

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA  
CURSO DE MESTRADO EM TECNOLOGIA E SOCIEDADE

MATIAS SEBASTIÃO PERUYERA

# USOS E APROPRIAÇÕES DE TECNOLOGIAS NO COTIDIANO DO JORNALISMO GUIADO POR DADOS

DISSERTAÇÃO

CURITIBA  
2015

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA  
CURSO DE MESTRADO EM TECNOLOGIA E SOCIEDADE

MATIAS SEBASTIÃO PERUYERA

**USOS E APROPRIAÇÕES DE TECNOLOGIAS  
NO COTIDIANO DO JORNALISMO GUIADO POR DADOS  
(COM FORMATAÇÃO APRESENTADA À BANCA)**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Tecnologia do Programa de Pós-Graduação em Tecnologia da Universidade Tecnológica Federal do Paraná. Linha de pesquisa: Mediações e Culturas.

Orientador: Prof. Luiz Ernesto Merkle.

CURITIBA  
2015

**CC BY:** a dissertação USOS E APROPRIAÇÕES DE TECNOLOGIA NO COTIDIANO DO JORNALISMO GUIADO POR DADOS, de Matias Sebastião Peruyera (2015), da UTFPR, Câmpus Curitiba, está licenciada com uma Licença /Creative Commons/ - Atribuição 4.0 Internacional.

Dissertação disponível no Repositório Institucional da UTFPR:  
<<http://repositorio.utfpr.edu.br/jspui/>>.

---

**Dados Internacionais de Catalogação na Publicação**

---

P471u Peruyera, Matias Sebastião  
2015 Usos e apropriações de tecnologias no cotidiano do jornalismo guiado por dados / Matias Sebastião Peruyera.-- 2015.  
216 p.: il.; 30 cm  
Texto em português, com resumo em inglês e espanhol.  
Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Tecnologia, Curitiba, 2015.  
Bibliografia: p. 199-206.  
1. Jornalismo - Aspectos sociais. 2. Jornalismo eletrônico. 3. Visualização da informação. 4. Jornalismo - Objetividade. 5. Jornalismo - Recursos de rede de computador. 6. Tecnologia apropriada. 7. Identidade social. 8. Comunicação de massa e tecnologia. 9. Inovações tecnológicas - Aspectos sociais. 10. Tecnologia - Dissertações. I. Merkle, Luiz Ernesto, orient. II. Universidade Tecnológica Federal do Paraná - Programa de Pós-graduação em Tecnologia. III. Título.

CDD 22 -- 600

---

**Biblioteca Central da UTFPR, Câmpus Curitiba**

## TERMO DE APROVAÇÃO

Título da Dissertação Nº 437

**Usos e Apropriações de Tecnologias no Cotidiano do Jornalismo Guiado por Dados**

por

**Matias Sebastião Peruyera**

Esta dissertação foi apresentada às 14h00 do dia **15 de junho de 2015** como requisito parcial para a obtenção do título de MESTRE EM TECNOLOGIA, Área de Concentração – Tecnologia e Sociedade, Linha de Pesquisa – Mediações e Culturas, Programa de Pós-Graduação em Tecnologia, Universidade Tecnológica Federal do Paraná. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho APROVADO (aprovado, aprovado com restrições, ou reprovado).

\_\_\_\_\_  
Prof.<sup>a</sup>. Dr.<sup>a</sup>. Marilda Pinheiro Queluz  
(UTFPR)

\_\_\_\_\_  
Prof. Dr. Leonelo Dell Anhol Almeida  
(PPGCA/UTFPR)

\_\_\_\_\_  
Prof. Dr. Marcelo Ruschel Träsel  
(PUC-RS)

\_\_\_\_\_  
Prof. Dr. Luiz Ernesto Merkle  
(UTFPR)  
*Orientador*



Visto da coordenação:

\_\_\_\_\_  
Prof.<sup>a</sup>. Dr.<sup>a</sup>. Faimara do Rocio Strauhs  
Coordenadora do PPGTE

**OBS: O documento original encontra-se arquivado na secretaria do PPGTE.**





Capa: imagem da capa da versão apresentada à banca examinadora. No ambiente de programação Processing, foi feito um programa para desenhar um quadrado para cada frase – ou um círculo, caso a frase tivesse as palavras “visualização”, “jornalismo” ou “jornalista”. Cores mais claras representam frases mais curtas, e cores mais próximas do preto representam frases mais longas.

*A meus amigos.*

## **AGRADECIMENTOS**

—Como tá o mestrado?

—Ah, tá massa. Tô me sentindo meio burro. Bah, bastante burro.

—Sensacional, sei a sensação, é ótimo.

Com esse diálogo, que tive com uma amiga no começo do mestrado, ilustro meu primeiro agradecimento, que vai para quem me proporcionou essa maravilhosa sensação de me sentir burro: professoras, professores, e colegas do Programa de Pós-Graduação em Tecnologia da UTFPR. Além do aprendizado, agradeço que tenham me mostrado que ainda há muitos caminhos a percorrer; se sentir burro diante desses caminhos é estimulante e desafiador. Agradeço especialmente ao meu orientador, professor Luiz Ernesto Merkle, por ter me mostrado vários desses caminhos e por ter me acompanhado neste.

À professora Marilda Pinheiro Queluz e aos professores Leonelo Dell Anhol Almeida e Marcelo Träsel, por colaborarem na banca de qualificação.

À professora Faimara do Rocio Strauhs, por ter me feito sentir em casa antes mesmo de entrar no mestrado, e à Bruna Bulla Correia, da secretaria do PPGTE, pelas várias ajudas.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pela bolsa de estudos.

Às companheiras e companheiros de caminhada, do PPGTE e fora dele: Catiane, Gonzatto, Fabiane, Liber, Kira, Rodrigo, Cátia, Camille, Suelen, Bruno, Luciano, Vinícius, Lena e Camila.

Às pessoas que entrevistei: Renata Silva Pinto, Guilherme Storck, Ricardo Ruas Silva, Bob Smith, João Guilherme Frey e José Lazaro Barros Junior do Livre.jor, além da professora Rosiane Correia de Freitas, do professor Ricardo Fotios e do professor Marcelo Träsel, que também participou da banca de qualificação.

Aos professores e professoras da Uninter, pelos vários empurrões iniciais, e aos colegas da Gazeta do Povo.

Aos amigos e amigas que estiveram por perto durante estes dois anos, e ao Basilio, por estar mais perto ainda.

“Me siento muy alto cuando miro un mapa”

*Alvaro Carmona*

## RESUMO

PERUYERA, Matias. **Usos e apropriações de tecnologias no cotidiano do jornalismo guiado por dados e nas visualizações**. 216 f. Dissertação (Mestrado em Tecnologia e Sociedade) – Programa de Pós-Graduação em Tecnologia, Universidade Tecnológica Federal do Paraná. Curitiba, 2015.

Esta dissertação aborda como profissionais do jornalismo, através das práticas do cotidiano, se apropriam de artefatos e tecnologias computacionais para trabalhar com Jornalismo Guiado por Dados e, especificamente, com visualizações de dados. Para isso, considera-se que é no cotidiano que as tecnologias são apropriadas por quem as usa, e que este uso leva as tecnologias a uma estabilização, na qual elas deixam de ser percebidas como um elemento estranho. Os modos de uso dos artefatos também são considerados enquanto elementos construtores de identidades. O objeto principal da pesquisa são seis profissionais que fazem uso de artefatos e técnicas do Jornalismo Guiado por Dados no seu dia a dia. Através de entrevistas e observações, foram levantadas algumas maneiras através das quais as práticas e apropriações de tecnologias constroem as identidades de quem faz uso delas e levam os artefatos à estabilização. A análise das práticas do cotidiano colaboraria para uma visão menos instrumentalista no desenho de artefatos e também no ensino de técnicas, legitimando assim os modos de uso de cada pessoa. Para entender essas práticas do cotidiano, são apresentados os conceitos de “táticas” e “estratégias” com o objetivo de situar as relações de poder do cotidiano e como as pessoas podem subvertê-las, e conceitos dos estudos em Ciência, Tecnologia e Sociedade como SCOT – construção social da tecnologia – e códigos técnicos, para analisar os processos que levam à estabilização de uma tecnologia. Relacionando esses conceitos com o jornalismo, é analisada a conotação de neutralidade dos grandes conjuntos de dados e comparando-a com a ideia de neutralidade da tecnologia. O Jornalismo Guiado por Dados é então relacionado a algumas taxonomias de gêneros e formatos jornalísticos, e é apresentado um breve histórico do uso do computador no jornalismo, para situá-lo dentro do jornalismo de modo geral e nas identidades jornalísticas que são construídas através das tecnologias consumidas no cotidiano. Por sua vez, é apresentado como os produtos jornalísticos derivados de grandes bases de dados propõem outro tipo de relação entre o jornalismo e o público, especificamente através da visualização de dados. São consideradas as possibilidades da visualização como forma de explorar e/ou comunicar grandes conjuntos de dados, assim como formas de leituras que elas propiciam. São descritos alguns processos para produzir visualizações de dados e mostrados exemplos de visualizações usadas no jornalismo, além de descrições de ferramentas de software usadas no Jornalismo Guiado por Dados. Como conclusão principal, defende-se que a legitimação das táticas, assim como uma visão menos instrumentalista e determinista do computar e de outras tecnologias, colaboraria para que mais pessoas se aproveitem dos recursos do Jornalismo Guiado por Dados e das visualizações.

**Palavras-chave:** Visualização de dados. Jornalismo Guiado por Dados. Cotidiano. Estudos CTS. Apropriação. Código técnico. Estabilização da tecnologia.

## ABSTRACT

PERUYERA, Matias. **Everyday uses and technology appropriation in data-driven journalism**. 216 p. Dissertation (Master in Technology and Society) – Post-Graduation Program in Technology. Universidade Tecnológica Federal do Paraná. Curitiba , 2015.

This dissertation looks at how journalism professionals, through everyday practices, appropriate artifacts and computational technologies to work with Data-Driven Journalism and, specifically, with data visualization. For this, we considered that it is in everyday life that technologies are appropriated for those who make use of them, and that this use leads technologies to a closure or stabilization, in which they are no longer perceived as a foreign element. We also consider the way people make use of artifacts as elements that construct identities. The main study subjects are six professionals that make use of artifacts and techniques related to Data-Driven Journalism in their everyday life. Through interviews and observations, we collected some ways through which people appropriate technology, and thus takes them to stabilization and build identities. The analysis of everyday practices would collaborate to a less instrumentalist approach in artifact design and technique teaching, thus legitimating the ways each person makes use of technologies. For a better understanding of those everyday practices, the concepts of "tactics" and "strategies" are introduced, in order to situate the everyday power relations and how people can subvert them, as well as concepts from the Science, Technology and Society studies – STS –, such as SCOT – social construction of technology – and technical codes, in order to analyze the processes that lead to the closure of technology. Relating these concepts to journalism, the neutral connotation of large data sets is analyzed and compared to the idea of technology neutrality. Data-Driven Journalism is then related to some taxonomies of journalistic genres and formats, and a brief history of computer use in journalism is presented to situate it within journalism and how journalism's identities are constructed through technology consumed in daily life. In turn, is presented how journalistic products derived from large databases propose another type of relationship between journalism and the public, specifically through data visualization. We describe some possibilities of visualization as a way to explore and/or communicate large data sets, as well as some different ways of reading they provide. We also describe some processes and tools for producing data visualization in journalism, as well as some software tools used in Data-Driven Journalism. The major conclusion of this study is that legitimating tactics, as well as a less instrumentalist and determinist approach to computing and other technologies, would help more people in making use of the artifacts and techniques of Data-Driven Journalism and data visualization.

**Keywords:** Data visualization. Data-driven Journalism. Everyday life. STS Studies. Appropriation. Technical code.



## RESUMEN

PERUYERA, Matias. **Usos y apropiaciones de tecnología en el cotidiano del periodismo guiado por datos.** 216 h. Disertación (Maestría en Tecnología y Sociedad) – Programa de Posgrado en Tecnología, Universidad Tecnológica Federal de Paraná. Curitiba, 2015.

Esta disertación trata de como profesionales del periodismo, a través de las prácticas cotidianas, se apropian de artefactos y tecnologías computacionales para trabajar con Periodismo Guiado por Datos y, específicamente, con visualizaciones de datos. Para eso, se considera que es en las situaciones del cotidiano que quienes usan las tecnologías se apropian de ellas, y que ese uso lleva las tecnologías a una estabilización, que es el momento en el cual dejan de ser percibidas como un elemento extraño. Los modos de uso de artefactos también son considerados en tanto que elementos constructores de identidades. El objeto principal de estudio son seis profesionales que hacen uso de artefactos y técnicas del Periodismo Guiado por Datos en sus quehaceres diarios. A través de entrevistas y observaciones, fueron levantadas algunas maneras a través de las cuales las prácticas y apropiaciones de tecnologías construyen las identidades de quienes hacen uso de ellas y llevan los artefactos a su estabilización. La análisis de las prácticas del cotidiano colaborarían para una visión menos instrumentista en el diseño de artefactos y también en el enseñar de las técnicas, legitimando así los modos de uso de cada persona. Para entender esas prácticas del cotidiano, son presentados los conceptos de “tácticas” e “estrategias” con el objetivo de situar las relaciones de poder del cotidiano y como pueden ser subvertida por las personas, y conceptos de los estudios em Ciencia, Tecnología e Sociedad, como SCOT – construcción social de la tecnología – y códigos técnicos, para analizar los procesos que llevan a la estabilización de una tecnología. Relacionando esos conceptos con el periodismo, se analiza la connotación de neutralidad de los grandes conjuntos de datos y se la compara con la idea de neutralidad de la tecnología. Se relaciona el Periodismo Guiado por Datos a algunas taxonomías de géneros y formatos periodísticos, y es presentado un breve histórico del uso de la computadora en el periodismo, para situarla dentro del periodismo de manera general y en las identidades periodísticas que son construídas a través de las tecnologías consumidas en el cotidiano. Por su vez, es presentado como los productos periodísticos derivados de grandes bases de datos proponen otro tipo de relación entre el periodismo e el público, específicamente a través de la visualización de datos. Son consideradas las posibilidades de la visualización como forma de explorar y/o comunicar grandes conjuntos de datos, así como formas de lecturas que ellas propician. Son descriptos algunos procesos para producir visualizaciones y mostrados ejemplos de visualizaciones usadas en periodismo, además de descripciones de herramientas de software usadas en Periodismo Guiado por Datos. Por conclusión principal, se defiende que la legitimación de las tácticas, así como una visión menos instrumentalista y determinista del computar y de otras tecnologías, colaboraría para que más personas se aprovechen de los recursos del Periodismo Guiado por Datos y de las visualizaciones.

