

# Relatório sobre a Base de Dados DomíniosGov

Beatriz de A. Campos<sup>1</sup>, Luiz H. Reis<sup>2</sup>

Departamento de Ciencia da Computação – Instituto de Matemática – Universidade Federal do Rio de Janeiro (UFRJ) – Rio de Janeiro – RJ – Brasil

**Abstract.** *This article aims to explain the operation and development of DomíniosGov database, created for the Banco de Dados I discipline the course of Computer Science (UFRJ) taught by Professor Giseli R. Lopes .*

**Resumo.** *Este artigo tem por objetivo explicar o funcionamento e elaboração da base de dados DomíniosGov, criada para a disciplina de Banco de Dados I do curso de Ciência da Computação (UFRJ) ministrado pela professora Giseli R. Lopes.*

## 1. Agrupamento dos Dados

Inicialmente, foi escolhido um conjunto de dados presente no site <http://dados.gov.br>, que disponibiliza dados sobre diferentes instituições subordinadas ao Governo Federal Brasileiro.

Após uma análise dos *datasets* disponíveis, foi escolhida a base 'Domínios Gov.br' que contém dados sobre os domínios (*sites*) de cunho governamental registrados no Registro.br e que estão autorizados pelo Ministério do Planejamento, Desenvolvimento e Gestão.

Os tipos de dados atualmente disponíveis no site para análise são:

- Domínio: corresponde ao endereço do site;
- Documento: CNPJ da empresa responsável pelo site;
- Nome: nome da empresa responsável;
- UF, Cidade e Cep: informações relativas a sede da empresa responsável;
- Data de Cadastro: referente a data de criação do domínio;
- Última Atualização: referente a data da última atualização do domínio;
- Ticket: Número do ticket vinculado ao registro da URL no cadastro do Registro.br.

Pra complementar os dados adquiridos no site Dados.gov e com o intuito de enriquecer as informações para a futura criação da base, foram coletados outros dados sobre os domínios disponibilizados no site utilizando o programa 'WhoIs'.

Os tipos de dados coletados utilizando o WhoIs são:

- nserver: DNS autoritativo do domínio;
- nslastaa: última atualização do DNS do domínio;
- person: pessoa responsável pelo domínio;
- nic-hdl-br: sigla referente a pessoa responsável pelo domínio;
- created: data de cadastro do responsável pelo domínio;
- changed: data de alteração do cadastro da pessoa responsável pelo domínio;
- tipo: tipo de função que uma 'person' pode exercer para determina do domínio (podem ser owner-c, admin-c, tech-c ou billing-c).

Afim de completar os requisitos do trabalho, também foi inserido um atributo 'telefone' na tabela 'Ents\_Resps', que representaria o telefone de comunicação com as empresas. A tabela de telefones contém dados fictícios e foi gerada de forma automática pela aplicação GenerateData (<http://www.generatedata.com/>).

## 2. Modelagem da Base

### 2.1. Projeto Conceitual

A modelagem conceitual da base foi gerada na forma Entidade-Relacionamento (ER) contendo 6 entidades e 6 relações.

A figura abaixo demonstra o modelo conceitual da base criado no BrModelo.

