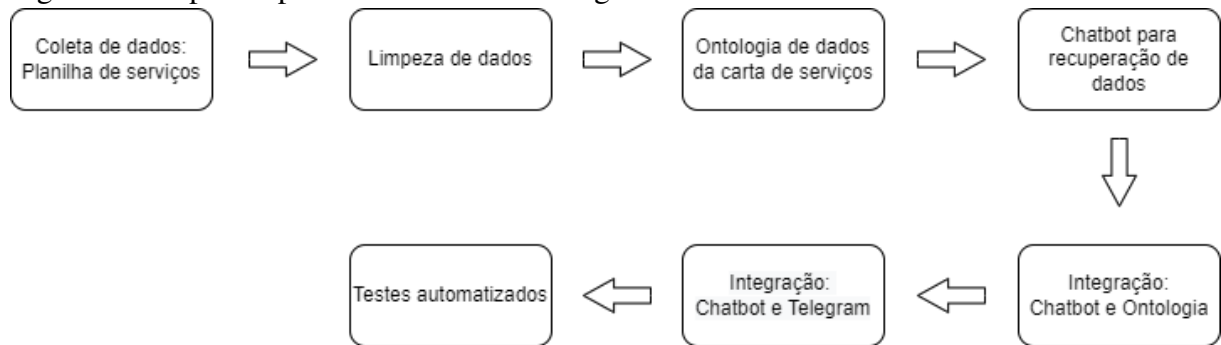
Trabalhos	Serviço Público	Limpeza de dados	Ontologia de dados	Chatbot	Testes Automatizados
Campos e Figueiredo (2021)	Sim	Sim	Não	Não	Não
Pereira <i>et al.</i> (2020)	Sim	Não	Sim	Não	Não
Pereira e Todesco (2020)	Sim	Não	Sim	Não	Não
GOUDOS <i>et al.</i> (2007)	Sim	Não	Sim	Não	Não
BARBOSA (2020)	Sim	Sim	Sim	Sim	Não
Trabalho proposto	Sim	Sim	Sim	Sim	Sim

Fonte: Elaborado pelo autor (2022).

4 PROCEDIMENTOS METODOLÓGICOS

Nesta seção, são descritas as etapas de execução para atingir os objetivos deste trabalho. Na Figura 3, está representado o passo a passo executado neste trabalho, que foi dividido em: Coleta de dados na planilha de serviços; Limpeza de dados; Ontologia de dados da carta de serviços; *Chatbot* para recuperação de dados; Integração: *Chatbot* e Ontologia; Integração do *Chatbot* e Telegram e Testes automatizados.

Figura 3 – Etapas de procedimentos metodológicos



Fonte: Elaborado pelo autor (2022).

4.1 Coleta de dados: planilha de serviços

O conjunto de dados da carta de serviços do estado do Ceará é a base de dados que foi utilizada para implementar este projeto. Os dados estão armazenados em um arquivo CSV, na qual foi preenchido por meio de uma aplicação em *python* que realizou a extração de dados com a técnica de *web scraping* do site oficial da carta de serviços ², com apoio das bibliotecas *scrapy*, *rdflib*, *pandas* e *unidecode*. Além disso, a realização de um levantamento quanto as cartas de serviços de outros estados para definir conceitos generalistas que possibilitem a implementação de qualquer outra carta de serviços na estrutura a ser definida para a ontologia de dados.

4.1.1 Limpeza de dados

Para facilitar o processamento de dados, foi realizada uma limpeza de dados buscando facilitar a interpretação dos dados pela máquina. Utilizou-se a linguagem *python* e as respectivas bibliotecas: *pandas*, *math* e *re* (para expressões regulares), visando alcançar os seguintes conceitos preestabelecidos anteriormente.

² Site oficial: cartadeservicos.ce.gov.br

- Transformação do texto em minúscula (Sensibilidade ao caso);
- Remoção de palavras de parada (*Stop Words*);
- Radicalização das palavras (*Lematização*);
- Remoção das palavras curtas (palavras até dois caracteres);
- Remoção de pontuações.

4.2 Ontologia de dados da carta de serviços

Primeiramente, faz-se necessário determinar quais conceitos serão definidos como *class*, *object properties* necessárias e a criação das *data properties* de acordo com a base de dados. A ontologia da carta de serviços, no qual é criado neste projeto e implementado em duas etapas: via *Protegé*, para desenvolvimento do conjunto *T-BOX* (Componente de terminologia) dos dados e via linguagem *python* a implementação do conjunto *A-BOX* (Componente de asserção), com apoio das bibliotecas *rdflib*, *pandas* e *unidecode*.

Buscando definir uma classificação coerente, foram averiguados Portais de Dados Abertos de outros países que trabalham com a publicação de dados abertos, como: Brasil, Finlândia, Inglaterra e o *data.gov.uk* (Portal de Dados Abertos do Reino Unido). Desse modo, a ontologia utiliza o novo arquivo CSV pré-processado para alimentar-se e preencher o componente *T-BOX* criado, além disso ela utiliza padrões já pré-existentes de outras ontologias: *Simple Knowledge Organization System (SKOS)* e *Resource Description Framework Schema (RDFS)*.

Por fim, após a criação da ontologia fez-se necessário a implantação de um *Endpoint Sparql* via *Apache Jena Fuseki* para a execução de consultas *Sparqls* testes e averiguar a funcionalidade na recuperação de dados buscados para facilitar a integração com outras aplicações.

4.3 Chatbot para recuperação de dados

Nesta fase do projeto, foi desenvolvido o *chatbot* por intermédio da linguagem *python* e apoio do *Framework Rasa* em sua versão 3.0. Além disso, na constituição do *chatbot* *Rasa* são estabelecidos padrões que auxiliam no desenvolvimento, como: *Intents*; *Entities*; *Actions*; *Stories*.

Os dados são previamente inseridos e pré-processados no arquivo ".yml" do projeto, onde são expressos quais são as respectivas intenções e entidades que devem ser identificadas para cada possível entrada. Por meio das intenções e entidades identificadas, torna-se possível

criar consultas *sparqls* automaticamente e relacionadas ao que foi detectado nas ações criadas.

4.3.1 Integração: Chatbot e Ontologia

O objetivo de criar um *chatbot* para integrar com a ontologia, atua diretamente no contexto da conversa, além de ampliar o alcance na interpretação dos objetivos do usuário. Devido à criação do *endpoint*, torna-se viável a implantação do mesmo, pois ao realizar a criação das consultas automatizadas é possível enviá-las e receber um *JavaScript Object Notation* (JSON) que pode facilmente ser retornado ao usuário via *python*, com apoio da biblioteca *Sparql Wrapper*.

4.3.2 Integração: Chatbot e Telegram

Nesta etapa, o chatbot foi integrado à API do Telegram com o *Rasa Endpoint*, usando as configurações definidas no *endpoint* de ações, conforme necessário na documentação do *framework* Rasa (RASA, 2020). O aplicativo Telegram é um serviço de mensagens simultâneas amplamente globalizado e de fácil acesso a todo e qualquer pessoa que possua um dispositivo com acesso a internet, sendo assim ele se encaixa perfeitamente na disponibilização de dados públicos ao cidadão.

4.4 Testes automatizados

Por fim, foram criados testes a fim de averiguar a precisão da identificação realizada pelo *chatbot*. Os testes são feitos na própria plataforma *rasa*, sendo necessário apenas um arquivo ".yml" onde as perguntas testes contidas serão solicitadas durante um período de tempo, com a base treinada anteriormente no *chatbot*. Com isso, é retornado gráficos e arquivos ".yml" com os resultados obtidos, onde pode-se averiguar o nível de precisão nas identificações e associações textuais realizadas pelo mesmo, de acordo com a base preestabelecida.

5 RESULTADOS

Nesta seção, são apresentados os resultados obtidos durante a implantação do projeto. As ferramentas utilizadas são a linguagem de programação *Python*, o *Protége*, o *Apache Jena Fuseki* e os *Frameworks Rasa* e *Scrapy*. O conteúdo de código-fonte do trabalho está disponível em repositórios no *Github*³.

5.1 Ontologia de dados: Metodologia Neon

A metodologia NeOn de Suárez-Figueroa (2010), utilizada para elaboração da ontologia, possui uma menor exigência criteriosa no processo de construção da ontologia. Além disso, a metodologia NeOn permite a elaboração de ontologias associadas à abordagem de desenvolvimento de software, fornecendo detalhes em cada etapa descrita. O desenvolvimento da ontologia da carta de serviços do estado do Ceará utiliza três das cinco etapas básicas da metodologia NeOn, sendo elas: 1- Fase inicial: definição de requisitos, 2- Fase de reuso e 3- Fase de implementação.