**Palabras clave:** Visualización de datos. Periodismo guiado por datos. Cotidiano. Apropiación.

## LISTA DE ILUSTRAÇÕES

Fig. 1: Captura de tela mostrando as ocorrências do termo procurado .....	47
Fig. 2: Captura de tela da planilha “Churras Calculator” .....	55
Fig. 3: Boneco representando a profissão de jornalista, geralmente dado de presente a jornalistas .....	61
Fig. 4: Visualização feita a partir da tag "sunset" .....	65
Fig. 5: Mapa dos lançamentos imobiliários de Curitiba.....	72
Fig. 6: Mapeamento de cruzamentos perigosos para ciclistas.....	73
Fig. 7: Matéria e visualização publicadas na página do Facebook do Livre.jor ....	82
Fig. 8: Tela do aplicativo “Is It Better to Rent or Buy?” com duas das 21 variáveis que podem ser escolhidas para obter um resultado personalizado .....	96
Fig. 9: Placa informativa com suporte para fotografar .....	99
Fig. 10: Trecho da visualização "Six Billion Tweets". Em destaque, a sede da UTFPR do centro de Curitiba.....	106
Fig. 11: Diagrama de dispersão com diversos personagens da atriz Meryl Streep.....	107
Fig. 12: Passes trocados entre jogadores da Seleção Brasileira .....	108
Fig. 13: Quatro níveis da visualização “Six Billion Tweets”: mundial, cidades brasileiras, Curitiba e arredores e centro de Curitiba, com a UTFPR em destaque.....	110
Fig. 14: Exemplos de mapas para "ler" (acima) e para "ver" (embaixo).....	112
Fig. 15: Utilização de visualizações no infográfico “La ballena Franca” (destaques nossos) .....	116
Fig. 16: Esquema das etapas para produzir uma visualização, segundo Bertin (1986), Bradshaw (2011), Fry (2008) e Silver (2014) .....	120
Fig. 17: Infográfico dos gastos dos 34 deputados estaduais do Paraná. No destaque, um deputado que gastou mais no final do ano e outro que gastava regularmente .....	126
Fig. 18: Exemplos de manuseio dos elementos da bola de futebol para comparar dados e visualizações das faces de diferentes peças.....	129
Fig. 19: Exemplar da visualização em forma de bola de futebol do jornal <i>Times of Oman</i> .....	129

Fig. 20: Primeira e última páginas do primeiro caderno do jornal argentino “La Voz del Interior” .....	130
Fig. 21: Gráfico em formato de barril publicado online pelo jornal The Guardian .....	131
Fig. 22: Gráfico sobre comportamento de cachorros, retirado de um artigo científico.....	132
Fig. 23: Infográfico redesenhado pela equipe do The New York Times.....	132
Fig. 24: Mapeamento e proposta de redesenho .....	133
Fig. 25: Fluxograma para jornalistas que queiram aprender – ou não – programação.....	135
Fig. 26: Tira satirizando o tempo de desenvolvimento de programas .....	139
Fig. 27: Trecho da base de dados dos lançamentos de Rivera .....	140
Fig. 28: Tela da visualização feita para explorar os dados dos arremessos .....	141
Fig. 29: Momento do vídeo no qual aparecem representados ao mesmo tempo todos os lançamentos feitos em 2009 .....	141
Fig. 30: Dois frames da animação na qual são mostrados todos os lançamentos .....	142
Fig. 31: Mapa de calor representando a concentração de lançamentos em duas regiões específicas.....	143
Fig. 32: Visualização da série "See Something or Say Something", de Eric Fischer, mostrando a cidade do Rio de Janeiro .....	144
Fig. 33: Infográfico com a quantidade de votos nos bairros de Curitiba para cada vereador eleito.....	146
Fig. 34: Basômetro .....	147
Fig. 35: Outra visão do basômetro.....	147
Fig. 36: Quatro quadros da visualização em vídeo "World War II in Europe: Every Day" .....	148
Fig. 37: Três abas da visualização "Crime sem castigo" .....	149
Fig. 38: Mapeamento de emigração .....	151
Fig. 39: Mapeamento de empregados com carteira assinada .....	151
Fig. 40: Mapeamento de concentração de pessoas solteiras .....	152
Fig. 41: Pirâmide de população .....	166

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>15</b>
1.1	CONTRIBUIÇÕES ESPERADAS .....	19
1.2	OBJETIVOS DA PESQUISA.....	20
1.3	MARCO TEÓRICO .....	21
1.4	METODOLOGIA.....	22
1.5	ESTRUTURA DA DISSERTAÇÃO.....	25
1.6	SOBRE AS VISUALIZAÇÕES ILUSTRATIVAS .....	26
<b>2</b>	<b>TECNOLOGIAS E JORNALISTAS .....</b>	<b>29</b>
2.1	COTIDIANO, ESTRATÉGIAS E TÁTICAS .....	29
2.2	TECNOLOGIA COMO CONSTRUÇÃO SOCIAL .....	37
2.3	SISTEMAS ABERTOS E FECHADOS .....	42
2.4	APROPRIAÇÕES DAS TECNOLOGIAS.....	44
2.4.1	Sistemas abertos: prevendo a apropriação .....	53
2.5	ESTABILIZAÇÃO DAS TECNOLOGIAS.....	56
2.6	CONSIDERAÇÕES.....	57
<b>3</b>	<b>JORNALISMO E O COMPUTAR .....</b>	<b>61</b>
3.1	O IDEAL DE OBJETIVIDADE DO JORNALISMO .....	62
3.1.1	O ideal de objetividade dos grandes conjuntos de dados .....	64
3.2	JORNALISMO GUIADO POR DADOS E GÊNEROS JORNALÍSTICOS .....	66
3.2.1	O caráter interpretativo do JGD .....	68
3.3	HISTÓRICO DO COMPUTAR NO JORNALISMO .....	74
3.4	POSSIBILIDADES DO JGD.....	77
3.5	JGD, ARTEFATOS E IDENTIDADES JORNALÍSTICAS .....	82
3.6	CONSIDERAÇÕES.....	89
<b>4</b>	<b>FORMATOS E PRÁTICAS DO JGD.....</b>	<b>93</b>
4.1	O PAPEL DOS GRANDES CONJUNTOS DE DADOS NO JORNALISMO.....	94
4.1.1	Relação entre público e jornalismo através da visualização .....	98
4.2	VISUALIZAÇÃO: CONCEITOS E POSSIBILIDADES .....	100
4.2.1	Observações sobre os conceitos de visualização .....	104
4.2.2	Processos de leitura de visualizações .....	109
4.2.3	Diferenças entre visualização de dados e infografia.....	114
4.2.4	Outros tipos de visualização .....	115
4.3	MÉTODOS DO JGD E DA VISUALIZAÇÃO .....	117
4.3.1	Etapas para produzir uma visualização .....	119
4.3.2	Obter e processar os dados .....	121
4.3.3	Visualizar os dados .....	124
4.3.4	Comunicar os dados .....	128

4.3.5	Considerações sobre programação .....	134
4.4	ALGUNS EXEMPLOS DE USO DE VISUALIZAÇÕES NO JORNALISMO.....	139
4.4.1	“How Mariano Rivera Dominates Hitters”, the new York Times .....	139
4.4.2	“See Something or Say Something”, de Eric Fischer .....	143
4.4.3	“Bases eleitorais”, Gazeta do Povo .....	145
4.4.4	“Basômetro” – O Estado de São Paulo .....	145
4.4.5	“World War II in Europe: Every Day” – Emperor Tigerstar .....	148
4.4.6	“Crime sem castigo” – Gazeta do Povo .....	148
4.4.7	Visualizações feitas por estudantes de jornalismo .....	150
4.5	CONSIDERAÇÕES.....	152
<b>5</b>	<b>ENTREVISTAS E OBSERVAÇÕES.....</b>	<b>157</b>
5.1	ENTREVISTADOS .....	157
5.2	OBSERVAÇÕES EM OFICINAS DE VISUALIZAÇÃO .....	161
5.3	MODOS DE USO.....	162
5.3.1	Considerações sobre os modos de uso observados .....	177
5.4	PROGRAMAÇÃO .....	182
5.4.1	Considerações sobre os modos de uso de programação.....	185
5.5	APRENDIZADO .....	187
5.5.1	Considerações sobre maneiras de aprendizado .....	189
5.6	VISUALIZAÇÕES .....	190
5.6.1	Considerações sobre modos de uso de visualizações .....	191
5.7	CONSIDERAÇÕES SOBRE AS ENTREVISTAS E OBSERVAÇÕES .....	192
<b>6</b>	<b>CONSIDERAÇÕES FINAIS E PERSPECTIVAS.....</b>	<b>195</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>199</b>
	<b>APÊNDICE A: GLOSSÁRIO DE FERRAMENTAS .....</b>	<b>207</b>
	FERRAMENTAS PARA EXTRAÇÃO DE DADOS .....	207
	FERRAMENTAS PARA PREPARAR, ORGANIZAR, LIMPAR E FAZER OPERAÇÕES COM OS DADOS .....	207
	FERRAMENTAS DE PROGRAMAÇÃO .....	208
	FERRAMENTAS PARA VISUALIZAÇÃO .....	208
	FERRAMENTAS PARA PRODUÇÃO DE MAPAS CARTOGRÁFICOS.....	209
	<b>APÊNDICE B: ROTEIRO DE ENTREVISTA SEMI ESTRUTURADA.....</b>	<b>210</b>
	<b>APÊNDICE C: QUESTÕES PARA PROFESSORES.....</b>	<b>212</b>
	<b>APÊNDICE D: DESCRIÇÃO DA PRODUÇÃO DAS VISUALIZAÇÕES ILUSTRATIVAS .....</b>	<b>213</b>





# 1 INTRODUÇÃO

Esta dissertação é resultado de várias ideias que foram sendo elaboradas ao longo de anos de trabalho com jornalismo e design, assim como de vivências e observações quase involuntárias feitas com colegas.

O design gráfico foi minha<sup>1</sup> primeira profissão, e aquela que me levou a trabalhar com infografia jornalística no jornal Gazeta do Povo, de Curitiba, no ano de 2002. Foi nessa função que começaram os questionamentos e pesquisas informais em infografia e suas áreas de interesse, como design da informação, tipografia, cartografia etc. Naquela época eu sequer sabia o que era um mestrado, mas muitas dessas pesquisas estão, anos depois, entre as referências bibliográficas desta dissertação.

A divisão de trabalho no jornal era bem estrita: repórteres, diagramadores, ilustradores e fotógrafos exerciam funções bem delimitadas. O departamento de infografia era exceção, graças ao caráter multidisciplinar da infografia e ao caráter multifuncional dos infografistas. Assim, as tarefas que não se encaixassem nas funções de outros departamentos terminavam na infografia, fosse a elaboração de uma fonte tipográfica especial, a diagramação de uma página que não se encaixasse nos padrões do projeto gráfico, reprodução de documentos, entre outras.

Essas tarefas incluíam a preparação de textos e de dados numéricos para a elaboração de gráficos. O departamento de infografia começou a assumir essa tarefa por causa de que muitas das pessoas que apuravam as informações – repórteres – para a produção de infográficos não tinham conhecimento de como entregar os dados, e às vezes nem de quais dados usar, ou quais dados eram adequados para a matéria em produção. Remediar essa carência se tornou fundamental para conseguir produzir

---

<sup>1</sup> Nesta seção, pelo caráter pessoal, é usada a 1ª pessoa do singular.

materiais adequados apesar da baixa qualidade dos dados que eram entregues pelos e pelas repórteres.

Nesta dissertação, considero como jornalista qualquer pessoa que trabalhe com atividades jornalísticas, seja produzindo matérias, dando forma a elas, filmando matérias televisivas ou preparando fotos para serem impressas. Não importa se a pessoa cursou uma faculdade de jornalismo ou se o diploma de jornalista é obrigatório ou não. Seja na função principal de produzir conteúdo ou de dar forma a ele, é necessário seguir a ética da profissão, defendendo o direito e dever de informar, mantendo boas relações com as fontes de informação, e não prejudicar os colegas. Nesse aspecto, considero jornalistas todos aqueles e aquelas que façam parte de um produto jornalístico ou com caráter jornalístico – por mais que, na prática, ser ou não jornalista é algo que varia de acordo com a conveniência do momento, seja para reafirmar uma relação de poder entre jornalistas e “outros” profissionais ou para justificar regimes de trabalho diferenciados. Foi considerando essa situação que cursei a faculdade de jornalismo, e também porque, pela convivência com a profissão, soube que seria um curso fácil.

Durante a faculdade tomei contato com a pesquisa acadêmica, o que me levou à publicação de alguns artigos sobre visualização de dados e jornalismo (PERUYERA, 2010A E 2010B) e à realização do trabalho de conclusão de curso sobre o mesmo assunto (PERUYERA, 2012). Antes disso, em 2007, formei um grupo no Centro Paralelo, em Curitiba, para pesquisar e produzir visualizações dentro de um contexto artístico. O grupo não teve produção significativa, mas colaborou com esta dissertação e com a obra “100 importâncias”, que por sua vez foi abordada em uma dissertação de mestrado, da fotógrafa Nicole Lima<sup>2</sup>, fundadora do Centro Paralelo e parceira na criação e discussões do grupo.

Enquanto isso, a função do departamento de infografia foi assumindo um caráter cada vez mais multidisciplinar, em sintonia com transformações similares em outros meios de comunicação e

---

<sup>2</sup> Cf. LIMA, Nicole C. Regimes de luz: a fotografia na produção de descontinuidades. 2011. 114 f. Dissertação (Mestrado em Artes Visuais) – CEART, Universidade do Estado de Santa Catarina, Florianópolis, 2011. Disponível em: <[http://www.tede.udesc.br/tde\\_busca/arquivo.php?codArquivo=2833](http://www.tede.udesc.br/tde_busca/arquivo.php?codArquivo=2833)>. Acesso em 20/8/2014.

na infografia jornalística de modo geral, que acompanhávamos fazendo as pesquisas informais já citadas. Fazíamos pequenas apurações e operações com dados, ao mesmo tempo que explorávamos o uso de aplicativos diversos para conseguir resultados que se aproximassem de nossos objetivos em tanto que jornalistas visuais.

Esses objetivos não se limitavam a, como muitos acreditavam, decorar a informação para torná-la “visualmente atraente” – o que Edward Tufte (1990, p. 34) chama de *chartjunk* e eu traduzo como “lixográfico”. O objetivo era, através da infografia, transmitir informação de maneiras eficazes que fossem além do texto, aproveitando as possibilidades da linguagem gráfica, como é o caso das visualizações que vão além das possibilidades do texto escrito.

O repertório de ferramentas e de maneiras de transmitir informação foi se ampliando, assim como os recursos disponíveis. A equipe aumentou de quatro para seis pessoas, foram surgindo novos softwares de visualização de dados, foram comprados computadores que cumpriam os requisitos mínimos para executar certos softwares, aprendemos a usar melhor planilhas eletrônicas e processadores para grandes volumes de texto, surgiram alternativas de software livre para aplicativos comerciais caros.

Essa exploração de ferramentas e técnicas foi crescendo com a ajuda de novos integrantes da equipe de infografia, mudanças na parte jornalística que valorizaram mais nosso trabalho e a percepção de como novos conhecimentos e técnicas podiam ampliar os conceitos de jornalismo. Graças à multidisciplinariedade dos designers gráficos integrantes da equipe, começamos a trabalhar com extração e limpeza de dados, como o trabalho de extração e re-divulgação dos gastos da Assembleia Legislativa do Paraná (GAZETA DO POVO, 2014), a introduzir programação na produção de infográficos (PERUYERA, 2010B), e a produzir visualizações com grandes conjuntos de dados (como os exemplos das figuras 5, 12, 17, 33 e 37).

Com esse tipo de trabalhos, a equipe de infografia estava fazendo jornalismo de dados, ou, como será preferido nesta dissertação, **Jornalismo Guiado por Dados**. O termo se refere às

atividades, técnicas e tecnologias necessárias para encontrar informações latentes em grandes conjuntos de dados, processo que pode ou não incluir a obtenção desses conjuntos de dados. Ou, nas palavras da professora Meredith Broussard, “encontrar histórias em números e usar números para contar histórias” (HOWARD, 2014, P. 5)<sup>3</sup>. Em paralelo, algumas e alguns jornalistas também pesquisavam técnicas de reportagem similares e começavam a incorporá-las ao seu trabalho.

Tendo acompanhado essas transformações, sempre observava as diferentes maneiras de trabalhar das e dos infografistas, repórteres e editores<sup>4</sup>, particularmente do computador<sup>5</sup>. Muitas vezes eu ensinava como fazer alguma tarefa, e aproveitava para ver de que maneira cada pessoa computava. Me intrigava particularmente por que algumas pessoas, diante de um erro, largavam o mouse como quem se queima com uma panela, enquanto outras pessoas conseguiam fazer uso de recursos pouco conhecidos pelos colegas, o que era visto, por muitos, como uma proeza. Dessas observações surgiram várias inquietações que tinham em comum a relação das pessoas com as ferramentas que utilizavam: por que havia pessoas que achavam que o termo “infografia” vinha de informática e não de informação? Por que eu já fui chamado de “o cara do Excel”? Por que éramos tão associados com as ferramentas que usávamos? Por que muitas vezes o mérito de conseguir algo era do computador, e não da pessoa? Por que havia pessoas que não se sentiam seguras para usar uma planilha e preferia transcrever os números depois de fazer os cálculos em uma calculadora de mão?

Essas inquietações retornaram à medida que avançava na pesquisa. Quando decidi fazer um mestrado, a intenção inicial era pesquisar o uso de visualizações de dados no jornalismo. Porém, o

---

<sup>3</sup> Tradução livre. No original: “finding stories in numbers and using numbers to tell stories”. O termo “stories” é usado como sinônimo de “matéria” ou de “reportagem”.

<sup>4</sup> Repórteres são jornalistas que apuram informações através de entrevistas e pesquisas, entre outras, e depois escrevem a matéria. Editores têm função de chefia, não saem a campo e não costumam realizar pesquisas nem entrevistas.

<sup>5</sup> Nesta dissertação, de modo geral, o termo “computador” não se refere somente ao computador como artefato e sim a um conjunto de ferramentas que inclui software, programas, outros artefatos computacionais como telefones celulares, ferramentas e bases de dados disponíveis na internet, as habilidades necessárias para trabalhar com essas tecnologias e inclusive elementos externos como lápis e papel. Em algumas vezes será preferido o termo computar, por implicar as pessoas que fazem uso do computador e das tecnologias próximas a ele.

interesse se concentrou nas práticas cotidianas, no faça você mesmo, no improviso, nos jeitinhos e nas gambiarras. Mais do que o estudo de visualizações sofisticadas e inalcançáveis para a maioria dos profissionais, me interessava saber como as pessoas fuçam, como elas dão jeitinhos, como elas “se viram”. As inquietações citadas no parágrafo anterior voltaram, mas dentro das discussões sobre determinismo tecnológico.

As habilidades do cotidiano me interessaram particularmente quando comecei a acompanhar discussões sobre as possibilidades do chamado Jornalismo Guiado por Dados. De modo geral, o tom das discussões e textos era sobre como o JGD seria um novo caminho para o jornalismo, e como as redações deveriam contratar programadores e treinar jornalistas para se manterem atualizadas e não perderem o bonde. Mas o que acontece com as e os jornalistas que estão dentro dessas organizações? E com quem não pertence a uma organização? Somente grandes redações poderão fazer JGD? Foi principalmente nessa discussão que senti que o estudo do cotidiano do jornalismo e dos modos de uso do computador traria alguma colaboração à discussão.

