

UNIVERSIDADE FEDERAL DO PARÁ INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS FACULDADE DE COMPUTAÇÃO

Jairo Nascimento de Sousa Filho

Modelo de Trabalho de Trabalho Acadêmico da Faculdade de Computação e Programa de Pós-Graduação em Ciência da Computação.

Belém

Jairo Nascimento de Sousa Filho

Modelo de Trabalho de Trabalho Acadêmico da Faculdade de Computação e Programa de Pós-Graduação em Ciência da Computação.

Monografia apresentada na Faculdade de Computação do Instituto de Ciências Exatas e Naturais como requisito parcial para obtenção do grau de Bacharel.

Universidade Federal do Pará

Orientador: Prof. Dr. Carlos Gustavo Resque dos Santos

Belém

2019

Solicite sua ficha catalográfica em: <http: bcficat.ufpa.br=""></http:>	

Jairo Nascimento de Sousa Filho

Modelo de Trabalho de Trabalho Acadêmico da Faculdade de Computação e Programa de Pós-Graduação em Ciência da Computação.

Monografia apresentada na Faculdade de Computação do Instituto de Ciências Exatas e Naturais como requisito parcial para obtenção do grau de Bacharel.

fonceito: Excelente!
selém, 1 de janeiro de 2019.
BANCA EXAMINADORA
Prof. Dr. Carlos Gustavo Resque dos Santos - Orientado UFPA
Nome Convidado 1 SIGLA INSTITUIÇÃO
Nome Convidado 2

SIGLA INSTITUIÇÃO



Agradecimentos

Os agradecimentos principais são direcionados à Gerald Weber, Miguel Frasson, Leslie H. Watter, Bruno Parente Lima, Flávio de Vasconcellos Corrêa, Otavio Real Salvador, Renato Machnievscz e todos aqueles que contribuíram para que a produção de trabalhos acadêmicos conforme as normas ABNT com LATEX fosse possível.

Agradecimentos especiais são direcionados ao Centro de Pesquisa em Arquitetura da Informação da Universidade de Brasília (CPAI), ao grupo de usuários latex-br e aos novos voluntários do grupo $abnT_EX2$ e que contribuíram e que ainda contribuirão para a evolução do $abnT_EX2$.

Resumo

Segundo a (ABNT, 2003), o resumo deve ressaltar o objetivo, o método, os resultados e as conclusões do documento. A ordem e a extensão destes itens dependem do tipo de resumo (informativo ou indicativo) e do tratamento que cada item recebe no documento original. O resumo deve ser precedido da referência do documento, com exceção do resumo inserido no próprio documento. (...) As palavras-chave devem figurar logo abaixo do resumo, antecedidas da expressão Palavras-chave:, separadas entre si por ponto e finalizadas também por ponto.

Palavras-chave: latex. abntex. editoração de texto.

Abstract

This is the english abstract.

 ${\bf Keywords: \ latex. \ abntex. \ text \ editoration.}$

Lista de ilustrações

Figura 1.	Usando o DTM Data Generator. Fonte: DTM Data Generator	30
Figura 2.	Usando o Redgate SQL Data Generator. Fonte: Red Gate SQL Data	
	Generator	31

Lista de quadros

Lista de tabelas

Lista de abreviaturas e siglas

ABNT Associação Brasileira de Normas Técnicas

abnTeX — ABsurdas Normas para TeX

Lista de símbolos

 Γ Letra grega Gama

 $\Lambda \qquad \qquad Lambda$

 \in Pertence

Sumário

1	INTRODUÇÃO	21
2	FUNDAMENTAÇÃO TEÓRICA	29
2.1	Dados Sintéticos	29
2.1.1	Trabalhos Relacionados	29
2.1.1.1	DTM Data Generator	29
2.1.1.2	Red Gate SQL Data Generator	31
2.1.1.3	Visual Studio (Premium) Data Generator	32
2.1.1.4	GEDIS Studio	32
2.1.1.5	dbForge Test Data Generator	32
2.1.1.6	Mockaroo	32
2.1.1.7	ApexSQL Generate	32
2.1.1.8	Datanamic Data Generator MultiDB	32
2.1.1.9	Upscene Advanced Data Generator	32
2.1.1.10	EMS Data Generator	32
2.2	Formato dos dados salvos	32
2.2.1	Arquivo	32
2.2.2	Web Service	33
3	ARQUITETURA DO PROJETO	35
3.1	Casos de uso do sistema	35
3.2	Ferramentas utilizadas	35
4	PROTÓTIPO	37
4.1	Tipos de Geradores de Dados	37
4.1.1	Sequencial	37
4.1.2	Aleatório	37
4.1.3	Funcional	37
4.1.4	Geométrico	37
4.2	Modos de Geração de Dados	37
4.2.1	Padrão	37
4.2.2	Streaming Data	37
4.2.3	Web Service	37
4.3	Estrutura de Interação Humano Computador	37
4.3.1	Interface do Usuário	37
4.3.2	Mensagens para o usuário	37