5.1.1 Fase inicial: definição de requisitos

Neste seção, foram levantados os requisitos para o desenvolvimento da ontologia da carta de serviços do estado do Ceará descritos no Quadro 1. Logo, define e documenta as propriedades que a ontologia deve atender de acordo com sua respectiva base de dados. Sendo assim, finaliza o que é necessário, dando foco ao andamento do projeto.

5.1.2 Fase de reuso

A fim de definir uma classificação coerente, foram visitados Portais de Dados Abertos de outros países que trabalham com a publicação de dados abertos em catálogo de dados, como: Brasil, Finlândia, Inglaterra e o *data.gov.uk* (Portal de Dados Abertos do Reino Unido). Com base nisso, a ontologia, inicialmente, utiliza padrões já preexistentes de outras ontologias: SKOS e RDFS.

³ <https://github.com/stars/jhonattan31/lists/projeto-tcc>

Quadro 1 – Especificação de Requisitos

1	Propósito Contribuir para divulgação de dados abertos da carta de serviços do estado do Ceará.
2	Escopo A ontologia deve atender à realidade dos serviços públicos do estado do Ceará
3	Linguagem de Implementação A ontologia formal será expressa em linguagem TTL, RDF/XML.
4	Usuários Pretendidos Cidadões cearenses comprometidos com o acesso aos dados abertos.
5	Usos Pretendidos Ser a base para o chatbot criado.
6	Requisitos a. Não-funcionais Descrever os serviços públicos fornecidos pelo estado do Ceará. Fornecer um conjunto de metadados públicos sobre serviços públicos. Possibilitar inferência e relações entre ontologias respectivas. b. Funcionais: Grupo das Questões de Competência Quais serviços funcionam 24 horas? \ Quais os requisitos para solicitar um serviço? Dentre os documentos necessários, quais possuem o CPF em comum? \ Qual a descrição de um serviço? \ Poderia informar, quanto ao tempo de atendimento para um serviço? \ Quais os serviços gratuitos atendem semanalmente?

Fonte: Elaborado pelo autor (2022).

5.1.3 Fase de implementação

A ontologia foi desenvolvida em duas etapas: via *Protegé*, para definição do conjunto *T-Box* e via *python*, com apoio da biblioteca *rdflib* para o conjunto *A-Box*. Com isso, é facilmente possível sua implementação, manutenção e alteração de dados em situações futuras.

5.2 Coleta de dados: planilha de serviços

O conjunto de dados utilizado inicialmente no projeto foi definido com base na carta de serviços do estado do Ceará, visando a geração de uma ontologia com conceitos padronizados que permita a expansão de outras ontologias para outras cartas de serviços. Portanto, foi feito um levantamento dos conceitos mais abrangentes para todos os estados e suas respectivas cartas de serviços. No entanto, mantiveram-se os conceitos que correspondessem ao que foi definido anteriormente, no Quadro 1 de requisitos, como base para o T-BOX da ontologia.

No Quadro 2, está exemplificada uma pequena parte da planilha de comparação, na qual pode-se encontrar seu conteúdo completo no apêndice A. A partir do quadro 2, pode-se inferir que muitos conceitos aparecem em praticamente todos os Estados, por exemplo: *horarioAtendimento*, *dscServico*, *etapasAtendimento* e entre outros. Ao final, percebeu-se que ao menos 50% dos conceitos estavam presentes nas 27 cartas de serviço dos estados e o Distrito Federal.

No Apêndice A pode-se verificar também a classificação 5 estrelas de Berners-Lee (2009) que constitui em avaliar os dados disponíveis na web de 1 a 5 estrelas, que é representada como:

- * : Dados disponibilizados na web sob uma licença aberta. Por exemplo, um arquivo PDF.
- ** : Dados acessados de forma estruturada. Por exemplo, um arquivo Excel.
- *** : Dados que permitem acesso em formatos não-proprietários. Por exemplo, um API.
- **** : Uso de URIs para representar dados para que outras informações possam ser vinculadas. Por exemplo, uma tripla RDF.
- ***** : Link de dados atrelados a outros links de dados gerando inferências de conhecimento com base em suas relações. Por exemplo, um tripla de RDF que possua links para diversas outras triplas RDF.

Dessa forma, em apenas dois sites, sendo eles a carta de Alagoas e a da Bahia, foram encontradas classificações de nível 3 estrelas por possuírem acesso estruturado às informações via API. Todos os outros disponibilizam informações apenas a mostra no site ou via download de arquivo *Portable Document Format*(PDF).

Quadro 2 – Tabela de relacionamento entre cartas

Atributos	Total Relacionadas
codServico	15
nomServico	27
orgao	23
secretaria	15
categoria	18
dscServico	27
poder	15
solicitante	15
prazoEntrega	15
servGratuito	20
requisitos	14
docNeces	20
etapasAtendimento	21
tempoAtendimento	18
tempAtendPriori	14
horarioAtendimento	27
endereco	27
acessoOnAcomp	23
unidPrestadora	15
areaRespon	15
servDigital	15
publicoAlvo	16
acessoPres	14
legislacao	19

Fonte: Elaborado pelo autor (2022).

Os dados estão armazenados em um arquivo CSV, na qual foi preenchido por meio de uma aplicação em *python* que realizou a extração de dados com a técnica de *web scraping* do site oficial da carta de serviços, com apoio do *Framework Scrapy*, o uso de expressões regulares e as bibliotecas *pandas* e *rdflib*.

5.2.1 *Limpeza de dados*

Para facilitar o processamento de dados, foi realizada uma pequena limpeza de dados buscando facilitar a interpretação dos dados pela máquina. Utilizou-se a linguagem *python* e as respectivas bibliotecas: *pandas*, *nltk*, *math* e *re* (para expressões regulares), visando alcançar os seguintes conceitos preestabelecidos anteriormente.

- Transformação do texto em minúscula (Sensibilidade ao caso);
- Remoção de palavras de parada (*Stop Words*);
- Radicalização das palavras (*Lematização*);
- Remoção das palavras curtas (palavras até dois caracteres);
- Remoção de pontuações.

De fato, os cinco conceitos definidos acima são consideravelmente simples, mas de grande importância na interpretação dos dados pela máquina, almejando amplificar o aprendizado da máquina quanto ao idioma. Sendo assim, o conceito de letras maiúsculas e minúsculas perde o sentido já que a máquina quer saber apenas o contexto/conteúdo da mensagem. Da mesma forma, as *stop words* (palavras de até dois caracteres) e pontuações foram retiradas do texto, principalmente de nomenclaturas dos serviços. Por fim, utilizou-se o conceito de lematização para diminuir a quantidade de variações de uma mesma palavra reduzindo-lá ao seu radical base, para facilitar na compreensão das informações pela máquina.

5.3 **Ontologia de dados: Implementação**

5.3.1 *Conjunto T-BOX*

O conjunto T-BOX é desenvolvido com a ajuda da ferramenta *Protége* que é um editor de ontologias gratuito, de código aberto e um *framework* para construção de sistemas inteligentes. O *Protége* é apoiado por uma forte comunidade de usuários acadêmicos, governamentais e corporativos que o utilizam na criação de soluções baseadas em conhecimento de diversos domínios. No Quadro 3, estão descritos os termos associados para o conjunto T-Box da ontologia

proposta no trabalho.

Quadro 3 – Descrição do T-BOX

Atributos	Descrição
nomServico	Nome denominado ao serviço
codServico	Código denominado ao serviço
estado	Nome do estado responsável pelo serviço
orgao	Nome do órgão responsável pela efetuação do serviço
secretaria	Nome do secretária responsável pela efetuação do serviço
categoria	Nome da categoria ao qual o serviço é atribuído.
dscServico	Descrição sobre o que o serviço pretende realizar
poder	Nome do poder responsável pelo serviço
solicitante	Definição de quem pode solicitar o serviço
prazoEntrega	Definição de qual a estimativa de entrega para a solução do serviço
servGratuito	Descrição quanto a gratuidade do serviço
requisitos	Definição dos principais requisitos para solicitar um determinado serviço
docNeces	Descrição dos principais documentos necessários para solicitar o serviço
etapasAtendimento	Definição das etapas para a realização do atendimento do serviço
tempoAtendiment	Definição do tempo mínimo esperado de atendimento para o serviço
tempAtendPriori	Definição do tempo mínimo esperado de atendimento para o serviço, em casos de prioridade
horarioAtendimento	Definição do horário de atendimento estabelecido para o serviço
endereco	Descrição do local para execução do serviço
acessoOnAcomp	Descrição do local para acompanhamento online do serviço
unidPrestadora	Nome da unidade que prestará o serviço
areaRespon	Nome da área responsável pelo serviço
servDigital	Descrição sobre a disponibilidade digital do serviço
publicoAlvo	Descrição do público alvo para o serviço
acessoPres	Descrição da forma de acesso presencial para o serviço
legislacao	Definição da legislação para o determinado serviço

Fonte: Elaborado pelo autor (2022).