## **1.1 CONTRIBUIÇÕES ESPERADAS**

Espera-se que a presente pesquisa sirva para realçar a importância das práticas cotidianas no aprendizado e exercício das atividades relacionadas ao Jornalismo Guiado por Dados e visualizações, seja no contexto de grandes redações ou para jornalistas independentes, ou ainda para pessoas que queiram realizar atividades similares em contextos não relacionados ao jornalismo.

Acredita-se que o estudo das possibilidades das práticas cotidianas colabore para a popularização e divulgação de técnicas de JGD e de visualizações, seja através do ensino – em faculdades de jornalismo, cursos de extensão, cursos online etc – ou através da criação de sistemas que propiciem o uso por um número maior de pessoas. Ou, ainda, através da legitimação das práticas cotidianas ou “gambiarras”, deixando-as em uma situação similar à de certas práticas jornalísticas “tradicionais” que são, mais do

que legitimadas, festejadas e consideradas parte do que é ser bom jornalista, como técnicas de entrevista e descobrir assuntos que possam render boas reportagens.

A pesquisa também pode colaborar para a análise de outros produtos jornalísticos que impliquem no uso de tecnologias ainda não estabilizadas, em situações nas quais a legitimação das práticas do cotidiano pode ajudar a que mais pessoas tenham acesso às tecnologias usadas para fazer esses produtos. É o caso de infográficos, aplicativos na internet ou para telefones celulares, ou ainda produtos como a chamada reportagem de forma longa<sup>6</sup>. Esse último tipo de produto, geralmente publicado na internet, combina várias mídias – textos, imagens estáticas, vídeos, áudio, entre outras. A produção dessas reportagens, assim como o JGD, exige a combinação de vários conhecimentos, como edição de vídeo, programação e escrita de textos adequados ao meio.

## 1.2 OBJETIVOS DA PESQUISA

O objetivo geral da pesquisa é:

- **Analisar maneiras que as pessoas têm de se apropriar de tecnologias computacionais e outras ferramentas necessárias ao Jornalismo Guiado por Dados e à produção de visualizações, e como essas apropriações podem colaborar nas atividades jornalísticas cotidianas.**

Os objetivos específicos são:

- Analisar como os modos de uso do cotidiano podem levar à estabilização das tecnologias e artefatos do Jornalismo Guiado por Dados e visualizações.
- Situar o Jornalismo Guiado por Dados e visualizações com relação aos gêneros, formatos e identidades jornalísticas.
- Descobrir possibilidades para a elaboração e publicação de visualização de dados e de que maneiras as pessoas que trabalham com jornalismo podem se aproveitar delas.

---

<sup>6</sup> Tradução livre do termo “Long-form journalism”. Um exemplo desse tipo de conteúdo está disponível em: <<http://www.gazetadopovo.com.br/vida-e-cidadania/especiais/ervamate/index.jpp>>. Acesso em 15/4/2015.



### 1.3 MARCO TEÓRICO

As motivações da pesquisa sugeriam a necessidade de autores que abordassem questões culturais e particulares de pessoas diante de relações de poder. Para isso, foram incorporados ao referencial teórico autores oriundos dos estudos culturais, como Michel de Certeau (2007), especificamente os conceitos de estratégia e tática. A estratégia é a força de quem tem poder, de quem tem um lugar, enquanto as táticas são combinações de elementos e saberes heterogêneos que aproveitam oportunidades para “contornar” as imposições da estratégia.

O consumo de artefatos em tanto que parte de um sistema maior que implica também em produção de identidades foi analisado com base em outro autor dos estudos culturais, Hugh Mackay (1997A; 1997B). A compreensão da tecnologia como cultura tira o protagonismo do artefato e põe o foco em quem o consome e nos processos cotidianos de apropriação e produção de significados.

Durante o estudo do cotidiano, foi vista a necessidade de estabelecer ligações entre a apropriação cotidiana e as tecnologias, para o qual foram usados conceitos oriundos dos estudos em Ciência, Tecnologia e Sociedade (FEENBERG, 1995; PINCH; BIJKER, 1997) e da Interação Humano-Computador (SPINUZZI, 2003). Os códigos técnicos (FEENBERG, 1995) são realizações de interesses na tecnologia. A apropriação é uma maneira que quem usa a tecnologia tem de interferir no desenho de uma tecnologia; é através da apropriação que artefatos tecnológicos alcançam sua estabilização (PINCH; BIJKER, 1997), que é quando um artefato é finalmente incorporado ao cotidiano<sup>7</sup>. Esses conceitos foram relacionados ao conceito de sistemas abertos (SPINUZZI, 2003), que são aqueles que facilitariam a apropriação por parte de quem os usa.

O estudo de teorias jornalísticas concentrou-se na obra de Nelson Traquina (2005) e o de gêneros jornalísticos se concentrou na tese de doutorado de Lia Seixas (2009) e na coletânea coordenada pelos pesquisadores Francisco Assis e José Marques de

---

<sup>7</sup> Esse processo também é chamado de fechamento – *closure*. Aqui foi preferido o termo estabilização, para evitar confusões com outros termos similares.

Melo (COSTA, 2010; COSTA; LUCHT, 2010). Percepções específicas sobre o trabalho com dados vieram do Manual de Jornalismo de Dados (BOUNEGRU; CHAMBERS; GRAY, 2012) e de um artigo e tese de doutorado de Marcelo Träsel (2013; 2014).

As visualizações foram analisadas com autores relacionados ao design da informação (BERTIN, 1986; TUFTE, 1990), da visualização (FRY, 2008; YAU, 2011) e do Jornalismo Guiado por Dados (BRADSHAW, 2011; SILVER, 2014).

#### 1.4 METODOLOGIA

Em um primeiro momento, principalmente entre Julho de 2013 e Outubro de 2014, foi feita pesquisa bibliográfica. A coleta de dados se concentrou no segundo semestre de 2014. Em sintonia com a bibliografia dos Estudos Culturais, que se interessam pelo particular, pelo pessoal, foram preferidos métodos qualitativos: entrevistas despadronizadas, questionários, observações e observações participantes.

##### *Entrevistas*

No segundo semestre de 2014, foi iniciada a busca por sujeitos para compor uma amostra não probabilística intencional (LAKATOS; MARCONI, 1985, P. 37–38), já que “o pesquisador está interessado na opinião (ação, intenção etc.) de determinados elementos da população, mas não representativos dela”. O método escolhido foi a entrevista padronizada (LAKATOS; MARCONI, 1985, P. 80–88), seguindo o roteiro do Apêndice B. As entrevistas foram registradas em áudio e os trechos mais importantes foram transcritos para uso na presente dissertação. As pessoas entrevistadas foram:

- **Renata Silva Pinto**, entrevistada em 7 de julho de 2014. Quando entrevistada, estudava jornalismo na Universidade Positivo, e trabalhava com JGD no Núcleo de Jornalismo Investigativo da universidade. Foi contatada durante oficina de Data Scraping na mesma universidade. Entrevista realizada via videoconferência, com o aplicativo Skype.
- **Guilherme Storck**, entrevistado em 24 de julho de 2014. Infografista, que trabalha com JGD e visualizações no jornal

Gazeta do Povo, de Curitiba. O contato foi estabelecido por ter sido colega de trabalho do autor. Entrevista realizada via áudio, com o aplicativo Skype.

- **Ricardo Ruas Silva**, entrevistado em 20 de agosto de 2014. Analista de sistemas, trabalhou com visualizações no Atlas do Desenvolvimento Humano no Brasil (PNUD, 2013). Entrevista realizada via videoconferência, com o aplicativo Skype.
- **João Guilherme Frey**, entrevistado em 16 de outubro de 2014. Jornalista, faz parte do projeto Livre.jor, que trabalha com jornalismo de dados públicos. Foi entrevistado pessoalmente, em encontro com seu colega José.
- **José Lazaro Barros Junior**, entrevistado em 16 de outubro de 2014. Jornalista, também faz parte do Livre.jor. Foi entrevistado pessoalmente, em encontro com seu colega João.

Além das pessoas entrevistadas, foram contatados mais quatro profissionais, sendo dois contatados por correio eletrônico e dois pessoalmente, reforçando o convite por correio eletrônico.

Apesar de não seguir a metodologia adotada, são incluídas referências à entrevista feita por correio eletrônico com Bob Smith, estudante de ensino médio estadunidense que produz mapeamentos em vídeo usando softwares básicos. Foi feito contato com ele através do sistema de mensagens do YouTube. A resposta foi enviada por ele no dia 10 de fevereiro de 2014.

### *Entrevistas com docentes*

Em outra frente de pesquisa, foram entrevistados docentes que abordam JGD em cursos de jornalismo. A intenção foi ouvir seus pontos de vista sobre a apropriação que os alunos e as alunas fazem das tecnologias e também para que relatassem experiências em sala de aula. Essas entrevistas foram realizadas por correio eletrônico, com questionário fechado (ver Apêndice C). O método foi sugerido pelos próprios entrevistados, para que pudessem respondê-lo em seu tempo livre. Foram feitas três entrevistas:

- **Rosiane Correia de Freitas**, da Universidade Positivo, de Curitiba. Enviou as respostas em 26 de agosto de 2014.
- **Marcelo Träsel**, da Pontifícia Universidade Católica, de Porto Alegre. Enviou as respostas em 27 de junho de 2014.

- **Ricardo Fotios**, dá aulas na Escola Superior de Propaganda e Marketing e na Universidade Metodista, ambas na cidade de São Paulo. Enviou as respostas em 5 de agosto de 2014.

### *Observações*

Na entrevista com os integrantes do Livre.jor foi combinado que seriam feitas observações com cada um deles por separado, em suas casas, que é onde eles realizam as atividades do Livre.jor. As observações foram registradas em caderno de campo, e referenciadas aqui como entrevistas, quando registrado o diálogo com o pesquisador, e também como notas de campo, quando foi algo observado.

A observação com José foi feita em 18 de novembro de 2014, no começo da tarde. Como ele segue um método de trabalho mais rígido, a maior parte do encontro consistiu na explicação desses métodos.

A observação com João foi feita em 4 de dezembro de 2014, no fim da tarde. Consistiu principalmente na apresentação de diferentes ferramentas e bancos de dados usados por ele.

Depois dessas observações, foi considerado fazer do Livre.jor o principal objeto de estudo. Mas pelo pouco tempo disponível, pela relevância dos dados coletados anteriormente, pela semelhança da rotina de trabalho e pelas atividades serem exercidas nas casas dos entrevistados, foi decidido que não seriam feitas novas observações.

### *Observação participante*

Foram incluídas observações coletadas durante a participação do autor como aluno nas oficinas de Data Scraping<sup>8</sup> ministradas na Universidade Positivo. As oficinas foram organizadas pela professora Rosiane Correia de Freitas, que concordou em que fossem feitas observações durante as aulas. A primeira observação foi feita no segundo encontro do grupo, no dia 26 de março de 2014, à noite. Foram acompanhadas cinco aulas, e as observações foram

---

<sup>8</sup> Também chamado de "raspagem", o termo se refere à extração de dados diretamente do código-fonte, na maioria dos casos em HTML, de páginas da internet.

registradas em um documento de texto salvo na plataforma Google Drive<sup>9</sup>.

A oficina de raspagem de dados deu origem a um Hackaton<sup>10</sup>, que aconteceu no feriado de 1º de novembro de 2014, no qual o autor também participou.

### *Outras observações*

Ao longo do trabalho também foram consideradas observações coletadas em oficinas de JGD e visualizações ministradas pelo autor na Universidade Positivo (ministrada em março de 2013) e na Universidade Federal do Paraná (ministrada em julho de 2013), assim como em uma oficina ministrada a alunas do Centro Universitário Uninter (em julho de 2013). Esta última foi feita em caráter experimental, já que foi considerado que esse tipo de oficina poderia ser o principal objeto de pesquisa, ideia que foi descartada pela pouca recepção por parte de estudantes e por não representar exatamente o cotidiano de quem trabalha com JGD, que é o foco principal do trabalho.

## **1.5 ESTRUTURA DA DISSERTAÇÃO**

Em primeiro lugar, no Capítulo 2, o cotidiano é situado como uma negociação de relações de poder. Para isso, são utilizados os conceitos de estratégias e táticas, de Michel de Certeau. Também é apresentada a Construção Social Da Tecnologia – SCOT, na sigla em inglês –, que considera que quem vai terminar de dar forma a artefatos tecnológicos, se apropriando deles apesar das relações de poder, são as pessoas que fazem uso deles, levando-os ao estágio de estabilização.

O Capítulo 3 trata do papel dos grandes conjuntos de dados e das tecnologias computacionais no jornalismo, questionando um possível caráter imparcial e científico do Jornalismo Guiado por Dados, a função dos grandes conjuntos de dados e das

---

<sup>9</sup> Plataforma para armazenamento e edições de arquivos de texto, planilhas, entre outros. Disponível em <<http://drive.google.com>>.

<sup>10</sup> Evento, tipicamente de duração de um dia, no qual profissionais da computação – e jornalismo, nesse caso – se reúnem para desenvolvimento de produtos, geralmente aplicativos, páginas da internet, etc.

visualizações nos diferentes gêneros jornalísticos, assim como um breve histórico da presença de computadores na atividade. Também são apresentadas noções de como artefatos tecnológicos têm importância na construção de identidades relacionadas ao jornalismo.

As visualizações são discutidas no Capítulo 4, que começa ponderando sobre como o JGD propicia outros tipos de relações entre jornalismo e público. Depois, são apresentadas possibilidades específicas da visualização não só como produto mas também como método dentro do JGD, considerando as suas possibilidades para a análise e processamento de grandes conjuntos de dados. São analisadas as diferentes formas de leitura proporcionadas por uma visualização, e apresentada uma proposta de trabalho comum ao JGD e às visualizações. São apresentados, de maneira descritiva, exemplos de uso de ferramentas citadas pelos entrevistados e por outras fontes, além de considerações sobre a suposta necessidade que jornalistas têm de aprender programação. Também são apresentados exemplos de visualizações de caráter jornalístico.

O Capítulo 5 compila as entrevistas e observações especificadas na seção que trata da Metodologia, assim como análises dos testemunhos tendo como base as teorias apresentadas nos capítulos anteriores.

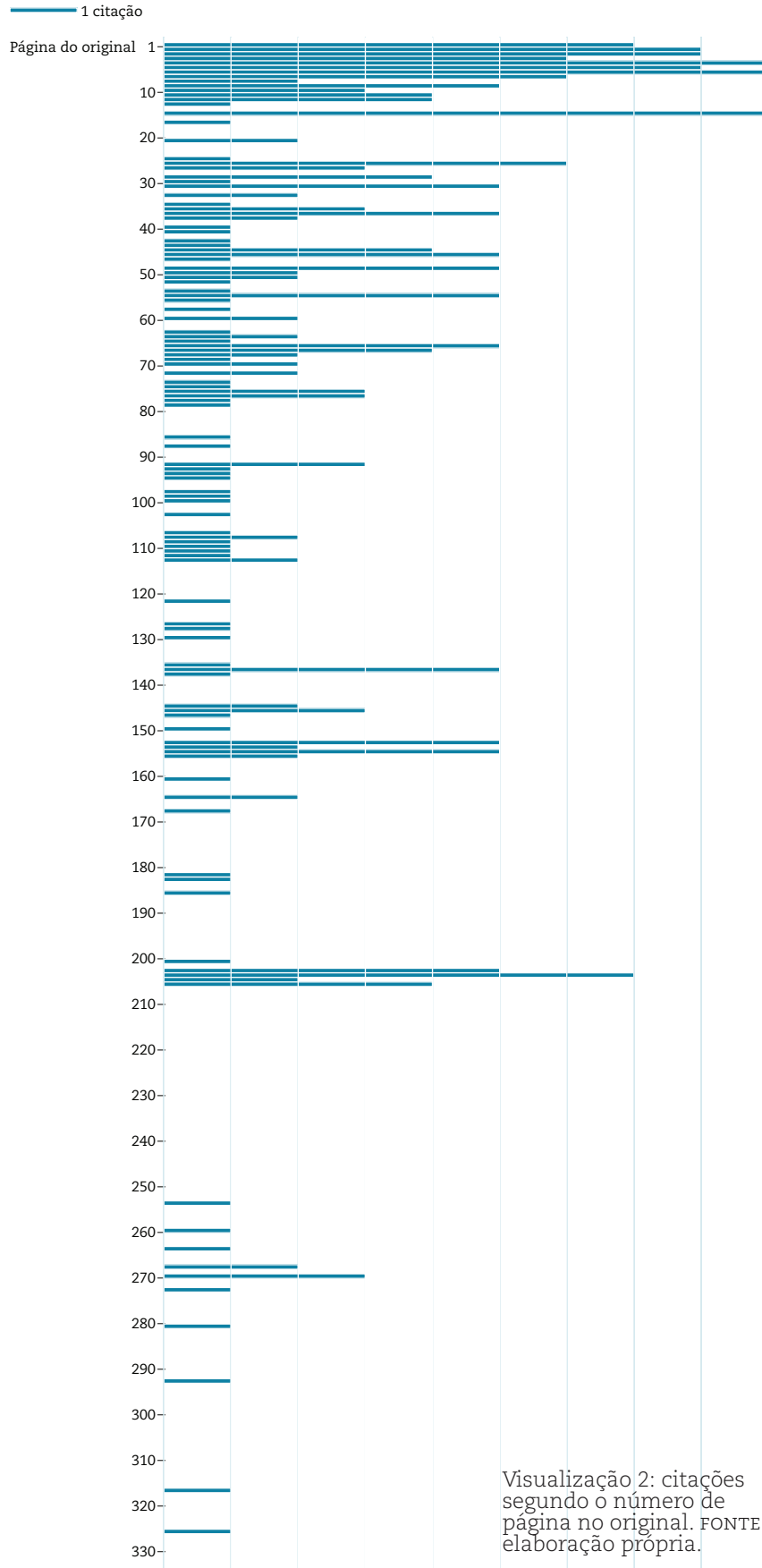
As Considerações Finais seguem apêndices com um glossário das ferramentas citadas ao longo do texto, questionários aplicados, e a descrição de como foram produzidas as “visualizações ilustrativas”.

