

# Dados Abertos - Caso de uso com a Secretaria de Saúde de Curitiba

Wagner Luís Pereira<sup>1</sup>, Nádia P. Kozievitch<sup>1</sup>, Rita Berardi<sup>1</sup>, Mariana G. Luz<sup>1</sup>

<sup>1</sup>Universidade Tecnológica Federal do Paraná (UTFPR)

Avenida Sete de Setembro, 3165

Departamento Acadêmico de Informática – DAINF – Curitiba – Brasil

wagnerluis1989@gmail.com, {nadiap, ritaberardi}@utfpr.edu.br,

marianagomesluz@alunos.utfpr.edu.br

**Abstract.** *Management reports are documents designed to present an overview of performance, perspectives, connecting and summarizing various data sources and indicators. Some reports of this type are generated by the municipal health department of Curitiba and, despite having some indicators in common, they do not present a sufficient standardization to allow a comparison between the data over time. This work maps the sources of the data contained in these reports, observing the health indicators. Entity Matching algorithms, which are able to locate similar terms, are present in the work to identify indicators that are maintained throughout the time.*

**Resumo.** *Relatórios de gestão são documentos elaborados para apresentar uma visão geral de desempenho, perspectivas, conectando e resumindo várias fontes de dados e indicadores. Alguns relatórios deste tipo são gerados pela secretaria municipal de saúde de Curitiba e que apesar de possuírem alguns indicadores em comum, não apresentam uma padronização suficiente que permita uma comparação entre os dados ao longo do tempo. Este trabalho mapeia as fontes dos dados contidas nestes relatórios, observando os indicadores de saúde. Algoritmos de Entity Matching, que são capazes de localizar termos semelhantes, foram utilizados no trabalho para identificar indicadores que se mantêm ao longo do tempo.*

## 1. Introdução

A disponibilização de dados abertos é cada vez mais comum no contexto global. Ao verificar o ranking de dados abertos nos países<sup>1</sup>, pode-se observar que existe uma tendência de que os países desenvolvidos tenham melhor pontuação conforme os critérios adotados pela pesquisa do referido ranking. Apesar disto, é possível visualizar também bons desempenhos dos latino-americanos México e Brasil, que apresentam pontuações semelhantes aos países desenvolvidos.

No contexto brasileiro, em grande parte o bom desempenho deve-se ao fato da abertura dos dados ser normatizada por lei, sendo seu marco regulatório a Lei Federal nº

---

<sup>1</sup>[https://opendatabarometer.org/?\\_year=2016&indicator=ODB](https://opendatabarometer.org/?_year=2016&indicator=ODB) acessado em: 18/06/2021

12.527/2011<sup>2</sup>. Esta lei é denominada Lei de Acesso à Informação (LAI), que tem como princípio cumprir o mandamento constitucional de assegurar aos cidadãos o direito de receber informações de órgão público para seu interesse particular ou coletivo e geral.

No caso de Curitiba, a regulamentação se dá pelo decreto 1.135/2012<sup>3</sup>. De acordo com este decreto a prefeitura tem uma política de disponibilização de dados abertos<sup>4</sup> em vigor desde outubro de 2014<sup>5</sup> e os disponibiliza via web.

Um exemplo dos dados disponibilizados pela prefeitura são os relatórios de gestão da secretaria municipal de saúde (SMS), acessíveis em regimes quadrimestral e anual<sup>6</sup>. Os relatórios não possuem uma formatação em um padrão específico e a disponibilização ocorre apenas em formato texto. Estes fatores não contribuem para que seja viável o uso dos relatórios de forma analítica, permitindo que ele apenas cumpra um papel burocrático. Os relatórios possuem, em muitos pontos, grandes diferenças estruturais por não possuírem um padrão específico replicado ao longo dos anos. Algumas delas são: entrada e saída de indicadores; alterações na fonte de dados do indicador; alteração na descrição do indicador; e alteração na formatação das informações que compõem o indicador.

Com a diferença estrutural os relatórios, torna-se um desafio gerar uma série histórica dos indicadores. Para exemplificar esta variação, pode-se observar a Figura 1 que têm o mesmo indicador em dois relatórios distintos da prefeitura, porém com nomenclaturas diferentes.