5.3.2 Conjunto A-BOX

O conjunto A-Box foi preenchido por meio de uma aplicação em python que realizou a extração de dados com a técnica de *web scraping* do site oficial da carta de serviços do Ceará, com apoio das bibliotecas scrapy, rdflib, pandas e unidecode. A partir do arquivo CSV limpo e seguindo o conceito do T-BOX, foram relacionadas as colunas com seus respectivos campos de dados na ontologia e a transferência de dados para o preenchimento da mesma.

Buscando fortalecer a implantação da ontologia para as cartas de serviços de outros estados, realizou-se o processo de implantação de outra ontologia, por exemplo: a carta de serviços do estado de Alagoas³. A ontologia foi implementada em conjunto com a base de dados da ontologia do estado do Ceará com os mesmos métodos.

Posto isso, integrou-se à ontologia 3666 serviços da carta cearense e 2232 servicios

³ <https://alagoasdigital.al.gov.br/>

da carta alagoana identificados no campo (nomServico) expreso no arquivo de dados. Mas também, a inserção de 25 atributos na base de dados do Ceará e 22 atributos do conjunto alagoano em seus respectivos serviços catalogados na base de dados. Desse modo, ao final dos dados preenchidos foram encontrados 140.754 células de dados expressos no CSV que foram absorvidas na ontologia de dados.

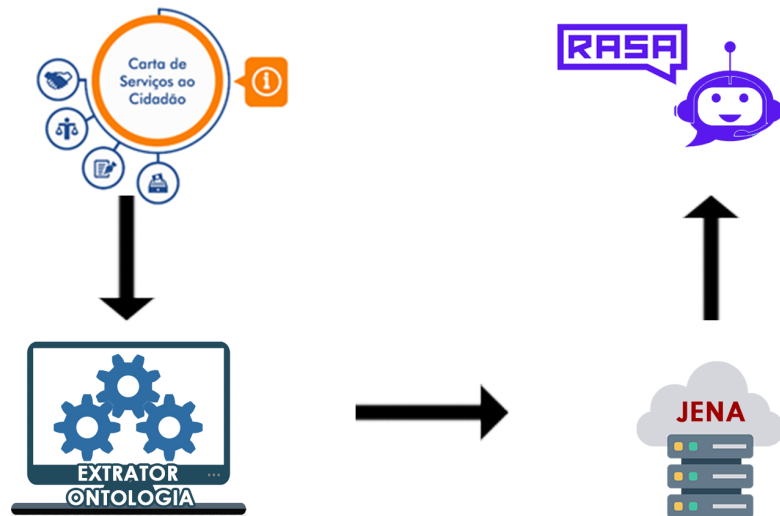
5.3.3 Aplicações de uma ontologia

Finalizada a alimentação do conjunto A-Box, a ontologia está pronta para ser disponibilizada em um servidor. Então, neste projeto utilizou-se o *Apache Jena Fuseki* como um servidor nomeado de *onto_service*. Além disso, trabalhou-se a implantação do arquivo TTL com as informações do composto T-Box e A-Box da ontologia. Implantada a ontologia, a mesma torna-se acessível por meio de uma biblioteca *python*, o *Sparql Wrapper*. Por intermédio dela, possibilita-se uma maneira de realizar consultas *SPARQL* diretamente ao servidor e receber as informações em um arquivo JSON que pode ser facilmente trabalhado como desejado pelo desenvolvedor.

De fato, o conceito de ontologia é uma forma de publicação dos dados prática para o contexto do governo eletrônico, pois possibilita a integração aos sistemas governamentais, quando possível, além da melhoria da classificação cinco estrelas de Berners-Lee (2009) do portal das carta de serviços. Tendo em vista o acesso do cidadão comum, é possível aplicar a ontologia em projetos que auxiliem no acesso da mesma, por exemplo: um *chatbot* ou uma página web para recuperar as informações da carta de serviços contidas na ontologia.

A Figura 4, apresenta o mecanismo para criar ontologias de dados e compartilhá-la em alguma aplicação para acesso simplificado. Inicialmente, define-se os campos de acesso que o extrator de ontologia deve capturar, por meio da técnica de *Web Scraping*, alimentando o conjunto A-Box da ontologia, com base nas definições T-Box estabelecidas no Quadro 3. Sendo assim, é permitido disponibilizar a ontologia de dados em um servidor, como o *Jena Fuseki*, para realizar consultas diretas na plataforma ou por meio de bibliotecas. Por fim, a ontologia pode ser empregada no sistema de integração de ontologias relacionadas, aplicada a um *chatbot* de dados e entre outras aplicações.

Figura 4 – Aplicação da Ontologia



Fonte: Elaborado pelo autor (2022).

5.3.4 Comparação de serviços de cartas diferentes

Com o conteúdo de duas cartas de serviços, situações podem ocorrer na qual os mesmos serviços, mas de estados diferentes, possuem dados distintos. Essa relação de informações é de grande importância, pois amplia o conhecimento da ontologia e, com isso, a inferência de informações trocadas entre as cartas. Na Figura 5 visualiza-se dois serviços equivalentes, mas de estados distintos, com dados divergentes para o campo **docNeces** e algumas variações na nomenclatura do campo **nomServico**, situado na `rdfs:label`.

Figura 5 – Serviços de estados diferentes

```

1 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 PREFIX servi: <http://servicos.gov.ce.br/>
3 SELECT ?uri ?nome ?estado ?doc
4 WHERE {
5   ?uri rdfs:label ?nome.
6   ?uri servi:estado ?estado.
7   ?uri servi:docNeces ?doc.
8   FILTER CONTAINS (?nome, "2 via")
9 }

```

QUERY RESULTS

Table Raw Response

Showing 1 to 2 of 2 entries

Search: Show 50 entries

	uri	nome	estado	doc
1	serv:2_via_carteira_motorista_nacional_habilitacao	"2 via carteira motorista nacional habilitação"	"Alagoas"	"Documento oficial com foto (RG, Carteira de Trabalho, Passaporte, Carteira de Identidade Profissional ou Reservista), CPF"
2	serv:2_via_cnh_carteira_de_motorista	"2 via cnh carteira motorista"	"Ceara"	"01 cpf e 02 documento de identidade rg com cpf"

Showing 1 to 2 of 2 entries

Fonte: Elaborado pelo autor (2022).

5.4 Chatbot para recuperação de dados

De fato, para o cidadão comum torna-se complicado acessar informações em uma ontologia de dados implementada, devido a necessidade de um pré-conhecimento para interpretar o que está na sua frente, além do conhecimento de *SPARQL* para realizar consultas na base de dados. Atentando-se para este ponto, foi proposto a criação de um *chatbot* como um mecanismo simplificado de acesso às informações da carta de serviços para o cidadão.

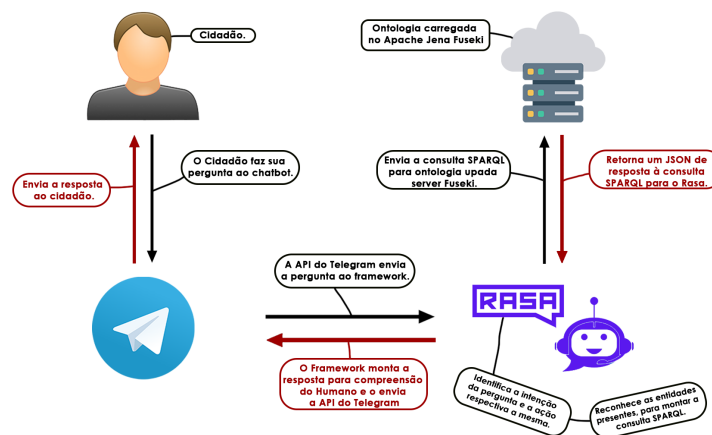
Com apoio do *framework* Rasa, trabalhou-se o desenvolvimento do *chatbot*. Além disso foram definidas 25 *entities* subjacentes aos 25 atributos da ontologia proposta pelo projeto, exemplificado no Quadro 3. Consoante a isso, no Quadro 4, estão descritos as 10 *intentions* e 10 *actions* que foram trabalhadas inicialmente para o treinamento de dados do *chatbot*.

