



Universidade de Brasília
Faculdade de Comunicação
Departamento de Jornalismo
Memorial de Projeto Final

MARÍLIA NESTOR SANTOS

**JORNALISMO E COMPUTAÇÃO:
Automação do lide**

Brasília DF
Fevereiro 2017

MARÍLIA NESTOR SANTOS

**JORNALISMO E COMPUTAÇÃO:
Automação do lide**

Memorial descritivo do produto apresentado à Universidade de Brasília, como requisito parcial para a obtenção do título de bacharel em Comunicação Social com habilitação em Jornalismo.

Orientadora: Márcia Marques

Brasília DF
Fevereiro 2017

MARÍLIA NESTOR SANTOS

**JORNALISMO E COMPUTAÇÃO:
Automação do lide**

Memorial descritivo do produto apresentado à Universidade de Brasília, como requisito parcial para a obtenção do título de bacharel em Comunicação Social com habilitação em Jornalismo.

BANCA EXAMINADORA

Prof. Dra Márcia Marques
ORIENTADORA

Prof. Dr. Benedito Medeiros Neto
MEMBRO

Prof. Dr. Márcio Carneiro dos Santos
MEMBRO

Prof. Dr. Sivaldo Pereira da Silva
SUPLENTE

Brasília DF
Fevereiro 2017

Agradecimentos

Agradeço a Deus e a Nossa Senhora, que me guiam e me mostram a cada dia como posso ser melhor e fazer o bem de diferentes maneiras com as mais variadas habilidades e ações, especialmente quando nada parece fazer sentido e começo a duvidar de mim e dEle. Também agradeço hodiernamente por terem me designado os homenageados dos próximos parágrafos.

Aos meus pais, Vital e Eliane, que vão sentir vontade de chorar quando lerem as primeiras seis palavras deste parágrafo e vão rir quando chegarem a este ponto final. Eles mesmos, que me moldaram de valores e ensinaram o humor (e a ironia) de cada dia, foram os combustíveis deste TCC, trabalho que não existiria não fosse pelos cafés da manhã cheios diálogos saudáveis, que poderiam ser chamados indiretas veladas, para eu terminar logo a graduação. Sem eles, não existiria luz. Exatamente, eu não existiria.

Ao meu bebê, também conhecido como filha 2 ou Milena, que me tornou filha única quando decidiu sair de casa antes da irmã mais velha e, como consequência, vai ter que gastar os cinco anos de psicologia portuguesa em mim porque não estou sabendo lidar com a sua falta todos os dias... Ou com os efeitos colaterais que sua falta causa no resto da família. Sério, o pessoal não está lidando bem, ficam descontando toda a carência em mim.

Falando nisso, agradeço a meus avós, em nome de quem agradeço a todos os tios, primos e membros agregados da família. Vocês, Cícero, Wilma e Iolanda, são a minha base e guia. Sou muito grata pela convivência, pelas conversas, pelas risadas, pelas indiretas e pela falta de noção, tantas vezes necessária. Sem vocês, eu não seria metade de quem sou. Sem vocês, eu também não existiria, então já nasci grata, né?

Ao meu Mozocrazy e petit gâteau. Obrigada pelo carinho, por me fazer uma pessoa (ainda) melhor e por ter me apresentado a computação quando eu já estava tendo espinhas e dor nas costas pela falta de rumo profissional, achando estar destinada a ser uma figurante desorientada da minha própria carreira. Obrigada, cara.

Aos meus orientadores, Márcia Marques e Benedito Medeiros, que acreditam mais em mim do que eu mesma. Sério, muito mais. Nem entendo essa confiança e futuro que veem em mim, além dessa paciência que deixaria Jó com inveja. Com

essa abordagem, me ajudam a ir além para buscar alcançar ou superar a expectativas. Muita gratidão.

Aos professores Márcio Carneiro, Guilherme Ramos e Suzana Guedes. Ao primeiro por não apenas ter me dito, mas por ter sido o exemplo que jornalista pode programar sim! Ao segundo por ter aceitado ensinar computação a uma aluna de jornalismo distraída no meio de uma turma empolgada de calouros de engenharia mecatrônica! À terceira, mas não menos importante, por ter me levantado quando eu estava duvidando do rumo profissional que escolhi seguir. Agradeço a todos.

A todos os meus amigos, que me deram força com conversas, risadas, almoços, jantares, lanchinhos, e bons *drinks* durante essa jornada. Vocês sabem bem que não há melhores combustíveis para mim!

Ou se tem chuva e não se tem sol
ou se tem sol e não se tem chuva!

Ou se calça a luva e não se põe o anel,
ou se põe o anel e não se calça a luva!

Quem sobe nos ares não fica no chão,
quem fica no chão não sobe nos ares.

É uma grande pena que não se possa
estar ao mesmo tempo nos dois lugares!

Ou guardo o dinheiro e não compro o doce,
ou compro o doce e gasto o dinheiro.

Ou isto ou aquilo: ou isto ou aquilo...
e vivo escolhendo o dia inteiro!

Não sei se brinco, não sei se estudo,
se saio correndo ou fico tranqüilo.

Mas não consegui entender ainda
qual é melhor: se é isto ou aquilo.

Cecília Meireles

Resumo

Este trabalho é o memorial de um algoritmo planejado e escrito por uma aluna de jornalismo em Python 3 para automação de lide de previsão do tempo. O código foi feito após um semestre de estudos sobre computação, linguagens de programação e suas aplicações no jornalismo, enquanto o memorial, baseado na metodologia *Design Science*, busca não apenas o embasamento teórico, mas também servir de incentivo e modelo a outros alunos, professores e profissionais da área.

Palavras-chave: jornalismo computacional, jornalismo automatizado, jornalismo algorítmico, jornalista-programador, automação de notícia, literacia de dados, tecnologia no jornalismo

Abstract

This is the memorial of an algorithm planned and written by a journalism student in Python 3 for the automation of weather forecasting. The code was made after a semester of studies on computing, programming languages and its applications in the field of journalism, while the memorial, based on the Science Design methodology, seeks not only a theoretical basis, but also intends to be an incentive and model for other students, teachers and professionals.

Palavras-chave: computational journalism, automation journalism, algorithmic journalism, news automation, journalist-programmer, data literacy, journalism technology

Lista de figuras

Figura 1.1: Lógica hierárquica para construção do lide	17
Figura 1.2: Lógica para construção do programa	22
Figura 1.3: Linhas 1 a 34: comentários iniciais (foto do programa)	23
Figura 1.4: Linhas 36 a 45: Bibliotecas (foto do programa)	24
Figura 1.5: Linhas 47 a 405: Funções de texto (foto do programa)	24
Figura 1.6: Linhas 407 a 427: Função de temperatura (foto do programa)	24
Figura 1.7: Linhas 429 a 465: Função de umidade relativa (foto do programa)	25
Figura 1.8: Linhas 467 a 556: Raspagem de dados - parte 1 (foto do programa)	25
Figura 1.9: Linhas 558 a 637: Raspagem de dados - parte 2 (foto do programa)	26
Figura 1.10: Linhas 639 a 676: Função para escolher a capital (foto do programa)	26
Figura 1.11: Linhas 680 a 682: Atribuição de valores (foto do programa)	26
Figura 1.12: Linhas 685 a 761: Ordem do texto (foto do programa)	27
Figura 1.13: Linhas 762 a 764: Fonte e fim do programa (foto do programa)	27
Figura 1.14: Foto da tela para resultado do programa: Lide da previsão do tempo para Fortaleza-CE 'hoje'	28
Figura 1.15: Foto da tela para resultado do programa: Lide da previsão do tempo para Belo Horizonte - MG 'amanhã'	29
Figura 5.1: Características do Big Data	35
Figura 5.2: Como algoritmos geram notícias	39
Figura 5.3: Previsões do tempo automatizadas do dia 6 de março de 1970	42
Figura 5.4: Extração de um dos textos da figura anterior	42

Sumário

Agradecimentos

Resumo

Abstract

1. Apresentação	10
1.1 Jornalismo Computacional	10
1.2 Automação no jornalismo	12
1.3 Descrição do produto	12
1.3.1 Estrutura do texto	14
1.3.2 Projeto do Algoritmo	22
1.3.3 Implementação	24
1.3.4 Análise do resultado	29
2. Problema de pesquisa	30
3. Justificativa	31
4. Objetivo de Pesquisa	33
4.1 Objetivo Geral	33
4.2 Objetivos Específicos	33
5. Referencial Teórico	34
5.1 Computação no jornalismo	34
5.1.1 Jornalista programador	36
5.2 Automação no jornalismo	38
5.2.1 Automação do lide	41
6. Metodologia	45
7. Conclusão	46
8. Referências bibliográficas	48

1. Apresentação

1.1 Jornalismo Computacional

A computação tem se espalhado cada vez mais profundamente em diversas áreas e transformado a rotina de diversos profissionais, inclusive dos que trabalham com notícia. Desde 2009, existe um campo no jornalismo dedicado unicamente à aplicabilidade da computação na área, chamado jornalismo computacional.

Como explicam Hamilton e Turner (2009, p.2), jornalismo computacional é a combinação de algoritmos, dados e um conhecimento de ciências sociais para complementar a função de contabilização do jornalismo. Segundo os autores, o jornalismo computacional se embasa em duas abordagens familiares: *computer assisted reporting* (CAR) e o uso de ferramentas das ciências sociais no jornalismo. Em suma, o jornalismo computacional busca capacitar repórteres para explorar a grande quantidade de informação, estruturada ou não, enquanto buscam e apuram as matérias (Hamilton; Turner, 2009, p.2).

Dois anos após publicarem a definição, Hamilton e Turner a aperfeiçoaram ao se juntarem à autora Sarah Cohen. Explicaram que, amplamente definido, o jornalismo computacional envolve mudar como histórias são encontradas, apresentadas, agregadas, rentabilizadas e arquivadas. O jornalismo pode, portanto, avançar inspirado em inovações em detecção de tópicos, análise de vídeos, personalização, agregação, visualização e *sensemaking*¹. (2011, p.1)

A abrangência do jornalismo computacional foi se desenvolvendo com o passar dos anos e pode-se dar como exemplo a ferramenta utilizada pela jornalista Meredith Broussard. Ela foi informada de que na Filadélfia, cidade estadunidense na Pensilvânia, não havia livros o suficiente nas escolas públicas, o que diminuía a competitividade dos alunos em exames nacionais frente aos que possuíam mais recursos. Ela, então, resolveu investigar.

A jornalista escreveu um segmento de *software*, derivado de outro software de inteligência artificial que fizesse os relatos. Em outras palavras, ela escreveu um conjunto de regras baseadas na lei local para que o sistema pudesse buscar os dados em fontes oficiais e relatar se estavam dentro ou fora dessas regras.

¹ Dervin, 2000, p.44

Em um artigo que apresentou em 2014 no *Computation+Journalism Symposium*², a própria Broussard explicou que máquinas 'inteligentes' não existem e que, geralmente, quando um cientista fala 'inteligência artificial' está se referindo a em campo de especialidade dentro de sua disciplina. A união entre a repórter e a computação resultou em uma série de reportagens investigativas chamada "*Stacked Up: Do Philly Students have the books they need?*"³ ("Empilhados: Os estudantes da Filadélfia têm os livros de que precisam?", tradução livre).

Além das reportagens, pode-se checar no site a quantidade de livros nas escolas das vizinhanças da Filadélfia em tempo real e o histórico das instituições não apenas em relação aos livros, mas também das notas nos testes nacionais. O site mostra ainda o cálculo do custo para remediar qualquer falta do material. A jornalista não precisa atualizar os dados, o sistema foi moldado para fazer isso sozinho.

Vale pontuar que o jornalismo computacional de que trato não é a matéria escrita do início ao fim por um computador, mas de um ser humano como protagonista do fazer jornalístico, utilizando-se das linguagens de programação e da lógica da computação como técnica.

Apesar de novas ferramentas serem interessantes aos jornalistas, o uso da computação abre questionamentos em relação à autoria e à credibilidade da história. No caso de Meredith Broussard, ela foi quem decidiu o que queria do algoritmo e escreveu o algoritmo e deu sentido às histórias, mas a evolução do uso pode resultar no menor protagonismo do profissional.

Torna-se necessário citar outro receio, bem frisado por Anderson (2012), em relação à visão da matemática e do algoritmo como futuro do jornalismo, esquecendo da abordagem sociológica necessária à humanização na história.

Levando os pontos supracitados em consideração, este trabalho trata a computação e suas possibilidades como ferramenta para o jornalista protagonista na escolha de fontes de informação, de pautas, de caminhos para tomar com a história e a maneira como a história é contada.

² <<http://computation-and-journalism.brown.columbia.edu/>> Acesso em 28 jan 2017

³ <<http://www.stackedup.org/#>> Acesso em 30 jan 2017

1.2 Automação no jornalismo

No caso deste memorial, especificamos a automação no jornalismo como códigos escritos de maneira a serem lidos por computadores e convertidos em textos narrativos informativos que não precisem do auxílio humano além do que foi determinado na concepção do código.

Um exemplo simples é um algoritmo escrito pelo professor Márcio Carneiro na linguagem Python, assim como o produto deste trabalho, para automatizar textos sobre o resultado do Campeonato Brasileiro de Futebol de 2013. O artigo em que o jornalista descreve esse resultado se chama "Narrativas automatizadas e geração de textos jornalísticos: a estrutura de organização do *lead*." Nele, o autor descreve como foi feita a modelagem do problema: "obter resultados dos jogos e informações complementares tais como local da partida e número da rodada; registrar essas informações em alguma estrutura simples de arquivo que pudesse posteriormente ser consultada para a construção do material; traduzir as próprias regras do torneio em termos de variáveis e relações para que a sintaxe do regulamento pudesse orientar a concatenação dos elementos do texto; gerar as frases a partir dos resultados das operações realizadas com os dados coletados nas partidas."