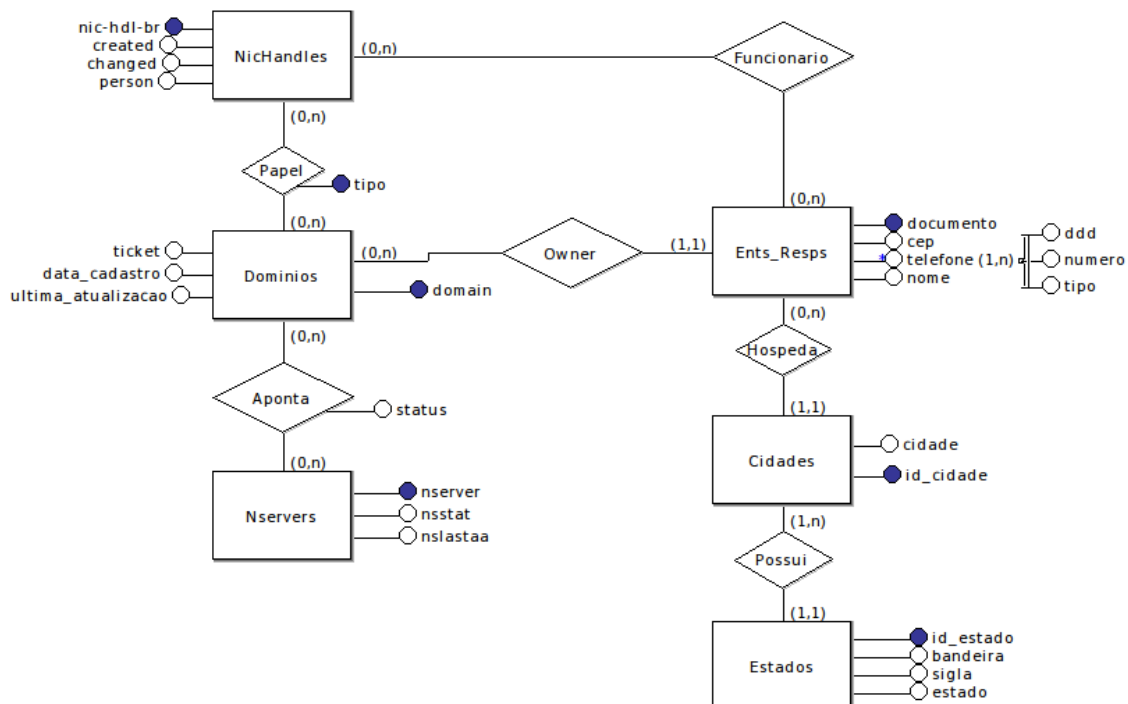


Imagem 1: Modelo Conceitual

## 2.2. Projeto Lógico

A modelagem lógica da base foi gerada utilizando o modelo Relacional. A figura abaixo demonstra o modelo lógico da base gerado pelo BrModelo, a partir do modelo ER.