Quadro 4 – Descrição das Intentions e Actions

Intent	Actions	Descrição
hor_func	action_hor_func	Identificar o horário de atendimento para um serviço x.
forma_acesso	action_acesso	Identificar se a forma de acessar um serviço pode ser presencial ou online.
req_serv	action_req_serv	Identificar os requisitos para um determinado serviço
doc_neces	action_doc_neces	Identificar os documentos necessários para solicitar um determinado serviço
serv_gratis	action_serv_gratis	Identificar se um determinado serviço é oferecido gratuitamente
serv_pres_on	action_pres_on	Identificar se um serviço presencial possibilita acompanhamento online
serv_online_desc	action_online_desc	Identificar se um serviço online, seu endereço de acesso e requisitos para solicitar
serv_pres_desc	action_pres_desc	Identificar se um serviço presencial, sua localização e requisitos para solicitar
temp_serv	action_temp_serv	Identificar o tempo de atendimento para um determinado serviço
desc_serv	action_desc_serv	Identificar a descrição para um determinado serviço

Fonte: Elaborado pelo autor (2022).

Figura 6 – Funcionamento da sistema proposto



Fonte: Elaborado pelo autor (2022).

Na Figura 6, está representado o funcionamento do sistema com sua integração completa e o percurso das informações da base de dados até o usuário comum. O usuário envia

sua dúvida ao *chatbot* por meio do Telegram e a API do Telegram envia os dados obtidos na mensagem para o *framework* Rasa. O mesmo realiza o processo de identificação e classificação da mensagem em *intents* e *entities*, além de determinar qual ação será efetuada e a montagem da consulta *SPARQL*, conforme o que foi predefinido no código. Após a construção da consulta, ela é enviada para o servidor *Jena Fuseki*. Caso a consulta seja efetuada com sucesso, o servidor retorna um JSON. Em seguida, o *framework* Rasa recebe o JSON, monta a mensagem para a compreensão do usuário e envia ao Telegram o seu conteúdo com a resposta para o cidadão.

5.4.1 Integração: Chatbot e Ontologia

Após a implantação da ontologia no servidor *Jena Fuseki*, trabalhou-se diretamente na identificação e montagem da consulta *SPARQL* correta a ser consultada na base de dados via *Sparql Wrapper*. Na Figura 7, localiza-se uma captura de tela da consulta *SPARQL* criada pela *intent req_serv* no *chatbot* e o resultado da mesma em um JSON recebido diretamente da API do *Jena Fuseki*, onde encontra-se a ontologia de dados. Já na figura 8, temos a tela do *Jena Fuseki*, onde foi executada a mesma consulta gerada pelo *chatbot* para efetivar que a aplicação está funcional.

Figura 7 – Consulta Sparql automatizada: Console da aplicação

```

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX serv: <http://servicos.gov.ce.br/>
SELECT ?uri ?nome ?req
WHERE {
  ?uri rdfs:label ?nome.
  FILTER CONTAINS (?nome, "núcleo de ação pela paz").
  ?uri serv:requisitos ?req.
}
{
  'head': {'vars': ['uri', 'nome', 'req']},
  'results': {'bindings': [
    {'uri': {'type': 'uri', 'value': 'http://servicos.gov.ce.br/nucleo_de_acao_pela_paz'},
      'nome': {'type': 'literal', 'value': 'núcleo de ação pela paz'},
      'req': {'type': 'literal', 'value': 'ser morador de vicente pinzón ou são miguel efetuar cadastro para participação nas atividades'}}
  ]}
}
^C
(rasa_teste) C:\Users\jh0nn\Documents\Estágio_Codes\rasa_ontologias_exemplo-master>

```

Fonte: Elaborado pelo autor (2022).

Figura 8 – Consulta Sparql automatizada: Tela do Jena Fuseki

The screenshot shows the Jena Fuseki web interface. At the top, there are input fields for the SPARQL endpoint (set to /onto_service/sparql), content type (set to JSON), and content type for the graph (set to Turtle). Below this, the SPARQL query is displayed in a text area:

```

1 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 PREFIX serv: <http://servicos.gov.ce.br/>
3 SELECT ?uri ?nome ?req
4 WHERE {
5   ?uri rdfs:label ?nome.
6   FILTER CONTAINS (?nome, "núcleo de ação pela paz").
7   ?uri serv:requisitos ?req.
8 }

```

Below the query, the 'QUERY RESULTS' section is visible. It has tabs for 'Table' and 'Raw Response'. The 'Table' view shows one result entry:

uri	nome	req
serv:nucleo_de_acao_pela_paz	"núcleo de ação pela paz"	"ser morador de vicente pinzón ou são miguel efetuar cadastro para participação nas atividades"

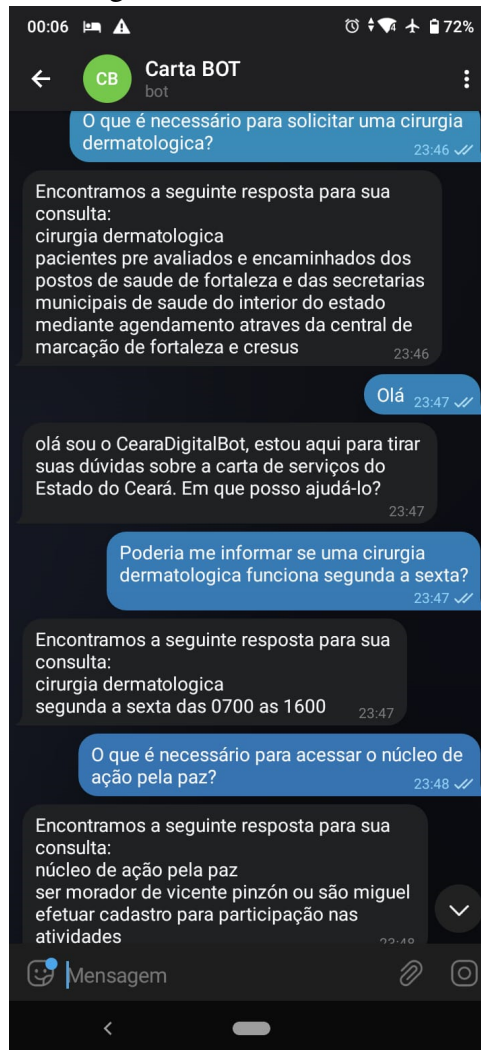
The interface also includes a search bar and a 'Show 50 entries' dropdown.

Fonte: Elaborado pelo autor (2022).

5.4.2 Integração: Chatbot e Telegram

Posteriormente, à conexão entre *bot* e ontologia inicia-se na etapa de disponibilização do *chatbot* online. Para evitar maiores complicações, a aplicação foi disponibilizada online por meio do aplicativo ngrok.exe que gera um túnel na porta do Rasa e disponibiliza o conteúdo online. Após isso, basta acessar o *host* do ngrok e disponibilizá-lo no *chatbot* para que ele estabeleça a conexão.

Figura 9 – Resposta da API do Telegram



Fonte: Elaborado pelo autor (2022).

Na Figura 9, pode-se encontrar um exemplo de como estão sendo recebidas as respostas pelo cidadão no Telegram. Na Imagem acima, encontra-se exemplos de perguntas para as intenções **req_serv** e **hor_func**, e as entidades **nomServico**, **horarioAtendimento** e **requisitos**.