Assim, de acordo com o texto, a partir de um endereço na internet que disponibilize os dados necessários, o computador consegue ler e coletar as informações automaticamente. No caso, foi usada a página do portal Terra⁴ dedicada a esportes que publicava os resultados e a tabela atualizada do campeonato a cada rodada.

Foi feita, então, a parte do código que salvava as informações importantes para o lide, associando-os a cada time. Esse exemplo facilitará a compreensão da descrição abaixo.

1.3 Descrição do produto

O produto deste trabalho é um algoritmo, uma sequência finita de instruções escritas em uma linguagem de programação, sendo que devem ser bem definidas e

⁴ <<https://www.terra.com.br/>> Acesso em 04 fev 2017

não ambíguas, uma vez que o computador não tem capacidade interpretativa, ele apenas cumpre ordens. Esse algoritmo serve para automação de um lide sobre previsão do tempo de Brasília e das capitais dos estados brasileiros.

É utilizada a raspagem de dados, método para extrair os dados em documentos, como em páginas da web, e torná-los usáveis, possíveis de serem processados pelo computador. A ideia final é mostrar como se faz esse algoritmo, em especial a lógica de texto que o jornalista pode utilizar para encaixar os dados raspados, para deixar claro que é uma habilidade acessível a todos os estudantes, professores e profissionais da área.

A previsão do tempo foi escolhida como conteúdo porque é um assunto neutro, fatídico e de fácil compreensão que pouco pode ser alterado por convicções, ideal para externar as possibilidades da ferramenta. Assim, a ideia é deixar bem claro um passo a passo que fique de simples reprodução ou adaptação até a um jornalista *freelancer* ou estudante, que não podem não possuir a estrutura de uma redação, e futuramente até embasar-se nele para escrever outro com finalidade distinta.

Dessa maneira, é possível que um algoritmo deste trabalho, tendo essa função, seja adaptado e utilizado para algo rentável. A razão pode ser compreendida em uma entrevista⁵ feita por Meredith Broussard, jornalista citada acima, com Justin Myers, editor de automação da *The Associated Press*⁶ (AP), em que ele afirma: “Se você escreve um sistema que poupa duas horas por mês, mas leva duas semanas para escrevê-lo, vai levar muito tempo para compensar o esforço.

Às vezes faço isso porque ajuda em projetos maiores mais à frente. Conteúdo que criamos frequentemente ou processos que que pessoa nas redações executam mais frequentemente são oportunidades mais interessantes.”⁷ (tradução livre). O

⁵ <<http://towcenter.org/robot-journalism-and-relative-size/>> Acesso em 03 fev 2017

⁶ <<https://www.ap.org/en-us/>> Acesso em 03 fev 2017

⁷ “*If you write a system that saves two hours a month, but it takes you two weeks to write it, it will take a very long time to recoup that effort. I do sometimes do this, because sometimes it helps set us up for larger projects down the road. Content that we create frequently, or processes that people in newsrooms perform more frequently, are more interesting opportunities.*”

diálogo foi publicada no Tow Center⁸, da Escola de Jornalismo da Universidade Columbia, no ano passado, 2016.

1.3.1 Estrutura do texto

Nesse ponto, estão os dados a serem raspados, os URLs⁹ de que vêm e a lógica estrutural do texto baseada nessas informações. Essa parte é especialmente importante a jornalistas porque, quanto melhor for o planejamento do texto, mais natural vai parecer ao leitor. Uma vez que o código depende da automação da raspagem de dados, abaixo está uma lista das URLs referentes às informações de cada capital. A fonte é o INMET, Instituto Nacional de Meteorologia.

ARACAJU - SE	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2800308
BELÉM - PA	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=1501402
BELO HORIZONTE - MG	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=3106200
BOA VISTA - RR	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=1400100
BRASÍLIA - DF	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=5300108
CAMPO GRANDE - MS	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=5002704
CUIABÁ - MT	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=5103403
CURITIBA - PR	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=4106902
FLORIANÓPOLIS - SC	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=4205407
FORTALEZA - CE	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2304400
GOIÂNIA - GO	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=5208707
JOÃO PESSOA - PB	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2507507

⁸ <<http://towcenter.org/>> Acesso em 03 fev 2017

⁹ URL é a sigla de *Uniform Resource Locator*, que é o endereço de um recurso disponível em uma rede.

MACAPÁ - AP	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=1600303
MACEIÓ - AL	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2704302
MANAUS - AM	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=1302603
NATAL - RN	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2408102
PALMAS - TO	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=1721000
PORTO ALEGRE - RS	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=4314902
PORTO VELHO - RO	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=1100205
RECIFE - PE	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2611606
RIO BRANCO - AC	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=1200401
RIO DE JANEIRO - RJ	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=3304557
SALVADOR - BA	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2927408
SÃO LUÍS - MA	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2111300
SÃO PAULO - SP	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=3550308
TERESINA - PI	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2211001
VITÓRIA - ES	http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=3205309

Fonte: Elaborado pela autora

Para o texto não ficar com aparência de robotizado, não poderia ser feita apenas uma estrutura básica. A ideia não é escrever um texto completo, mas apenas um parágrafo simples que exemplifique o objetivo do trabalho, então as

variáveis foram separadas em grupos e subgrupos, de acordo com o glossário do INMET¹⁰ e às informações disponíveis nas 'URLs fonte'.

Nessa situação, na estrutura básica de um lide¹¹, as perguntas 'Quem?' e 'Onde?' se referem à mesma informação (a cidade), 'Quando?' é uma variável, 'O quê?' são os dados raspados e 'Por quê' e 'Como' são perguntas ignoradas graças à natureza do conteúdo. Portanto, ajustando as variáveis a uma lógica a ser utilizada no algoritmo, fica assim:

Variáveis raspadas:

- "Nome_cidade" para o nome de cada cidade
- "Temp_min" para a temperatura mínima
- "Temp_max" para a temperatura máxima
- "Umid_min" para a umidade mínima
- "Umid_max" para a umidade máxima
- "Texto_manha" para o texto da manhã
- "Texto_tarde" para o texto da tarde
- "Texto_noite" para o texto da noite

Outras variáveis::

- "Qnd": O dia ou o dia seguinte começando com letra maiúscula
- "qnd": O dia ou o dia seguinte começando com letra minúscula

Grupo 1: Cidade + Quando	Grupo 2: Temperatura	Grupo 3: Umidade	Grupo 4: Textos	Grupo 5: Fonte
Subgrupo 1.1 início do lide	Subgrupo 2.1 frase completa	Subgrupo 3.1 frase completa	Subgrupo 4.1 frase completa	Fonte
Subgroup 1.2 no fim da primeira frase	Subgrupo 2.2 início da frase ou do lead	Subgrupo 3.2 início da frase	Subgrupo 4.2 início da frase	
	Subgrupo 2.3 fim de frase	Subgrupo 3.3 fim de frase	Subgrupo 4.3 fim da primeira frase	

Fonte: Elaborado pela autora

Um vez que cada grupo representa um tipo de informação, a ordem foi ajustada de acordo com a importância. Por isso, foram criadas regras que devem ser seguidas para respeitar a hierarquia, representadas pela imagem abaixo e com uma explicação posterior:

¹⁰ <<http://www.inmet.gov.br/portal/index.php?r=home/page&page=glossario>> Acesso em 03 fev 2017

¹¹ O quê? Quem? Como? Quando? Onde? Por quê?

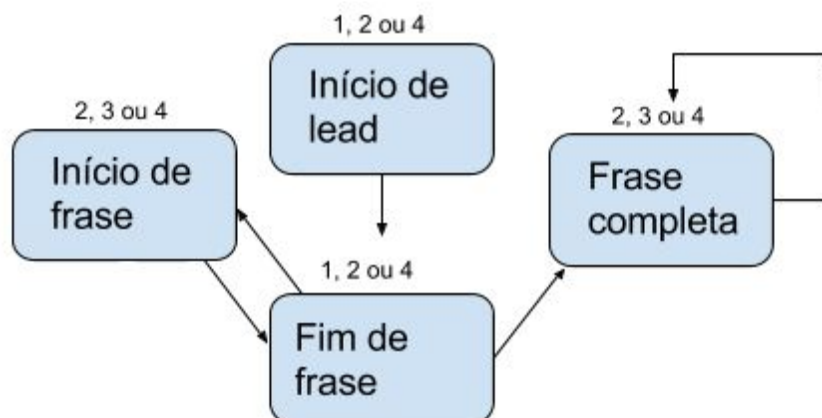


Figura 1.1: Lógica hierárquica para construção do lide (Elaborado pela autora)

- Grupo 1 deve sempre estar na primeira frase.
- Apenas os grupos 1, 2 e 4 podem começar o lide.
- Grupos 1 e 2 nunca estarão no fim do lide.
- Grupo três nunca estará na primeira frase.
- Grupo 5 sempre será a última frase e terá um comando de 'pula linha' antecedendo.
- A barra de espaço entre um ponto e uma palavra estará sempre no grupo anterior.
- Cada subgrupo terá cinco blocos de texto para facilitar a programação, mesmo que os textos se repitam.

Esses são os blocos de textos dos subgrupos:

Grupo 1.1 no início de lide	Grupo 1.2 no fim da frase
Em (*Nome_cidade*), (*qnd*)	em (*Nome_cidade*) (*qnd*).
A previsão de (*qnd*) para (*Nome_cidade*) é	de (*Nome_cidade*) (*qnd*).
(*Qnd*) em (*Nome_cidade*),	(*qnd*) em (*Nome_cidade*).
(*Qnd*) (*Nome_cidade*) tem expectativa de	(*qnd*) em (*Nome_cidade*).
(*Qnd*) (*Nome_cidade*) espera	(*qnd*) em (*Nome_cidade*).

Grupo 2.1 frase completa	Grupo 2.2 início de frase	Grupo 2.3 no fim de frase
Temperatura mínima prevista de (*Temp_min*) °C e máxima de (*Temp_max*) °C.	Temperatura mínima prevista de (*Temp_min*) °C e máxima de (*Temp_max*) °C,	temperatura mínima de (*Temp_min*) °C e máxima de (*Temp_max*) °C.
Temperatura máxima de (*Temp_max*) °C e	Temperatura máxima de (*Temp_max*) °C e	temperatura máxima de (*Temp_max*) °C e

temperatura mínima de (*Temp_min*) °C.	temperatura mínima de (*Temp_min*) °C,	temperatura mínima de (*Temp_min*) °C.
Temperatura variando entre (*Temp_min*) °C e (*Temp_max*) °C.	Temperatura variando entre (*Temp_min*) °C e (*Temp_max*) °C,	temperatura variando entre (*Temp_min*) °C e (*Temp_max*) °C.
Máxima prevista de (*Temp_max*) °C e mínima de (*Temp_min*) °C.	Máxima prevista de (*Temp_max*) °C e mínima de (*Temp_min*) °C,	máxima de (*Temp_max*) °C e mínima de (*Temp_min*) °C.
Mínima prevista de (*Temp_min*) °C e máxima de (*Temp_max*) °C.	Mínima prevista de (*Temp_min*) °C e máxima de (*Temp_max*) °C,	mínima de (*Temp_min*) °C e máxima de (*Temp_max*) °C.

Grupo 3.1 frase completa	Grupo 3.2 no início de frase	Grupo 3.3 no fim de frase
Umidade variando entre (*Umid_min*)% e (*Umid_max*)%.	Umidade variando entre (*Umid_min*)% e (*Umid_max*)%,	com umidade variando entre (*Umid_min*)% e (*Umid_max*)%.
Umidade mínima prevista de (*Umid_min*)% e máxima de (*Umid_max*)%.	Umidade mínima prevista de (*Umid_min*)% e máxima de (*Umid_max*)%,	com umidade mínima prevista de (*Umid_min*)% e máxima de (*Umid_max*)%.
Umidade máxima de (*Umid_max*)% e mínima de (*Umid_min*) %.	Umidade máxima de (*Umid_max*)% e mínima de (*Umid_min*) %,	com umidade máxima de (*Umid_max*)% e mínima de (*Umid_min*) %.
Umidade prevista de (*Umid_max*)% e (*Umid_min*)%.	Umidade prevista de (*Umid_max*)% e (*Umid_min*)%,	com umidade prevista de (*Umid_max*)% e (*Umid_min*)%.
Umidade prevista entre (*Umid_min*)% e (*Umid_max*)%.	Umidade prevista entre (*Umid_min*)% e (*Umid_max*)%,	com umidade prevista entre (*Umid_min*)% e (*Umid_max*)%.

Se (*Texto_manha*) == (*Texto_tarde*) e (*Texto_manha*) == (*Texto_noite*)		
Grupo 4.1 completo	Grupo 4.2 início de lide	Grupo 4.3 fim da primeira frase
(*Texto_manha*) pela manhã e permanece assim ao longo do dia e da noite.	Amanhece o dia (*Texto_manha*) ((Grupo 1 fim de frase)) e permanece assim até a noite.	amanhece o dia (*Texto_manha*) e permanece assim até a noite.
Céu (*Texto_manha*) permanece assim durante todo dia e noite.	(*Texto_manha*) durante todo dia e noite	permanece (*Texto_manha*) durante todo dia e noite.
(*Texto_manha*) durante todo dia e noite.	(*Texto_manha*) é o céu durante todo o dia	tem céu (*Texto_manha*) de manhã, à tarde e à noite.
De manhã, à tarde e à noite,	De manhã, céu	permanece durante dia e noite

céu se mantém (*Texto_manha*).	(*Texto_manha*), permanece assim ao longo do dia	com céu (*Texto_manha*).
De manhã, à tarde e à noite, céu se mantém (*Texto_manha*).	(*Texto_manha*) durante todo dia e noite	permanece durante dia e noite com céu (*Texto_manha*).