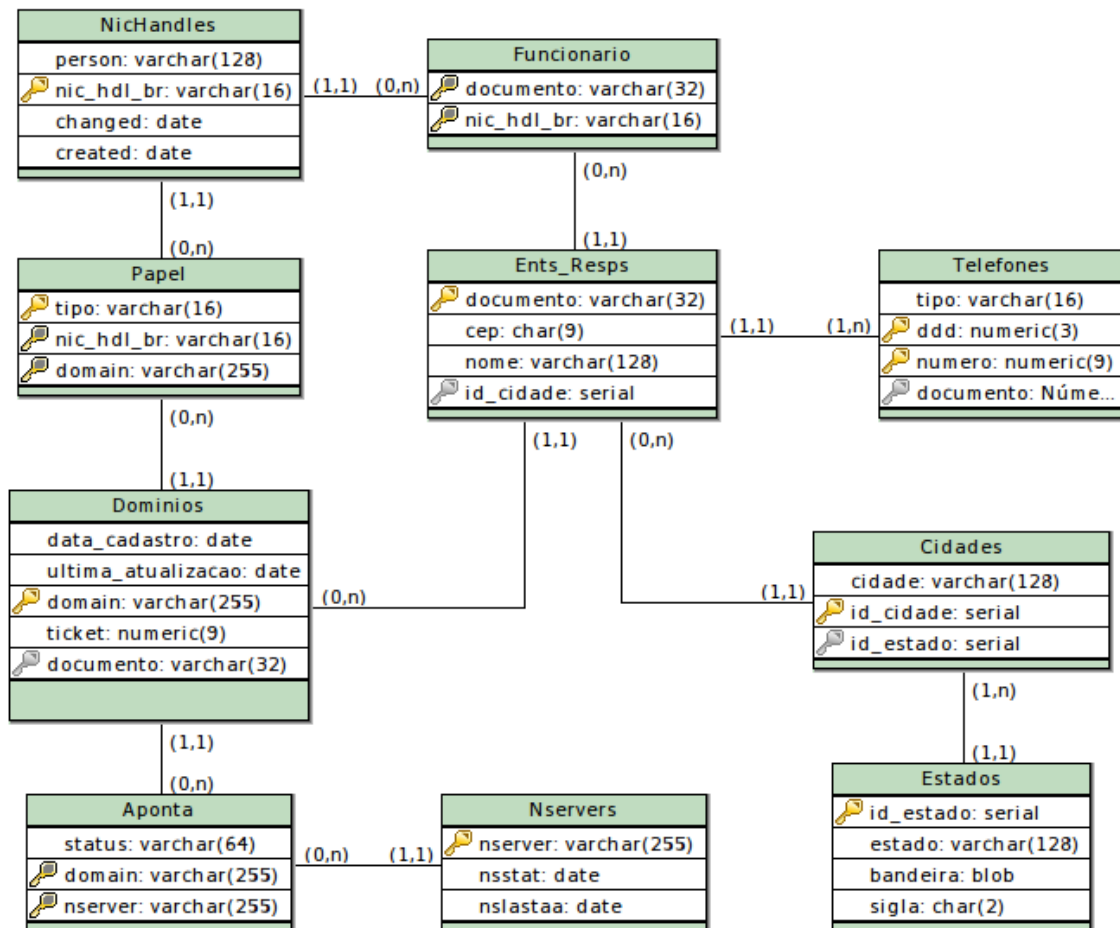


Imagem 2: Modelo Lógico

## 2.3. Projeto Físico

O modelo físico da base foi criado utilizando a linguagem SQL e o software MySQL / MariaDB.

O *script* de criação da base de dados, gerado a partir da modelagem lógica, está em anexo ao relatório, bem como os *scripts* gerados para povoar o banco de dados.

### 3. Aplicação WEB

Para a demonstração do funcionamento da base de dados, foi criada uma aplicação web disponível no endereço ['http://ec2-52-32-171-79.us-west-2.compute.amazonaws.com:8080/'](http://ec2-52-32-171-79.us-west-2.compute.amazonaws.com:8080/).

Em anexo encontra-se o *script* contendo todas as *queries* abaixo.

#### 3.1. Consultas disponíveis

As seguintes consultas estão disponíveis:

- **Consulta 1 (Simples):** Selecionar o nome de todos os domínios cujos valores de ticket são maiores que 20000.

```
SELECT domain as dominios
FROM Dominios
WHERE ticket > 20000;
```

- **Consulta 2 (Junção 2 tabelas):** Ordenar, de forma crescente, as empresas que possuem o maior número de funcionarios como administradores de domínios.

```
SELECT e.nome as Empresa, COUNT(DISTINCT f.nic_hdl_br) as Num_Func
FROM Ents_Resps e INNER JOIN Funcionario f ON e.documento=f.documento
GROUP BY nome
ORDER BY Num_Func;
```

- **Consulta 3 (Junção 2 tabelas com junção externa):** Selecionar as datas de criação e alteração dos domínios, seus endereços, o documento da empresa responsável e o papel exercido pelos funcionários, bem como suas siglas.

```
SELECT * FROM Dominios LEFT OUTER JOIN Papel ON Papel.domain=Dominios.domain
UNION ALL
SELECT * FROM Dominios RIGHT OUTER JOIN Papel ON Papel.domain=Dominios.domain
WHERE Dominios.domain=Papel.domain;
```

- **Consulta 4 (Junção 3 tabelas):** Selecionar o nome dos funcionários e a empresa na qual ele possui ligação.

```
SELECT DISTINCT p.person, aux.nome
FROM NicHandles p INNER JOIN (SELECT f.nic_hdl_br, e.nome FROM Funcionario f INNER JOIN
Ents_Resps e ON e.documento=f.documento) as aux ON p.nic_hdl_br=aux.nic_hdl_br;
```