4.3.3	Atalhos do Teclado
4.3.4	Ajuda
5	TESTE
5.1	Teste de usabilidade do sistema
6	RESULTADOS
6.1	Teste de usabilidade
7	CONCLUSÃO 43
	REFERÊNCIAS 45

1 Introdução

Este documento e seu código-fonte são exemplos de referência de uso da classe abntex2 e do pacote abntex2cite. O documento exemplifica a elaboração de trabalho acadêmico (tese, dissertação e outros do gênero) produzido conforme a ABNT NBR 14724:2011 Informação e documentação - Trabalhos acadêmicos - Apresentação.

A expressão "Modelo Canônico" é utilizada para indicar que abnTEX2 não é modelo específico de nenhuma universidade ou instituição, mas que implementa tão somente os requisitos das normas da ABNT. Uma lista completa das normas observadas pelo abnTEX2 é apresentada em (ARAUJO, 2015a).

Sinta-se convidado a participar do projeto abnTEX2! Acesse o site do projeto em ">http://www.abntex.net.br/>. Também fique livre para conhecer, estudar, alterar e redistribuir o trabalho do abnTEX2, desde que os arquivos modificados tenham seus nomes alterados e que os créditos sejam dados aos autores originais, nos termos da "The LATEX Project Public License".

Encorajamos que sejam realizadas customizações específicas deste exemplo para universidades e outras instituições — como capas, folha de aprovação, etc. Porém, recomendamos que ao invés de se alterar diretamente os arquivos do abnTEX2, distribua-se arquivos com as respectivas customizações. Isso permite que futuras versões do abnTEX2 não se tornem automaticamente incompatíveis com as customizações promovidas. Consulte (ARAUJO, 2015b) para mais informações.

Este documento deve ser utilizado como complemento dos manuais do abnTEX2 (ARAUJO, 2015a; ARAUJO, 2015c; ARAUJO, 2015d) e da classe memoir (WILSON; MADSEN, 2010).

Esperamos, sinceramente, que o abnTEX2 aprimore a qualidade do trabalho que você produzirá, de modo que o principal esforço seja concentrado no principal: na contribuição científica.

Equipe abnTFX2

Lauro César Araujo

^{1 &}lt;http://www.latex-project.org/lppl.txt>

2 Fundamentação Teórica

2.1 Dados Sintéticos

O conceito de geração de dados sintéticos vieram por volta de 1993, por Rubin. (RUBIN, 1993) Em suma, seu objetivo era tonar anônimo os domicílios que participaram do censo daquela época. A partir desse fato, confidencialidade dos dados se tornou muito necessário, o que ajudou na popularização dos dados sintéticos. Portanto, dados sintéticos foi definido como "qualquer dado produzido o qual possa ser aplicado a uma dada situação que não foi obtido por mensuração direta.". (EDUCATION, 2016)

A necessidade de dados sintéticos podem ser de várias formas, desde a escassez de dados reais ou indisponibilidade; para teste de dados não usuais; para evitar lidar com questões de privacidade dos dados; teste de aplicação sem precisar modificar dados da aplicação de produção; criar teste de estresse da aplicação com *Big Data* antes de criar versão para produção; bem como não precisar adicionar os dados de teste manualmente. (KUMAR, 2019)

A aplicabilidade dos dados sintéticos é ilimitada, e é bastante explorada por setores cujos dados são sensíveis como a financeiro (LOPEZ-ROJAS; AXELSSON, 2012) e de saúde. (BERGEAT et al., 2014) Também são muito bem aplicáveis para exaustívos testes de segurança, os quais são necessários vários casos de teste, já que o pesquisador tem controle suficiente processamento (fórmulas matemáticas ou regras de geração) e saída do dado, como um sistema de detecção de fraudes. (BARSE; KVARNSTROM; JONSSON, 2003)