Nascidos Vivos 2012								
1º QUADRIMESTRE 2012: 8539								
2º QUADRIMESTRE 2012: 8429								
3º QUADRIMESTRE 2012: 7319								

PERFIL DE NASCIDOS VIVOS:								
ANO	1º QUADRIM 2012		2º QUADRIM 2012		3º QUADRIM 2012		TOTAL 2012	
	Nº	%	Nº	%	Nº	%	Nº	%
TOTAL	8539	100,0	8429	100,0	7319	100,0	24248	100,0
RISCO	1709	20,0	1678	19,9	1327	18,1	4228	19,7
ALTO RISCO	180	2,1	176	2,1	169	2,3	451	2,1

Fonte: SINASC

**Figura 1. Número de nascidos vivos de 2012 - 3º quadrimestre[SMS 2012].**

Essas mesmas dificuldades ocorreriam ao condensar estes relatórios para a geração de outros, pois estes são criados apenas para atender a demanda interna da prefeitura. Além disso, há uma dificuldade na integração dos dados da prefeitura com os relatórios da base nacional SargSus<sup>7</sup>. Outro ponto é que realizar a comparação entre

<sup>2</sup><https://www12.senado.leg.br/transparencia/arquivos/sobre/cartilha-lai/> acessado em: 18/06/2021

<sup>3</sup><https://mid.curitiba.pr.gov.br/2014/00147194.pdf> acessado em: 15/06/2021

<sup>4</sup><https://mid.curitiba.pr.gov.br/2014/00147194.pdf> acessado em: 15/06/2021

<sup>5</sup><https://www.curitiba.pr.gov.br/noticias/politica-de-dados-abertos-de-curitiba-esta-entre-as-mais-bem-avaliadas-do-mundo/36584> acessado em: 28/11/2021

<sup>6</sup><http://www.saude.curitiba.pr.gov.br/a-secretaria/relatorios-de-gestao.html> acessado em: 29/06/2021

<sup>7</sup><https://sargsus.saude.gov.br/sargsus> acessado em: 04/09/2018

Número de Nascidos Vivos e óbitos infantil por Distrito Sanitário. Curitiba, 1º quadrimestre 2018.			
Distrito Sanitário	Nascidos Vivos	Total Óbitos infantil	Coefficiente de Mortalidade
Bairro Novo	601	3	5,0
Boa Vista	1035	12	11,6
Boqueirão	720	4	5,6
Cajuru	785	1	1,3
CIC	815	8	9,8
Matriz	476	3	6,3
Portão	654	5	7,6
Pinheirinho	554	4	7,2
Santa Felicidade	550	4	7,3
Tatuquara	554	7	12,6
<b>Total do</b>	<b>6744</b>	<b>51</b>	<b>7,6</b>

Fonte: SMS / CE / Coordenação Eventos Vitais - SIM e SINASC, em 02/05/2018. \* dados preliminares

**Figura 2. Número de nascidos vivos de 2018 - 1º quadrimestre [SMS 2018].**

relatórios de forma manual é um processo extremamente custoso e demorado, exigindo etapas como o mapeamento dos indicadores, já que eles podem muitas vezes ser alterados, mudando sua nomenclatura, posição e até disponibilidade nos relatórios. Nesta direção, algoritmos de *Entity Matching*, que são capazes de localizar termos semelhantes, podem ser utilizados para identificar indicadores que se mantêm ao longo do tempo.

Este trabalho complementa a pesquisa de [Pereira and Kozievitch 2021], tendo como objetivo apresentar a problemática de comparabilidade dos dados de relatórios, observando os indicadores de saúde. Algoritmos de *Entity Matching* foram utilizados no trabalho para identificar indicadores que se mantêm ao longo do tempo.

## 2. Trabalhos Relacionados

*Entity matching*, que também pode ser entendido como identificação de duplicatas ou resolução de entidades, é uma tarefa crucial para integração e limpeza de dados [Cohen et al. 2000, Hernández and Stolfo 1995, Rahm and Do 2000]. Ele pode ser definido como uma tarefa de identificação de entidades (objetos, instâncias de dados), as quais se referem à mesma entidade no mundo real. Essas entidades podem ser localizadas em fontes de dados únicas ou distribuídas, citando como exemplos, um banco de dados do DataSUS ou nos bancos de dados de diversas prefeituras distintas pelo país.