Se (*Texto_manha*) == (*Texto_tarde*) e (*Texto_manha*) != (*Texto_noite*)		
Grupo 4 completo	Grupo 4 início de lide	Grupo 4 fim da primeira frase
Céu amanhece (*Texto_manha*) e permanece assim ao longo do dia, mas fica (*Texto_noite*) à noite.	Amanhece com céu (*Texto_manha*) ((Grupo 1 fim de frase)) e permanece assim ao longo do dia, mas fica (*Texto_noite*) à noite.	amanhece o céu (*Texto_manha*) e permanece assim ao longo do dia, mas fica (*Texto_noite*) à noite.
Céu (*Texto_manha*) permanece assim durante todo dia. À noite, céu (*Texto_noite*).	Céu (*Texto_manha*) durante todo dia e noite (*Texto_noite*)	permanece (*Texto_manha*) durante todo dia, mas à noite fica (*Texto_noite*).
(*Texto_manha*) durante todo o dia. À noite, céu (*Texto_noite*).	(*Texto_manha*) é o céu durante todo o dia, enquanto à noite ele fica (*Texto_noite*)	tem céu (*Texto_manha*) de manhã e à tarde, mas noite termina com céu (*Texto_noite*).
De manhã e à tarde, céu se mantém (*Texto_manha*), mas fica (*Texto_noite*) durante a noite.	Céu (*Texto_manha*) durante todo dia e noite (*Texto_noite*)	amanhece o céu (*Texto_manha*) e permanece assim ao longo do dia, mas fica (*Texto_noite*) à noite.
De manhã, à tarde, céu se mantém (*Texto_manha*), mas acaba (*Texto_noite*) à noite.	Amanhece com céu (*Texto_manha*) ((Grupo 1 fim de frase)) e permanece assim ao longo do dia, mas fica (*Texto_noite*) à noite.	permanece durante todo o dia com céu (*Texto_manha*). À noite, céu fica (*Texto_noite*).

Se (*Texto_manha*) != (*Texto_tarde*) e (*Texto_manha*) == (*Texto_noite*)		
Grupo 4 completo	Grupo 4 início de lide	Grupo 4 fim da primeira frase
Céu amanhece (*Texto_manha*), se torna (*Texto_tarde*) ao longo do dia e volta a ficar (*Texto_noite*) à noite.	Amanhece com céu (*Texto_manha*) ((Grupo 1 fim de frase)), mas fica (*Texto_tarde*) ao longo do dia e volta a ficar (*Texto_noite*) à noite.	amanhece o céu (*Texto_manha*), que vai ficar (*Texto_tarde*) ao longo do dia, mas volta a ficar (*Texto_noite*) à noite.
Céu amanhece (*Texto_manha*), mas fica (*Texto_tarde*) assim durante	Céu (*Texto_manha*) durante de manhã, (*Texto_tarde*) à tarde e noite (*Texto_noite*)	permanece (*Texto_manha*) durante manhã e noite, mas tarde pode ser (*Texto_tarde*).

todo dia. À noite, céu (*Texto_noite*).		
(*Texto_manha*) durante todo o dia e noite, mas céu fica (*Texto_tarde*) na parte da tarde.	(*Texto_manha*) é o céu durante toda manhã e noite, mas ao longo da tarde ele fica (*Texto_tarde*).	tem céu (*Texto_manha*) de manhã, mas fica (*Texto_tarde*) e volta a ficar (*Texto_noite*) à noite.
De manhã, céu se mantém (*Texto_manha*), mas deve ficar (*Texto_tarde*) na parte da tarde e voltar a ficar (*Texto_noite*) durante a noite.	Céu (*Texto_manha*) durante toda a manhã, tarde com céu (*Texto_tarde*) e noite volta ao (*Texto_noite*)	amanhece o céu (*Texto_manha*), mas se torna (*Texto_tarde*) e volta a ficar (*Texto_noite*) à noite.
(*Texto_manha*) durante todo o dia e noite, mas céu fica (*Texto_tarde*) na parte da tarde.	Céu (*Texto_manha*) durante toda manhã e noite, (*Texto_tarde*) durante a tarde	amanhece o céu (*Texto_manha*), que vai ficar (*Texto_tarde*) ao longo do dia, mas volta a ficar (*Texto_noite*) à noite.

Se (*Texto_manha*) != (*Texto_tarde*) e (*Texto_tarde*) == (*Texto_noite*)		
Grupo 4 completo	Grupo 4 início de lide	Grupo 4 fim da primeira frase
Céu amanhece (*Texto_manha*), mas fica (*Texto_tarde*) durante a tarde e permanece assim ao longo da noite.	Amanhece com céu (*Texto_manha*) ((Grupo 1 fim de frase)), mas fica (*Texto_noite*) durante tarde e noite.	amanhece o céu (*Texto_manha*), mas fica (*Texto_noite*) durante o resto do dia.
De manhã, céu (*Texto_manha*). Durante a tarde, céu (*Texto_tarde*) que permanece assim à noite.	Céu (*Texto_manha*) durante a manhã e (*Texto_noite*) no resto do dia	permanece (*Texto_manha*) durante toda a manhã, mas à tarde fica (*Texto_noite*).
Céu (*Texto_manha*) durante toda manhã. À tarde e à noite, céu (*Texto_noite*).	(*Texto_manha*) é o céu durante a manhã, enquanto ele fica (*Texto_noite*) e permanece assim	tem céu (*Texto_manha*) de manhã, mas tarde e noite terminam com céu (*Texto_noite*).
De manhã, céu se mantém (*Texto_manha*), mas fica (*Texto_noite*) durante a tarde e a noite.	Céu (*Texto_manha*) durante toda a manhã, mas fica (*Texto_noite*)	permanece o céu (*Texto_manha*) durante a manhã, mas fica (*Texto_noite*) mais tarde.
De manhã, à tarde, céu se mantém (*Texto_manha*), mas acaba (*Texto_noite*) no resto do dia.	Céu (*Texto_manha*) durante a manhã e (*Texto_noite*) no resto do dia	amanhece o céu (*Texto_manha*), mas fica (*Texto_noite*) durante o resto do dia.

Se (*Texto_manha*) != (*Texto_tarde*) e (*Texto_manha*) != (*Texto_noite*) e (*Texto_tarde*) != (*Texto_noite*)		
Grupo 4 completo	Grupo 4 início de lide	Grupo 4 fim da primera frase
Céu amanhece (*Texto_manha*), se torna (*texto_tarde*) ao longo do dia e fica (*Texto_noite*) à noite.	Amanhece com céu (*Texto_manha*) ((Grupo 1 fim de frase)), mas fica (*Texto_tarde*) ao longo do dia e (*Texto_noite*) à noite.	amanhece o céu (*Texto_manha*), que vai ficar (*Texto_tarde*) ao longo do dia e (*Texto_noite*) à noite.
Céu amanhece (*Texto_manha*), mas fica (*Texto_tarde*) assim durante todo dia. À noite, céu (*Texto_noite*).	Céu (*Texto_manha*) durante a manhã, à tarde fica (*Texto_tarde*) e noite (*Texto_noite*)	permanece (*Texto_manha*) durante a manhã, fica (*Texto_tarde*) à tarde. Noite (*Texto_noite*).
Céu amanhece (*Texto_manha*), mas fica (*Texto_tarde*) assim durante todo dia. À noite, céu (*Texto_noite*).	(*Texto_manha*) é o céu durante toda manhã, mas ao longo da tarde ele fica (*Texto_tarde*) e à noite permanece (*Texto_noite*)	tem céu (*Texto_manha*) de manhã, mas fica (*Texto_tarde*) ao longo do dia e (*Texto_noite*) à noite.
Céu amanhece (*Texto_manha*), se torna (*texto_tarde*) ao longo do dia e fica (*Texto_noite*) à noite.	Céu (*Texto_manha*) durante toda a manhã, tarde com céu (*Texto_tarde*) e noite no (*Texto_noite*)	amanhece o céu (*Texto_manha*), mas se torna (*Texto_tarde*) e fica (*Texto_noite*) à noite.
(*Texto_manha*) durante toda a manhã, mas céu fica (*Texto_tarde*) na parte da tarde. (*Texto_noite*) à noite.	(*Texto_manha*) é o céu durante toda manhã, mas ao longo da tarde ele fica (*Texto_tarde*) e à noite permanece (*Texto_noite*)	amanhece o céu (*Texto_manha*), que vai ficar (*Texto_tarde*) ao longo do dia e acaba (*Texto_noite*) à noite.

Grupo 5 Completo
(pula linha) Fonte: INMET

Fonte: Elaborado pela autora

1.3.2 Projeto do algoritmo

Na figura abaixo, são expressadas em fluxograma as ações a serem automatizadas, seguindo a sequência das linhas de instrução do algoritmo, com explicação abaixo. É necessário notar que não é isso que o programa necessariamente faz, esse é apenas o planejamento, de maneira que pode ser alterado, assim como o foi.

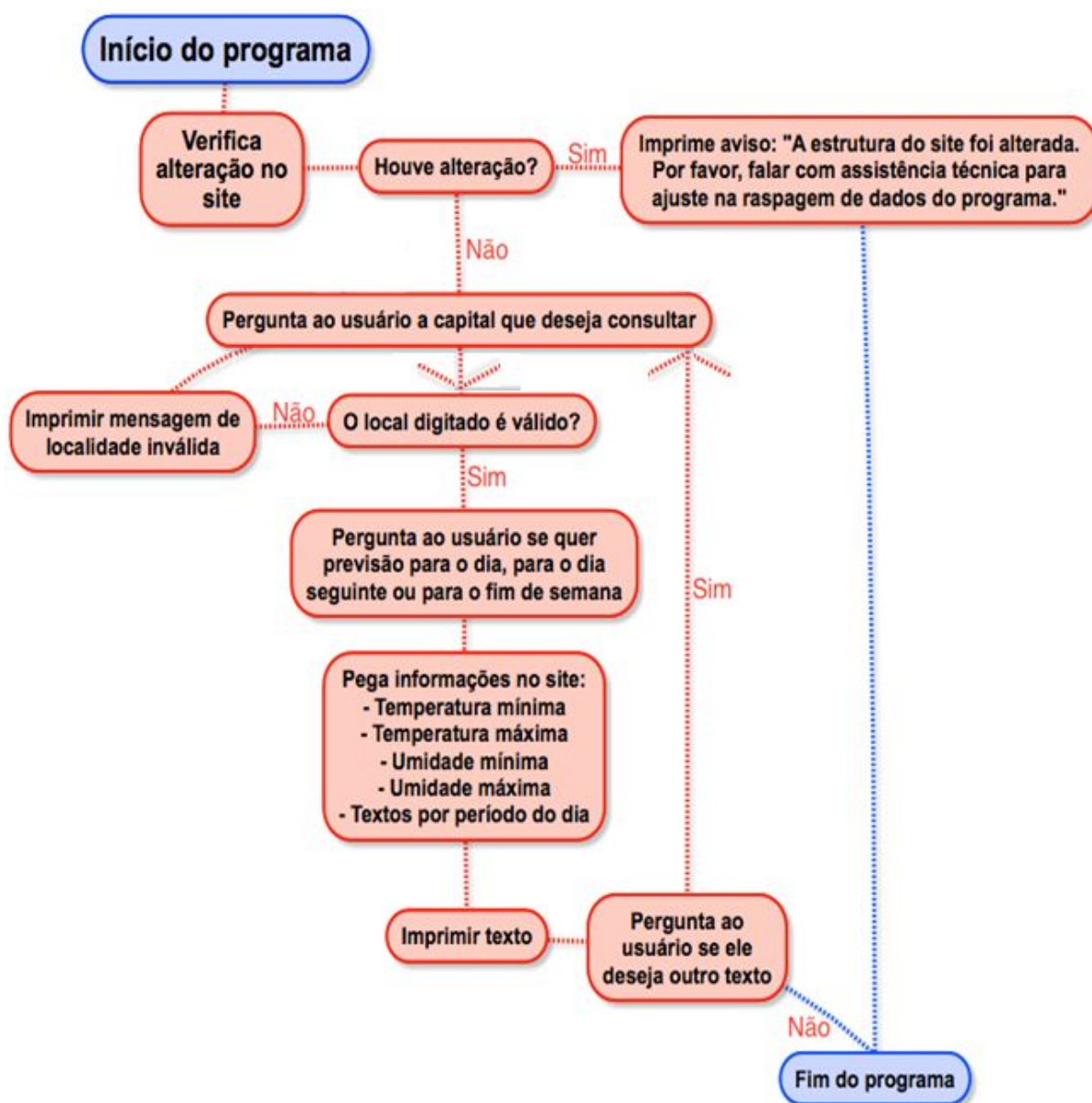


Figura 1.2: Lógica para construção do programa (Elaborado pela autora)

- Abrir uma página na internet;
- Saber se ela foi alterada ou não, de maneira que o código continua a raspagem ou transmite um aviso de que não é possível continuar a ação;
- Caso a página permaneça inalterada, pergunta qual informação deseja consultar;
- Ao receber a resposta, checa se é válida ou não, de maneira a continuar caso seja ou repetir a pergunta até obter resposta válida;
- Pergunta se o usuário deseja informação de ontem, hoje ou amanhã;
- Pega as informações do site, escolhe uma sequência de texto para impressão;
- Imprime o texto;
- Finaliza.