## **1.6 SOBRE AS VISUALIZAÇÕES ILUSTRATIVAS**

Com o objetivo de situar o autor também como produtor de visualizações, foram feitas as visualizações ilustrativas que antecedem cada capítulo. Essas visualizações, de caráter lúdico, foram feitas com base em informações retiradas do conteúdo da presente dissertação. Para diferenciá-las das figuras referenciadas no texto, optou-se por usar o termo “Visualização” em vez de “Fig.”.



Os métodos usados para fazer essas visualizações são descritos no Apêndice D. O autor não defende que esses métodos sejam os mais adequados para esses propósitos, e alerta inclusive para imprecisões e problemas técnicos. Porém, a heterogeneidade dos métodos e a falta de técnica e conhecimento que podem ser visíveis na execução de alguns códigos diz respeito à maneira táctica de se apropriar das ferramentas, aqui defendida.



# 2 TECNOLOGIAS E JORNALISTAS

Este capítulo considera o cotidiano como a situação na qual as tecnologias são apropriadas através da negociação entre hierarquias de poder, considerando que quem faz uso de tecnologias não o faz de uma maneira passiva e sim através do aproveitamento de brechas, ocasiões, combinações de elementos heterogêneos, no que Michel de Certeau (2007) chama de táticas. Essa apropriação cotidiana também é parte importante da tecnologia, já que segundo a *Social Construction of Technology – SCOT* –, o processo de desenho de um artefato não termina quando determinado por quem o projeta e sim quando as pessoas se apropriam e criam significados para ele através do uso. Também são consideradas possíveis abordagens no projeto de tecnologias que propiciem que quem as use possa modificá-las através da apropriação, dentro do conceito de “sistemas abertos” (SPINUZZI, 2003).

## 2.1 COTIDIANO, ESTRATÉGIAS E TÁTICAS

O jornalismo tem que lidar diariamente com situações nas quais relações de poder ficam evidentes. Por exemplo, há instituições de pesquisa que priorizam jornais de circulação nacional, confiando a eles informações embargadas<sup>11</sup> para que possam produzir matérias mais aprofundadas, em detrimento de jornais menores. Esses meios menores recebem o material somente quando divulgado oficialmente, geralmente em uma sexta-feira, o que impede uma análise mais cuidadosa. Também há páginas de divulgação de documentos oficiais que não facilitam a consulta, como relatado pelos integrantes do Livre.jor [BARROS

---

<sup>11</sup> O termo “embargado” se refere aos materiais entregues a jornalistas sob a condição que não sejam publicados antes de uma certa data. A prática foi observada pelo autor durante seu trabalho no jornal Gazeta do Povo.

JUNIOR; FREY, 2014]<sup>12</sup>. Ou o entrevistado que concede uma entrevista em um lugar barulhento, o que impede a gravação da entrevista, deixando assim o ou a repórter vulnerável a questionamentos caso o entrevistado não goste de algo que foi publicado<sup>13</sup>. Em organizações maiores, há relações de hierarquia entre editores e repórteres, ou entre o departamento de jornalismo e os acionistas, ou entre jornalistas e departamentos de informática que impõem padrões, como a escolha dos computadores e aplicativos que serão disponibilizados aos funcionários.

A esta pesquisa interessam especialmente as relações de poder entre quem produz jornalismo e as tecnologias usadas para essa produção. Para situar essas relações em um contexto maior, será trazido da Interação Humano-Computador (SPINUZZI, 2003, p. 30) a divisão dos níveis nos quais as atividades podem ser analisadas. Resumidamente, o nível macroscópico é o da organização, do contexto; mesoscópico é o da função do usuário dentro da organização; e o microscópico é o operacional, da instância do clique (SPINUZZI, 2003, p. 30). Muitas abordagens da IHC tendem a se concentrar em somente um deles (KUUTTI; BANNON *apud* SPINUZZI, 2003, p. 36)<sup>14</sup>, mas para a compreensão do funcionamento do sistema é importante considerar os três níveis, além de diferenciar “como eles interagem, como eles se co-constituem mutuamente, e como as inovações, em qualquer nível, afetam os outros níveis” (SPINUZZI, 2003, p. 27)<sup>15</sup>. Por exemplo, uma ação no nível microscópico – uma operação, um ato inconsciente (SPINUZZI, 2003, p. 34) é feita sob certas condições, que pertencem ao nível mesoscópico, e ambos são influenciados pelo contexto no qual são feitas – o nível macroscópico, que por sua vez é influenciado pelos outros níveis. Dentro desses três níveis, interessam particularmente a esta pesquisa os níveis nos quais as

---

<sup>12</sup> Quando colocadas entre colchetes, as referências são das entrevistas feitas pelo autor da pesquisa.

<sup>13</sup> Comportamento narrado por um jornalista ao autor.

<sup>14</sup> KUUTTI, Kari; BANNON, Liam. Searching for unity among diversity: Exploring the “interface” concept. In: Human factors in computing systems: CHI '93 conference proceedings, 263–268. New York: ACM, 1993.

<sup>15</sup> Tradução livre. No original: “discern how they interact, how they coconstitute each other, and how innovations at any given level affect the other levels.”

formas de uso das pessoas se fazem presentes, que são os níveis meso e micro, sem desconsiderar as implicações do nível macro.

Retornando às dinâmicas das relações de poder, para começar a compreendê-las serão utilizados os conceitos usados por Michel de Certeau em sua obra “A Invenção do cotidiano” (2007): estratégias e táticas. De Certeau parte da análise de espaços panópticos<sup>16</sup> – espaços de controle da população, como prisões e hospitais – para entender como as regras estabelecidas para o funcionamento e controle desses lugares organizaram um “espaço do discurso” (DE CERTEAU, 2007, P. 113), no qual se manifestam – ou situam – hierarquias de poder. A organização desse espaço do discurso é perpetrada por procedimentos não discursivos e tecnologias que passaram por um “desenvolvimento privilegiado” (DE CERTEAU, 2007, P. 114). Certeau prefere descrever esses procedimentos através de alguns termos usados por Foucault: dispositivos, instrumentos, técnicas, mecanismos, máquinas (2007, P. 112). O resultado são situações de poder nas quais o exercício desse poder pode ser evidente e ter a figura de uma pessoa ou uma organização; mas também pode ser um poder

opaco, sem dono, sem lugar privilegiado, sem superior ou inferior, sem atividade repressiva nem dogmatismo, eficaz, quase autonomamente, pela capacidade tecnológica para distribuir, classificar, analisar e individualizar espacialmente o objeto tratado. (DE CERTEAU, 1990, P. 77)<sup>17</sup>

Nesse contexto, De Certeau identifica duas forças que se opõem nas relações de poder: a **estratégia**, que é o que detém algum tipo de poder, e a **tática**, que é a força que vai tentar se opor a esse poder.

A estratégia pressupõe identificar um sujeito, o que pode acontecer mesmo em uma situação na qual o poder é “opaco”:

Chamo de estratégia o cálculo (ou a manipulação) das relações de forças que se torna possível a partir do momento em que um sujeito de querer e poder (uma empresa, um exército, uma cidade, uma instituição científica) pode ser isolado. (DE CERTEAU, 2007, P. 99)

<sup>16</sup> De Certeau parte da análise desses espaços feita por Michel Foucault, especificamente no livro *Vigiar e Punir*.

<sup>17</sup> Tradução livre. No original: "opaque, sans possesseur, sans lieu privilégié, sans supérieurs ni inférieurs, sans activité répressive ni dogmatisme, efficace d'une façon quasi autonome par sa capacité technologique de distribuer, classer, analyser et individualiser spatialement l'objet traité."

Esse isolamento é facilitado por outra característica da estratégia, que é ter um lugar próprio:

A estratégia postula um lugar suscetível de ser circunscrito como algo próprio e ser a base de onde se podem gerir as relações com uma exterioridade de alvos ou ameaças (os clientes ou os concorrentes, os inimigos, o campo em torno da cidade, os objetivos e objetos da pesquisa etc.). (DE CERTEAU, 2007, P. 99)

Apesar de que a estratégia é um lugar “ordenado pelas técnicas de organização de sistemas” (1990, P. 57)<sup>18</sup>, em tanto que *lugar*, pode ser percorrido, navegado, explorado, rodeado, ocupado. Essas ações são conseguidas através das táticas, que são forças exteriores, sem lugar; são as forças “do outro”. Certeau define tática como

a ação calculada que é determinada pela ausência de um próprio. Então nenhuma delimitação de fora lhe fornece a condição de autonomia. A tática não tem por lugar senão o do outro. E por isso deve jogar com o terreno que lhe é imposto tal como o organiza a lei de uma força estranha. (DE CERTEAU, 2007, P. 100)

Se a estratégia tem um lugar, a tática tem um tempo. Já que a tática “não dispõe de base onde capitalizar os seus proveitos, preparar suas expansões e assegurar uma independência em face das circunstâncias” (DE CERTEAU, 2007, P. 46). O equivalente ao “próprio” da tática é a vitória conseguida através da espera por oportunidades para poder jogar com os acontecimentos e através da combinação de “elementos heterogêneos e móveis” (2007, P. 47).

É através dessas oportunidades que a tática consegue contornar relações de poder, como quando diante da dificuldade de acesso a dados oficiais é usada uma combinação de ferramentas e conhecimentos para obtê-los de maneiras não oficiais, como é a raspagem de páginas oficiais. Outra situação observada durante a pesquisa foi uma pessoa que se apressou a copiar em seu computador informações que foram disponibilizadas inadvertidamente pelo governo em uma ocasião específica.

Justamente pelo seu “não-lugar”, a tática tem uma mobilidade, uma flexibilidade para aproveitar – ou inclusive criar – uma

---

<sup>18</sup> Tradução livre. No original: “ordonné par les techniques organisatrices de systèmes”.

oportunidade (DE CERTEAU, 2007, P. 61), tirando partido de “forças que lhe são estranhas” (DE CERTEAU, 2007, P. 46) através da criatividade das pessoas, que “instaura *pluralidade* e criatividade” para obter “efeitos imprevistos” (DE CERTEAU, 2007, P. 93, GRIFOS NO ORIGINAL).

O cotidiano “se inventa com mil maneiras de *caça não autorizada*” (DE CERTEAU, 2007, P. 38, GRIFOS NO ORIGINAL), e catalogar algumas dessas maneiras é a intenção do segundo volume de “A invenção do cotidiano” (DE CERTEAU; GIARD; MAYOL, 1996), que narra pesquisas sobre o modo de vida das pessoas: como elas ocupam suas casas, como elas cozinham, onde compram comida e como usam os objetos do dia a dia. Ele cita o caso de Irene, que comprou assadeiras bonitas para levar à mesa, dispensando assim o uso de travessas (DE CERTEAU; GIARD; MAYOL, 1996, P. 327).

É esse tipo de situação e tática que interessa à presente dissertação, mas no âmbito do JGD e das visualizações: o cotidiano de quem trabalha com obtenção de dados, possivelmente através de táticas, como esses dados são cruzados usando conhecimentos obtidos em diversas fontes, como as ferramentas são subvertidas e como esses dados são utilizados e visualizados. As técnicas usadas no JGD também propiciam o surgimento de novas táticas, através de novas oportunidades que se tornam evidentes, permitindo assim mais possibilidades de apropriação de tecnologias e um maior aproveitamento de ocasiões a serem contornadas.

É em situações como essas que as táticas aparecem, através das maneiras de fazer, das diferentes combinações de procedimentos, de ações minúsculas através das quais o consumidor contorna as imposições das estratégias e rescreve as regras que lhe foram impostas, e se espalha “nas regiões definidas e ocupadas pelos sistemas de ‘produção’” (DE CERTEAU, 2007, P. 39). As táticas dão ao usuário aquilo que não deveria lhe pertencer:

[u]ma criança ainda rabisca e suja o livro escolar; mesmo que receba um castigo por esse crime, a criança ganha um espaço, assina aí sua existência de autor. (DE CERTEAU, 2007, P. 94)

O espaço que essa criança ganhou ao rabisca e sujar seu livro é um pequeno sucesso que provoca uma euforia no usuário (DE CERTEAU, 2007, P. 47). Para quem começa a trabalhar com JGD,

esses pequenos sucessos são estímulos para continuar o aprendizado. Durante os encontros da oficina de programação para *Data Scraping*, na Universidade Positivo, foi possível localizar várias dessas “euforias” das e dos jornalistas:

O primeiro encontro se caracterizou pela surpresa que os e as jornalistas tiveram com as possibilidades da programação, mesmo com funções básicas. O professor mostrou o comando *puts*, que resultou em um “Hello World”<sup>19</sup>, e também que o site retornava resultados de operações simples, como 2+2, o que ocasiona surpresa nos alunos e alunas. Ele insere os comandos *idade=19* e depois *puts=idade*. O programa retorna o valor da variável *idade*, o que provocou uma interjeição de surpresa e admiração. “Isso pra dados é muito bom!”, exclamou uma aluna. Os alunos e alunas também tentavam adivinhar o funcionamento da linguagem, sugerindo o uso de aspas e colchetes para escrever comandos, e perguntavam se poderiam fazer operações matemáticas com as variáveis. (NOTAS DO DIÁRIO DE CAMPO, 26/3/2014, 2º ENCONTRO DA OFICINA DE DATA SCRAPING, UNIVERSIDADE POSITIVO)

É através de pequenas vitórias do cotidiano como essas que as pessoas marcam uma presença na tecnologia. E essa presença é parte importante do desenvolvimento de artefatos tecnológicos, como visto a seguir.

Mackay situa quem consome tecnologias em um papel ativo em “dar forma a artefatos tecnológicos e culturais e seus significados” (MACKAY, 1997B, P. 4)<sup>20</sup>. Com isso, ele situa a tecnologia não como algo passivo e isolado (MACKAY, 1997A, P. 265) e sim como artefato cultural, situado em diferentes contextos. Isso evita uma visão determinista de atribuir à tecnologia os “efeitos” causados pelo uso e (re)apropriação, possivelmente por meio de táticas, por parte de quem as consome (MACKAY, 1997A, P. 261):

o computador doméstico nos mostra que a forma da tecnologia não é inevitável, (pre)determinada ou imutável. A tecnologia é cultural – não somente porque existe em um contexto cultural, com artefatos tecnológicos rodeados por cultura, mas é cultural de qualquer maneira, em seu desenho, seu significado, seu uso, e, conseqüentemente, em sua própria forma (MACKAY, 1997A, P.268)<sup>21</sup>

Os Estudos Culturais concebem que produção e consumo se constituem mutuamente e não são atividades que se opõem

<sup>19</sup> Frase típica usada em exemplos básicos de programação.

<sup>20</sup> Tradução livre. No original: “shaping technological and cultural artifacts and their meanings.”

<sup>21</sup> Tradução livre. No original: “the home computer shows us that the form of the technology is not inevitable, (pre-)determined or immutable. The technology is cultural – not simply in that it exists in a cultural context, that technological artifacts are surrounded by culture, but is cultural through and through, in its design, its meaning, its use and thus its very form.”



(MACKAY, 1997B, P. 4) e sim que podem ser concebidas como um único ato:

O consumo não é o final de um processo, mas o começo de outro, e também uma forma de produção (e então podemos nos referir ao “trabalho” de consumir). (MACKAY, 1997B, P. 7)<sup>22</sup>

Essa continuidade de processos está representada no “circuito cultural”, nome dado às articulações entre os cinco processos – consumo, produção, regulação, representação e identidade – que podem ser considerados para a compreensão de artefatos culturais (DU GAY *et al.*, 2003, P. 3). Esses cinco processos são considerados momentos de igual importância e com ligações complexas, mesmo quando presentes em um mesmo ato:

Nossa noção do “circuito cultural” nos força a considerar o consumo e a produção de maneira simétrica, e considerar a complexidade (que não é unilinear nem de mão única) das ligações entre os dois. E uma abordagem dos estudos culturais nos alerta para o significado *simbólico* (em oposição ao simplesmente funcional) das tecnologias. (MACKAY, 1997A, P. 269, GRIFOS NO ORIGINAL)<sup>23</sup>

Nesse “significado simbólico das tecnologias” é possível considerar um dos processos do circuito cultural, que é a construção de identidades, que pode acontecer de várias maneiras, inclusive através do consumo de tecnologias (MACKAY, 1997B, P. 2). Ao conceber jornalistas como “consumidores” de tecnologias computacionais, também são concebidos como “produtores” não só de produtos jornalísticos ou visualizações, mas de identidades e significados.