Como existem grandes diferenças nas fontes de dados e de entidades onde se busca fazer o *match*, não existe um "melhor algoritmo de *Entity Matching*". Com isso, uma mesma abordagem de *match* pode ter resultados muito diferentes dependendo do formato e domínio dos dados que serão trabalhados. Alguns trabalhos [Guha et al. 2004, Sarawagi and Kirpal 2004] provaram que não existe um algoritmo de *string matching* universal, ao invés disso, em geral se encontram mais resultados positivos ao combinar-se vários métodos.

Nesta direção, o algoritmo de Levenshtein [Levenshtein 1965] faz o cálculo da distância de duas *strings* com base em sua similaridade. Para isso são observados caracteres que foram substituídos, inseridos ou apagados.

Já o algoritmo "*Soundex*" foi desenvolvido pela primeira vez por Russel e Odell

[Russell and Odell 1918], que logo em seguida solicitou sua patente. Este algoritmo possui diversas implementações, mas elas na maioria dos casos apresentam um padrão onde uma *string* é convertida em um código que consiste na sua primeira letra, seguida por três dígitos. A ideia é que os dígitos sejam atribuídos de acordo com um agrupamento pré-determinado de consoantes, e que esses grupos consonantais compartilhem características fonéticas, ou seja, soam de forma semelhante.

Por outro lado, o conceito de Dados Abertos [W3C 2012] tem sido cada vez mais explorado, e possui normas fundamentais que são: disponibilidade e acesso, reuso, redistribuição e participação universal. Essas características visam a livre utilização para desenvolvimento de quaisquer aplicações, independentemente de seu propósito final.

Nesta direção, os dados conectados estão ligados à área de inteligência artificial por usar um sistema que usa ontologias, sendo estas definidas por Gruber [Gruber 1993] como uma especificação explícita e formal de uma conceituação compartilhada. A linguagem para construção de ontologias *Ontology Web Language (OWL)* [W3C 2012] teve sua última atualização em 2012. Além das ontologias é utilizado um esquema de triplas *framework RDF* [W3C 2014], compostas por um sujeito e um objeto ligados por predicado que indica uma ação.

Dados Abertos Conectados é utilizado em [Pierin and Sichman 2018], por exemplo, para a recuperação de informação, interpretando e gerando triplas RDF, por meio de padrões identificados documentos. Após isso, as informações são verificadas pela ontologia OWL criada e, com isso, são minimizadas as inconsistências, e os dados são disponibilizados através de um *endpoint SPARQL*, cumprindo, assim, os pré-requisitos para as 5 estrelas do padrão de Dados Abertos Conectados.

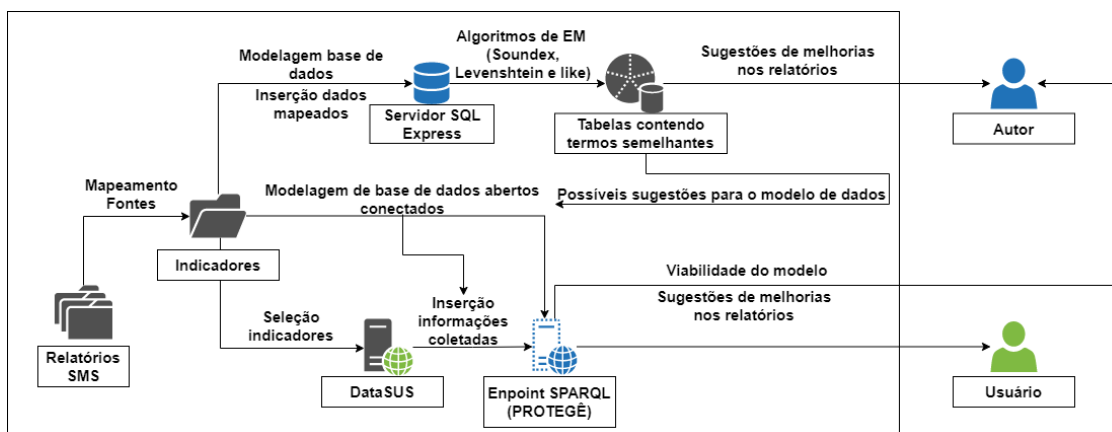
O uso de dados de qualidade no contexto das cidades inteligentes pode ser observado no artigo de [Araújo et al. 2017], onde é feito o uso de algoritmos de *Entity Matching* para aumentar a qualidade dos dados utilizados para tomada de decisões sobre o transporte público. Para isso, são integradas duas bases de dados de Curitiba (IPPUC e URBS), tendo como resultado a redução de ruídos nos dados.