1.3.3 Implementação

As figuras abaixo são fotos da tela do algoritmo em anexo, escrito em Python3 e com a devida indentação. É importante perceber que são mais de 764 linhas de código e, portanto, as estruturas repetidas que se diferenciam da anterior apenas pelos dados não estão na imagem, mas podem ser vistas no anexo.

```

1  '''
2  Universidade de Brasilia, UnB
3  Faculdade de Comunicacao Social, FAC
4  Departamento de Jornalismo
5
6  Produto de Trabalho de Conclusao de Curso, TCC
7  Título: "JORNALISMO E COMPUTAÇÃO: Automacao do lide"
8
9  Aluna: Marilia Nestor Santos
10 Matricula: 11/0132254
11 Orientadora: Marcia Marques
12
13 Brasilia, 07 de fevereiro de 2017
14
15 **** Informacoes necessarias:
16 --- Este programa foi escrito para automatizar um lide de clima e tempo
17
18 --- Especificações tecnicas
19 Programa escrito em python3.5
20 Para rodar, foi utilizado o editor de texto Text Wrangler e o terminal do computador
21 O computador e um MacBook Pro mid 2012, que utiliza macOS Sierra 10.12.3
22
23 -- Variaveis utilizadas com exemplos:
24 nome_cidade = "Brasilia"
25 quando = "hoje"
26 Quando = "Hoje"
27 temp_min = "14"
28 temp_max = "28"
29 umidade_min = "55"
30 umidade_max = "70"
31 texto_manha = "Parcialmente nublado"
32 texto_tarde = "Parcialmente nublado a nublado com possibilidade de chuva em areas isoladas"
33 texto_noite = "Parcialmente nublado a nublado com possibilidade de chuva em areas isoladas"
34 '''

```


Figura 1.3: Linhas 1 a 34: comentários iniciais. (Elaborado pela autora)

```
36 import string
37 import datetime
38 import random
39 import doctest
40 import requests
41 import urllib
42 import urllib.request
43 import string
44 from bs4 import BeautifulSoup
45 from random import randint
```

Figura 1.4: Linhas 36 a 45: Bibliotecas (Elaborado pela autora)

```
47 def Grupo_1_1_IniciarLide (nome_cidade, quando, Quando):
48     aleatorio = (randint(0,4))
49     if aleatorio == 0:
50         print ("Em "+nome_cidade+" "+quando+" ", end = "")
51     elif aleatorio == 1:
52         print ("A previsao de "+quando+" para "+nome_cidade+" ", end = "")
53     elif aleatorio == 2:
54         print (Quando+" em "+nome_cidade+" ", end = "")
55     elif aleatorio == 3:
56         print (Quando+" "+nome_cidade+" tem expectativa de ", end = "")
57     elif aleatorio == 4:
58         print (Quando+" "+nome_cidade+" espera ", end = "")
59
60 def Grupo_1_2_no_fim_da_frase (nome_cidade, quando, Quando):
61     aleatorio = (randint(0,2))
62     if aleatorio == 0:
63         print ("em "+nome_cidade+" "+quando+". ", end = "")
64     elif aleatorio == 1:
65         print ("de "+nome_cidade+" "+quando+". ", end = "")
66     elif aleatorio == 2:
67         print (quando+" em "+nome_cidade+" ", end = "")
68
69 def Grupo_2_1_frase_completa (temp_min, temp_max):
70     aleatorio = (randint(0,4))
71     if aleatorio == 0:
72         print ("Temperatura minima prevista de "+temp_min+"°C e maxima de "+temp_max+"°C. ", end = "")
73     elif aleatorio == 1:
74         print ("Temperatura maxima de "+temp_max+"°C e temperatura minima de "+temp_min+"°C. ", end = "")
75     elif aleatorio == 2:
76         print ("Temperatura variando entre "+temp_min+"°C e "+temp_max+"°C. ", end = "")
77     elif aleatorio == 3:
78         print ("Maxima prevista de "+temp_max+"°C e minima de "+temp_min+"°C. ", end = "")
79     elif aleatorio == 4:
80         print ("Minima prevista de "+temp_min+"°C e maxima de "+temp_max+"°C. ", end = "")
81
82 def Grupo_2_2_inicio_frase (temp_min, temp_max):
83     aleatorio = (randint(0,4))
84     if aleatorio == 0:
85         print ("Temperatura minima prevista de "+temp_min+"°C e maxima de "+temp_max+"°C, ", end = "")
86     elif aleatorio == 1:
87         print ("Temperatura maxima de "+temp_max+"°C e temperatura minima de "+temp_min+"°C, ", end = "")
88     elif aleatorio == 2:
89         print ("Temperatura variando entre "+temp_min+"°C e "+temp_max+"°C, ", end = "")
90     elif aleatorio == 3:
91         print ("Maxima prevista de "+temp_max+"°C e minima de "+temp_min+"°C, ", end = "")
92     elif aleatorio == 4:
93         print ("Minima prevista de "+temp_min+"°C e maxima de "+temp_max+"°C, ", end = "")
```

Figura 1.5: Linhas 47 a 405: Funções de texto (Elaborado pela autora)

```
407 def FuncaoTemperatura (theurl): # Essa funcao busca todas as temperaturas da pagina
408
409     thepage = urllib.request.urlopen(theurl)
410     soup = BeautifulSoup(thepage, "html.parser")
411     lista = ["ola", "mundo", "aqui", "estamos", "te"]
412     i = 0
413     for temps in soup.findAll('div', {"id": "quadro1_interno_circulo_img"}):
414         lista[i] = temps.find('input')
415         i = i + 1
416
417     lista[0] = str(lista[0]); lista[1] = str(lista[1]); lista[2] = str(lista[2]); lista[3] = str(lista[3]); lista[4] = str(lista[4])
418     a = lista[0][114]; b = lista[0][115]
419     temp_min_hoje = a+b
420     a = lista[1][114]; b = lista[1][115]
421     temp_max_hoje = a+b
422     a = lista[2][114]; b = lista[2][115]
423     temp_min_amanha = a+b
424     a = lista[3][114]; b = lista[3][115]
425     temp_max_amanha = a+b
426
427     return (temp_min_hoje, temp_max_hoje, temp_min_amanha, temp_max_amanha)
```

Figura 1.6: Linhas 407 a 427: Função de temperatura (Elaborado pela autora)

```

429 def FuncaoUmidadeRelativa(theurl): # Essa funcao busca todas as temperaturas da pagina
430
431     thepage = urllib.request.urlopen(theurl)
432     soup = BeautifulSoup(thepage, "html.parser")
433     lista = ["ola", "mundo"]
434     i = 0
435     for temps in soup.findAll('div', {"id":"quadro1_interno_dados_txt_umidade_min"}):
436         lista[i] = temps.find('b')
437         i = i + 1
438
439     umidade_min_hoje = str(lista[0])
440     umidade_min_hoje = str(umidade_min_hoje)
441     umidade_min_hoje = umidade_min_hoje.replace("<b>", "") # reescreve a string obtida de uma maneira mais bonita
442     umidade_min_hoje = umidade_min_hoje.replace("</b>", "")
443     umidade_min_amanha = str(lista[1])
444     umidade_min_amanha = str(umidade_min_amanha)
445     umidade_min_amanha = umidade_min_amanha.replace("<b>", "") # reescreve a string obtida de uma maneira mais bonita
446     umidade_min_amanha = umidade_min_amanha.replace("</b>", "")
447
448
449     i = 0
450     for temps in soup.findAll('div', {"id":"quadro1_interno_dados_txt_umidade_max"}):
451         lista[i] = temps.find('b')
452         i = i + 1
453
454     umidade_max_hoje = str(lista[0])
455     umidade_max_hoje = str(umidade_max_hoje)
456     umidade_max_hoje = umidade_max_hoje.replace("<b>", "") # reescreve a string obtida de uma maneira mais bonita
457     umidade_max_hoje = umidade_max_hoje.replace("</b>", "")
458
459     umidade_max_amanha = str(lista[1])
460     umidade_max_amanha = str(umidade_max_amanha)
461     umidade_max_amanha = umidade_max_amanha.replace("<b>", "") # reescreve a string obtida de uma maneira mais bonita
462     umidade_max_amanha = umidade_max_amanha.replace("</b>", "")
463
464
465     return (umidade_min_hoje, umidade_max_hoje, umidade_min_amanha, umidade_max_amanha)

```

Figura 1.7: Linhas 429 a 465: Função de umidade relativa (Elaborado pela autora)

```

467 def RaspagemDeDados (capital_escolhida, quando):
468     if capital_escolhida == "1":
469         capital_escolhida = "ARACAJU - SE"
470         theurl = 'http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=2800308'
471     elif capital_escolhida == "2":
472         capital_escolhida = "BELEM - PA"
473         theurl = 'http://www.inmet.gov.br/portal/index.php?r=tempo2/verProximosDias&code=1501402'
474     elif capital_escolhida == "3":
475
476     elif capital_escolhida == "4":
477
478     elif capital_escolhida == "5":
479
480     elif capital_escolhida == "6":
481
482     elif capital_escolhida == "7":
483
484     elif capital_escolhida == "8":
485
486     elif capital_escolhida == "9":
487
488     elif capital_escolhida == "10":
489
490     elif capital_escolhida == "11":
491
492     elif capital_escolhida == "12":
493
494     elif capital_escolhida == "13":
495
496     elif capital_escolhida == "14":
497
498     elif capital_escolhida == "15":
499
500     elif capital_escolhida == "16":
501
502     elif capital_escolhida == "17":
503
504     elif capital_escolhida == "18":
505
506     elif capital_escolhida == "19":
507
508     elif capital_escolhida == "20":
509
510     elif capital_escolhida == "21":
511
512     elif capital_escolhida == "22":
513
514     elif capital_escolhida == "23":

```

Figura 1.8: Linhas 467 a 556: Raspagem de dados - parte 1 (Elaborado pela autora)

```

558 ▶ elif capital_escolhida == "24": ☐
561 ▶
562 ▶ elif capital_escolhida == "25": ☐
565 ▶
566 ▶ elif capital_escolhida == "26": ☐
569 ▶
570 ▶ elif capital_escolhida == "27": ☐
573 ▶
574 ▶
575 thepage = urllib.request.urlopen(theurl)
576 soup = BeautifulSoup(thepage, "html.parser")
577
578 if quando == "hoje":
579     #Umidade
580     umidade_minima, umidade_maxima, lixo1, lixo2 = FuncaoUmidadeRelativa(theurl)
581
582
583     #Quadro manha
584     texto_manha = soup.find('div', {"id": "quadro1_interno_manha_quadro_txt"}).find('i')
585     texto_manha = str(texto_manha)
586     texto_manha = texto_manha.replace("<i>", "") # reescreve a string obtida de uma maneira mais bonita
587     texto_manha = texto_manha.replace("</i>", "")
588
589
590     #Quadro Tarde
591     texto_tarde = soup.find('div', {"id": "quadro1_interno_tarde_quadro_txt2"}).find('i')
592     texto_tarde = str(texto_tarde)
593     texto_tarde = texto_tarde.replace("<i>", "") # reescreve a string obtida de uma maneira mais bonita
594     texto_tarde = texto_tarde.replace("</i>", "")
595
596
597     #Quadro Noite
598     texto_noite = soup.find('div', {"id": "quadro1_interno_noite_quadro_txt3"}).find('i')
599     texto_noite = str(texto_noite)
600     texto_noite = texto_noite.replace("<i>", "") # reescreve a string obtida de uma maneira mais bonita
601     texto_noite = texto_noite.replace("</i>", "")
602
603
604     #Quadro Temperatura maxima
605     temp_min, temp_max, lixo1, lixo2 = FuncaoTemperatura(theurl)
606
607 elif quando == "amanha":
608     #Umidade ☐
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637 ▶ return (capital_escolhida, temp_max, temp_min, texto_manha, texto_tarde, texto_noite, umidade_minima, umidade_maxima)

```

Figura 1.9: Linhas 558 a 637: Raspagem de dados - parte 2(Elaborado pela autora)

```

639 ▶ def Escolher_capital ():
640     print("Este programa informara a umidade, temperatura e o tempo ao longo do dia.")
641     print("Escolha uma das capitais a seguir pelo numero:")
642     print("1 - ARACAJU - SE")
643     print("2 - BELEM - PA")
644     print("3 - BELO HORIZONTE - MG")
645     print("4 - BOA VISTA - RR")
646     print("5 - BRASILIA - DF")
647     print("6 - CAMPO GRANDE - MS")
648     print("7 - CUIABA - MT")
649     print("8 - CURITIBA - PR")
650     print("9 - FLORIANOPOLIS - SC")
651     print("10 - FORTALEZA - CE")
652     print("11 - GOIANIA - GO")
653     print("12 - JOAO PESSOA - PB")
654     print("13 - MACAPA - AP")
655     print("14 - MACEIO - AL")
656     print("15 - MANAUS - AM")
657     print("16 - NATAL - RN")
658     print("17 - PALMAS - TO")
659     print("18 - PORTO ALEGRE - RS")
660     print("19 - PORTO VELHO - RO")
661     print("20 - RECIFE - PE")
662     print("21 - RIO BRANCO - AC")
663     print("22 - RIO DE JANEIRO - RJ")
664     print("23 - SALVADOR - BA")
665     print("24 - SAO LUIS - MA")
666     print("25 - SAO PAULO - SP")
667     print("26 - TERESINA - PI")
668     print("27 - VITORIA - ES")
669     capital_escolhida = input("Escolha uma das capitais a seguir pelo numero: ") # capital_escolhida deve ser um numero entre 1 e 27
670     quando = input("Qual dia(hoje ou amanha)? ")
671     if quando == "hoje":
672         Quando = "Hoje"
673     elif quando == "amanha":
674         Quando = "Amanha"
675
676     return (capital_escolhida, quando, Quando)