- **Consulta 5 (Junção 3 tabelas):** Seleccionar os nomes e os tickets dos domínios e o nome de seus respectivos nameservers.

```
SELECT ticket, dominio, nameserver
FROM Dominios INNER JOIN
(SELECT ap.domain as dominio, ns.nserver nameserver FROM Aponta ap INNER JOIN Nservers ns
ON ns.nserver=ap.nserver) as nomes
ON nomes.dominio=Dominios.domain;
```

- **Consulta 6 (Conjuntos):** Seleccionar o número de documento das empresas responsáveis por domínios criados durante os anos 2014 e 2015.

```
(SELECT documento FROM Dominios WHERE data_cadastro like '2015%')
UNION
(SELECT documento FROM Dominios WHERE data_cadastro like '2014%');
```

- **Consulta 7 (Agregação):** Seleccionar a quantidade de domínios sob responsabilidade de uma mesma empresa.

```
SELECT e.nome, COUNT(DISTINCT d.domain) as qnt_dominios
FROM Ents_Resps e INNER JOIN Dominios d ON e.documento=d.documento
GROUP BY nome;
```

- **Consulta 8 (Agregação):** Seleccionar o nome das pessoas (person) que são responsáveis por mais de 2 domínios diferentes, em qualquer tipo de papel.

```
SELECT p.person as nome, count(d.domain) as qnt_dominios
FROM NicHandles p INNER JOIN Papel d ON p.nic_hdl_br=d.nic_hdl_br
GROUP BY p.person
HAVING count(d.domain) > 2;
```

- **Consulta 9 (Consulta Aninhada):** Seleccionar o nome e o documento das empresas que estão sediadas no Rio de Janeiro.

```
SELECT nome, documento
FROM Ents_Resps
WHERE id_cidade in (SELECT id_cidade FROM Cidades WHERE id_estado in (SELECT id_estado
FROM Estados WHERE sigla like 'RJ'));
```

- **Consulta 10 (Consulta Aninhada):** Selecionar os domínios cujos nameservers foram atualizados pela ultima vez entre 2010 e 2015.

```
SELECT domain
FROM Aponta
WHERE nserver in (SELECT nserver FROM Nservers WHERE nslastaa BETWEEN '2009-12-31' AND '2016-01-01');
```

- **Consulta 11 (Modelo Relatório):** Média de domínios por empresa e Total de domínios existentes.

```
SELECT avg(T.qnt_dominios) as media_empresa, sum(T.qnt_dominios) as total_dominios
FROM ( SELECT e.nome, COUNT(DISTINCT d.domain) as qnt_dominios
FROM Ents_Resps e INNER JOIN Dominios d ON e.documento=d.documento
GROUP BY nome) as T;
```

- **Consulta 12 (Modelo Relatório):** Estados ordenados pela quantidade de empresas responsáveis por domínios.

```
SELECT T.estado as estados, COUNT(e.nome) as qnt_empresas, T.bandeira
FROM Ents_Resps e INNER JOIN (SELECT id_cidade, estado, bandeira FROM Cidades INNER JOIN Estados ON Estados.id_estado=Cidades.id_estado) as T ON T.id_cidade=e.id_cidade
GROUP BY estados
ORDER BY qnt_empresas DESC;
```

#### 4. Ferramentas Utilizadas

As ferramentas utilizadas para a criação da base e da aplicação web foram:

1. MySql – para a criação da base;
2. Pearl – para a criação dos scripts de varredura;
3. BrModelo – para a elaboração do modelo conceitual e auxílio na criação do modelo lógico.

## **5. Divisão do Trabalho**

Os dois membros do grupo participaram conjuntamente da elaboração de todas as partes do trabalho, porém com diferentes ênfases, como listado abaixo:

### **Beatriz de Andrade:**

- Criação das consultas em sql.
- Criação do relatório final.

### **Luiz Reis:**

- Criação dos scripts em Pearl.
- Criação do servidor para hospedagem da base.

## **6. Referências**

HEUSER, C.A., (2009) “Projeto de Banco de Dados”, Bookman, 6ª edição.