Já o trabalho de [Rautenberg 2019] propõe uma ontologia para citação e classificação de documentos. No texto o autor cita a importância do uso de modelos já estabelecidos para aumentar a interoperabilidade e redução redundâncias, encorajando o reuso dos dados compartilhados. Com isso ele se utiliza de modelos ontológicos para montar seu modelo além de descrever requisitos específicos para que seu modelo possa ser reaproveitado.

As principais publicações que serviram de referência para a realização deste trabalho foram as de [Lopes et al. 2016], denominada “integração de dados na saúde pública”, a de [Victorino et al. 2017], “*big data* com dados governamentais”; a de [Alencar et al. 2018], “vitrine de Currículos Lattes”, e o de [Araújo et al. 2017], “integração de bases em cidades inteligentes”. A escolha para estas publicações se justifica por eles trazerem contribuições para a maioria dos principais conceitos abordados no presente trabalho. O diferencial proposto é baseado no uso de dados abertos para mapear indicadores de saúde, utilizando algoritmos de *Entity Matching*.

### 3. Desenvolvimento

Este trabalho tem como objetivo mapear as fontes dos dados contidas nos relatórios de saúde, observando os indicadores de saúde, utilizando algoritmos de *Entity Matching*.



**Figura 3. Arquitetura do projeto**

As seguintes tecnologias foram utilizadas: A) *Software*: SQL Express 2019 v15.0.2000.5<sup>8</sup>, SQL Server Management Studio (SSMS) v17.3<sup>9</sup>, Protege v5.5.0<sup>10</sup>; B) *Algoritmos*: Soundex nativo do SQL Server, Distância de Levenshtein: No processo de *Entity Matching*; C) *Linguagens*: OWL 2.0, Transact-SQL; D) *Frameworks*: RDF.

A Figura 3 ilustra a arquitetura do projeto. O mapeamento dos indicadores (ilustrado na Tabela 1) foi realizado manualmente, gerando uma planilha com mais de 300 itens, com informações referentes aos indicadores contidos nos relatórios estudados (descrição, fonte, relatório (descrito por ano e quadrimestre), 3 página(s) no relatório indicado).

**Tabela 1. Mapeamento indicadores**

Relatório	Informação
12-3	RAMOS DE ATIVIDADES MAIS DENUNCIADOS
12-3	SITUAÇÕES AMBIENTAIS MAIS SOLICITADAS
12-3	MEDIDAS ADMINISTRATIVAS EFETIVADAS PELOS DISTRITOS SANITÁRIOS E CENTRO DE SAÚDE AMBIENTAL
12-3	PLANTÃO DE FINAL DE SEMANA DE VIGILÂNCIA SANITÁRIA
12-3	PROGRAMA DE AVALIAÇÃO SANITÁRIA DE FARMÁCIAS DE MANIPULAÇÃO DE CURITIBA - PAS-FARMA
12-3	LEI ANTIFUMO

Apesar de serem mapeados as várias fontes e indicadores disponíveis nos relatórios (como ilustra a Tabela 1), somente dois foram selecionados na etapa "Seleção indicadores" ("Total de Leitos SUS Curitiba" e "Rede Física de Serviços no SUS Curitiba") nesta primeira etapa do trabalho. A sua escolha deve-se a disponibilidade no sistema do

<sup>8</sup><https://www.microsoft.com/pt-br/sql-server/sql-server-downloads> acessado em: 08/07/2021

<sup>9</sup><https://docs.microsoft.com/pt-br/sql/ssms/release-notes-ssms?view=sql-server-ver15#1791>