Consumir – ou não – certas tecnologias pode situar jornalistas em algum ponto entre os estereótipos citados por Cairo (2012) de jornalistas que são “escritor-humanista” ou o “pesquisador-cientista”, ou talvez levá-lo a ser um “repórter de computador” (MEYER, 1999, P. 5). Não é a intenção deste trabalho analisar essas identidades em particular, mas sim entender como elas – e outras

<sup>22</sup> Tradução livre. No original: “Consumption is not the end of a process, but the beginning of another, and thus itself a form of production (and hence we can refer to the ‘work’ of consumption).”

<sup>23</sup> Tradução livre. No original: “Our notion of the ‘cultural circuit’ forces us to consider symmetrically the realms of consumption and production, and to take account of the complex (not unilinear or one-way) linkages between the two. And a cultural studies approach alerts us to the *symbolic* (as opposed to simply functional) significance of Technologies.”

– podem ser produzidas através do consumo da tecnologia, e não da tecnologia por si só (ver seção 3.5).

É necessário então entender o contexto no qual as tecnologias são consumidas. No caso desta pesquisa, esse contexto é o cotidiano de quem trabalha com JGD: seus “modos de fazer” e como se apropriam das tecnologias computacionais no dia a dia, seja dentro de uma organização ou de maneira independente.

A importância do estudo do consumo em situações do cotidiano foi defendida por Mackay (1997B, p. 6), especificamente quando feita através de estudos de casos e métodos etnográficos. O cotidiano é um conceito chave dos Estudos Culturais, já que

o termo [cotidiano] expressa a noção que devemos conceber da cultura, no sentido antropológico, como comum. “Cultura” se refere àquilo que todos fazemos, do que todos fazemos parte [...]. Então focar no “cotidiano” dirige nossa atenção para “o comum” – nossos processos cotidianos de produzir significados e compreender o mundo. (MACKAY, 1997B, p. 7)<sup>24</sup>

Esse caráter etnográfico colaborou para que os pesquisadores encontrassem que os consumidores,

mais do que serem passivos e facilmente manipulados, [os pesquisadores] encontraram que consumidores jovens eram ativos, criativos e críticos na sua apropriação e transformação de artefatos materiais. Em um processo de bricolagem, eles se apropriaram, reacentuaram, rearticularam ou transcodificaram o material da cultura de massa para suas próprias finalidades, através de um espectro de práticas do cotidiano criativas e simbólicas. Através desses processos de apropriação, identidades são construídas. (MACKAY, 1997B, p. 6)<sup>25</sup>

O foco dos Estudos Culturais, como já visto com De Certeau, é o “impredizível, o improvisado e com as atividades rotineiras e o controle das pessoas comuns durante seu dia a dia” (MACKAY, 1997B, p. 7)<sup>26</sup>. É no cotidiano, nos níveis meso e microscópico (SPINUZZI, 2003, p. 30) que aparecem ações impredizíveis, táticas,

<sup>24</sup> Tradução livre. No original: “the term expresses the notion that we should conceive of culture, in an anthropological sense, as ordinary. ‘Culture’ refers to what we all do, what we all take part in [...]. So a focus in ‘everyday life’ directs our attention to ‘the ordinary’ – our everyday processes of making meanings and making sense of the world.”

<sup>25</sup> Tradução livre. No original: “Rather than being passive and easily manipulated, they found that young consumers were active, creative and critical in their appropriation and transformation of material artifacts. In a process of bricolage, they appropriated, re-accented, rearticulated or trans-coded the material of mass culture to their own ends, through a range of everyday creative and symbolic practices. Through such processes of appropriation, identities are constructed.”

<sup>26</sup> Tradução livre. No original: “the unpredictable, the improvised and with the routine activities and control of ordinary people as they go about their day-to-day lives.”

negociações com a estratégia. E é através dessas ações que a tecnologia termina de ser desenhada, mesmo depois dos artefatos terem saído das pranchetas.

## 2.2 TECNOLOGIA COMO CONSTRUÇÃO SOCIAL

As ações do cotidiano, os modos de uso, não são considerados no instrumentalismo, que é uma visão segundo a qual a tecnologia é neutra e, conseqüentemente, controlada pela humanidade:

O instrumentalismo afirma que a tecnologia é neutra: como um meio transparente, não acrescenta nada substancial aos fins que ela serve mas apenas acelera sua realização, ou os realiza em uma escala maior, ou sob novas condições. Como a tecnologia é neutra, a decisão de usá-la pode ser feita apenas com bases racionais, ou seja, melhorias mensuráveis na eficiência. (FEENBERG, 1995, P. 22)<sup>27</sup>

Mensurar a tecnologia pela eficiência leva a ignorar o contexto no qual ela é utilizada, priorizando o nível macro, da organização como um todo, em detrimento dos níveis meso e micro. No jornalismo, essa medição de produtividade cresceu nas últimas três décadas, usando métodos inadequados para um trabalho intelectual (PICARD, 1998, P. 71–72) e chegando a situações como a descrita por uma pessoa que trabalhava no jornal O Estado de São Paulo:

"no meio da redação do @Estadao instalaram uma grande lâmpada que acende toda hora que a audiência do portal cai"  
#meninojornalismotámorto (SAMPAIO, 2015, ASPAS NO ORIGINAL).

Apesar dessas situações nas quais há manifestações da estratégia, da hierarquia que implica na necessidade de produção que possa ser mensurada, deixando as pessoas que trabalham com jornalismo no papel de meros produtores, há lugar para que as pessoas tornem as tecnologias suas. Mesmo considerando que a população é “fragmentada e desqualificada” (1995, P. 138), as tecnologias podem ser apropriadas de diferentes maneiras:

<sup>27</sup> Tradução livre. No original: “Instrumentalism holds that technology is neutral: like a transparent medium, it adds nothing substantive to the ends it serves but merely accelerates their realization, or realizes them on a larger scale, or under new conditions. Because technology is neutral, the decision to employ it can be made on purely rational grounds, that is, measurable improvements in efficiency.”

O construtivismo pressupõe que há muitas soluções diferentes para os problemas técnicos. [...] os estudos contemporâneos da tecnologia desafiam esse olhar [determinista e instrumentalista] e sugerem que há muitos fatores, além da eficiência, que têm um papel na escolha do desenho. A eficiência não é decisiva para explicar o sucesso ou o fracasso de desenhos alternativos, já que nos inícios de uma linha de desenvolvimento geralmente há uma disputa entre várias opções viáveis. (FEENBERG, 2005, P. 114)<sup>28</sup>

A Construção Social Da Tecnologia – Social Construction Of Technology, abreviada como SCOT – é uma abordagem que concebe o processo de desenvolvimento de artefatos tecnológicos como um processo multidirecional, no qual a participação de quem usa as tecnologias não é vista de uma forma passiva e sim como participante no processo de desenho, em sintonia com os autores já apresentados (DE CERTEAU, 2007; MACKAY, 1997; SPINUZZI, 2003). Esse conceito se opõe a um desenvolvimento linear da tecnologia, que supõe que a decisão de como artefatos serão usados é de quem os projeta (PINCH; BIJKER, 1997, P. 22).

A multidirecionalidade da SCOT se dá através da participação das pessoas, que “escolhem” entre variantes e alternativas disponíveis, no que Pinch e Bijker chamam de flexibilidade interpretativa (PINCH; BIJKER, 1997, P. 40). As pessoas – e também quem projeta os artefatos – podem atribuir vários significados à tecnologia. No caso analisado por Pinch e Bijker, as bicicletas com rodas grandes, no final do século 19, eram vistas como rápidas, mas também como inseguras – e foi este último significado que fez com que as pessoas preferissem a variante mais segura, com rodas do mesmo tamanho, desenho que direcionou o desenvolvimento posterior da bicicleta. Nesse exemplo, a tecnologia em si não é determinante; o que direcionou o desenvolvimento foram as interpretações dos grupos sociais (FEENBERG, 2010, P. 171).

Com essa participação e apropriação das pessoas, a tecnologia alcança o estágio de estabilização, que é quando os significados e funções de um artefato são finalmente definidos, se tornando uma

---

<sup>28</sup> Tradução livre. No original: “El constructivismo presupone que hay muchas soluciones diferentes para los problemas técnicos. [...] los estudios contemporâneos de la tecnología desafían esa mirada [determinista y instrumentalista] y sugieren que hay muchos factores, además de la eficiencia, que juegan un papel en la elección del diseño. La eficiencia no es decisiva para explicar el éxito o el fracaso de diseños alternativos, ya que en los inicios de una línea de desarrollo usualmente compiten varias opciones viables.”

“caixa preta” estável (FEENBERG, 1995, P. 154). A estabilização será descrita com mais detalhes na seção 2.5.

A estabilização da tecnologia é consolidada então nos códigos técnicos, que “definem o objeto em termos estritamente técnicos, em conformidade com o significado social que o próprio objeto adquiriu” (FEENBERG, 2010, P. 173). Os códigos técnicos também são interesses que são realizados de uma maneira “tecnicamente coerente” (FEENBERG, 2005, P. 114) durante as etapas nas quais uma tecnologia é projetada, e que podem inclusive serem justificados de maneira racional. Além de definir os artefatos, os códigos técnicos “refletem os valores hegemônicos e crenças que prevalecem no processo de design” (FEENBERG, 1995, P. 4)<sup>29</sup>, o que se relaciona com o que De Certeau (2007, P. 114) chama de “desenvolvimento privilegiado” de tecnologias, processo no qual são privilegiados certos interesses no desenvolvimento de artefatos.

Esses códigos técnicos já passaram por um processo de aceitação cultural e são, portanto, invisíveis e não passíveis de questionamentos sobre o que há dentro dos artefatos (FEENBERG, 1995, P. 5). Essa falta de questionamento sobre os artefatos pode levar a situações como a do Minitel, um dispositivo “fofo” que carregava códigos técnicos que Feenberg (1995, P. 162) considera maliciosos a ponto dele comparar o Minitel a um Cavalo de Troia<sup>30</sup>. Mais recentemente, tecnologias como telefones celulares ou redes sociais na internet poderiam merecer a mesma comparação, já que seus códigos técnicos implicam na apropriação de informações de quem usa esses artefatos em troca da praticidade de enviar mensagens ou compartilhar fotografias, por exemplo.

Os códigos técnicos podem ou não surgir deliberadamente. Um projeto que dê preferência a certo grupo de usuários pode ter códigos técnicos que dificultem o uso por parte de outro grupo. Em alguns casos isso pode ser deliberado, como é o caso de

---

<sup>29</sup> Tradução livre. No original: “reflect the hegemonic values and beliefs that prevail in the design process”

<sup>30</sup> O Minitel foi um serviço de vídeo-texto, disponível na França entre a década de 1980 e 2012. Nas palavras de Feenberg: “disfarçado em um dispositivo telefônico ‘fofo’, o Minitel é uma espécie de Cavalo de Troia para códigos técnicos racionalistas” (no original: “Disguised as a ‘cute’ telephonic device, the Minitel is a kind of Trojan horse for rationalistic technical codes”).

ferramentas perigosas cujo uso por crianças é propositadamente dificultado.

Nos modos de uso do computador, no computar, esses códigos técnicos podem se manifestar, por exemplo, em um design que propicie que o computador seja usado em uma escrivadinha, ou em um editor de texto que simule a impressão de um artigo acadêmico, já que foi considerado que provavelmente esse é o objetivo de quem usa um aplicativo desses. Elas também podem se manifestar de maneiras mais sutis, como na decisão da família tipográfica usada para uma interface, que pode favorecer ou prejudicar a leitura por parte de alguns grupos de pessoas.

Há códigos técnicos que resultam da priorização de um aspecto do produto durante seu projeto, como é o caso do uso observado da plataforma Tableau, usada no JGD principalmente para produzir visualizações. A plataforma Tableau consta de um aplicativo e de servidores que hospedam os infográficos nele gerados. Quem quiser contornar essas exigências precisará apelar para alguma solução não prevista – ou deliberadamente impedida – durante o projeto da plataforma. Para fazer o infográfico “Bases eleitorais” (ver seção 4.4.3), os mapas foram feitos no Tableau, mas foram salvos em arquivos PDF e então editados em programas de ilustração vetorial. O código técnico do Tableau também exigiu, durante alguns anos, que fosse usado em computadores com Windows – agora também está disponível para a plataforma Mac.

Em outros casos, o código técnico permite outros modos de uso, como é o caso do QGIS, aplicativo de cartografia que, por ser *open source* e gratuito, permite que mais pessoas possam usá-lo, em vários sistemas operacionais, além de permitir o desenvolvimento de *plug-ins* para usos não previstos pelos desenvolvedores.

A divulgação de dados do governo também é uma situação ligada ao JGD cujos artefatos podem ser analisados dessa maneira. É possível que os códigos técnicos de artefatos como portais da transparência e diários oficiais de diferentes esferas do governo realizem interesses de dificultar a utilização dos dados que eles são obrigados, por lei, a divulgar. Também é possível que sejam realizações de interesses de outras pessoas, como de quem projetou o sistema, que priorizou outros aspectos. Mas para João e

José, do Livre.jor, a dificuldade nesse acesso é deliberada: “pouquíssimos órgãos fazem do jeito mais fácil” [BARROS JUNIOR; FREY, 2014A]. “A maioria [dos sites] tem *captcha*<sup>31</sup>, e nem pra consulta de lista, pra consulta individual, é horrível”, diz José [BARROS JUNIOR; FREY, 2014A]. Outras situações são descritas por eles:

[o Diário Oficial da União<sup>32</sup>] me dá o link em azul e uma breve descrição do post. Eu não consigo ter ideia do conteúdo. Pra saber o que é, tenho que ir no link, ele me joga pra um PDF, ele me joga pra página do PDF em que ele supõe que esteja o conteúdo que não é buscável o PDF. Então eu tenho que achar visualmente... é um inferno. [BARROS JUNIOR; FREY, 2014A]

No [portal da transparência do] governo do estado, você tem que fazer um cadastro pra cada consulta e tem que saber o nome completo do funcionário, sem errar [nenhuma letra], senão você perde a busca e tem que tentar de novo. É deliberadamente difícil. [BARROS JUNIOR; FREY, 2014A]

A necessidade de cadastro, nesse e em outros sistemas, dificulta o acesso não só pelo próprio desenho do sistema<sup>33</sup> como também pela ameaça implícita em saber os dados de quem consulta, no que José e João consideram uma intimidação:

Por que o Estado precisa pedir meu endereço e meu telefone pra eu ter acesso aos dados? Pra me intimidar? Isso pega a ideia que estão sendo transparentes e torce ela de tal forma que vira uma faca apontada pro teu peito.

[...]

Se você é um curioso e quer saber de uma pessoa, você vai descobrir [as informações]. Agora, se você é alguém que quer tirar alguma informação maior desses dados, pra usar esses dados, aí você vai se ferrar. [BARROS JUNIOR; FREY, 2014A]

Segundo João e José, uma tática comum para contornar essa exigência de cadastro é a utilização de dados de pessoas famosas. Eles citaram a ironia de um jornalista que usa o CPF do governador do estado do Paraná para fazer consultas no sistema.

Situações como essa representam novas formas de opressão – mediadas por computador – que seriam possíveis desdobramentos da sociedade “epistemologicamente deficiente” à qual Feenberg

<sup>31</sup> Combinações de letras e/ou números cujas formas foram deliberadamente distorcidas, exigindo que sejam lidas e digitadas por uma pessoa, impedindo assim o acesso automatizado.

<sup>32</sup> Disponível em: <<http://portal.in.gov.br>>. Acesso em: 20/1/2015.

<sup>33</sup> Essa exigência de cadastro está prevista no artigo 10 da Lei de acesso à informação, disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/112527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm)>. Acesso em 21/1/2015.

(1995, p. 139) se refere. Novas formas de opressão implicam em novas resistências, como subverter esse controle através da abertura das “caixas pretas”:

[O computador] pode se tornar o instrumento “dos sonhos” para controlar e regular exclusivamente pelo princípio da performatividade.... Mas também pode ajudar [pessoas ou grupos de pessoas], fornecendo-lhes a informação que eles geralmente precisam para tomar decisões. A linha a seguir para que a computerização siga a segunda opção é, a princípio, bastante simples: dê ao público acesso livre à memória e aos bancos de dados. (LYOTARD *apud* FEENBERG, 1995, p. 131)<sup>34</sup>

Outra forma de opressão mediada por computador seria a impossibilidade de se apropriar de um sistema; não só da maneira narrada por João e José, mas também quando não se consegue fazer o uso se deseja fazer de um artefato porque não foi projetado para permitir diferentes modos de uso.