```

Figura 1.10: Linhas 639 a 676: Função para escolher a capital (Elaborado pela autora)

```

680 (qual_capital, quando, Quando) = Escolher_capital() #funcao que pede a cidade
681
682 qual_capital, temp_max, temp_min, texto_manha, texto_tarde, texto_noite, umidade_min, umidade_max = RaspagemDeDados(qual_capital, quando)

```

Figura 1.11: Linhas 680 a 682: Atribuição de valores (Elaborado pela autora)

```

685 um_ja_foi = 0
686 dois_ja_foi = 0
687 tres_ja_foi = 0
688 quatro_ja_foi = 0 #Variaveis usadas para indicar quais dados ja foram escolhidos na parte randomica
689 # grupo 1: nome da cidade
690 # grupo 2: temperatura maxima e minima
691 # grupo 3: umidade maxima e minima
692 # grupo 4: textos sobre o tempo da manha, tarde e noite
693
694 aleatorio = (randint(1, 4))
695
696 if aleatorio == 3: #Nao queremos que a informacao comece por umidade
697     aleatorio = 4
698
699
700 if aleatorio == 1:
701     um_ja_foi = 1
702     Grupo_1_1_IniciarLide(qual_capital, quando, Quando)
703
704 elif aleatorio == 2:
705     dois_ja_foi = 1
706     Grupo_2_2_inicio_frase(temp_min, temp_max)
707 elif aleatorio == 4:
708     quatro_ja_foi = 1
709     Grupo_4_2_inicio_de_lide(texto_manha, texto_tarde, texto_noite)
710
711 # Ja temos a primeira frase INCOMPLETA na tela e so temos mais TRES informacoes a serem usadas.
712
713 if um_ja_foi == 0: # o grupo 1 (nome da cidade) precisa sair ou na primeira ou na segunda frase
714     um_ja_foi = 1
715     Grupo_1_2_no_fim_da_frase(qual_capital, quando, Quando)
716 else: # aqui, o grupo 1 foi sorteado na variavel aleatoria. Precisamos agora completar a primeira frase
717     aleatorio = (randint(2, 4))
718
719     if aleatorio == 3:
720         aleatorio = 2 # novamente, nao queremos que umidade seja a primeira informacao
721
722     if aleatorio == 2:
723         dois_ja_foi = 1
724         Grupo_2_3_no_fim_de_frase(temp_min, temp_max)
725     elif aleatorio == 4:
726         quatro_ja_foi = 1
727         Grupo_4_3_fim_da_frase(texto_manha, texto_tarde, texto_noite)
728
729 # Ja temos a primeira frase COMPLETA na tela e so temos mais DUAS informacoes a serem usadas. (umidade e mais outra)
730
731 aleatorio = (randint(0, 3)) # usamos outras variavel aleatoria para randomizar as duas ultimas informacoes
732 if aleatorio == 0: # se der zero, usaremos duas frases completas pra finalizar a previsao do tempo
733     Grupo_3_1_frase_completa(umidade_min, umidade_max)

```

Figura 1.12: Linhas 685 a 761: Ordem do texto (Elaborado pela autora)

```

762 print("")
763 print("Fonte: INMET")
764 print("")

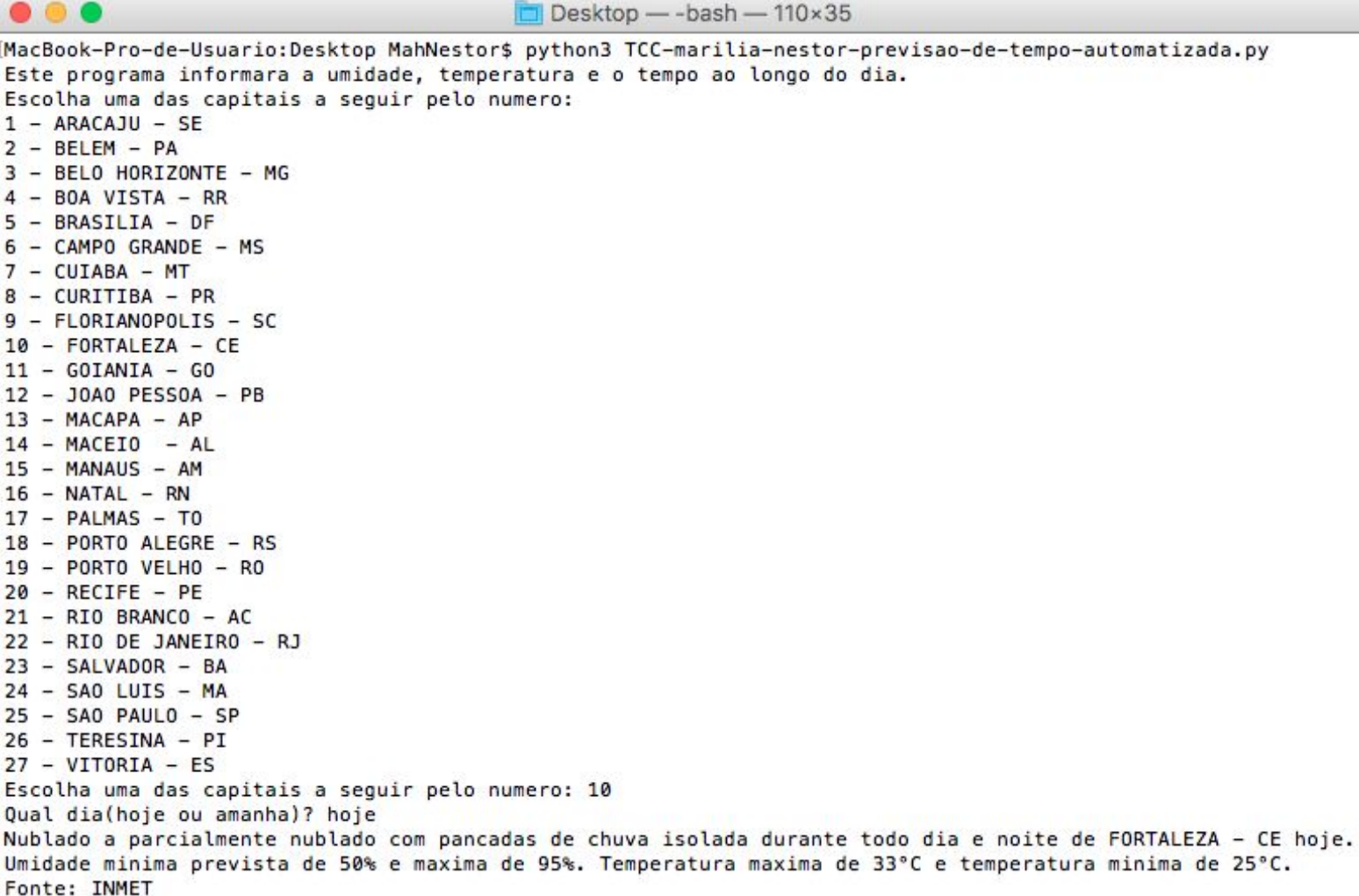
```

Figura 1.13: Linhas 762 a 764: Fonte e fim do programa (Elaborado pela autora)

1.3.4 Análise do resultado

Apesar da pouca elaboração estética visual, o algoritmo atingiu o objetivo final.

As figuras abaixo são fotos da tela do computador com resultado do programa:



```
Desktop — -bash — 110x35
[MacBook-Pro-de-Usuario:Desktop MahNestor$ python3 TCC-marilia-nestor-previsao-de-tempo-automatizada.py
Este programa informara a umidade, temperatura e o tempo ao longo do dia.
Escolha uma das capitais a seguir pelo numero:
1 - ARACAJU - SE
2 - BELEM - PA
3 - BELO HORIZONTE - MG
4 - BOA VISTA - RR
5 - BRASILIA - DF
6 - CAMPO GRANDE - MS
7 - CUIABA - MT
8 - CURITIBA - PR
9 - FLORIANOPOLIS - SC
10 - FORTALEZA - CE
11 - GOIANIA - GO
12 - JOAO PESSOA - PB
13 - MACAPA - AP
14 - MACEIO - AL
15 - MANAUS - AM
16 - NATAL - RN
17 - PALMAS - TO
18 - PORTO ALEGRE - RS
19 - PORTO VELHO - RO
20 - RECIFE - PE
21 - RIO BRANCO - AC
22 - RIO DE JANEIRO - RJ
23 - SALVADOR - BA
24 - SAO LUIS - MA
25 - SAO PAULO - SP
26 - TERESINA - PI
27 - VITORIA - ES
Escolha uma das capitais a seguir pelo numero: 10
Qual dia(hoje ou amanha)? hoje
Nublado a parcialmente nublado com pancadas de chuva isolada durante todo dia e noite de FORTALEZA - CE hoje.
Umidade minima prevista de 50% e maxima de 95%. Temperatura maxima de 33°C e temperatura minima de 25°C.
Fonte: INMET
```

Figura 1.14: Foto da tela para resultado do programa: Lide da previsão do tempo para Fortaleza-CE 'hoje' (Elaborado pela autora)

```
Desktop — -bash — 107x35
MacBook-Pro-de-Usuario:Desktop MahNestor$ python3 TCC-marilia-nestor-previsao-de-tempo-automatizada.py
Este programa informara a umidade, temperatura e o tempo ao longo do dia.
Escolha uma das capitais a seguir pelo numero:
1 - ARACAJU - SE
2 - BELEM - PA
3 - BELO HORIZONTE - MG
4 - BOA VISTA - RR
5 - BRASILIA - DF
6 - CAMPO GRANDE - MS
7 - CUIABA - MT
8 - CURITIBA - PR
9 - FLORIANOPOLIS - SC
10 - FORTALEZA - CE
11 - GOIANIA - GO
12 - JOAO PESSOA - PB
13 - MACAPA - AP
14 - MACEIO - AL
15 - MANAUS - AM
16 - NATAL - RN
17 - PALMAS - TO
18 - PORTO ALEGRE - RS
19 - PORTO VELHO - RO
20 - RECIFE - PE
21 - RIO BRANCO - AC
22 - RIO DE JANEIRO - RJ
23 - SALVADOR - BA
24 - SAO LUIS - MA
25 - SAO PAULO - SP
26 - TERESINA - PI
27 - VITORIA - ES
Escolha uma das capitais a seguir pelo numero: 3
Qual dia(hoje ou amanha)? amanha
Minima prevista de 21°C e maxima de 24°C, de BELO HORIZONTE - MG amanha. Umidade prevista entre 85% e 96%.
Encoberto com chuvisco durante todo dia e noite.
Fonte: INMET
```

Figura 1.15: Foto da tela para resultado do programa: Lide da previsão do tempo para Belo Horizonte - MG 'amanhã' (Elaborado pela autora)

2. Problema de Pesquisa

Se antes a língua inglesa foi um diferencial na formação intelectual do indivíduo, hoje ela está na base curricular das escolas brasileiras. A tendência é que o mesmo aconteça com o ensino de computação. Com o movimento de aceleração que a internet deu à globalização, a popularidade do uso de computadores e dispositivos que acessam a rede aumentou e cada vez mais cedo e com mais frequência ocorre o contato 'ser humano-máquina'.

Esse movimento afetou o jornalismo com décadas de mudanças constantes nas redações motivadas pela internet e pelas novas tecnologias. Mesmo enquanto matéria prima do jornalismo, a informação é o resultado de processamento de dados. Portanto, vê-se necessária a literacia de dados¹² por parte de jornalistas para o aperfeiçoamento das abordagens à informação, a recuperação da credibilidade dos veículos e a solução para o financiamento do jornalismo. Nesse momento de transição forçada, diversos jornais de adaptaram, padeceram ou agonizam com a resistência ao novo.

Dentre as práticas para uso de dados adotadas pelo jornalismo, esta problematização foca no jornalismo automatizado, que se refere a processos algorítmicos que convertem dados em textos informativos narrativos limitados a nenhuma intervenção humana além das escolhas iniciais de programação. Dentro desse tema, vêm as inquietações trazidas pelo "Guia do Jornalismo Automatizado" ("*Guide to Automated Journalism*", tradução livre), em que o autor fala:

- "O jornalismo humano e automatizado provavelmente se tornarão estreitamente integrados e formarão um 'casamento homem-máquina.'"¹³
- "É melhor aconselhar os jornalistas a desenvolver habilidades que os algoritmos não podem realizar, como análises aprofundadas, entrevistas e relatórios investigativos."¹⁴
- "O jornalismo automatizado deve substituir jornalistas que apenas cobrem assuntos de rotina, mas também gerará novos empregos no desenvolvimento de algoritmos geradores de notícias."¹⁵

Este trabalho, portanto, dá luz às novas possibilidades para um terreno expansivo de conteúdo de notícias que excede em muito as capacidades de produção de jornalistas humanos.

¹² Literacia de dados é a capacidade de ler, criar e comunicar dados como informações.

¹³ "*Human and automated journalism will likely become closely integrated and form a 'man-machine marriage.'*"

¹⁴ "*Journalists are best advised to develop skills that algorithms cannot perform, such as in-depth analysis, interviewing, and investigative reporting.*"

¹⁵ "*Automated journalism will likely replace journalists who merely cover routine topics, but will also generate new jobs within the development of news-generating algorithms.*"

3. Justificativa

Quando a grande recorrência das palavras ‘crise’, ‘desemprego’ e ‘passaralho’ se encontram com o fim da graduação, o medo é natural. Após pesquisas sobre as tendências do jornalismo na próxima década, encontrei alguns caminhos interessantes, como o *crowdfunding* e o empreendedorismo, mas me identifiquei mais com uma ferramenta: a computação.

Comecei a pesquisar e percebi a falta de literatura acadêmica no Brasil acerca do assunto. Nos primeiros meses de pesquisa, achei poucos artigos e teses nacionais, enquanto a discussão ganha cada vez mais espaço fora do território nacional, especialmente nos Estados Unidos e no Reino Unido, onde é possível encontrar graduações que unem o jornalismo à ciência da computação.