2.1.1 Trabalhos Relacionados

2.1.1.1 DTM Data Generator

DTM Data Generator (SOFT., 2019) é uma plataforma de geração de dados sintéticos que existe de 1998. Esta possui suporte para geração de dados em arquivos, em banco de dados, também para *Big Data*. Possui suporte multiplataforma, através do modo *multiplatform runtime*, contudo é limitado quando comparado à versão Windows, o qual suporta a versão para servidor também. É válido destacar que é um software essencialmente pago, isto é, existem versões gratuitas - demonstrações, para ser mais exato - mas limitadas. Além disso, há categorias de versões pagas, que vão desde limitações de geração (Standart - Professional) à vantagens mais técnicas (Professional - Entrerprise).

O DTM Data Generator possui uma vasta coleção de funcionalidades, as quais

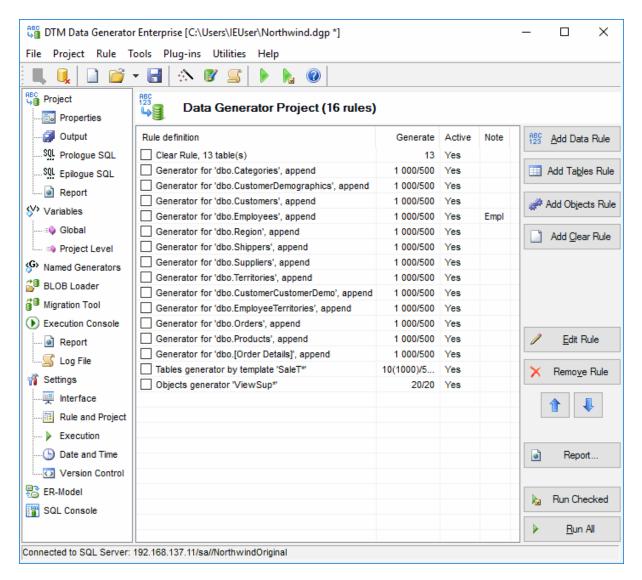


Figura 1. Usando o DTM Data Generator. Fonte: DTM Data Generator

liberadas de acordo com as versões pagas. Adotando a versão mais cara, a lista de features é composta por geração de dados em JSON, XML, CSV ou geração por separador customizado. Também permite gerar dados por arquivo DSN (Database Source Name), gerar dados por linha de comando, e gerar um arquivo SQL para não seja necessário conexão com banco de dados.

É possível gerar cerca de 9.2 sextilhões de registros por *rule*, modos de atualizar dados existentes (adicionar, substituir e *Data Scrambling*), e suporte para bibliotecas de dados realistas. A plataforma disponibiliza entrada de dados através de SQL, XML, JSON, pela WEB através de HTTP ou FTP, XLSM, arquivos de texto e scripts em Python. Também é possível visualizar e testar os dados gerados, bem como gerá-los nos principais arquivos de texto (TSV, CSV, "DSV", JSON, XML) e banco de dados. (MS SQL Server, Oracle, DB2, MySQL, PostgreSQL, Informix, Sybase, SQLite e Firebird)

Há uma suíte de produtos relacionados fornecidos pela DTM soft. Além do gerador

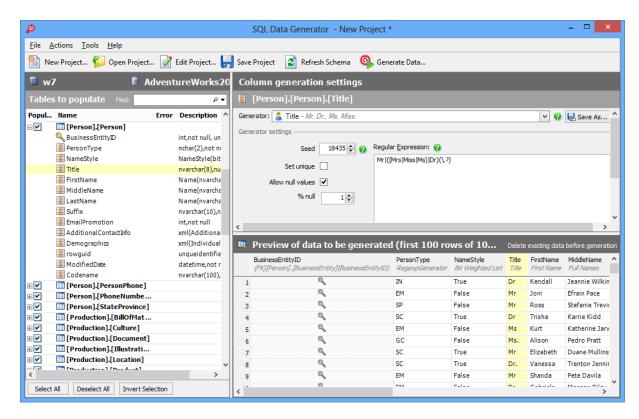


Figura 2. Usando o Redgate SQL Data Generator. Fonte: Red Gate SQL Data Generator

de dados, há o gerador de dados XML para teste de aplicação (DTM Test XML Generator); um gerador de planilhas Excel (DTM Data Generator for Excel) testador exaustivo - teste de estresse - de banco de dados (DTM DB Stress); Bem como editor, visualizador (DTM Data Editor), comparador e sincronizador de banco de dados (DTM Data Comparer) entre outros. 1