<sup>10</sup><https://protege.stanford.edu/products.php> acessado em: 08/07/2021



```

SELECT
A.ID,
A.DESCRICAO,
B.ID,B.DESCRICAO,
      DBO.EDIT_DISTANCE_WITHIN(A.DESCRICAO, B.DESCRICAO, 25)
FROM
INF_LIKE INNER JOIN INFORMACAO A ON ID1 = A.ID
INNER JOIN INFORMACAO B ON ID2 = B.ID
WHERE
      DIFFERENCE(A.DESCRICAO, B.DESCRICAO) > 3
      AND DBO.EDIT_DISTANCE_WITHIN(A.DESCRICAO, B.DESCRICAO, 25) <> -1

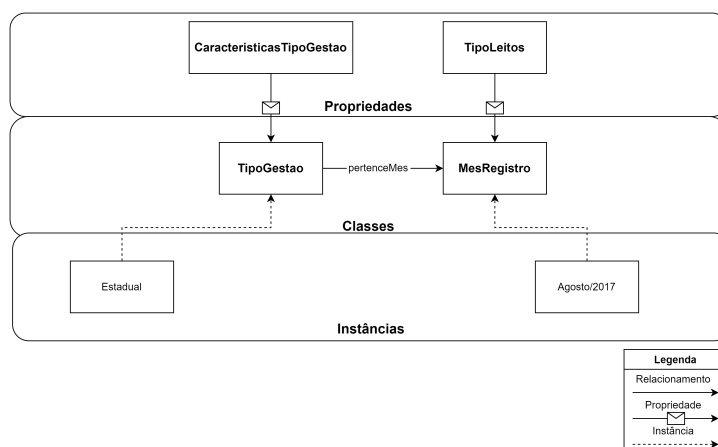
```

**Figura 5. Consulta exemplo Soundex em conjunto com distância de Levenshtein.**

ID1	S1	DESCRIÇÃO 1	ID2	S2	DESCRIÇÃO 2	Dist.
7	A200	Ações desenvolvidas no quadrimestre contra o mosquito Aedes Aegypti – SMS Curitiba	6	A200	Ações desenvolvidas no quadrimestre contra o mosquito Aedes Aegypti	15
23	A436	Alterações no quadro próprio de profissionais, segundo motivo de desligamento (Estatutário e CLT) SMS – Curitiba/ABRIL 2020	22	A436	Alterações no quadro próprio de profissionais, segundo motivo de desligamento (Estatutário e CLT) SMS – Curitiba	11
27	A542	Análises de amostras de água da SANEPAR realizadas por parâmetro analisado e total no período em Curitiba	26	A542	Análises de amostras de água da SANEPAR realizadas por parâmetro analisado e total no período	12
29	A542	Análises de amostras de água de Fontes Alternativas em Curitiba	28	A542	Análises de amostras de água de Fontes Alternativas	12
36	A353	ATENDIMENTO À POPULAÇÃO EM RELAÇÃO À FAUNA SINANTRÓPICA - COMENTÁRIOS	35	A353	ATENDIMENTO À POPULAÇÃO EM RELAÇÃO À FAUNA SINANTRÓPICA	14
45	A353	Atendimentos de Fisioterapia, Terapia Ocupacional e Fonoaudiologia SUS/Curitiba	44	A353	Atendimentos de Fisioterapia, Terapia Ocupacional e Fonoaudiologia	13
47	A353	Atendimentos realizadas pelo SAMU 192 2º Quadrimestre 2019	46	A353	Atendimentos realizadas pelo SAMU 192	21

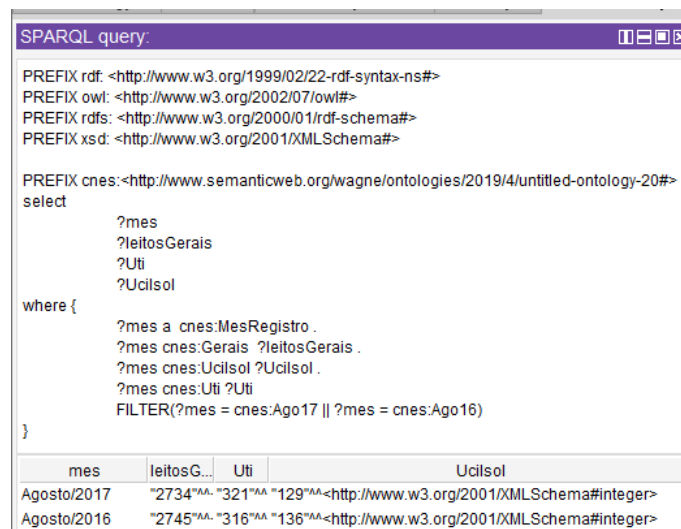
**Tabela 2. Exemplo de dicionário dos indicadores, resultante da consulta da Figura 5 (Soundex + Levenshtein).**

manualmente a uma base local com a mesma estrutura do DataSus, com o objetivo de permitir consultas SPARQL. Na Figura 7, encontra-se um exemplo desse tipo de consulta, onde foram selecionados dados com as colunas mês, leitos gerais, unidades de terapia intensiva e unidades de isolamento hospitalar com filtro para registros referentes aos meses de agosto de 2016 e agosto de 2017.