Contudo, o que uma das observações que me assustou foi a falta de uso por parte dos veículos brasileiros, restrito ao grandes, como os do Grupo Globo¹⁶, ou específicos, como o Nexo¹⁷. Por mais que fossem necessários recursos para inserir essas novas possibilidades, ainda não entendia o motivo de não haver mais empresas que as usassem.

Entendi quando comecei a conversar sobre esse assunto na faculdade com professores e outros alunos. Se o meu medo foi a faísca para começar a estudar, a resistência encontrada foi o combustível. O que encontrei foi muito parecido com a descrição feita por Berret e Phillips no livro *“Teaching Data and Computational Journalism”* (“Ensinando Jornalismo de Dados e Computacional”, tradução livre), quando pontuam como identificar os assuntos a ensinar aos alunos em uma espécie de guia para professores em sala de aula. Eles mencionam: a programação de computadores, os recursos, a proficiência no assunto e o compromisso dos estudantes:

“A resistência à matemática é muito mais ampla do que o jornalismo, mas terá de ser abordado se o ensino de jornalismo de dados for levado a sério. Isso se aplica a professores e alunos, alguns dos quais podem ter optado pelo jornalismo, em parte, porque pensaram que iria exigir pouca ou nenhuma matemática.”¹⁸ (Berret e Phillips, 2016, p. 46)

A ideia de explorar a área é ser um exemplo para mostrar as possibilidades para alunos, professores e profissionais do jornalismo, além de instigar a discussão acadêmica acerca das novas técnicas e tecnologias no país. Mesmo com a ascensão do jornalismo de dados, essa tarefa acaba nas mãos de cientistas da computação alheios à área que percebem a necessidade do serviço.

¹⁶ <<http://www.grupoglobo.globo.com/>> Acesso em 03 fev 2017

¹⁷ <<https://www.nexojornal.com.br/>> Acesso em 03 fev 2017

¹⁸ “Resistance to math is an issue far broader than the field of journalism, but it will need to be addressed if teaching data journalism is to be taken seriously. This applies to both teachers and students, some of whom may have chosen to pursue journalism in part because they thought that it would require little or no math.”

As ferramentas da computação no jornalismo podem ser auxiliares no momento como encontrar novas pautas, clarear informações, facilitar o *fact checking*, auxiliar na batalha contra notícias falsas, poupar profissionais de tarefas repetitivas e incrementar a maneira de contar histórias (Lecomte, 2015; Diakopoulos, 2013). Em resumo, pode ser uma auxiliar do jornalismo ético e auxiliar da democracia, chamado bom jornalismo, cada vez mais caro aos veículos de comunicação.

Acredito que o bom jornalismo não tem prazo de validade. Creio também que nem todo jornalista precisa aprender a programar. A boa pauta, a história escrita com ética e qualidade não morre, o mundo tem necessidade dela independente da velocidade da modernização, mas a relutância com novas tecnologias leva a uma 'espiral da resistência' e contribui para a apropriação alheia do nosso ramo.

4. Objetivo de Pesquisa

4.1 Objetivo Geral:

Mostrar como jornalistas podem usar a automação

4.2 Objetivos Específicos

- Estruturar e desenvolver um programa de computador para produção automatizada de lides de clima e tempo.
- Contribuir com a produção acadêmica de jornalismo computacional no Brasil.

5. Referencial Teórico

5.1 Computação no jornalismo

Dentro do que envolve o campo do jornalismo computacional, conceituado na apresentação deste trabalho, aqui se trata das mudanças em como histórias são encontradas e apresentadas (Cohen; Hamilton; Turner, 2011, p.1). Um bom exemplo desse tipo de aplicação é o Dollars for Docs¹⁹, uma base de dados que acabou sendo uma solução diferente da ProPublica²⁰, uma redação estadunidense independente, sem fins lucrativos que produz jornalismo investigativo de interesse público.

Os repórteres buscaram de onde vinha o dinheiro para os médicos do país, coletaram os dados durante os anos 2009 e 2013 e fizeram uma página na internet para caso o cidadão queira pesquisar pelo seu médico e apurar que empresas farmacêuticas pagam para que ele realize palestras, por exemplo, já que esse pode ser o motivo de ele receitar aquele remédio mais caro e não outra opção mais viável financeiramente.

Essa solução acabou sendo mais efetiva e prática do que se fosse feita uma grande reportagem em jornal impresso, por exemplo, uma vez que nem todos os médicos seriam contemplados pelo texto e não seria tão prático e funcional ao cidadão. São feitas atualizações periódicas da informação por meio de dados abertos.

Dados abertos são informações que podem ser acessadas legalmente por qualquer indivíduo. A definição aceita pelo governo brasileiro é para "quando qualquer pessoa possa livremente usá-los, reutilizá-los e redistribuí-los, estando sujeito a, no máximo, a exigência de creditar sua autoria e compartilhar pela mesma licença"²¹. Segundo a própria fonte, essa é uma adaptação do conceito da *Open Knowledge Foundation*²².

Esse ponto é importante porque é necessário que a captação da informação esteja dentro da lei, uma vez que não é difícil para uma pessoa com maior conhecimento sobre programação quebrar as barreiras legais, inclusive se forem jornalistas, o que pode ter consequências jurídicas.

Depois dessa conceituação, torna-se interessante explicar também o que é *big data*, uma vez que esclarecer isso facilita a compreensão da necessidade da literacia de dados na redação. De acordo com um estudo feito por Martha Stone para o Instituto Reuters para estudos de Jornalismo, *Big Data* é um termo para "uma variedade de estratégias e táticas que envolvem conjuntos de dados maciços, e tecnologias que fazem sentido no meio dessa resma de bloqueadores de

¹⁹ <<https://projects.propublica.org/docdollars/>> Acesso em 05 fev 2017

²⁰ <<https://www.propublica.org/>> Acesso 05 fev 2017

²¹ <<http://dados.gov.br/dados-abertos/>> Acesso em 28 jan 2017

²² <<http://opendefinition.org>> Acesso em 28 jan 2017

significação de dados."²³ (2014, p.1) Na imagem abaixo, é possível ter uma ideia do motivo pelo qual se chama *big data* e suas características:

Defining Big Data

The 3Vs that define *Big Data* are Variety, Velocity and Volume.

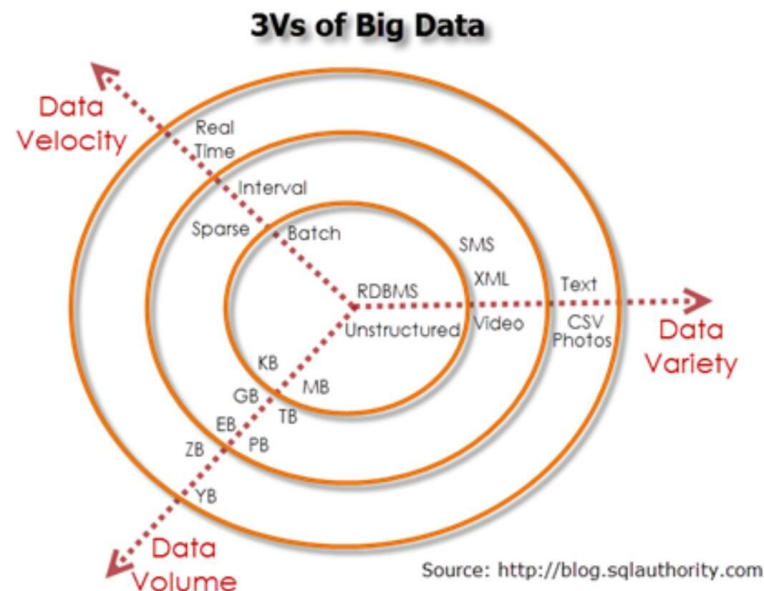


Figura 5.1: Características do Big Data

Fonte: Aplicativo "Get started with Big Data"

A imagem mostra a expansão dos 3 Vs:

- Variedade: expressa a quantidade de conteúdo e formatos para os dados.
- Velocidade: demonstra como eles devem ser analisados rápido, especialmente no caso de notícias
- Volume: demonstra como não é possível se guardar dados inúteis, uma vez que espaço para dados pode sugar bastante recurso.

A essa lógica, Stone adiciona um 'V' importante para a notícia: o valor. Para ela, os dados precisam ter potencial jornalístico ou empresarial. Para ressaltar a importância desses pontos, pode ser esclarecida também a diferença entre *Little* e *Big Data*, sendo a primeira medida até *gigabytes*, e a segunda em *terabytes*, *petabytes*, *zettabytes* para cima, tornando-se inviável a um computador pessoal.

Com essa quantidade de conteúdo, é necessário que as empresas de mídia desenvolvam um banco de dados estruturado de vídeo, áudio, fotografia, textos para as diferentes mídias, estatísticas de audiência, entre outros, de maneira a tornar o que já têm rapidamente acessível e a facilitar a entrada de novos dados.

²³ "Big Data is an umbrella term for a variety of strategies and tactics that involve massive data sets, and technologies that make sense out of these mind boggling reams of data."

A automação pode ser grande simplificadora nesse processo, em especial na captura, análise e visualização de dados. Não apenas para o banco de dados interno, mas para *open data* em rede e análises de audiência que auxiliem a entender o público e, dessa maneira, direcionar melhor o conteúdo a diferentes meios, como Stone explica.

Assim, compreende-se que o jornalismo precisa de *open data*, mas "*Does Open Data Need Journalism?*" ("Dados abertos precisam do jornalismo?", tradução livre). Esse foi o artigo de 2015 escrito por Jonathan Stoneman após fazer estudos de caso no Quênia, Estados Unidos e Reino Unido, comparando aos locais em que dados abertos não foi uma ideia que decolou: Croácia e Burquina Faso.

Como conclusão, ele diz que a ideia de dados abertos deve ser diferenciada da liberdade de requisitar informações. Isso porque os dados disponíveis nem sempre são transparentes e completos, tornando necessário que alguém os utilize para dar significado a eles (p. 21), como, por exemplo, explicar que o número de homicídios diminuiu 30% de um ano para outro, em vez de apenas falar que ocorreram 55.

Jornalistas poderiam também pressionar governos por mais abertura e transparência, uma vez que o autor observa ser mais fácil que procurem uma pauta quando requerem um conteúdo pela lei de acesso à informação de cada país do que pelos portais do governo. (p. 22)

Além disso, é levantada a questão de que o jornalista precisa traduzir informações de um grupo para outro, já que elas são acessadas pelos cidadãos apenas quando sentem a necessidade, independente de serem importantes ou não. (p. 22) Em suma, dados precisam de visibilidade e o jornalismo está aí para isso, já que não adianta eles serem publicados para ficarem inutilizados e invisíveis.

5.1.1 O Jornalista-programador

Por que programar? Pode ser citado novamente o caso de Meredith Broussard, mas existe uma infinidade de *softwares* e programas para auxiliarem os jornalistas a captarem e publicarem dados sem a necessidade de conhecimento computacional. No entanto, a necessidade vem da mentalidade e da independência que a lógica de linguagens de programação empresta a quem a estuda e a que a mentalidade do jornalismo retorna às ferramentas da computação.

"Imagine uma fluência computacional e estética semelhante à melhor escrita do jornalista"²⁴, sugere um texto²⁵ feito por um dos organizadores do Simpósio Computação + Jornalismo 2013²⁶ (*Computation+Journalism Symposium 2013*), Nick

²⁴ "*Imagine a computational fluency and aesthetic akin to the best journalist writing*"

²⁵ <<http://www.niemanlab.org/2013/02/finding-tools-vs-making-tools-discovering-common-ground-between-computer-science-and-journalism/>> Acesso em 03 fev 2017

²⁶ <<http://computation-and-journalism.com/symposium2013/>> Acesso em 04 fev 2017

Diakopoulos. Ele toca nessa questão fazendo analogia com ferrageiros e forjadores de ferramentas.

Seguindo o dito popular brasileiro "Não dê o peixe, ensine a pescar", Diakopoulos afirma que as ferramentas existentes, feitas por pessoas de fora do jornalismo, apenas resolvem problemas imediatos e superficiais, enquanto quem fabrica o próprio instrumento tende a ter um mecanismo mais apropriado à sua necessidade.

O jornalista reforça que não há problema que as ferramentas existentes sejam usadas, já elas foram feitas para isso mesmo, mas insiste que "Os jornalistas pensam em termos de histórias, os tecnólogos pensam em termos de produtos. As histórias são específicas, mas os produtos são generalizáveis."²⁷

O movimento inicial da programação nas redações foi de fora para dentro, quando foram contratadas pessoas para trabalhar com as novas tecnologias. Contudo, após algum tempo, jornalistas que já possuíam experiência com CAR²⁸ começaram a se inteirar dessas novas linguagens de programação. (Parasie et al, p. 861)

O The New York Times²⁹, jornal estadunidense fundado em 1851, serve de exemplo hoje a outros jornais no mundo como modelo de adequação aos novos meios. Enquanto tantos outros veículos ainda não acompanhavam o desenvolvimento das mídias sociais, ele criou uma área apenas para notícias interativas no fim dos anos 2000, como explica Cindy Royal em seu artigo "*The journalist as programmer: A case study of the New York Times interactive news technology department*" ("O jornalista como programador: Um estudo de caso do departamento de tecnologia em notícias interativas", tradução livre):

"As raízes da proposta foram os CAR, reportagens assistidas por computador, mas a visão era reduzir a burocracia associada à criação de cada projeto e elevar o status dos codificadores ao trabalhar com repórteres e editores. O departamento exigiria um grupo de pessoas especiais, que fossem tão fluentes em jornalismo quanto em codificação, que pudessem entender as necessidades editoriais, desenvolver a funcionalidade e envolver os usuários. (...) suas responsabilidades reais são muito mais amplas e abrangem uma ênfase editorial mais forte do que os nomes indicam."³⁰ (p. 10)

²⁷ "Journalists think in terms of stories, technologists think in terms of products. Stories are specific but products are generalizable."