2.1.1.2 Red Gate SQL Data Generator

O SQL Data Generator (LTD, 2019) é um software que compõe uma suíte de ferramentas (chamada de SQL Toolbelt) da Red Gate. O software é exclusivo para o ecossistema Windows, com suporte do Windows 7 ao 10, à versão para servidores do Windows, ao SQL Server (2008 ao 2017), .NET e Oracle. Este produto é distribuido através de licenças pagas e vitalícias, com atualizações gratuitas e, no mínimo 1 ano de suporte gratuito.,

O SQL Toolbelt tem funcionalidades bem delimitadas e a função do Data Generator é popular um banco de dados. A população acontece ao escolher, primeiramente, uma tabela do banco. A partir disso, escolhe-se um gerador para cada coluna da tabela. Um gerador tem classificação fortemente baseada na realidade, isto é, possui geradores como palavras relacionadas à compras, pagamentos, pessoas (primeiro e último nome), dado geográficos e afins. Contudo, também disponibiliza a geração a partir de expressões regulares Regex generator e scripts de python. Por se tratar de banco de dados, também há checagem e

tratamento de constraints, Foreign keys e Dependencies. O SQL Data Generator também permite lidar com arquivos XML, quer seja para geração de valores XML, como utilizar como dados de entrada, além de mesclá-los com o Regex generator.

Quanto ao SQL Toolbelt oferecido pela Red Gate, ele conta com 2 modalidades, o completo com 14 programas e o essentials com 10. Entre os mais relevantes, pode-se citar o SQL Data Compare, SQL Data Generator, SQL Test, SQL Backup Pro e SQL Scripts Manager. 2

- 2.1.1.3 Visual Studio (Premium) Data Generator
- 2.1.1.4 GEDIS Studio
- 2.1.1.5 dbForge Test Data Generator
- 2.1.1.6 Mockaroo
- 2.1.1.7 ApexSQL Generate
- 2.1.1.8 Datanamic Data Generator MultiDB
- 2.1.1.9 Upscene Advanced Data Generator
- 2.1.1.10 EMS Data Generator

2.2 Formato dos dados salvos

2.2.1 Arquivo

JSON (BRAY, 2017) (CROCKFORD, 2003) (Javascript Object Notation, ou em português Notação de Objecto Javascript), lançado em 2002, é uma formatação leve para troca de dados. O uso é facilitado tanto para seres humano quanto para máquina. O JSON é um formato de texto que é independente de linguagem, mas foi baseado no objeto provido do Javascript (ECMA-262, 1999).

Quanto aos tipos de dados suportados, o JSON (BRAY, 2017) é uma sequência de tokens. Os tipos de tokens aceitos é do tipo *object*, *array*, *string*, *number* e nomes literais como *false*, *true* e *null*.

CSV (SHAFRANOVICH, 2005) (comma-separated values, ou em português Valores Separados por Vírgula) é tipo de texto MIME (Internet Media) (FREED J. KLENSIN, 1996) que utiliza a encodificação de caracteres US-ASCII (HAUSENBLAS E. WILDE, 2014). Ao longo dos anos, seu uso foi consolidado para exportar dados entre vários softwares de tabelas (Microsoft suíte para Apple Suíte, por exemplo). A padronização do CSV demorou a ocorrer e por isso, vários outros estilos surgiram, a exemplo, o uso do CSV com ponto-e-vírgula (;). Outros estilos foram criados a ponto de ser chamado de arquivo

DSV (RAYMOND, 2003). Por conseguinte, outro estilo que teve notoriedade na troca de dados entre bancos de dados ou tabelas de dados foi o TSV (KORPELA, 2000). A idea é similar ao CSV, porém é utilizado uma tabulação em vez de vírgula.