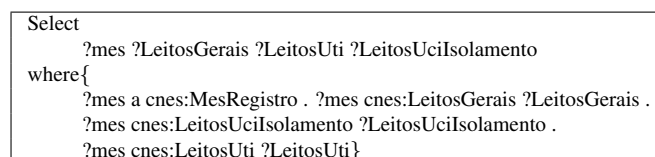
**Figura 6. Mapeamento das Classes.**

A Figura 8 exemplifica a consulta para o indicador de leitos SUS. O resultado



**Figura 7. Exemplo consulta SPARQL.**

da consulta foi tratado (em um processo manual, mas que poderia ser automatizado em SPARQL), gerando a Tabela 3, com estrutura similar às geradas nos relatórios. Note que as etapas anteriores de seleção de indicadores, "matching" e integração com SPARQL permitiram um resultado unificado de relatórios com estruturas diferenciadas.



**Figura 8. Consulta exemplo SPARQL**

Competência	Gerai	UTI	UCI e Isolamento
2016.2	2745	316	136
2016.3	2745	321	136
2017.2	2734	321	129
2017.3	2734	321	136
2017.3	2731	321	129
2018.1	2736	320	129

**Tabela 3. Tabela Leitos SUS**

De maneira resumida, o trabalho unificou os dados de relatórios da Secretaria da Saúde de Curitiba, mesmo não possuindo padrão (como as estruturas e fontes) entre os mesmos. Foi necessário bastante trabalho manual (como na comparação dos indicadores, ou ajuste de parâmetros na comparação do *match*). Para os órgãos responsáveis por estes dados, podemos citar as seguintes sugestões (que facilitarão a integração realizada neste trabalho): (i) Mapear e padronizar os indicadores e fontes (com seção, subseção e fonte); (ii) Indicar e mapear alterações de nome desses indicadores para o relatório anterior; (iii) Buscar um modelo padronizado; (iv) Disponibilizar os dados tabulados em formatos abertos; e (v) Criar e alimentar os dados e atualizar o modelo como forma de pré tabulação dos dados.



Detalhes adicionais a respeito deste trabalho podem ser encontrados online <sup>12</sup> e em [Pereira 2021].

#### 4. Conclusão

Relatórios de gestão podem não possuir uma formatação padrão, com a disponibilização apenas em formato texto. Além disso, podem possuir alteração entre entrada e saída de indicadores; fonte de dados do indicador; descrição do indicador; formatação das informações que compõem o indicador, além de estruturas diversificadas entre relatórios.

Nesta direção, o objetivo deste trabalho foi apresentar a problemática de comparabilidade dos dados de relatórios de gestão da Secretária de Saúde da prefeitura de Curitiba. Algoritmos de *Entity Matching* (capazes de localizar termos semelhantes) foram utilizados com dados de 16 relatórios compreendendo o período de 2012 até o segundo quadrimestre de 2020. O resultado indicou que, apesar do trabalho manual, é possível integrar os dados. O trabalho ainda listou sugestões para a melhoria da integração, como o uso de um modelo padronizado em conjunto como formatos abertos.

Como trabalhos futuros, podemos citar o refinamento do processo de *Entity Matching* para localizar de forma mais eficiente os indicadores que se mantêm ao longo do tempo, e a integração de outros tipos de dados.

#### 5. Agradecimentos

Os autores gostariam de agradecer a Universidade Tecnológica Federal do Paraná (DI-REC 10/2021), o Instituto de Pesquisa e Planejamento Urbano de Curitiba (IPPUC) e a Prefeitura de Curitiba.