### 2.3 SISTEMAS ABERTOS E FECHADOS

Alguns códigos técnicos podem ser originados por diferentes abordagens do projeto, particularmente a maneira como as pessoas que vão usar o artefato são consideradas. Spinuzzi (2003, p. 13–18) critica abordagens do desenho de interfaces que põem usuárias e usuários em um papel de vítima. Nessas abordagens, projetistas são considerados como detentores do conhecimento necessário não somente para projetar um artefato mas também para dizer como usá-lo. Os consumidores não saberiam das possibilidades da tecnologia, afinal, “seu trabalho é fazer seu trabalho, não projetar sistemas” (BEYER; HOLTZBLATT *apud* SPINUZZI, 2003, p. 18)<sup>35</sup>. Essa concepção que põe o projetista de artefatos como detentor do conhecimento e quem usa os artefatos como mero receptor é relacionada por Spinuzzi à existência de sistemas

<sup>34</sup> Tradução livre. No original: “[The computer] could become the ‘dream’ instrument for controlling and regulating exclusively by the performativity principle.... But it could also aid [people or groups of people] by supplying them with the information they usually lack for making knowledgeable decisions. The line to follow for computerization to take the second of these two parts is, in principle, quite simple: give the public free access to the memory and data banks.” (LYOTARD, Jean-François. *The Postmodern Condition: A Report on Knowledge*. Minneapolis: University of Minnesota Press, 1991, p. 67).

<sup>35</sup> Tradução livre. No original: “[i]t’s their job to do their job, not to design systems” (BEYER, Hugh; HOLTZBLATT, Karen. *Contextual Design: Defining customer-centered systems*. San Francisco: Morgan Kaufmann, 1998).



nos quais o espaço para inovações do usuário é deliberadamente reduzido, no que Spinuzzi (2003, p. 6) chama de “sistema fechado”.

Sistemas fechados têm por princípio que quem deve manter o controle sobre a maneira que o sistema é usado são os projetistas. Spinuzzi nota que essa hierarquização entre o projetista e quem usa o artefato, apesar de parecer centrada no sistema, pode ser classificada como centrada no usuário. A intenção dos designers seria a construção de uma “ecologia artificial” planejada desde o topo em vez de permitir que o sistema se desenvolva naturalmente (SPINUZZI, 2003, p. 204). É importante ressaltar que sistemas fechados, na presente dissertação, não são considerados necessariamente ruins, e estima-se que em algumas situações essa abordagem de projeto possa ser considerada mais adequada que um sistema aberto.

Em um dos exemplos levantados para sua pesquisa, Spinuzzi descreve como Barbara, uma policial que em vez de usar um sistema como indica o manual – usando um mapa impresso de grandes dimensões para consultar coordenadas que devem ser informadas a um programa de computador – prefere anotar as coordenadas usadas com mais frequência em um pedaço de papel (SPINUZZI, 2003, p. 1). O modo como Barbara contorna a pouca praticidade do sistema diz respeito à atitude que usuárias e usuários podem ter quando se deparam com um problema cuja solução não foi prevista por quem projetou o sistema, o que implica, como no exemplo, a subversão através do uso de recursos externos ao sistema, como o post-it (SPINUZZI, 2003, p. 202). A apropriação continua acontecendo, apesar do sistema não a facilitar. Essa tentativa de regulação também complica a interseção de diferentes tarefas dentro de um mesmo sistema (SPINUZZI, 2003, p. 204), fazendo com que quem o usa tenha que encontrar outras formas de se apropriar dele, como nas observações feitas por Spinuzzi com trabalhadores de um departamento de trânsito:

Para fazer seu trabalho, os trabalhadores frequentemente devem adaptar gêneros adicionais, reinterpretar partes do sistema fechado à luz de suas próprias atividades, e às vezes até

abandonar por completo o sistema fechado. (SPINUZZI, 2003, P. 204)<sup>36</sup>

A discussão de Spinuzzi sobre sistemas abertos ou fechados pode ser ampliada para pensar além de sistemas computacionais. A decisão sobre que o grau de abertura de um sistema – ou artefato – é um dos fatores que compõem seu código técnico.

O Diário Oficial do governo do estado do Paraná, por exemplo, faz uso de uma linguagem jurídica que representa uma barreira para o acesso geral. Essa escolha de linguagem, junto às escolhas de como o texto é divulgado, formato de arquivo etc, são manifestações do código técnico que fazem pensar que o sistema como um todo não teve como prioridade ser usado de outra maneira – ou seja, seria um sistema fechado.

Grande parte das atividades do *Livre.jor* implicam na interpretação da linguagem usada nos diários oficiais para encontrar situações que possam ter interesse jornalístico e explicar, em linguagem mais acessível, o que há de interessante no que foi encontrado [BARROS JUNIOR; FREY, 2014]. Nessa interpretação da linguagem – e no uso feito por José e João dos sistemas – há uma negociação do código técnico que torna acessível algo que não foi concebido com essa intenção. É importante notar que, nesse caso, se trata de um profissional que trabalhou vários anos com jornalismo político, durante os quais aprendeu essa terminologia. Esse conhecimento é, nos termos de De Certeau (2007, P. 47), um dos “elementos heterogêneos e móveis” cuja combinação caracteriza as táticas.

## 2.4 APROPRIAÇÕES DAS TECNOLOGIAS

Nesse cenário de estratégias como manifestações de códigos técnicos, seja em ferramentas ou na obtenção de grandes conjuntos de dados, é fácil situar as e os jornalistas/consumidores como vítimas. Essa ideia já foi contraposta às visões de De Certeau (2007) e Spinuzzi (2003), e inclusive situando as pessoas como

---

<sup>36</sup> Tradução livre. No original: “To get their jobs done, workers often must adapt additional genres, reinterpret parts of the closed system in light of their own activities, and sometimes even abandon the closed system altogether.”

parte do processo de desenho de um artefato (PINCH; BIJKER, 1997; FEENBERG, 1995). É no cotidiano que acontece o ato de consumir e, em consequência, a apropriação e ressignificação de tecnologias, no que De Certeau chama de “fabricação”:

A “fabricação” que se quer detectar é uma produção, uma poética – mas escondida, porque ela se dissemina nas regiões definidas e ocupadas pelos sistemas de “produção” (televisiva, urbanística, comercial etc.) e porque a extensão sempre mais totalitária desses sistemas não deixa aos “consumidores” um lugar onde possam marcar o que *fazem* com os produtos. (DE CERTEAU, 2007, P. 39)

Essas práticas estão situadas em um espaço, mas principalmente em um tempo, em uma oportunidade, em aproveitar, através de táticas, brechas na estratégia, desconstruindo assim a ideia de passividade que está por trás do “pudico nome de consumidores” (DE CERTEAU, 2007, P. 38).

Spinuzzi fala na vitimização do usuário como um lugar comum na literatura de design de informação, especificamente quando se fala em design centrado no usuário: “o João ou a Maria que é oprimido por uma tirania injusta e que precisa ser resgatado” (SPINUZZI, 2003, P. 1)<sup>37</sup>. Ou, retomando o caso da policial Barbara:

Barbara não está esperando para ser resgatada. [...] Ela pega ferramentas disponíveis, as adapta de maneiras idiossincráticas, e faz acontecer. Através dessas inovações “invisíveis” [...] ela subverte o sistema de informação, inventando suas próprias maneiras de transformá-los de acordo a suas necessidades. Trabalhadores como Barbara tendem a criar suas próprias práticas, ferramentas e textos constantemente, às vezes cooperando com os sistemas de informação existentes, às vezes competindo com eles. (SPINUZZI, 2003, P. 2)<sup>38</sup>

Essas “práticas, ferramentas e textos”, típicas do uso cotidiano, é o que tira as pessoas do suposto papel de vítima. As apropriações são influenciadas pelo repertório, atitudes e referências culturais do consumidor (MACKAY, 1997A, P. 271). Por exemplo, Renata [PINTO, 2014] se sente confortável para usar a planilha Excel, e pode usar seu “repertório” quando precisa realizar novas tarefas. Isso

<sup>37</sup> Tradução livre. No original: “the everyday Joe or Jane who is oppressed by an unjust tyranny and in need of rescue.”

<sup>38</sup> Tradução livre. No original: “Barbara is not waiting around to be rescued. [...] She picks up available tools, adapts them in idiosyncratic ways, and makes do. Through these ‘invisible’ innovations [...] she subverts the information system, inventing her own ways to turn it to her needs. Workers like Barbara tend to create their own practices, tools, and texts constantly, sometimes in cooperation with the existing information systems, sometimes in competition with them.”

aconteceu, por exemplo, na ocasião em que precisou contar o número de ocorrências em uma planilha, o que exigiu o aprendizado de uma função nova, para depois enviar os dados a outra planilha, na qual usou uma função que já conhecia (NOTAS DO DIÁRIO DE CAMPO, 1º/11/2014, HACKATON NA UNIVERSIDADE POSITIVO).

Outro exemplo no qual convivem vários tipos de táticas é o procedimento de leitura que José, do Livre.jor, faz do Diário Oficial do estado do Paraná. Aqui são consideradas táticas por serem a maneira que José negocia com o poder implícito na estratégia, nos códigos técnicos das diferentes tecnologias que ele usa para conseguir seus resultados. Ele combina os seus conhecimentos da linguagem própria dos diários oficiais e de política paranaense com a escolha de ferramentas que lhe permitem, por exemplo, contabilizar o número de ocorrências de um termo e visualizar, na barra de rolagem do navegador Google Chrome (Fig. 1), se todos os termos estão na mesma seção do documento:

José começa pela leitura do índice, e já pela quantidade de páginas de cada seção ele decide em qual parte prestar atenção. No exemplo do dia 18 de novembro de 2014, eram 40 páginas de finanças. Ele está acostumado a “passar o olho” por essas páginas em busca de expressões que signifiquem alguma coisa para ele. Passa rapidamente pela seção de decretos, já que ele sabe que provavelmente não haverá nada de interesse. “Se fosse a assembleia [legislativa], eu leria com mais calma”. Ele encontra um ato de mudança de natureza da despesa, cujo código ele desconhece. “Se eu quiser achar isso tenho que olhar a pedra de Rosetta” que é como ele se referiu ao Manual Técnico da Receita Federal. Ele acha o código através do comando CTRL-F. Ele puxa o site da Receita Federal para uma janela separada, “para poder ter o ctrl-tab [e alternar as janelas]”. Acha o código que procurava e vê que se refere à Justiça.

Continuando com a leitura, lhe chama a atenção uma transferência no valor de R\$ 85 milhões. “Posso supor que é fundo de transferência do governo federal”. Deixa para ver isso depois. “É inviável olhar os diários todos no detalhe”, diz.

“Chegamos onde tem coisa quente: MP [Ministério Público].” José lê os arquivos PDF do diário usando o navegador Google Chrome, porque “me dá opções de busca melhores”. Com o comando CTRL-F, ele localiza as ocorrências do termo “inquerito civil” – em minúsculas e sem acento – e aparecem 774 ocorrências.

Um dos motivos pelos quais José lê os PDFs no Chrome é que, quando feita uma busca por um termo, aparecem pontos amarelos na barra de rolagem, representando cada ocorrência. O acúmulo de pontos amarelos indica que não há ocorrências do termo fora da seção referente ao Ministério Público.

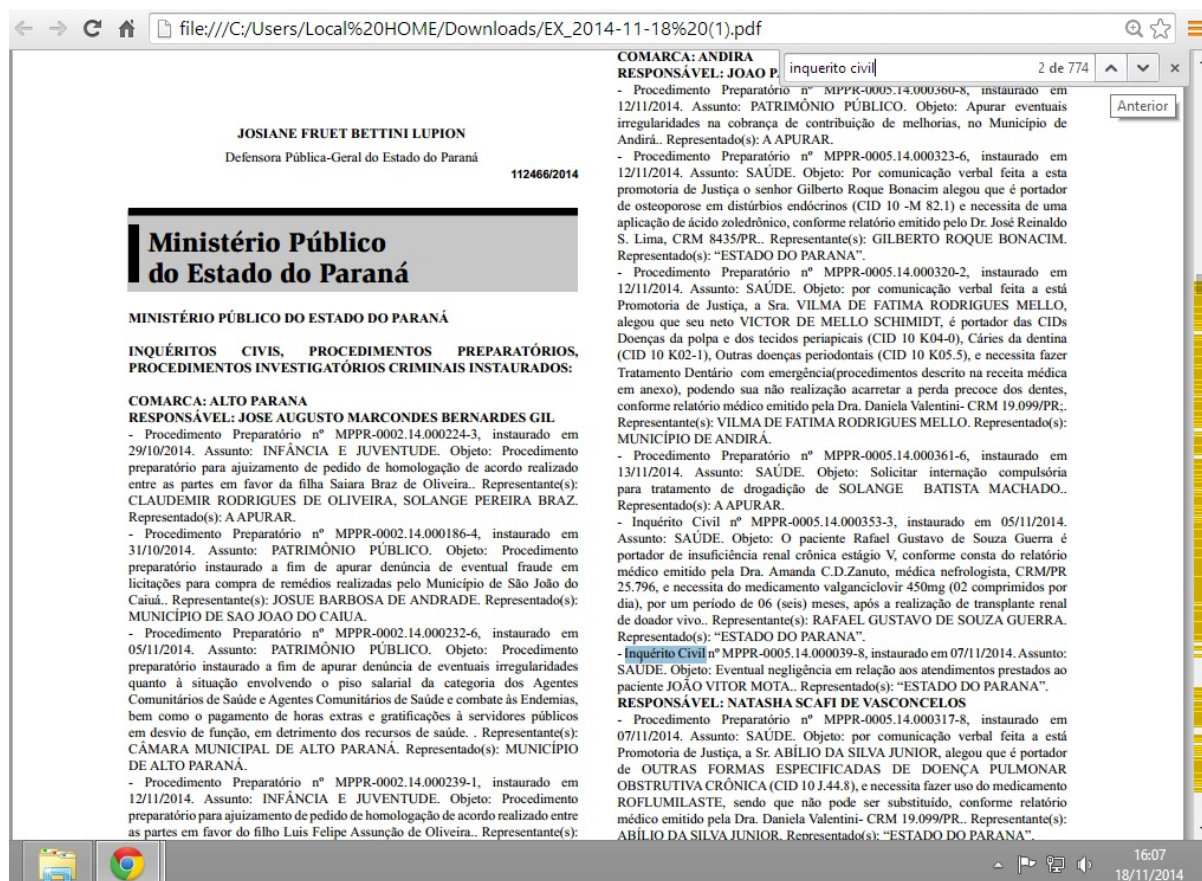


Fig. 1: Captura de tela mostrando as ocorrências do termo procurado. Fonte: BARROS JUNIOR, 2014B.

Das 774 ocorrências, ele quer extrair o número de inquéritos que foram abertos. Para isso ele seleciona a primeira ocorrência e usa o atalho CTRL-SHIFT-G para ir para a última, e segura o atalho para ir de trás pra frente até encontrar o ponto em que a listagem se divide, que ele identifica pela diagramação e tipografia. Ele vê que a última ocorrência da primeira parte é a "545 de 774", e anota esse número num papel. Ele repete o processo com a busca por "procedimento preparatorio", que retorna 492 resultados, e "procedimento investigatorio criminal", que retorna 77 resultados. Nesses dois casos, ele não precisa separar, somente visualiza se as ocorrências, marcadas em amarelo, se acumulam na parte da barra de rolagem referente ao Ministério Público. No mesmo papel em que anotou, faz a soma dos três tipos de inquérito (ele calcula 1124, mas o total é 1114) investigações do Ministério Público.

Quando encontra algo que considera que vai ser a notícia do dia do Livre.Jor, José faz uma captura da tela e a edita no Microsoft Paint – programa de desenho incluso com o sistema operacional Windows – e salva o arquivo com um nome que representa a data da publicação, no formato ano-mês-dia (141118), para facilitar possíveis buscas futuras. Ele inclui indicações de que partes da captura foram omitidas, e publica a imagem junto de um texto explicativo na página do Livre.jor no Facebook. (NOTAS DO DIÁRIO DE CAMPO, ENTREVISTA COM JOSÉ, CURITIBA, 18/11/2014)

A combinação de elementos feita por José para chegar ao resultado desejado – assim como outras combinações similares

narradas pelas pessoas entrevistadas – atenta para o caráter sócio-construtivista das táticas, relacionado com a abordagem de Feenberg (1995, p. 155), segundo a qual as funções dos artefatos são descobertas não só durante o desenvolvimento mas também durante o uso dos objetos. Para os estudos CTS, a apropriação é um elemento importante da abordagem construtivista, que, em oposição ao determinismo tecnológico, não se atém às características técnicas e de eficiência para explicar o sucesso ou fracasso de um artefato; o que conta é a sua relação com o contexto social (FEENBERG, 1995, p. 154).