2.2.2 Web Service

(GROUP, 2004) Um Web service é definido como um software sistema criado para suportar interoperabilidade entre máquinas através da rede computadores. Também possui uma interface descrita em um formato processável por máquinas (WSDL) e um protocolo para comunicação (SOAP). (GROUP, 2004) Essa era a arquitetura utilizada em 2004. Atualmente é predominante o uso de REST que em vez de exportar serviços como o SOAP, exporta os dados em si e não necessita do WSDL. (STACKIFY, 2017)

3 Arquitetura do projeto

- 3.1 Casos de uso do sistema
- 3.2 Ferramentas utilizadas

4 Protótipo

- 4.1 Tipos de Geradores de Dados
- 4.1.1 Sequencial
- 4.1.2 Aleatório
- 4.1.3 Funcional
- 4.1.4 Geométrico
- 4.2 Modos de Geração de Dados
- 4.2.1 Padrão
- 4.2.2 Streaming Data
- 4.2.3 Web Service
- 4.3 Estrutura de Interação Humano Computador
- 4.3.1 Interface do Usuário
- 4.3.2 Mensagens para o usuário
- 4.3.3 Atalhos do Teclado
- 4.3.4 Ajuda

5 Teste

5.1 Teste de usabilidade do sistema

6 Resultados

6.1 Teste de usabilidade

7 Conclusão

Sed consequat tellus et tortor. Ut tempor laoreet quam. Nullam id wisi a libero tristique semper. Nullam nisl massa, rutrum ut, egestas semper, mollis id, leo. Nulla ac massa eu risus blandit mattis. Mauris ut nunc. In hac habitasse platea dictumst. Aliquam eget tortor. Quisque dapibus pede in erat. Nunc enim. In dui nulla, commodo at, consectetuer nec, malesuada nec, elit. Aliquam ornare tellus eu urna. Sed nec metus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas.

Phasellus id magna. Duis malesuada interdum arcu. Integer metus. Morbi pulvinar pellentesque mi. Suspendisse sed est eu magna molestie egestas. Quisque mi lorem, pulvinar eget, egestas quis, luctus at, ante. Proin auctor vehicula purus. Fusce ac nisl aliquam ante hendrerit pellentesque. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Morbi wisi. Etiam arcu mauris, facilisis sed, eleifend non, nonummy ut, pede. Cras ut lacus tempor metus mollis placerat. Vivamus eu tortor vel metus interdum malesuada.

Sed eleifend, eros sit amet faucibus elementum, urna sapien consectetuer mauris, quis egestas leo justo non risus. Morbi non felis ac libero vulputate fringilla. Mauris libero eros, lacinia non, sodales quis, dapibus porttitor, pede. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Morbi dapibus mauris condimentum nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Etiam sit amet erat. Nulla varius. Etiam tincidunt dui vitae turpis. Donec leo. Morbi vulputate convallis est. Integer aliquet. Pellentesque aliquet sodales urna.

Referências

- ARAUJO, L. C. A classe abntex2: Modelo canônico de trabalhos acadêmicos brasileiros compatível com as normas ABNT NBR 14724:2011, ABNT NBR 6024:2012 e outras. [S.l.], 2015. Disponível em: http://www.abntex.net.br/. Citado na página 27.
- ARAUJO, L. C. Como customizar o abnTeX2. 2015. Wiki do abnTeX2. Disponível em: https://github.com/abntex/abntex2/wiki/ComoCustomizar. Acesso em: 27 abr 2015. Citado na página 27.
- ARAUJO, L. C. O pacote abntex2cite: Estilos bibliográficos compatíveis com a ABNT NBR 6023. [S.l.], 2015. Disponível em: http://www.abntex.net.br/. Citado na página 27.
- ARAUJO, L. C. O pacote abntex2cite: tópicos específicos da ABNT NBR 10520:2002 e o estilo bibliográfico alfabético (sistema autor-data). [S.l.], 2015. Disponível em: http://www.abntex.net.br/. Citado na página 27.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 6028: Resumo apresentação. Rio de Janeiro, 2003. 2 p. Citado na página 11.
- BARSE, E. L.; KVARNSTROM, H.; JONSSON, E. Synthesizing test data for fraud detection systems. In: IEEE. 19th Annual Computer Security Applications Conference, 2003. Proceedings. [S.l.], 2003. p. 384–394. Citado na página 29.
- BERGEAT, M. et al. A french anonymization experiment with health data. In: . [S.l.: s.n.], 2014. Citado na página 29.
- BRAY, T. The JavaScript Object Notation (JSON) Data Interchange Format. 2017. Internet Engineering Task Force (IETF). Disponível em: https://tools.ietf.org/html/rfc8259. Acesso em: 31 jul 2019. Citado na página 32.
- CROCKFORD, D. *ECMA-404 The JSON Data Interchange Standard*. 2003. Json.org. Disponível em: https://json.org/json-pt.html. Acesso em: 31 jul 2019. Citado na página 32.
- EDUCATION, M.-H. The McGraw-Hill Dictionary of Scientific and Technical Terms, Seventh Edition (McGraw-Hill Dictionary of Scientific & Technical Terms). McGraw-Hill Professional, 2016. ISBN 0071608990. Disponível em: . Citado na página 29.
- FREED J. KLENSIN, J. P. N. Multipurpose Internet Mail Extensions (MIME) Part Four: Registration Procedures. 1996. Internet Engineering Task Force (IETF). Disponível em: https://tools.ietf.org/html/rfc2048. Acesso em: 31 jul 2019. Citado na página 32.
- GROUP, W. W. Web Services Architecture. 2004. Www.w3.org. Disponível em: https://www.w3.org/TR/ws-arch/>. Acesso em: 02 ago 2019. Citado na página 33.