#### Referências

- Alencar, A., Xavier, D., Chaves, L. C., and Souza, D. (2018). Publicação e consumo de dados abertos conectados acadêmicos. *Revista Principia - Divulgação Científica e Tecnológica do IFPB*, 1(42):136–145.
- Araújo, T. B., Cappiello, C., Kozievitch, N. P., Mestre, D. G., Pires, C. E. S., and Vitali, M. (2017). Towards reliable data analyses for smart cities. In *Proceedings of the 21st International Database Engineering Applications Symposium, IDEAS 2017*, page 304–308, New York, NY, USA. Association for Computing Machinery.
- Cohen, W. W., Kautz, H., and McAllester, D. (2000). Hardening soft information sources. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 255–259, Boston, MA, USA. Association for Computing Machinery.
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220.
- Guha, S., Koudas, N., Marathe, A., and Srivastava, D. (2004). Merging the results of approximate match operations. In *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*, pages 636–647, Toronto, Canada. VLDB Endowment.

---

<sup>12</sup><https://github.com/wlpereira89/tcc/tree/master/database%20relat%C3%B3rios>

- Hernández, M. A. and Stolfo, S. J. (1995). The merge/purge problem for large databases. *ACM Sigmod Record*, 24(2):127–138.
- Levenshtein, V. (1965). Levenshtein distance. <https://xlinux.nist.gov/dads/HTML/Levenshtein.html>.
- Lopes, G., Vidal, V. M. P., and Oliveira, M. (2016). Construção de linked data mashup para integração de dados da saúde pública. In *SBBD*, Salvador, BA, Brasil. SBC.
- Pereira, W. (2021). Geração de Indicadores da Secretaria de Saúde de Curitiba com base no modelo de Dados Abertos Conectados. Monografia (Bacharel em Informática), UTFPR.
- Pereira, W. L. and Kozievitch, N. P. (2021). Geração de indicadores da secretaria de saúde de curitiba com base no modelo de dados abertos conectados. In *WPCCG 2021*, Ponta Grossa - PR - Brasil. UTFPR.
- Pierin, F. L. and Sichman, J. S. (2018). Integraweb: uma arquitetura baseada em mapeamentos semânticos. In *Proceedings of the XI Seminar on Ontology Research in Brazil and II Doctoral and Masters Consortium on Ontologies*, pages 116–127, São Paulo, SP, Brasil. CEUR-WS.
- Rahm, E. and Do, H. H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.*, 23(4):3–13.
- Rautenberg, S. (2019). Web semântica e bibliografia: um estudo para a publicação de registros bibliográficos como dados abertos conectados. *Em Questão*, 25:313–330.
- Russell, R. and Odell, M. (1918). Soundex patent 01 261 167.
- Sarawagi, S. and Kirpal, A. (2004). Efficient set joins on similarity predicates. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, pages 743–754, Paris, France. Association for Computing Machinery.
- SMS (2012). Relatório quadrimestral da secretaria de saúde 2012 3º quadrimestre, disponível em [https://saude.curitiba.pr.gov.br/images/a\\_secretaria/Acompanhamento%20das%20A%C3%A7%C3%B5es%20SUS\\_3%C2%BA%20quadrim12\\_final.pdf](https://saude.curitiba.pr.gov.br/images/a_secretaria/Acompanhamento%20das%20A%C3%A7%C3%B5es%20SUS_3%C2%BA%20quadrim12_final.pdf).
- SMS (2018). Relatório primeiro 1º quadrimestre de 2018 da secretaria de saúde de curitiba, disponível em [https://saude.curitiba.pr.gov.br/images/SMS\\_%20relatorio%201%C2%BA%20Quadrimestre%202018%20com%20errata%20\\_atualizado%20em%2008.08.2018.pdf](https://saude.curitiba.pr.gov.br/images/SMS_%20relatorio%201%C2%BA%20Quadrimestre%202018%20com%20errata%20_atualizado%20em%2008.08.2018.pdf).
- Victorino, M. d. C., Shiessl, M., Oliveira, E. C., Ishikawa, E., Holanda, M. T. d., and Hokama, M. d. L. (2017). Uma proposta de ecossistema de big data para a análise de dados abertos governamentais conectados. *Informação & Sociedade: Estudos*, 27(1).
- W3C, O. W. G. (2012). *W3C Recommendation*. Cambridge, MA, EUA.
- W3C, S. W. (2014). *Resource Description Framework (RDF)*. Cambridge, MA, EUA.