5.5 Testes automatizados

Por fim, foram criados testes a fim de averiguar a precisão da identificação realizada pelo *chatbot*. Os testes são feitos na própria plataforma Rasa, sendo necessário apenas um arquivo ".yml" onde as perguntas testes contidas serão solicitadas durante um período de tempo, com a base treinada anteriormente no *chatbot*. Após isso, são retornadas matrizes de confusão e arquivos ".yml" com os dados de *Accuracy*, *F1-Score* e *Precision*. Com isso, pode-se averiguar o nível de precisão nas intenções e entidades realizadas pelo mesmo, de acordo com a base preestabelecida. Além disso, foram definidas cerca 1400 perguntas para execução dos testes automatizados, com 200 perguntas diferentes para 7 tipos de serviços aleatórios da base de dados. De acordo com os testes realizados conseguimos o seguinte resultado de:

Quadro 5 – Resultado dos testes automatizados

<i>Evaluation Results on CONVERSATION level:</i>			<i>Evaluation Results on Entity level:</i>		<i>Evaluation Results on Intent level:</i>	
rasa.core.test	Correct	469 / 1379	Correct:	null	Correct:	2345 / 2758
rasa.core.test	Accuracy:	0.340	Accuracy:	0.477	Accuracy:	0.849
rasa.core.test	F1-Score:	0.328	F1-Score:	0.460	F1-Score:	0.808
rasa.core.test	Precision:	0.345	Precision:	0.445	Precision:	0.804

Fonte: Elaborado pelo autor (2022).

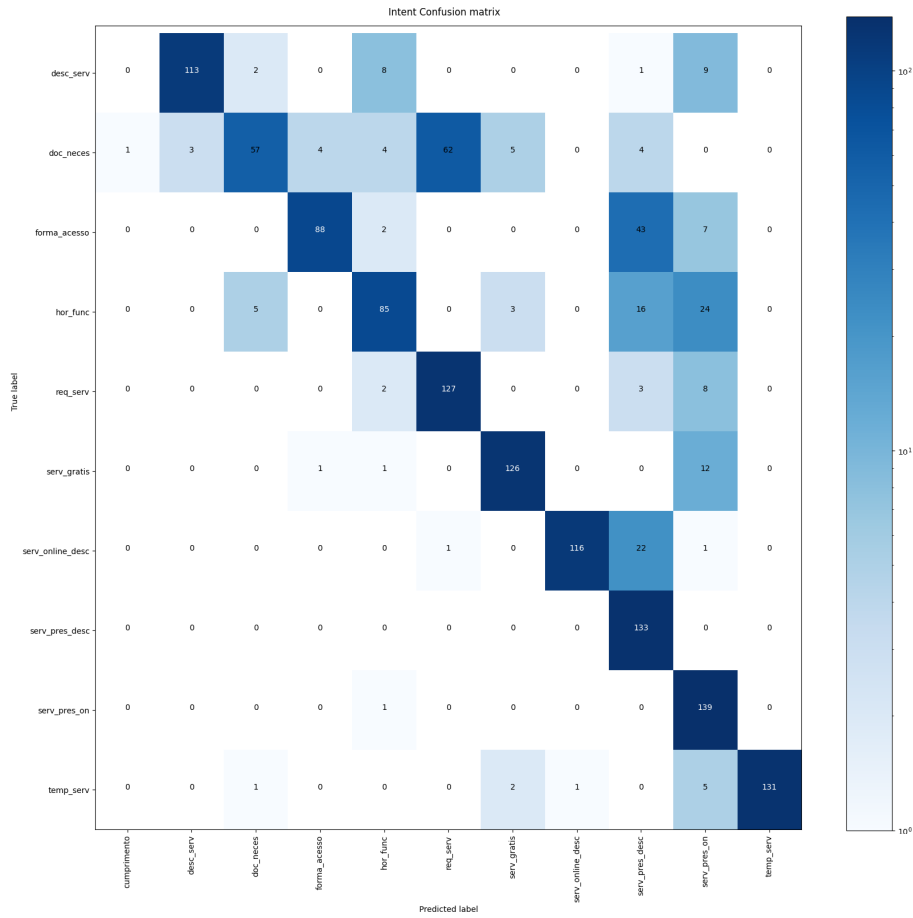
No Quadro 5, conclui-se que nos resultados a nível de conversação geral obteve-se um desempenho de *Accuracy* de 0,340, *f1-score* de 0.460 e *Precision* de 0.445, consistindo em um desempenho relativamente baixo evidenciado por problemas nos extratores de entidades e intenções utilizados. O *DietClassifier* estava configurado para extrair apenas intenções e o *RegexEntity* para extrair entidades. Porém, ambos estavam confundindo extrações de entidades mesmo quando setado para ambos cuidarem de etapas diferentes do processamento de informações, como citado anteriormente.

No Quadro 5, conclui-se que nos resultados a nível de entidade alcançou-se um desempenho de *accuracy* de 0.477, *f1-score* de 0.460 e *Precision* de 0.445, também considerado baixo com base nas mesmas situações do nível de conversação citado acima.

Ao final dos testes, ainda no Quadro 5, conquistou-se resultados a nível de intenção obtendo uma *Accuracy* de 0,849, *f1-score* de 0,808 e *Precision* de 0.804, que podem ser considerados relativamente altos, sendo assim pode-se inferir que o *Diet Classifier* conseguiu associar as intenções de maneira correta, com base no resultado apresentado na figura 10.

Na Figura 10, pode-se averiguar o resultado obtido na matriz de confusão para as intenções com relação aos testes efetuados. Além disso, pode-se notar um bom provei-

Figura 10 – Matriz de Confusão de Intenções



Fonte: Elaborado pelo autor (2022).

tamento na identificação das intenções, especificamente nas intenções de **serv_online_desc**, **serv_pres_desc**, **temp_serv** e **desc_serv**, além dos principais erros de identificação situados nas intenções de **hor_func**, **doc_neces**, **serv_gratis** e **forma_acesso**.

6 CONCLUSÕES E TRABALHOS FUTUROS

A ontologia e *chatbot* desenvolvida pelo projeto atua como uma proposta da publicação e acesso aos dados públicos da Carta de Serviços do Estado do Ceará. A ontologia foi desenvolvida em duas etapas: via *Protegé*, para definição do conjunto *T-Box* e via *python*, com apoio da biblioteca *rdflib* para o conjunto *A-Box*. Já o *Chatbot* foi desenvolvido via *python*, com apoio do *framework* *Rasa*.

O projeto atua na conceituação das definições das cartas de serviços do Brasil, onde implantar ontologias para todos os estados facilita uma padronização e inferências de dados para ambas as partes. Além disso, propor uma maneira que interligue cada estado e seus respectivos serviços entre si e viabilizar ao cidadão diversas maneiras de encontrar o que ele busca.

Inicialmente, definiu-se os requisitos para criar as ontologias de dados relacionando-os com os dados encontrados nas demais cartas dos estados. Além disso, implementou-se duas cartas de serviços visando mostrar a possibilidade de generalização com os atributos definidos para implantar demais cartas quando se fizer necessário. Com a ontologia aplicada no servidor *Jena Fuseki*, abrem-se as possibilidades para criar aplicações que possam consumir da ontologia. O projeto foi desenvolvido com apoio do *framework* *Rasa* integrado à API do Telegram, o que facilita o acesso de maneira mais simplificada para o cidadão quanto as informações das cartas de serviços. Ao final, foi criado um arquivo de testes automatizados com 1400 questões para testar no modelo de treinamento do *chatbot* e executado com o comando de testes do *framework* *Rasa*.

Como trabalhos futuros, pretende-se ampliar o número de intenções e a qualidade da base de treinamento para acentuar curva de aprendizado do *chatbot* sobre a carta de serviços, além de atua-lá por completo. Além disso, propor extração das ontologias de mais estados com base nos requisitos definidos e evoluir o *chatbot* de dados para agregar as novas ontologias criadas. Por fim, aumentar a qualidade de respostas na implementação da API do Telegram e implementar o *chatbot* para a API do Whatsapp, visando ampliar a facilidade de acesso ao *chatbot* para os cidadãos.