- HAUSENBLAS E. WILDE, J. T. M. *ECMA-404 The JSON Data Interchange Standard*. 2014. Internet Engineering Task Force (IETF). Disponível em: https://tools.ietf.org/html/rfc7111#page-3. Acesso em: 31 jul 2019. Citado na página 32.
- KORPELA, J. *Tab Separated Values (TSV): a format for tabular data exchange.* 2000. Http://jkorpela.fi. Disponível em: http://jkorpela.fi/TSV.html. Acesso em: 31 jul 2019. Citado na página 33.
- KUMAR, V. 15 Best Test Data Generation Tools In 2019. 2019. Https://www.rankred.com. Disponível em: https://www.rankred.com/test-data-generation-tools/>. Acesso em: 17 ago 2019. Citado na página 29.
- LOPEZ-ROJAS, E. A.; AXELSSON, S. Money laundering detection using synthetic data. In: LINKÖPING UNIVERSITY ELECTRONIC PRESS. *The 27th annual workshop of the Swedish Artificial Intelligence Society (SAIS); 14-15 May 2012; Örebro; Sweden.* [S.l.], 2012. p. 33–40. Citado na página 29.
- LTD, R. G. S. *SQL Data Generator*. 2019. Https://www.red-gate.com/. Disponível em: https://www.red-gate.com/products/sql-development/sql-data-generator/. Acesso em: 18 ago 2019. Citado na página 31.
- RAYMOND, E. S. Data File Metaformats. Chapter 5. Textuality. 2003. Http://www.catb.org. Disponível em: http://www.catb.org/~esr/writings/taoup/html/ch05s02.html. Acesso em: 31 jul 2019. Citado na página 33.
- RUBIN, D. B. Statistical disclosure limitation. *Journal of official Statistics*, v. 9, n. 2, p. 461–468, 1993. Citado na página 29.
- SHAFRANOVICH, Y. Common Format and MIME Type for Comma-Separated Values (CSV) Files. 2005. Internet Engineering Task Force (IETF). Disponível em: https://tools.ietf.org/html/rfc4180#page-2. Acesso em: 31 jul 2019. Citado na página 32.
- SOFT., D. *DTM Database Tools*. 2019. Http://www.sqledit.com/. Disponível em: http://www.sqledit.com/. Disponível em: http://www.sqledit.com/. Disponível em: http://www.sqledit.com/. Disponível em: http://www.sqledit.com/. Disponível em: http://www.sqledit.com/dg/index.html. Acesso em: 17 ago 2019. Citado na página 29.
- STACKIFY. SOAP vs. REST: The Differences and Benefits Between the Two Widely-Used Web Service Communication Protocols. 2017. Stackify.com. Disponível em: https://stackify.com/soap-vs-rest/. Acesso em: 02 ago 2019. Citado na página 33.
- WILSON, P.; MADSEN, L. The Memoir Class for Configurable Typesetting User Guide. Normandy Park, WA, 2010. Disponível em: http://mirrors.ctan.org/macros/latex/contrib/memoir/memman.pdf>. Acesso em: 19 dez. 2012. Citado na página 27.