A apropriação, seja de caráter tático ou dentro de limites de caráter estratégico, é uma maneira de descobrir ou redefinir as funções dos artefatos. A tendência determinista de pensar que as funções de um objeto são predeterminadas se deve a que, no caso de tecnologias já estabelecidas, pareceria que as funções dos artefatos foram predeterminadas e não socialmente construídas (FEENBERG, 1995, p. 155). Já no caso de uma tecnologia nova,

às vezes não há uma definição clara da função no começo. Como resultado, não há uma distinção clara entre diferentes tipos de significados associados com a tecnologia: uma bicicleta construída para ser segura é funcionalmente e conotativamente diferente. De fato, conotações de um design podem ser funções vistas do ponto de vista de um outro [design] . (FEENBERG, 1995, p. 156)<sup>39</sup>

Essas diferentes “conotações de um design” podem ser interpretadas como uma “vida dupla” da tecnologia:

a tecnologia tem uma vida dupla, uma de acordo com as intenções dos designers e os interesses do poder e outra que os contradiz – agindo pelas costas dos seus arquitetos para resultar em consequências não planejadas e possibilidades imprevistas. (MACKAY, 1997A, p. 274)<sup>40</sup>

Essas associações de significados a que Feenberg se referiu acontecem através da apropriação cotidiana, colaboram para uma decodificação dos códigos técnicos, para uma abertura dos

<sup>39</sup> Tradução livre. No original: “there is often no clear definition of function at first. As a result, there is no clear distinction between different types of meanings associated with the technology: a bicycle built for safety are both functionally and connotatively different. In fact, connotations of one design may be functions viewed from the angle of the other.

<sup>40</sup> Tradução livre. No original: “technology leads a double life, one which conforms to the intentions of designers and interests of power and another which contradicts them – proceeding behind the backs of their architects to yield unintended consequences and unanticipated possibilities.”

artefatos. Assim, o código técnico é decodificado e os artefatos, modificados:

Protestos ou reapropriações bem sucedidas modificam o código técnico para refletir os interesses excluídos em estágios iniciais do processo de design. (FEENBERG, 1995, P. 10)<sup>41</sup>

O código técnico pode ser legitimado, negociado, adaptado, oposto, ignorado ou até aproveitado pelo consumidor. Essa negociação também acontece em esferas menores, como o caso específico de sistemas computacionais que são impostos a um usuário:

Padrões impostos vão causar gambiarras. Porque padrões impostos não têm preocupações para cada contingência local, os usuários vão adaptar os formulários, sistemas de informação e agendas padronizados, entre outros, para que atendam às suas necessidades. (BOWKER; STAR *apud* SPINUZZI, 2003, P. 3)<sup>42</sup>

Gambiarras também são maneiras de negociar os códigos técnicos e modificar tecnologias, através de usos inovadores que não necessariamente respeitam o uso previsto no projeto (FEENBERG, 1995, P. 8). Essas modificações podem acontecer de diferentes maneiras, como através de mudanças no hardware ou software do artefato, ou simplesmente usando a tecnologia de uma maneira não prevista por quem a projetou, no que De Certeau (2007, P. 41) chama de uma “criatividade dispersa, tática e bricoladora”.

Essas práticas de modificação de artefatos levam ao que Feenberg chama de “democratização da mudança tecnológica”<sup>43</sup>. Esse processo implica em

dar acesso ao processo de projeto aos atores que não tenham capital financeiro, cultural ou político. Não há razão para pensar que a participação deles seja prejudicial, já que atores não-técnicos já estão envolvidos; a democratização simplesmente aumentaria a sua quantidade e variedade. (FEENBERG, 1995, P. 7)<sup>44</sup>

<sup>41</sup> Tradução livre. No original: “Successful protest or reappropriation modifies the technical code to reflect interests excluded at earlier stages in the design process.”

<sup>42</sup> Tradução livre. No original: “Imposed standards will produce workarounds. Because imposed standards cannot account for every local contingency, users will tailor the standardized forms, information systems, schedules, and so forth to meet their needs.” (BOWKER, Geoffrey; STAR, Susan. *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press, 1999)

<sup>43</sup> Tradução livre. No original: “democratization of technical change”.

<sup>44</sup> Tradução livre. No original: “granting actors who lack financial, cultural, or political capital access to the design process. There is no reason of principle to think that their participation

Apesar desse tipo de prática ser desencorajada por discursos que defendem a centralização do processo de design por especialistas (FEENBERG, 1995, P. 8), Feenberg advoga que a democratização da tecnologia sim é possível, e que o envolvimento do público – no caso, através da apropriação – não somente deve ser aceito como legitimado, inclusive no âmbito profissional (FEENBERG, 1995, P. 8). A democratização da mudança tecnológica é importante, mais ainda considerando um contexto no qual a vida social está cada vez mais rodeada por sistemas técnicos e as modificações da tecnologia por parte dos usuários se tornam cada vez mais importantes em um aspecto político (FEENBERG, 1995, P. 9).

Uma situação interessante, apontada por Spinuzzi, é que os usos não previstos podem ser interpretados pelo designer como “sintomas de problemas e pontos de partida para esforços de redesenho globais” (SPINUZZI, 2003, P. 16)<sup>45</sup> que terminam sendo uma maneira de “oficializar” as inovações não oficiais:

às vezes esses métodos implicam examinar as inovações dos usuários, coletar avaliações, e inclusive colaborar com os usuários para reprojeter artefatos. Mas no fim, o objetivo é transformar um conjunto bagunçado de soluções ad hoc e não-oficiais em uma única, organizada, coerente, oficial – e estática – solução generalizada. (SPINUZZI, 2003, P. 22)<sup>46</sup>

E que essa “oficialização” tem um foco na eficiência do sistema como um todo, mais do que no usuário:

designers deveriam se pôr no lugar dos trabalhadores. Mas estas histórias e cenários, e suas soluções, tendem a focar no empoderamento funcional: como melhorar a eficiência e produtividade dos trabalhadores através do redesenho de artefatos e práticas. (SPINUZZI, 2003, P. 13)<sup>47</sup>

Para Spinuzzi, essa “oficialização”, aqui interpretada como uma reapropriação das táticas, transforma soluções informais

would be detrimental, since nontechnical actors are already involved; democratization would simply increase their number and variety.”

<sup>45</sup> Tradução livre. No original: “symptoms of problems and as starting points for global redesign efforts.”

<sup>46</sup> Tradução livre. No original: “these methods sometimes entail examining user innovations, collecting feedback, and even collaborating with users to redesign artifacts. But in the end, the goal is to transform a messy set of ad hoc, unofficial solutions into a single, neat, coherent, official – and static – generalized solution.”

<sup>47</sup> Tradução livre. No original: “designers should put themselves in the workers’ shoes. But these stories and scenarios, and their solutions, tend to focus on functional empowerment: how to improve the workers’ efficiency and productivity by redesigning artifacts and practices”.



dinâmicas em algo estático. Esse processo pode aparecer não só no design do artefato mas também na forma de manuais e documentação de uso (SPINUZZI, 2003, P. 10).

Outra tensão entre tática e estratégia pode se manifestar na distribuição do conhecimento gerado pelas táticas. As táticas podem ser divulgadas de maneiras opostas à centralização de documentações oficiais e manuais. Os usuários que fazem esse tipo de divulgação foram apelidados de “jardineiros” (NARDI *apud* SPINUZZI, 2003, P. 207)<sup>48</sup>, e são quem, através de conexões informais, distribuem conhecimento também “informal”, como truques, dicas e outras formas de conhecimento. Esses “jardineiros” podem alcançar uma importância tal a ponto de terem como função principal a difusão desse conhecimento (TERVEEN *apud* SPINUZZI, 2003, P. 207)<sup>49</sup>. Nardi (*apud* SPINUZZI, 2003, P. 206) também cita o orgulho de quem cria soluções e as compartilha ou divulga.

A distribuição “tática” de conhecimento também pode ser “oficializada” de maneira similar ao que acontece com as gambiarras. Da mesma maneira, pode ser resistida pelos usuários, que, diante de instruções de difícil compreensão, priorizarão formas não oficiais de distribuição de conhecimento, como conselhos de colegas ou dicas achadas na internet, por exemplo. Guilherme [STORCK, 2014], por exemplo, diz que seu perfil de aprendizado é “meio autodidata”, o que ele define como o “esforço em buscar informação e ir aplicando ela, fazendo tentativa e erro até chegar [a um resultado].” Normalmente ele procura esse aprendizado quando tem uma tarefa específica. Partindo da necessidade de cumprir uma tarefa, ele a divide em passos e procura a solução para cada parte:

Vou fazendo buscas sobre alguém que explique como fazer isso. Então uma coisa que eu percebo nesse processo inteiro é que é muito bom você aprender a fazer buscas. Porque não é muito fácil “ah, como eu faço isso”, mas “isso” é difícil de explicar às vezes. E às vezes as pessoas têm outros termos, elas dizem de outro jeito, mas até você achar a informação que você quer, demora um pouco... é bom aprender a fazer buscas. [STORCK, 2014]

<sup>48</sup> NARDI, Bonnie; O'DAY, Vicki. Information ecologies: Using technology with heart. Cambridge, MA: MIT Press, 1999. P. 139–151.

<sup>49</sup> TERVEEN, Loren; SELFRIDGE, Peter; LONG, M. David. Living design memory: Framework, implementation, lessons learned. Human-Computer Interaction, 10, 1–37. 1995.

Guilherme cita a documentação dos processos como de importância para poder refazer tarefas que já foram feitas, ou para futura referência:

durante o processo eu vou anotando: joguei no navegador, tentei baixar e não deu certo por causa disso; tentei limpar o negócio e usei esse site. Essa anotação foi sempre importante, tanto pra eu ter noção de onde que eu estou, tanto como para quando eu precisar fazer o negócio de novo; eu preciso voltar lá e fazer a mesma tarefa e que bom que eu encontro uma anotação velha. [STORCK, 2014]

Renata [PINTO, 2014] também combina conhecimentos obtidos de diferentes maneiras. O “pontapé inicial” foi em um curso de jornalismo investigativo nos Estados Unidos, onde aprendeu Tableau, Excel e Access. Depois, ela fez um curso de Excel avançado no CIEE/PR<sup>50</sup>, no qual era a única jornalista. Além desse curso e dos encontros de jornalismo, Renata cita plataformas de aprendizado online. Ela começou vários Moocs<sup>51</sup>, mas acabou desistindo: “não consegui terminar assim. Me cadastrei em quatro cursos de graça e eu não consegui manter o ritmo.” Para o aprendizado pela internet, ela prefere a plataforma Codeacademy, na qual pode fazer exercícios mais curtos:

comecei a brincar lá, é bem didático. Daí você começa, aquela coisa, “Hello World”. Aí você vai um pouquinho, ah, faz string, chama add 3, soma a com b dá 4, aí vai. Não precisa separar uma hora pra fazer, pode entrar, se exercitar um pouco e sair, que o curso estará lá quando voltar. [PINTO, 2014]

Para aprender, ela ainda cita recursos online, como a lista de email do Necir – que usa para tirar dúvidas específicas –, um grupo online de usuários de Python, além do “nosso querido amigo Google.” Segundo ela, há um senso comunitário entre quem trabalha com JGD: “gosto bastante dessa comunidade jornalística de dados [...] o pessoal é super aberto [...] elas não se importam em passar os dados e dicas e tudo o mais” [PINTO, 2014].

---

<sup>50</sup> Centro de Integração Empresa–Escola do Paraná.

<sup>51</sup> Massive Open Online Course (Curso Online Aberto e Massivo).

### 2.4.1 Sistemas abertos: prevendo a apropriação

Tecnologias podem propiciar diferentes graus e tipos de apropriação. Para alcançar a democratização da mudança tecnológica a que Feenberg (1995, p. 7) se refere, seria desejável que os artefatos fossem projetados de modo tal que permitissem que mais pessoas pudessem não só usá-los, mas se apropriar deles, modificar seus códigos técnicos, levando os artefatos mais perto da estabilização.

O código técnico dos artefatos pode fazer com que eles sejam mais ou menos fáceis de serem usados de maneiras não previstas por quem os projetou. Spinuzzi (2003, p. 6) usa os termos “sistema aberto” e “sistema fechado” para se referir a que tanto um sistema é projetado para permitir essas maneiras de uso não previstas. Spinuzzi também relaciona essa abertura ou fechamento de um sistema aos conceitos de design centrado no usuário e design centrado no sistema. Robert Johnson (*apud* SPINUZZI, 2003, p. 7)<sup>52</sup> compara os dois em termos de empoderamento do usuário: o design centrado no sistema se caracteriza por ser formalista, racionalista, determinista, modernista, em oposição ao caráter cooperativo e construcionista do design centrado no usuário. O design centrado no usuário tem entre seus princípios que a “realidade é mutante, que não há verdades absolutas, e que o conhecimento é construído através do conhecimento criado pela comunidade e pela ação” (JOHNSON *apud* SPINUZZI, 2003, p. 8)<sup>53</sup>. O design centrado no usuário implica que ele tenha liberdade para criar e compartilhar suas próprias soluções, facilitando a apropriação da tecnologia<sup>54</sup>.

É possível projetar sistemas a partir de outra abordagem, que dê espaço para que o usuário se aproprie de uma outra maneira. Um projeto aberto, segundo Spinuzzi (2003, p. 204), consistiria de um núcleo central que pudesse ser considerado pelos usuários

---

<sup>52</sup> JOHNSON, Robert. *User-centered technology: A rhetorical theory for computers and other mundane artifacts*. New York: SUNY Press, 1998.

<sup>53</sup> Tradução livre. No original: “reality is mutable, that there are no certain truths, and that knowledge is constructed through communally created knowledge and action”.

<sup>54</sup> Para Spinuzzi, o design centrado no sistema não é predominante e sim um modelo que é citado em oposição ao design centrado no usuário como uma espécie de mau exemplo (SPINUZZI, 2003, p. 6).

como uma base para se apropriarem da forma mais conveniente para eles. Apesar de que o sistema continuaria tendo um caráter centralizador, ele prevê as apropriações dos usuários (SPINUZZI, 2003, p. 205). Outra característica de um sistema aberto é que ele seja modular, o que permite contemplar usuários com diferentes graus de conhecimento (SPINUZZI, 2003, p. 205).

Baseado na pesquisa de Bonnie Nardi, Spinuzzi cita a planilha de cálculo como exemplo de sistema aberto, já que “provê abertura para que os trabalhadores contribuam com suas próprias inovações, e oferece várias maneiras de conectar essas inovações” (NARDI *apud* SPINUZZI, 2003, p. 205)<sup>55</sup>. A planilha tem uma funcionalidade central, de fácil acesso para pessoas que queiram começar a usar o software sem precisar de muito aprendizado, e as diferentes dificuldades das funções permitem um aprendizado gradual, aplicando novos conhecimentos e adquirindo experiência no processo (NARDI *apud* SPINUZZI, 2003, p. 206) ou ainda que cada usuário saiba usar os “módulos” que precisar. Os conjuntos de funções também são modulares, agrupando funções úteis para diferentes áreas, como estatística, engenharia, finanças, trigonometria etc. Também há várias maneiras de executar as tarefas, como botões ou atalhos de teclado. Usuários que saibam de programação podem escrever macros para adaptar o programa às suas necessidades (NARDI *apud* SPINUZZI, 2003, p. 205–206). Permite a troca de dados e arquivos com outros softwares, facilitando a interseção de diferentes tarefas e permitindo a distribuição de planilhas que possam ser úteis a outras pessoas. A possibilidade de impressão das planilhas também facilita apropriações externas ao computador, como preenchê-las e marcá-las com lápis (SPINUZZI, 2003, p. 207). As planilhas têm tanto documentação oficial como tutoriais e dicas circulando de maneiras não oficiais.

Considerando o conceito de Nardi (*apud* SPINUZZI, 2003, p. 205), que afirma que planilhas eletrônicas são “sistemas de programação para o usuário [que permitem criar] aplicativos úteis

---

<sup>55</sup> Tradução livre. No original: “provides openings for workers to contribute their own innovations, and it offers several ways to connect those innovations” (NARDI, Bonnie. *A small matter of programming: Perspectives on end user computing*. Cambridge, MA: MIT Press, 1993).

em pouco tempo sem conhecimento anterior de programação<sup>56</sup>, um exemplo de uma possível apropriação da planilha de cálculo como sistema aberto é um arquivo que circula via email e está disponível em vários sites com o título “Churras Calculator” (CHURRAS, 2001). É uma planilha de cálculo feita no Microsoft® Excel, que somente permite a edição de duas células para indicar o número de convidados – homens e mulheres – para um churrasco. Inseridos os dados, a lista de compras dos itens para o churrasco são atualizados, assim como avisos sobre se o churrasco seria “de família” ou “da galera” e também alertas sobre a possível disparidade entre convidados e convidadas (Fig. 2). O arquivo encontrado data de 2001; atualmente há sites e aplicativos para smartphones que fazem a mesma operação<sup>57</sup>. Outras apropriações similares são jogos de perguntas e respostas, também programados no Excel<sup>58</sup>.