7 RECONHECIMENTO

Este trabalho é parcialmente financiado pelo projeto Governo Digital do Estado do Ceará - 04772314/2020 FUNCAP

REFERÊNCIAS

- BARBOSA, S. D. J. **Um agente conversacional pergunta-resposta com recomendações baseadas em uma ontologia de domínio**. Tese (Doutorado) – PUC-Rio, 2020.
- BERNERS-LEE, T. **Linked data**. Reino Unido: Cambridge W3C, 2009. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 01 jul. 2022.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. **ScientificAmerican.com**, [S.I], 05 2001.
- BEYSOLOW, T. I. **Applied Natural Language Processing with Python: Implementing machine learning and deep learning algorithms for natural language processing**. [S.I]: Apress, 2018.
- BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília: Centro Gráfico, 1988.
- CAMPOS, M. L. M.; CAMPOS, M. L. d. A.; CAMPOS, L. M. **Web Semântica e a Gestão de Conteúdos Informacionais**. Salvador/Brasília: Bibliotecas Digitais: Saberes e Práticas, EDUFBA/IBICT, 2005. v. 1.
- CAMPOS, S. L.; FIGUEIREDO, J. Uso de técnicas de processamento de linguagem natural para identificação de similaridade de serviços públicos. In: . WORKSHOP DE COMPUTAÇÃO APLICADA EM GOVERNO ELETRÔNICO, 9. Porto Alegre, 2021. **Anais [...]**: SBC, 2021. p. 83–94. ISSN 2763-8723.
- CONSORTIUM, W. W. W. **Resource Description Framework (RDF)**. United States: [s.n], 2004. Disponível em: <https://www.w3.org/RDF/>. Acesso em: 07 jul. 2022.
- GLEZ-PEÑA, D.; LOURENÇO, A.; LÓPEZ-FERNÁNDEZ, H.; REBOIRO-JATO, M.; FDEZ-RIVEROLA, F. Web scraping technologies in an api world. **Briefings in Bioinformatics**, [S.I], 04 2013.
- GOUDOS, S. K.; LOUTAS, N.; PERISTERAS, V.; TARABANIS, K. Administration domain ontology for a semantic web services egovernment framework. In: IEEE. **IEEE International Conference on Services Computing, (SCC 2007)**. Salt Lake City, UT, USA, 2007. p. 270–277.
- GROUP, S.; AGENCY, M.; ENTERPRISE, D.; AUDIENCE, S. **The state of chatbots report**. Massachusetts: US Census, 2018. Disponível em: <https://www.drift.com/blog/Chatbots-report/>. Acesso em: 26 jan. 2022.
- GRUBER, T. R. **What is an Ontology?** USA: University Stanford, 2002. Disponível em: <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>. Acesso em: 26 jan. 2022.
- GUIZZARDI, G. **Ontological foundations for structural conceptual models**. Tese (Doutorado) – University of Twente, [S.I], oct 2005.
- GÓMEZ-PÉREZ, A.; FERNÁNDEZ-LÓPEZ, M.; CORCHO, O. **Methodologies and Methods for Building Ontologies**. London: Springer London, 2004. 107–197 p.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining, southeast asia edition: Concepts and techniques**. Burnaby: Morgan Kaufmann, 2006. v. 2.

MANOLA, F.; MILLER, E.; MCBRIDE, B. *et al.* Rdf primer. **W3C recommendation**, Citeseer, [S.I.], v. 10, n. 1-107, p. 6, 2004.

MITCHELL, T. M. **Machine Learning**. New York: McGraw-Hill, 1997. v. 1. ed. ISBN 978-0-07-042807-2.

PEREIRA, A.; CAPPELLI, C.; BAIÃO, F.; NUNES, V.; DIIRR, B. Ontotrans: An ontology on transparency. In: . WORKSHOP DE COMPUTAÇÃO APLICADA EM GOVERNO ELETRÔNICO, 8. Porto Alegre, 2020. **Anais [...]**: SBC, 2020. p. 84–95. ISSN 2763-8723.

PEREIRA, G. C.; MONTEIRO, I. T.; VASCONCELOS, D. R.; BRAZ, L.; SILVA, C. H. C. Classificação taxonômica de categorias de serviços públicos para aplicações digitais. In: . WORKSHOP DE COMPUTAÇÃO APLICADA EM GOVERNO ELETRÔNICO, 9. Porto Alegre, 2021. **Anais [...]**: SBC, 2021. p. 119–130. ISSN 2763-8723.

PEREIRA, L.; TODESCO, J. Ogdpub - uma ontologia para publicação de dados abertos governamentais. In: . WORKSHOP DE COMPUTAÇÃO APLICADA EM GOVERNO ELETRÔNICO, 8. Porto Alegre, 2020. **Anais [...]**: SBC, 2020. p. 72–83. ISSN 2763-8723.

PLANEJAMENTO, S. G.; OUVIDORIA, C. G. E.; CONTAS, T. E. C. **Carta de Serviços ao Cidadão**. Ceará: Portal do Governo do Estado do Ceará, 2017. Disponível em: <https://www.seplag.ce.gov.br/gestao/carta-de-servicos-ao-cidadao/>. Acesso em: 03 fev. 2022.

RASA. **Rasa Open Source**. EUA: [s.n], 2020. Disponível em: <https://rasa.com/docs/>. Acesso em: 04 fev. 2022.

SONGINI, M. L. **Put in Plain Language**: The high portable, object-oriented python language moves into enterprise application development. United States: [s.n], 2005. Disponível em: <https://www.computerworld.com/article/2557590/python-software-foundation-s-python-put-in-plain-language.html>. Acesso em: 04 maio. 2022.

SUÁREZ-FIGUEROA, M. C. **Methodology for Building Ontology Networks**: Specification, scheduling and reuse. Tese (Doutorado) – Informática, June 2010. Ontology Engineering Group. Disponível em: <http://oa.upm.es/3879/>. Acesso em: 07 jan. 2022.

TURING, A. M. Computing machinery and intelligence (1950). **Mind**, v. 59, n. 236, p. 433–464, nov. 1950. Disponível em: <https://cir.nii.ac.jp/crid/1570009751286409472>. Acesso em: 01 fev. 2022.

UNIÃO, C.-G. Lei n° 12.527, de 18 de novembro de 2011: Lei de acesso à informação. **Diário Oficial da União**, Brasília, v. 221, p. 1, 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 28 dez. 2021.

UNIÃO, C.-G. Lei n° 13.460, de 26 de junho de 2017: Carta de serviço ao usuário. **Diário Oficial da União**, Brasília, v. 121, p. 4, 2017. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2017/lei/113460.htm. Acesso em: 07 jun. 2022.

USCHOLD, M.; GRUNINGER, M. Ontologies: principles, methods and applications. **The Knowledge Engineering Review**, Cambridge University Press, v. 11, n. 2, p. 93–136, 1996.

USCHOLD, M.; KING, M. **Towards a methodology for building ontologies**. [S.I]: Citeseer, 1995. v. 11.

**APÊNDICE A – PLANILHA DE ANÁLISE DAS CARTAS DE SERVIÇO
BRASILEIRAS**

Encontra-se abaixo, a planilha comparativa das cartas de serviço brasileiras. Nela estão representados os 26 Estados e o Distrito Federal.

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
	Atributos	Total Relacionadas	AC
Data da informação		01/07	15/06
Classificação 5 estrelas	27	*	* * *
Sítio dos Dados Abertos	-		<u>Link</u>
Sítio do Portal	-	<u>Link 1</u> e <u>Link 2</u>	<u>Link</u>
codServico	15		Codigo Servico
nomServico	27	Servico	Servico
orgao	23	Orgao	Órgão
secretaria	15		Secretaria
categoria	18	Categoria	Categoria
dscServico	27	Descrição Servico	O Que É Serviço
poder	15	Poder Executivo	Poder
solicitante	15	Usuario	Solicitante
prazoEntrega	15	Qual Prazo	Tempo Estimado Para Realizar Esse Serviço
servGratuito	20	Pagamento pelo Servico	Custo Deste Serviço
requisitos	14	Condições para Solicitar	Descrição Da Etapa
docNeces	20	Documentação Solicitados	Documentos Necessários
etapasAtendimento	21		Meios De Realização Desta Etapa
tempoAtendimento	18		Tempo De Atendimento
tempAtendPriori	14		
horarioAtendimento	27	Horario Atendimento / Onde posso solicitar	Horário
endereco	27	Endereco / Onde posso solicitar	Endereço Da Unidade De Atendimento
acessoOnAcomp	23		E-Mail
unidPrestadora	15	Unidade Prestadora / Responsavel pela informação	Nome Da Unidade De Atendimento
areaRespon	15	Responsavel pela informação	Area Reponsavel / Nome Da Unidade De Atendimento
servDigital	15	Endereço eletrônico	
publicoAlvo	16		Público Alvo
acessoPres	14		Forma Acesso
legislacao	19		Legislação
Total por Estado		15	22