²⁸ Computer assisted reporting

²⁹ <<https://www.nytimes.com/>> Acesso em 04 fev 2017

³⁰ "The roots of the proposal were in computer-assisted reporting, but the vision was to reduce the bureaucracy associated with creating each project and to elevate the status of the coders in working with reporters and editors. The department would require a group of special people, those who were as fluent in journalism as they were in coding, who could understand the editorial needs, develop the functionality and engage users. (...) their

No estudo, a autora ainda explica que o time foi formado, em sua maioria, por autodidatas de outras áreas, já que a paixão pelo jornalismo era uma característica buscada e esse não é um traço comum em engenheiros de *software* ou cientistas da computação. (p.12)

Isso é especialmente interessante no jornalismo porque estudos sugerem que estudantes de ciências da computação e matemática possuem dominância do lado esquerdo do cérebro, enquanto os de música, arte, comunicação oral e jornalismo mostraram dominância do lado direito do cérebro. (Monfort et al, p. 170)

Ainda assim, um estudo mais detalhado feito posteriormente explica que essa dominância serve apenas para linguagens procedurais, como COBOL³¹, enquanto linguagens de orientação a objeto, como Python e Java³², são bem absorvidas por ambos os lados do cérebro. Existe até a facilidade de linguagens de script, como HTML³³ e XLM³⁴, terem mais compatibilidade com o lado direito. (White et al, p 63)

5.2 Automação no jornalismo

Jornalismo automatizado "pode ser resumido pelo fato de que hoje parte do conteúdo jornalístico publicado já não é mais escrito por humanos e sim por máquinas, via software, através de ferramentas e soluções que envolvem desde simples listas de palavras até complexos modelos de inteligência artificial". A conceituação do professor Márcio Carneiro se completa pelo questionamento do que a desconexão de texto e jornalista pode acarretar. (2016, p. 166)

Dois pontos importantes de serem ressaltados, mas que não serão detalhados aqui, é que redações costumam utilizar softwares de processamento de linguagem natural para produção das notícias, como Narrative Science³⁵, utilizado pela ProPublica³⁶, e Automated Insights³⁷, utilizado pela AP³⁸, por exemplo (Lecompte, 2015). No entanto, isso não será utilizado neste trabalho.

O outro ponto é que os processos automatizados não se referem apenas a textos, mas também à edição de áudio e vídeo, alertas, análises, personalização de produtos e outros processos repetitivos. Contudo, sendo o foco neste momento a notícia textual em linguagem natural, a figura abaixo pode esclarecer como o processo costuma funcionar para um algoritmo gerar notícias:

actual responsibilities are much broader and encompass a stronger editorial emphasis than the names indicate."

³¹ "Common Business Oriented Language", linguagem orientada para processamento de banco de dados comerciais.

³² <<https://www.oracle.com/br/java/index.html>> Acesso em 03 fev 2017

³³ <http://www.w3schools.com/html/html_intro.asp> Acesso em 03 fev 2017

³⁴ <<https://www.w3.org/XML/>> Acesso em 03 fev 2017

³⁵ <<https://www.narrativescience.com/>> Acesso em 03 fev 2017

³⁶ <<https://www.propublica.org/>> Acesso em 03 fev 2017

³⁷ <<https://automatedinsights.com/>> Acesso em 03 fev 2017

³⁸ <<https://www.ap.org/en-us/>> Acesso em 03 fev 2017



Figura 5.2: Como algoritmos geram notícias

Fonte: "Guide to automated journalism", Andreas Graefe

Tradução livre:

1. Coleta dados
 - Fontes pré definidas e/ou mineração de dados
 - Dados novos ou históricos, dados de contextualização
2. Identifica eventos interessantes
 - Análise estatística (anexos, tendências, correlações)
 - Regras pré definidas por assunto (por ex., corridas por time determinam um campeão)
3. Prioriza visões
 - Valor notícia
 - Regras pré definidas (por ex., resultado é mais importante que lesões)
4. Gera a narrativa
 - Especificação de estrutura e *storyline*
 - Guia de estilo oficial do veículo
5. Publica a história
 - Por meio do sistema de gerenciamento de conteúdo do veículo
 - Automaticamente ou após revisão editorial

Seguindo cada especificação citada na figura acima, algoritmos são capazes de criar histórias para um mesmo assunto de maneira rápida e com menos erros do

que um jornalista humano. É válido frisar que algoritmos não erram, eles seguem exatamente os comandos e as ordens dadas. Quando o resultado não foi o esperado, quem errou foi o programador.

Lecompte (2015), resolveu conversar com editores, jornalistas e empresas que aderiram à automação no jornalismo para saber sobre o choque, a experiência e os resultados alcançados. O fruto foi um texto³⁹ em que expõe esses diálogos e chega a conclusões como: "Empresas de notícias menores poderiam, na verdade, ser as mais beneficiadas pelo jornalismo automatizado. Elas estão bem posicionadas para utilizar dados locais e construir histórias."⁴⁰ (tradução livre) Para ela, a automação pode ser utilizada para aumentar a cobertura e a quantidade de produtos de pequenas organizações.

Os temores que esse desenvolvimento pode causar a jornalistas foram ainda mais justificados em 2015, quando o NYT⁴¹ publicou um jogo⁴² com o título "Um humano ou um computador escreveu isso?"⁴³ (tradução livre). Apesar de estar em inglês, esse jogo demonstra bem como já se percebe a dificuldade de distinção de ambos os tipos de texto, o que pode assustar qualquer jornalista.

Para acalmar os profissionais da notícia, Graefe (2016) aponta dois lados da automação. Segundo ele, além das questões citadas acima, algoritmos também podem usar os mesmos dados para contar histórias em diferentes línguas com diferentes pontos de vista e têm o potencial para criar notícias por demanda, quando um usuário fizer perguntas sobre os dados.

No entanto, há certas limitações a serem levadas em consideração, como o fato do algoritmo depender de dados e suposições, sujeitos a manipulação e erros. Além disso, um código não sabe tomar decisões, fazer perguntas, explicar novos fenômenos ou estabelecer causalidade sozinho, o que limita a capacidade de observar a sociedade. Por fim, mesmo com o exposto pelo jogo do NYT, a qualidade da escrita é inferior à humana, apesar de isso ser menos verdade à medida que as tecnologias avançam.

³⁹ <<http://niemanreports.org/articles/automation-in-the-newsroom/>> Acessado em 04 fev 2017

⁴⁰ "Small news organizations could, in fact, have the most to gain from automated journalism. They are well-positioned to draw on local data to write stories"

⁴¹ <<https://www.nytimes.com/>> Acesso em 03 fev 2017

⁴²<https://www.nytimes.com/interactive/2015/03/08/opinion/sunday/algorithm-human-quiz.html?smid=pl-share&_r=0> Acesso em 03 fev 2017

⁴³ "Did a Human or a Computer Write This?"

5.2.1 Automação do lide

Tacyana Arce, quando escreveu sobre o lide automatizado como "possibilidade de tratamento da informação para o jornalismo impresso diário" (2009), concluiu como uma maneira suprimir a necessidade de aceleração do processo de produção de notícias e possibilitar melhor tratamento das matérias produzidas pelos próprios jornais:

"Não sendo preocupação atual das empresas jornalísticas (melhor dizendo, sendo uma prioridade que se deixa de lado diante da dificuldade de se tratar, literalmente da noite para o dia, uma imensa produção diária de notícias), a ausência de um adequado tratamento da informação cria dificuldades para os jornalistas, que não conseguem acessar bancos de dados de matérias produzidas anteriormente e, com isso, sofrem um grande desgaste toda vez que precisam recorrer a produções antigas para produzir novo material." (2009, p. 16)

A prática da automação de lide não é nova. Isso pode ser facilmente provado por um artigo escrito há mais de quatro décadas que, não por coincidência, trata exatamente do mesmo assunto deste trabalho: a automação de lide em previsão do tempo. Chamado "*Computer-produced weather forecasts*" ("Previsões meteorológicas produzidas por computador", tradução livre), foi escrito por Harry Glahn em 1970 para apresentar exemplos de previsões meteorológicas redigidas inteiramente por computador, com o objetivo de serem usados no Centro Meteorológico Nacional da ESSA Meteorologia.

MAR 6, 1970

TDL EXPERIMENTAL FORECASTS

GOOD MORNING. THE SYSTEMS DEVELOPMENT OFFICE BRINGS YOU THE LATEST FORECAST FOR WASHINGTON, D. C. AND VICINITY. STRONG NORTHERLY WINDS THIS MORNING 20 MPH WITH GUSTS TO 30 MPH BECOMING NORTHERLY 5 MPH BY EVENING. COLDER TODAY, MAXIMUM TEMPERATURE 48 DEGREES. MOSTLY CLOUDY THIS MORNING WITH DECREASING CLOUDINESS THIS AFTERNOON. ONLY 2 PERCENT PROBABILITY OF PRECIPITATION TODAY.

GOOD MORNING. THE ESSA WEATHER BUREAU BRINGS YOU THE LATEST FORECAST FOR ATLANTA AND VICINITY. MOSTLY SUNNY THIS MORNING WITH A FEW MORE CLOUDS THIS AFTERNOON. COLDER TODAY, MAXIMUM TEMPERATURE 62 DEGREES. NORTHERLY WINDS 15 MPH THIS MORNING BECOMING NORTHEASTERLY 5 MPH BY LATE AFTERNOON. ONLY 2 PERCENT PROBABILITY OF PRECIPITATION TODAY.

GOOD MORNING. THE TECHNIQUES DEVELOPMENT LABORATORY BRINGS YOU THE LATEST FORECAST FOR ST. LOUIS AND VICINITY. PARTLY CLOUDY WITH LITTLE CHANGE IN TEMPERATURE TODAY, HIGH OF 58 DEGREES. EASTERLY WINDS 5 MPH THIS MORNING BECOMING NORTHEASTERLY 10 MPH BY LATE AFTERNOON. ONLY 2 PERCENT PROBABILITY OF PRECIPITATION TODAY.

GOOD MORNING. THE SYSTEMS DEVELOPMENT OFFICE BRINGS YOU THE LATEST FORECAST FOR NEW YORK JFK AND VICINITY. PARTLY CLOUDY AND COLDER TODAY, MAXIMUM TEMPERATURE 43 DEGREES. NORTHERLY WINDS 15 MPH, WITH ONLY 5 PERCENT PROBABILITY OF RAIN AND 5 PERCENT PROBABILITY OF SNOW.

FIG. 1. Automated forecast for 6 March 1970.

Figura 5.3: Previsões do tempo automatizadas do dia 6 de março de 1970

Fonte: "Computer-produced worded forecasts", Glahn. *Bull. Amer. Meteor. Soc.*, vol. 51, p. 1127

GOOD MORNING. THE SYSTEMS DEVELOPMENT OFFICE BRINGS YOU THE LATEST FORECAST FOR WASHINGTON, D. C. AND VICINITY. STRONG NORTHERLY WINDS THIS MORNING 20 MPH WITH GUSTS TO 30 MPH BECOMING NORTHERLY 5 MPH BY EVENING. COLDER TODAY, MAXIMUM TEMPERATURE 48 DEGREES. MOSTLY CLOUDY THIS MORNING WITH DECREASING CLOUDINESS THIS AFTERNOON. ONLY 2 PERCENT PROBABILITY OF PRECIPITATION TODAY.

Figura 5.4: Extração de um dos textos da figura anterior

Fonte: "Computer-produced worded forecasts", Glahn. *Bull. Amer. Meteor. Soc.*, vol. 51, p. 1127

Tradução livre: "Bom dia. O escritório de desenvolvimento de sistemas traz para você as últimas previsões do tempo para Washington D. C e Vicinity. Ventos do norte fortes nesta manhã 20 a 30 milhas por hora com rajadas, passando para 5 milhas por hora no fim da tarde. Mais frio hoje, máxima de 48 F. Manhã

predominantemente nublada com diminuição de nuvens pela tarde. Apenas 2 por cento de possibilidade de precipitação hoje."

A primeira figura está ali para mostrar que as diferentes possibilidades de textos para os dados também não é nova. Nessa última figura, o primeiro texto da figura anterior, são identificadas as mesmas perguntas feitas no produto deste trabalho, também chamado lide jornalístico (JORGE, 2008. p. 132):

- Quem: Washington D. C e Vicinity
- Onde: Washington D. C e Vicinity
- O quê: Dados
- Quando: Hoje
- Como: pergunta ignorada graças à natureza do conteúdo.
- Por quê: pergunta ignorada graças à natureza do conteúdo.

A diferença do exibido nessas imagens para este trabalho é a apropriação do jornalista à técnica e, portanto, a visão do jornalismo ao produto final, como a mentalidade de que fala a professora Thais de Mendonça Jorge no "Manual do Foca" (2008):

"Pense nas pirâmides do Egito: a Grande Pirâmide de Quéops, em Guizé, mede 138 metros de altura e é considerada uma das sete maravilhas do mundo. O segredo da pirâmide está na base bem-feita, a partir da qual os egípcios equilibraram matematicamente a estrutura. Se a posição da pirâmide for invertida, a base vem para cima e continua sendo a porção mais importante. Assim funciona o sistema pirâmide invertida no jornalismo, que está sentado no tripé."(p. 166)

Nessa teoria da pirâmide invertida, a autora destaca que o "lide é a peça mais importante do sistema, no qual um dos fatos é o mais importante ou interessante." (p. 167) Esse pensamento ainda é reforçado por Guillermo Franco em 2010, quando ele o divide em três subníveis (p. 64):

- Texto linear colocado em uma só página da web: publicado em uma só página.
- Texto linear dividido tematicamente em uma só página da web: assunto principal é exposto no primeiro parágrafo, mas os subtemas são apresentados por intertítulos, permitindo ao usuário uma leitura não linear.
- Texto linear dividido tematicamente em uma só página Web: introdução ao assunto em uma página, seguida por links que levam aos subtemas, de maneira que o usuário determina a hierarquia da informação.