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
Atributos	Total Relacionadas	AP	AM
Data da informação		03/07	03/07
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-		
Sítio do Portal	-	<u>Link</u>	<u>Link 1 e Link 2</u>
codServico	15		
nomServico	27	Servico	Servico
orgao	23	Órgão	Orgao responsavel
secretaria	15		
categoria	18	Categoria	Categoria
dscServico	27	O Que Este Servico	O que é
poder	15	Poder	Poder
solicitante	15		Usuarios
prazoEntrega	15	Prazo Maximo Atendimento	
servGratuito	20	Custo Deste Serviço	Valor cobrado
requisitos	14		
docNeces	20	Documentação Necessaria	Documentos necessarios
etapasAtendimento	21	Etapas De Atendimento	Etapas para ser atendido
tempoAtendimento	18	Tempo De Espera	
tempAtendPriori	14		Tempo Prioridade /SuperProridade
horarioAtendimento	27	Horario Atendimento	Horarios de atendimento
endereço	27	Endereço / Canais De Acesso	Local de acesso
acessoOnAcomp	23	Email / Canais De Acesso	Email / Telefone
unidPrestadora	15		
areaRespon	15		
servDigital	15	Site	Site
publicoAlvo	16	Quem Pode Utilizar	Quem pode utilizar
acessoPres	14	Como Acessar	Acesso presencial
legislacao	19	Legislações	
Total por Estado		17	16

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
	Atributos	Total Relacionadas	BA
Data da informação		01/06	
Classificação 5 estrelas	27	* * *	*
Sítio dos Dados Abertos	-	<u>Link</u>	
Sítio do Portal	-	<u>Link</u>	<u>Link</u>
codServico	15	Id	Codigo Servico
nomServico	27	Servico	Servico
orgao	23	Órgão	Orgao Responsavel
secretaria	15	Nome_Secretaria	Secretaria
categoria	18	Categoria	Categoria
dscServico	27	Descricao	Servico
poder	15	Poder	Poder
solicitante	15		Solicitante
prazoEntrega	15	Prazo_Maximo_Estimado	Prazo Entrega
servGratuito	20		Servico Gratuito
requisitos	14	Procedimentos	Requisitos
docNeces	20		Documentos Necessarios
etapasAtendimento	21	Etapas	Etapas Atendimento
tempoAtendimento	18	Previsao_Espera	Tempo Atendimento
tempAtendPriori	14	Prioridades_Atendimento	Tempo Atendimento Prioridade
horarioAtendimento	27	Horario_Atendimento	Horario Atendimento
endereco	27	Endereco / Meios_Contato	Endereco
acessoOnAcomp	23	Email / Meios_Contato	Acesso Online
unidPrestadora	15		Unidade Prestadora
areaRespon	15		Area Responsavel
servDigital	15		Servico Digital
publicoAlvo	16	Publico_Alvo	Publico Alvo
acessoPres	14	Mecanismos_Acessos	Acesso Presencial
legislacao	19		Legislacao
Total por Estado		17	24

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
Atributos	Total Relacionadas	DF	ES
Data da informação		15/06	03/07
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-	Link	
Sítio do Portal	-	Link	Link
codServico	15	Id Pedido	Id Servico
nomServico	27	Servico	Servico
orgao	23	Orgão	
secretaria	15		
categoria	18	DescSubCategoriaAssunto	
dscServico	27	DescStatusPedido	O Que É
poder	15		Poder
solicitante	15	IdOrigemSolicitacao - DescOrigemSolicitacao	
prazoEntrega	15	DatPrazo	
servGratuito	20		Pagamentos E Taxas
requisitos	14		
docNeces	20		Etapas de Atendimento
etapasAtendimento	21	DatResposta	Etapas de Atendimento
tempoAtendimento	18	TempoAtendimento / HorFuncionamento	Tempo Para Ser Atendido
tempAtendPriori	14		
horarioAtendimento	27	HorFuncionamento	Horario De Atendimento
endereco	27	LocalAcesso	Endereço / Local
acessoOnAcomp	23	ContatoEmail	Email / Telefone Recepção
unidPrestadora	15		
areaRespon	15		
servDigital	15		Ouvidoria Site
publicoAlvo	16		
acessoPres	14	FormaAcesso	Formas de Acesso
legislacao	19		
Total por Estado		13	13

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
Atributos	Total Relacionadas	GO	MA
Data da informação		15/06	15/06
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-		
Sítio do Portal	-	<u>Link</u>	<u>Link</u>
codServico	15	IdServico	NumServico
nomServico	27	Servico	Servico
orgao	23	Órgão	
secretaria	15		Secretaria
categoria	18		
dscServico	27	DescServico	DescServico / InformaçõesPessoais
poder	15	Poder	
solicitante	15		
prazoEntrega	15		Prazo para o serviço
servGratuito	20	CustoTotal	CustoServico
requisitos	14		
docNeces	20	DocNeces	DocNeces
etapasAtendimento	21	EtapasAtendimento	EtapasServico
tempoAtendimento	18	TempoAtendimento	TempoAtendimento
tempAtendPriori	14	AtendPrioridade	Prioridade
horarioAtendimento	27	TempoAtendimento	PrazoAtendimento
endereco	27	AtendimentoPresencial	AcompanhamentoPresencial
acessoOnAcomp	23	AtendimentoOnline	AcompanhamentoOnline
unidPrestadora	15	Unidade Responsavel / AreaResponsavel	Unidade Responsavel / AreaResponsavel
areaRespon	15	AreaResponsavel	AreaResponsavel
servDigital	15	CanaisPrestacaoWEB / AgendamentoEletronico	AcessoInternet / AgendamentoEletronico
publicoAlvo	16	PúblicoAlvo	PúblicoAlvo
acessoPres	14	FormasAcesso	
legislacao	19		Legislacao
Total por Estado		18	18

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.			
	Atributos	Total Relacionadas	MT	MS
Data da informação			03/07	03/07
Classificação 5 estrelas	27		*	*
Sítio dos Dados Abertos	-			
Sítio do Portal	-		Link 1 e Link 2	Link
codServico	15			
nomServico	27		Serviço	Servico
orgao	23		Órgão	Orgao
secretaria	15			Secretaria
categoria	18			Categoria do Serviço
dscServico	27		O que é	Descricao
poder	15		Poder	
solicitante	15			Quem Pode Solicitar
prazoEntrega	15		Prazo do serviço	Prazo Maximo
servGratuito	20		Servico Gratuito	
requisitos	14			Caracteristicas
docNeces	20			Documentos
etapasAtendimento	21			Etapas de Processamento
tempoAtendimento	18		Horario de Atendimento	PrazoAtendimento
tempAtendPriori	14		Prioridade de Atendimento	Prioridade
horarioAtendimento	27		Horario de Atendimento	PrazoAtendimento
endereço	27		Localização	Endereço
acessoOnAcomp	23		Email / Telefone	Email / Telefone
unidPrestadora	15			EquipeResponsavel
areaRespon	15			Area Responsavel / EquipeResponsavel
servDigital	15			
publicoAlvo	16		Usuarios	
acessoPres	14			
legislacao	19		Legislação especifica	Legislação
Total por Estado			13	18

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
Atributos	Total Relacionadas	MG	PA
Data da informação		14/06	03/07
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-		
Sítio do Portal	-	<u>Link 1</u> e <u>Link 2</u>	<u>Link</u>
codServico	15	IdServ	Codigo
nomServico	27	Servico	Servico
orgao	23	OrgaoVinculado	Orgao
secretaria	15	Secretaria	Secretaria
categoria	18	Tipo	Categoria do Serviço
dscServico	27	DescServ	Descricao
poder	15	Poder	
solicitante	15	Categoria	Quem Pode Solicitar
prazoEntrega	15		
servGratuito	20		Taxa
requisitos	14		Pré-requisito
docNeces	20		Documentos
etapasAtendimento	21	Etapas	Etapas do serviço
tempoAtendimento	18		
tempAtendPriori	14		
horarioAtendimento	27	PrazoAtendimento	PrazoAtendimento
endereco	27	Formas de Acesso	Local para realizar
acessoOnAcomp	23	Formas de Acesso	Email / Telefone
unidPrestadora	15	UnidadeResponsável	EquipeResponsavel
areaRespon	15	Area Responsavel / UnidadeResponsável	Area Responsavel / EquipeResponsavel
servDigital	15		
publicoAlvo	16		
acessoPres	14	Formas de Acesso	Forma Acesso / Local para realizar
legislacao	19	Legislação	Base Legal
Total por Estado		16	18

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
Atributos	Total Relacionadas	PB	PR
Data da informação		03/07	03/07
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-		
Sítio do Portal	-	<u>Link</u>	<u>Link</u>
codServico	15		
nomServico	27	Servico	Nome Do Servico
orgao	23	Orgao responsavel	
secretaria	15		Secretaria
categoria	18	Categoria	Categoria
dscServico	27	O que é	Desc Servico
poder	15	Poder	Poder Executivo
solicitante	15	Usuarios	Usuario
prazoEntrega	15		
servGratuito	20	Valor cobrado	Custo Deste Serviço
requisitos	14		
docNeces	20	Documentos necessarios	Documentação
etapasAtendimento	21	Etapas para ser atendido	
tempoAtendimento	18		
tempAtendPriori	14	Tempo Prioridade	
horarioAtendimento	27	Horarios de atendimento	Horario Atendimento
endereco	27	Local de acesso	Endereco
acessoOnAcomp	23	Email / Telefone	Email
unidPrestadora	15		Unidade Responsavel / AreaResponsavel
areaRespon	15		Area Responsavel
servDigital	15	Site	Site Oficial
publicoAlvo	16	Quem pode utilizar	
acessoPres	14	Forma Acesso / Local de acesso	Endereco
legislacao	19		Legislações
Total por Estado		16	16

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
Atributos	Total Relacionadas	PE	PI
Data da informação		03/07	03/07
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-		
Sítio do Portal	-	<u>Link</u>	<u>Link</u>
codServico	15	Id	
nomServico	27	Servico	Nome Do Servico
orgao	23	Órgão	
secretaria	15	Id Secretaria / Nome Secretaria	Secretaria
categoria	18	Categoria	Categoria
dscServico	27	Descricao	O Que Este Servico
poder	15	Poder	Poder
solicitante	15		
prazoEntrega	15	Prazo Maximo Estimado	Prazo Maximo Atendimento
servGratuito	20		Custo Deste Serviço
requisitos	14	Documentos / Procedimentos	Documentação Necessaria
docNeces	20	Etapas / Procedimentos	Etapas De Atendimento
etapasAtendimento	21	Etapas	
tempoAtendimento	18	Previsao Espera	
tempAtendPriori	14	Prioridades Atendimento	Horario Atendimento
horarioAtendimento	27	Previsao Espera	Endereco / Canais De Acesso
endereco	27	Localizacao	Email / Canais De Acesso
acessoOnAcomp	23	Contato	
unidPrestadora	15		
areaRespon	15		
servDigital	15		Site
publicoAlvo	16	Publico Alvo	Quem Pode Utilizar
acessoPres	14	Localizacao	Canais De Acesso
legislacao	19	Legislacoes	Legislações
Total por Estado		19	16

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
Atributos	Total Relacionadas	RJ	RN
Data da informação		14/06	03/07
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-		
Sítio do Portal	-	<u>Link</u>	<u>Link</u>
codServico	15	IdServico	IdServ
nomServico	27	NomServico	Servico
orgao	23	Órgão	OrgaoVinculado
secretaria	15	Secretaria	Secretaria
categoria	18		Tipo / Categoria
dscServico	27	InformaçõesNecessarias	DescServ
poder	15		
solicitante	15		Usuario
prazoEntrega	15		
servGratuito	20	CustoServico	CustoServico
requisitos	14	DocNeces	Requisitos
docNeces	20	EtapasAtendimento	
etapasAtendimento	21	PrazoMaximo	Etapas
tempoAtendimento	18	TempoAtendimento	
tempAtendPriori	14	FormaAcesso	PrazoAtendimento
horarioAtendimento	27	LocaisAcesso	PrazoAtendimento
endereco	27	LocaisAcesso	Formas de Acesso
acessoOnAcomp	23		
unidPrestadora	15		Unidade Responsavel
areaRespon	15		UnidadeResponsavel / AresResponsavel
servDigital	15		
publicoAlvo	16	PúblicoAlvo	
acessoPres	14		
legislacao	19	Legislação	Legislação
Total por Estado		15	16

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
Atributos	Total Relacionadas	RS	RO
Data da informação		03/07	01/07
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-		
Sítio do Portal	-	Link	Link 1 e Link 2
codServico	15		
nomServico	27	Nome Do Serviço	Servico
orgao	23	Órgão	Órgão
secretaria	15		
categoria	18	Categoria	Categoria
dscServico	27	Descserv	O que é
poder	15	Poder	
solicitante	15	Solicitante	Solicitante
prazoEntrega	15	Tempo Estimado Para Realizar Esse Serviço	
servGratuito	20	Custo Deste Serviço	Pagamentos e Taxas
requisitos	14	Descrição Da Etapa	Requisitos necessarios
docNeces	20	Documentos Necessários	
etapasAtendimento	21	Meios De Realização Desta Etapa	
tempoAtendimento	18		Prazo Máximo para ser atendido
tempAtendPriori	14		Atendimento Prioridade
horarioAtendimento	27	Horário	Horario de atendimento
endereco	27	Endereço Da Unidade De Atendimento	Endereço / Local
acessoOnAcomp	23	E-Mail	Email / Telefone Recepção
unidPrestadora	15	Nome Da Unidade De Atendimento	
areaRespon	15	Area Responsavel / Nome Da Unidade De Atendimento	
servDigital	15		Ouvidoria site
publicoAlvo	16	Público Alvo	
acessoPres	14		
legislacao	19	Legislação	
Total por Estado		18	13

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.		
	Atributos	Total Relacionadas	RR
Data da informação		04/07	14/06
Classificação 5 estrelas	27	*	*
Sítio dos Dados Abertos	-		
Sítio do Portal	-	<u>Link 1</u> e <u>Link 2</u>	<u>Link</u>
codServico	15	NumServico	Codigo Servico
nomServico	27	NomServico	Nome Servico
orgao	23	Orgao	Órgão
secretaria	15		
categoria	18		Categoria / Tema
dscServico	27	DescServico	Características
poder	15		
solicitante	15	Solicitante	Solicitante
prazoEntrega	15		Tempo De Solicitação
servGratuito	20	CustoServico	
requisitos	14		Finalidade
docNeces	20	DocNeces	Requisitos Exigidos Do Usuário
etapasAtendimento	21	EtapasServico	Etapas Do Processo
tempoAtendimento	18	TempoAtendimento	Tempo de Execução
tempAtendPriori	14		
horarioAtendimento	27	PrazoAtendimento	Horario Atendimento
endereco	27	AcompanhamentoPresencial	Email / Canal De Atendimento
acessoOnAcomp	23	AcompanhamentoOnline	Contato / Canal De Atendimento
unidPrestadora	15	Unidade Responsavel / AreaResponsavel	Unidades Prestadoras
areaRespon	15	AreaResponsavel	Area Responsavel / Unidades Prestadoras
servDigital	15	AgendamentoEletronico	Url Do Site Serviço Digital
publicoAlvo	16	PúblicoAlvo	
acessoPres	14		
legislacao	19	Legislacao	Legislação
Total por Estado		17	18

LEGENDA:	Em caso de informações de campos diferentes acopladas em um único campo, foi inserido (/) e repetido o nome do respectivo campo.			
	Atributos	Total Relacionadas	SP	SE
Data da informação		14/06	03/07	03/07
Classificação 5 estrelas	27	*	*	*
Sítio dos Dados Abertos	-	Link		
Sítio do Portal	-	Link 1 e Link 2	Link	Link 1 e Link 2
codServico	15	Código do Serviço		
nomServico	27	Nome do Serviço	Serviço	Servico
orgao	23	Órgão	Órgão	Órgão
secretaria	15	Secretaria	Secretaria	Secretaria
categoria	18			
dscServico	27	O que é	O que é	DescServico
poder	15			
solicitante	15			
prazoEntrega	15	Prazo Máximo	Prazo do serviço	
servGratuito	20		Servico Gratuito	CustoServico
requisitos	14	Requisitos		
docNeces	20	Principais Etapas		DocNeces
etapasAtendimento	21	Documentos		EtapasServico
tempoAtendimento	18	Tempo Atendimento / Horário de Atendimento	Horario de Atendimento	TempoAtendimento
tempAtendPriori	14		Prioridade de Atendimento	
horarioAtendimento	27	Atendimento	Horario de Atendimento	PrazoAtendimento
endereco	27	Endereço (prest)	Localização	AcompanhamentoPresencial
acessoOnAcomp	23	E-mail (info) / Telefone (info)	Email / Telefone	AcompanhamentoOnline
unidPrestadora	15	Unidade de Gestão		Unidade Responsavel / AreaResponsavel
areaRespon	15	Area Responsavel / Unidade de Gestão		AreaResponsavel
servDigital	15	Home Page Informações		AcessoInternet
publicoAlvo	16		Usuarios	PúblicoAlvo
acessoPres	14			
legislacao	19	Legislação	Legislação especifica	Legislacao
Total por Estado		17	13	16