Apesar desse pensamento ser para um texto, a ideia pode ser adaptada para a estrutura de um lide. Sendo o jornalista a decidir a ordem da importância de acordo com sua experiência, critérios do meio onde publica e manual de redação do veículo.

6. Metodologia

Inspirada pelo trabalho do professor Márcio Carneiro dos Santos (2016), a metodologia escolhida para este trabalho foi *Design Science* (DS), desenvolvida por Herbert Simon no livro "*The science of the artificial*" ("As ciências do artificial"), nome como a metodologia era conhecida anteriormente.

Simon faz a divisão do mundo em dois, natural e artificial. O primeiro, ele explica, tem como função tornar maravilhoso o lugar-comum: "mostrar que a complexidade, corretamente vista, é apenas uma máscara para a simplicidade; Para encontrar o padrão escondido no caos aparente." (p.1)

O mundo artificial, em que diz estarmos muito mais envolvidos, é feito pelo homem, de maneira a cada elemento do meio mostrar evidências do toque humano, como o ar condicionado, que artificializa a temperatura do ambiente (p.2). Simon propõe que, enquanto as ciências naturais estudam a natureza, a *Design Science* se propõe a estudar as ciências do artificial, ressaltando que artificial se desgruda do caráter pejorativo da palavra para algo que não se opõe ao 'natural', pois não possui a licença para ignorar ou violar a lei natural.

Com olhares para a engenharia, ele fala que se volta mais para como a preocupação se volta para "como as coisas devem ser como elas devem ser com o objetivo de atingir uma finalidade, e funcionarem"⁴⁴ (p. 5). Enfim, ele chama de "artefatos" esses objetos do artificial, que são adaptados às metas e propósitos humanos. "Eles são o que eles são, a fim de satisfazer o nosso desejo de voar ou comer bem. À medida que nossos objetivos mudam, o mesmo acontece com os nossos artefatos e vice-versa."⁴⁵ (p. 3)

Apesar do viés de concepção da metodologia, pode-se ver a aplicação às áreas da comunicação social em Dresch et. al.:

"A questão da relevância do conhecimento produzido e a tensão na relação teoria-prática exigem, na verdade, um novo foco de pesquisa: pesquisas efetivamente direcionadas ao projeto de artefatos que sustentem melhores soluções para os problemas existentes." (2015)

Sendo o objeto do estudo a computação ferramenta que, no contexto, pode ser chamada artefato. A estudante passou seis meses aprendendo linguagens e lógica de computação e suas aplicações no jornalismo para aplicar a este trabalho.

⁴⁴ "...how things ought to be how they ought to be in order to attain goals, and to function."

⁴⁵ "They are what they are in order to satisfy our desire to fly or to eat well. As our aims change, so too do our artifacts and vice versa"

7. Conclusão

A computação tem sido auxiliar dos jornalistas em diferentes fases do processo de produção de notícias, mas esse nível de intimidade vem crescendo e assustando a muitos na área. Apesar disso, o jornalista permanece autoridade na criação da notícia nos momentos em que define sua linha editorial, o que será publicado, o viés da história e seu público.

A divisão do trabalho homem-máquina continua mudando com a tecnologia e, enquanto alguns observadores veem o potencial de automatizar tarefas rotineiras na redação para aumentar a qualidade das notícias e dar tempo aos jornalistas para matérias mais aprofundadas, a tecnologia eliminar trabalhos na redação é o que domina o debate, mesmo após a AP ter diminuído a quantidade de erros e aumentado em dez vezes sua quantidade de histórias com o uso de programas de computador, mas ter afirmado não haver cortado nenhum posto de trabalho por esse motivo⁴⁶.

Vale lembrar que o jornalismo automatizado apenas funciona quando estão disponíveis dados confiáveis e estruturados e que a resistência não impede a tecnologia de se desenvolver, por isso a necessidade do jornalista se colocar como protagonista desse processo, em vez de apenas sofrer as consequências.

Este trabalho sem dúvidas foi uma experiência marcante. Ele nasceu de um medo, partiu para uma esperança, se tornou um desafio e espero que a existência dele possa causar um bom impacto para outras pessoas. O mais interessante foi o choque com um novo aprendizado, que permite fazer tanto em tão pouco tempo. Ao efetivamente estudar computação, o olhar sobre as ferramentas utilizadas no cotidiano, tanto da vida pessoal quanto da profissão, se diferencia devido à mentalidade desenvolvida no processo.

Como o foco neste trabalho é o jornalista, não dá tanta consideração a outras implicações que possui a automação no jornalismo, mas podem ser exploradas em outras pesquisas, como o aumento de produção de notícia que a prática gera podem aumentar também:

- A dificuldade para os leitores de checar a veracidade e a credibilidade da história;
- O fardo de separar o que é relevante ou não, a manipulação de números para determinada visão de uma história;
- A personalização de notícias e reportagens;

⁴⁶< <http://www.poynter.org/2015/robot-writing-increased-aps-earnings-stories-by-tenfold/315931/>> Acesso em 05 fev 2017

- A alienação do público por fragmentação da opinião pública;
- As alterações nas práticas jornalísticas;
- Os questionamentos sobre o que é o jornalismo e como ele deve operar.

Espero que este trabalho incomode, guie, ajude ou abra os olhos de alguém.

8. Referências bibliográficas

ANDERSON, C.W. "Towards a Sociology of Computational and Algorithmic Journalism", *New Media & Society*, 2012. Disponível em: <<http://nms.sagepub.com/content/early/2012/12/06/1461444812465137>> Acesso em 01 Fev 2017

ARCE, Tacyana. O lead automatizado: uma possibilidade de tratamento da informação para o jornalismo impresso diário. *Revista Exacta*, Belo Horizonte, v. 2, n. 3, 2009. Disponível em: <<http://revistas.unibh.br/index.php/dcet/article/view/253/143> > Acesso em: 03 Jan 2017

BERRET, C., PHILLIPS, C. *Teaching Data and Computational Journalism*. 2016. Disponível em: <<https://www.gitbook.com/book/columbiajournalism/teaching-data-computational-journalism/details>> Acesso em 01 Fev 2017.

BOUNEGRU L., CHAMBERS L., GRAY J. eds. *The Data Journalism Handbook*. Sebastopol, Calif.: O'Reilly Media, 2012. Disponível em: <<http://datajournalismhandbook.org/>> Acesso em: 01 Fev 2017

BROUSSARD, M. *Artificial Intelligence for Public Affairs Reporting*, 2014. Disponível em <http://compute-cuj.org/cj-2014/cj2014_session1_paper1.pdf> Acesso em: 01 Fev 2017

CLERWALL, Christer. Enter the Robot Journalist: User's perception of automated content. In: *Journalism Practice. Special Issue – Future of Journalism in an age of digital media and economic uncertainty*. Volume 8, Issue 5. New York: Taylor&Francis Online, 2014. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/17512786.2014.883116> > Acesso em: 20 jan 2017

COHEN, S., HAMILTON J. T., TURNER F. *Computational Journalism*. 2011 Disponível em: <<http://cacm.acm.org/magazines/2011/10/131400-computational-journalism/fulltext>> Acesso em 22 Jan 2017

COHEN, S., LI, C., YANG J., YU C. *Computational Journalism: A Call to Arms to Database Researchers*. 2011 Disponível em: <<http://static.googleusercontent.com/media/research.google.com/pt-BR//pubs/archive/37381.pdf>> Acesso em 22 Jan 2017

DANIEL, A., FLEW, T., SPURGEON, C. "The Promise of Computational Journalism." *Journalism Practice*, 2012, p.157–171. Disponível em: <http://eprints.qut.edu.au/39649/1/39649_flew_2011008025.pdf> Acesso em 27 Jan 2017

DERVIN, B., *Chaos, Order and Sense-making: A Proposed Theory for Information Design*. In: JACOBSON, R. *Information Design*, 2000. The MIT Press, pp. 35-47.

DIAKOPOULOS, N. "A Functional Roadmap for Innovation in Computational Journalism." 2011. Disponível em: <http://www.nickdiakopoulos.com/wp-content/uploads/2007/05/CJ_Whitepaper_Diakopoulos.pdf> Acesso em 19 Jan 2017

_____. *Finding Tools Vs. Making Tools: Discovering Common Ground between Computer Science and Journalism*. 2013. Disponível em: <<http://www.niemanlab.org/2013/02/finding-tools-vs-making-tools-discovering-common-ground-between-computer-science-and-journalism/>> Acesso em 19 Jan 2017

DOS SANTOS, Márcio Carneiro. *Comunicação Digital e Jornalismo de Inserção*. Labcom Digital, 2016.

_____. *Narrativas Automatizadas e a Geração de Textos Jornalísticos/ A Estrutura de Organização do Lead Traduzida em Código*. *Brazilian Journalism Research*, v.12, n. 1, p. 160-185, 2016. Disponível em: < <https://bjr.sbpjor.org.br/bjr/article/view/757> > Acesso em: 29 Jan 2016.

FRANCO, Guillermo. "Como escrever para a web: elementos para a discussão e construção de manuais de redação online." *Centro Knight* (2010). Disponível em: <<https://knightcenter.utexas.edu/es/ebook/como-escribir-para-la-web-como-escribir-para-web>> Acesso em: 29 Jan 2016.

GLAHN, H.R. Computer-produced worded forecasts. *Bull. Amer. Meteor. Soc.*, vol. 51, pp. 1126-1131, 1970. Disponível em: <[http://journals.ametsoc.org/doi/abs/10.1175/1520-0477\(1970\)051%3C1126%3ACPWF%3E2.0.CO%3B2](http://journals.ametsoc.org/doi/abs/10.1175/1520-0477(1970)051%3C1126%3ACPWF%3E2.0.CO%3B2)> Acesso em 27 Jan 2017

GRAEFE, A. *Guide to automated journalism*. Tow Center for Digital Journalism, 2016. Disponível em:< <http://towcenter.org/research/guide-to-automated-journalism/> >

GRAEFEA, A., HAIMA, M., HAARMANNB, B., BROSIUS, H., *Perception of Automated Computer-Generated News: Credibility, Expertise, and Readability*. 2015. Disponível em: <https://www.researchgate.net/profile/Andreas_Graefe/publication/289529002_Perc

option_of_Automated_Computer-Generated_News_Credibility_Expertise_and_Readability/links/568ed9ea08aeaa1481b042ed.pdf> Acesso em 27 Jan 2017

HAMILTON J. T., TURNER F. Accountability Through Algorithm: Developing the Field of Computational Journalism. 2009. Disponível em: <<http://web.stanford.edu/~fturner/Hamilton%20Turner%20Acc%20by%20Alg%20Final.pdf>> Acesso em 27 Jan 2017

HERBERT, Simon A. The science of the artificial. 3ed. Cambridge: MIT Press, 1996. Disponível em: <https://monoskop.org/images/9/9c/Simon_Herbert_A_The_Sciences_of_the_Artificial_3rd_ed.pdf> Acesso em 02 fev 2017

JORGE, Thaís M. "Manual do foca." *São Paulo: Contexto* (2008).

LACERDA, Daniel Pacheco; DRESCH, Aline; PROENCA, Adriano and ANTUNES JUNIOR, José Antonio Valle. Design Science Research: método de pesquisa para a engenharia de produção. *Gest. Prod.* [online]. 2013, vol.20, n.4, pp.741-761. Epub Nov 26, 2013. ISSN 0104-530X. Disponível em: <http://www.scielo.br/scielo.php?pid=S0104-530X2013000400001&script=sci_abstract&tlng=pt> Acesso em 02 fev 2017

LECOMPTE, C. Automation in the Newsroom: How algorithms are helping reporters expand coverage, engage audiences, and respond to breaking news. *Nieman reports*, 2015. Disponível em: <<http://niemanreports.org/articles/automation-in-the-newsroom/>> Acesso em 02 fev 2017

MONFORT, M. MARTIN, S. A. FREDERICKSON W. "Information-processing differences and laterality of students from different colleges and disciplines." 1990. *Perceptual & Motor Skills*, 70(1), 163- 172.

ROYAL, C. "The journalist as programmer: A case study of the New York Times interactive news technology department." *International Symposium on Online Journalism*. 2010. Disponível em: <<https://online.journalism.utexas.edu/2010/papers/Royal10.pdf>> Acesso em 01 Fev 2017.

SIVITANIDES, M., WHITE, G. A Theory of the Relationships between Cognitive Requirements of Computer Programming Languages and Programmers' Cognitive Characteristics. *Journal of Information Systems Education*; West Lafayette 13.1 (2002): 59-66. Disponível em: <<http://jise.org/Volume13/Contents-13-1.html>> Acesso em 01 Fev 2017.

STONE, M. L. Big Data for Media. Reuters Institute for the Study of Journalism, 2014. Disponível em: <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Big%20Data%20For%20Media_0.pdf> Acesso em: 28 Abr 2016

STONEMAN, J. Does Open Data Need Journalism? Reuters Institute for the Study of Journalism, 2015. Disponível em: <<https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Stoneman%20-%20Does%20Open%20Data%20need%20Journalism.pdf>> Acesso em: 28 Jan 2017