	A	B	C	D	E	F	G	H	I	J	K	L
1	<b>Churras Calculator</b>											
2	CopyRight, Tabajara Inc.											
3												
4	Número de Homens	5	Desequilibrado: Chama mais Homens !									
5	Número de Mulheres	9										
6	<b>Inventário</b>											
7	Lingüiça	3.05 kg										
8	Carne	4.30 kg										
9	Cerveja	6 Caixa de 12 Latas	=	66 Latas								
10	Refrigerante	3 Garrafas 2l	=	6 Litros								
11	Pães	24 Unidades										
12	Farofa Temperada	1 Sacos										
13	Carvão	2 Sacos										
14	Guardanapos	2 Embalagens c/50										
15	Alcool	1 Litro										
16	Pano de Prato	2 Unidades										
17	Facas Churrasqueiro	3 Unidades										
18	Pratos	4 Unidades										
19	Garfos	3 Unidades										
20	Copos plásticos	1 Embalagens c/20	=	20 Unidades								
21	Sal Grosso	1 Sacos										
22	Limão	10 Unidades										
23	Pinga	2 Litros										
24	Açucar	1 Sacos 1kg										
25	<b>Simulador de Animação</b>											
26	"Animal!" - Diversão, risadas, alegria e descontração garantida!											
27	"Churras da galera" - boas diversão, piadas.											
28	Churrasco de família....											
29												

É só preencher os  
dois campos  
Amarelos, o resto é  
automático....

Fig. 2: Captura de tela da planilha “Churras Calculator”. Fonte: CHURRAS, 2001.

<sup>56</sup> Tradução livre. No original: “end-user programming systems [...] useful applications in a short time without prior programming instruction”.

<sup>57</sup> Exemplo disso pode ser encontrado no site <<http://www.sadia.com.br/s-para-voce/>> e no aplicativo Churrascômetro, disponível em <<https://itunes.apple.com/br/app/churrascometro/id326229525>>.

<sup>58</sup> Vários exemplos desses jogos podem ser baixados no site <[http://www.quizmasters.biz/3rd\\_Party/Excel.html](http://www.quizmasters.biz/3rd_Party/Excel.html)>.

## 2.5 ESTABILIZAÇÃO DAS TECNOLOGIAS

Um exemplo que situa o computador como uma tecnologia que ainda não é considerada parte do que é ser jornalista é a utilização do termo RAC – reportagem assistida por computador – para se referir a técnicas de jornalismo que fazem uso do computador. Segundo Meyer (1999, p. 4), o jornalismo deveria abandonar o termo RAC e deixar de diferenciar os profissionais somente por fazerem uso do computador, colaborando assim para que o computador, como em outras profissões, se torne “natural” para a profissão.

Essa naturalização do computador defendida por Meyer será relacionada aqui com o conceito de estabilização. Segundo Feenberg, a estabilização acontece quando uma tecnologia tem seus significados e funções finalmente definidos e se transforma numa caixa preta estável (FEENBERG, 1995, p. 154). Quando uma tecnologia chega a este ponto, o seu caráter social deixa de ser evidente e seu desenvolvimento se limita à parte técnica (FEENBERG, 1995, p. 154). Além disso, a estabilização de uma tecnologia implica em que seus códigos técnicos, que “define o objeto em estritos termos técnicos de acordo com os significados sociais mais gerais que tenha adquirido” (FEENBERG, 1995, p. 156)<sup>59</sup>, foram modificados ou aceitos. A estabilização de um sistema técnico não é determinado pelos designers, mas sim por um processo social, através da apropriação, como descrito na seção 2.2. Até um artefato alcançar a estabilização, ele pode ter uma “vida dupla”,

já que o artefato ainda não está fechado e não tem amarras institucionais que o amarrem definitivamente a um de seus muitos usos. Então, ambiguidades na definição de uma nova tecnologia devem ser resolvidas através do desenvolvimento técnico em si. Designers, compradores e usuários têm seus papéis no processo através do qual o significado de uma nova tecnologia é finalmente decidido. (FEENBERG, 1995, p. 156)<sup>60</sup>

<sup>59</sup> Tradução livre. No original: “define the object in strictly technical terms in accordance with the most general social meanings it has acquired.”

<sup>60</sup> Tradução livre. No original: “These ambiguities are not merely conceptual, since the device is not yet ‘closed’ and no institutional lock-in ties it decisively to one of its several uses. Thus ambiguities in the definition of a new technology must be resolved through technical development itself. Designers, purchasers, and users all play a role in the process by which the meaning of a new technology is finally fixed.”

A estabilização é então um processo negociado entre esses vários atores. O designer ou desenvolvedor de uma tecnologia e também quem a consome, que geralmente são os menos favorecidos no processo – jornalistas, no caso da presente pesquisa. A participação dos consumidores no processo de desenho dos artefatos é fundamental para conseguir a democratização da mudança tecnológica (FEENBERG, 1995, P. 7). Essa participação não necessariamente se dá durante o processo de design, mas depois do produto chegar ao público, já que os

sistemas técnicos não podem ser considerados finalizados até que eles tenham superado testes sociais que os tenham exposto a uma ampla gama de influências do público e interesses excluídos na fase do design. O fato de que, ao menos nesses casos, os sistemas técnicos tenham passado por grandes mudanças depois do lançamento sugerem um processo [de design] falho. (FEENBERG, 1995, P. 7)<sup>61</sup>

Há duas maneiras de que um artefato alcance a estabilização: a primeira é a estabilização retórica, que é quando os problemas “desaparecem”. Esse sumiço dos problemas não exige que eles sejam de fato solucionados: basta que deixem de ser percebidos como problemas (PINCH; BIJKER, 1997, P. 44). A outra maneira de chegar à estabilização é a redefinição do problema (PINCH; BIJKER, 1997, P. 45). Pinch e Bijker exemplificam com o pneu da bicicleta, que deixou de ser visto como inseguro quando foi visto que aumentava a velocidade das bicicletas em competições.

## 2.6 CONSIDERAÇÕES

Os autores apresentados situam quem faz uso das tecnologias como participante do projeto de um artefato. É através do uso, no cotidiano, que os artefatos vão adquirir suas funções e significados, que podem ser diferentes do previsto por quem projeta esses artefatos.

Sendo assim, é preciso legitimar as soluções criadas pelo consumidor no cotidiano. Essas ações invisíveis, mesmo quando

---

<sup>61</sup> Tradução livre. No original: “technical systems cannot be considered finished until they have withstood social tests that expose them to a wide range of public influences and concerns excluded in the design phase. The fact that, in these cases at least, the technical systems underwent major changes after release suggests a flawed process.”

parecerem rudimentares, devem ser reconhecidas como uma forma válida de consumo, já que é nessa forma de consumir que as pessoas conseguem atingir seus objetivos, seja ter uma travessa a menos para lavar, andar de bicicleta de forma mais segura, ou conseguir dados difíceis de obter e produzir uma visualização com eles. Através dessas ações, possivelmente táticas, é que os códigos técnicos serão negociados e os artefatos, estabilizados.

No jornalismo, alguns aspectos das tecnologias computacionais parecem estar estabilizados, como é o uso do computador como editor de texto, ou para enviar e receber mensagens. Já a variedade de tecnologias do JGD citadas em entrevistas e enumeradas em manuais sugerem que elas – e consequentemente o JGD como um todo – estão numa situação similar à da bicicleta no final do século 19 (PINCH; BIJKER, 1997, P. 28-47), com seus códigos técnicos ainda abertos, pelo menos para o grupo social de jornalistas; ferramentas usadas – e ainda não estabilizadas – no JGD podem estar estabilizadas para outros grupos, como aplicativos de georreferenciamento entre cartógrafos, por exemplo.

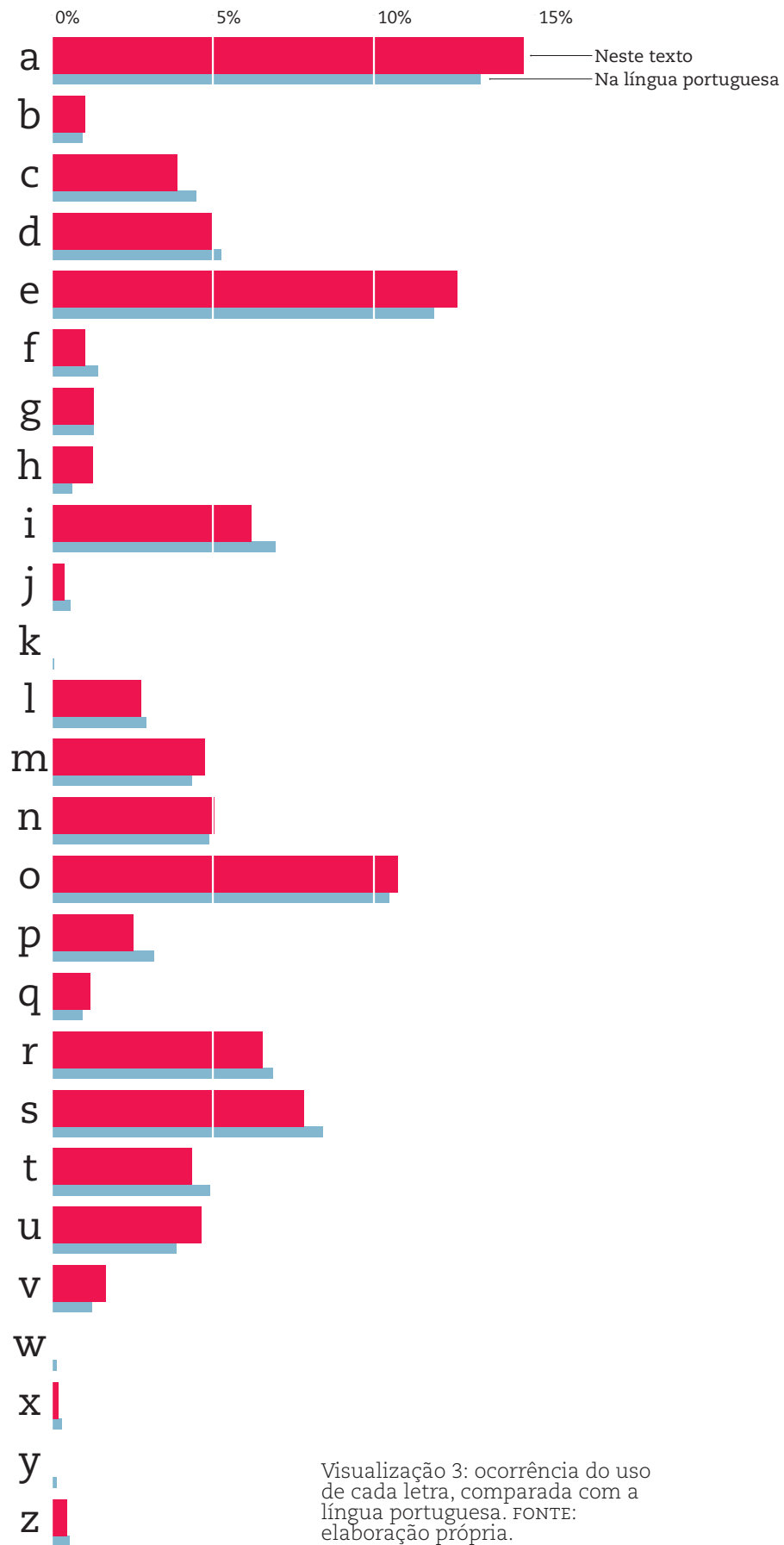
Mapear os processos e apropriações necessárias para que o JGD alcance sua estabilização pode ser particularmente complexo, considerando que, em níveis meso e micro, cada jornalista monta seu próprio sistema para trabalhar com JGD no qual podem conviver artefatos com diferentes códigos técnicos e que oferecem diferentes maneiras de serem usados. Por outra parte, há várias maneiras de mapear e considerar artefatos. A programação, por exemplo, pode ser mapeada como um todo, mas também podem ser consideradas diferentes linguagens ou ainda aspectos de cada linguagem.

Já em um nível macro, pode ser considerada a estabilização do JGD como um todo dentro do jornalismo. Feenberg narra como o Minitel se tornou “um sistema coerente [que] surgiu a partir da ação de diversas forças” (FEENBERG, 2010, P. 169), o que é possível que também aconteça com o JGD – de outras maneiras, considerando a situação peculiar do Minitel.

Considerando o histórico do JGD (ver seção 3.3), é possível pensar que ele, como um todo, estaria próximo da estabilização



para o grupo social dos jornalistas, trabalhem eles com JGD ou não. Já em níveis meso e micro, a quantidade de variantes de artefatos sugere que, mesmo para o grupo de pessoas que trabalha com JGD, o código técnico ainda está aberto, permitindo negociações entre artefatos mais abertos ou fechados – nos termos de Spinuzzi (2003) – e novas variantes que possam surgir. Acredita-se que, mais do que definir se essas tecnologias vão se estabilizar em algum momento, a compreensão da tecnologia como socialmente construída colabora para uma visão não determinista, na qual o computador não é um artefato que produz por si só, mas é usado por pessoas, e que seja dessa maneira que ele não seja visto como um elemento estranho ao jornalismo.



Visualização 3: ocorrência do uso de cada letra, comparada com a língua portuguesa. FONTE: elaboração própria.

# 3 JORNALISMO E O COMPUTAR



Fig. 3: Boneco representando a profissão de jornalista, geralmente dado de presente a jornalistas.

Alguns estereótipos situam o ser jornalista como pessoas que mantêm distância do que consideram tecnologia. Algumas tecnologias se encontram estabilizadas, como blocos de notas e câmeras fotográficas, enquanto outras, como certos aspectos do computador, não parecem ter lugar em identidades mais tradicionais (Fig. 3).

Com o objetivo de entender a função dos grandes conjuntos de dados no jornalismo, neste capítulo será discutido brevemente como o JGD tem uma conotação de objetividade que remete às primeiras teorias jornalísticas. Uma nova desconstrução dessas teorias se faz necessária para conceber o JGD não como uma garantia e sim como uma tentativa de alcançar a objetividade. Ainda para refletir sobre a situação do JGD no jornalismo, são feitas considerações sobre a relação do JGD com os diferentes gêneros jornalísticos, mas sem intenção de encaixá-lo em algum. É

apresentado um breve histórico do uso de computadores em redações, e como jornalistas constroem suas identidades com o uso de computadores e outras tecnologias.

### 3.1 O IDEAL DE OBJETIVIDADE DO JORNALISMO

O ideal de objetividade do jornalismo é ilustrado por Traquina (2005, p. 52) com a frase “*Comment is free, but facts are sacred*”, extraída de um ensaio publicado em 1921 por C. P. Scott, então editor do jornal britânico *The Guardian*. Várias décadas depois, parte da frase – *facts are sacred* – foi usada como subtítulo do *blog* de Jornalismo Guiado por Dados e visualizações do próprio *The Guardian*<sup>62</sup>, além de ser o título do livro de Simon Rogers, ex editor desse *blog*<sup>63</sup>. A frase que em 1921 era uma declaração de princípios da publicação hoje é usada para declarar princípios similares aplicados a um novo tipo de “fatos”: os dados estatísticos e as visualizações feitas a partir deles.

Essa sacralização dos fatos e dos dados estatísticos remete à Teoria do Espelho, que é “a teoria mais antiga e responde que as notícias são como são porque a realidade assim as determina” (TRAQUINA, 2005, p. 146). Segundo a Teoria do Espelho, a função do jornalismo é transmitir os fatos acontecidos exatamente da maneira que aconteceram, sem distorções por parte do jornalista, configurando um “novo jornalismo<sup>64</sup> na era do positivismo, [na qual] vive-se um culto dos fatos” (TRAQUINA, 2005, p. 52). Nesse “novo jornalismo” a função do jornalista era de ser um comunicador sem “interesses específicos a defender” que o desviassem da “missão de *informar, procurar a verdade, contar o que aconteceu, doa a quem doer*” (TRAQUINA, 2005, p. 147, GRIFOS NO ORIGINAL).

Traquina cita o surgimento da fotografia como inspirador da ideia de objetividade, dando ao jornalismo um papel “de ser as

---

<sup>62</sup> Data Blog. Disponível em <<http://www.theguardian.com/news/datablog>>. Acesso em 19/8/2014.

<sup>63</sup> Cf. ROGERS, Simon. **Facts are Sacred: The Power of Data**. Londres, Faber&Faber, 2013.

<sup>64</sup> Não se refere ao movimento de jornalismo literário *New Journalism* explicado em nota na página 72.

‘lentes’ da sociedade” (TRAQUINA, 2005, P. 38), o que teria definido a meta do jornalismo:

O realismo fotográfico tornou-se assim o farol orientador da prática jornalística, como podemos verificar num texto escrito em 1855: Porque um repórter deve ser uma mera máquina que repete, apesar de uma orientação editorial. Ele não deve conhecer nenhum dono mas só o seu dever, e esse dever é o de fornecer a verdade exata. A sua profissão é superior, e nenhum amor por lugar ou popularidade deve desviá-lo de fornecer a verdade na sua integridade. A política do jornal tem sido de reportar *ipsi verbis*. (TRAQUINA, 2005, P. 52)

Traquina situa que, na concepção da Teoria do Espelho, o repórter é uma espécie de artefato que transforma acontecimentos em jornalismo; em tanto que tecnologia, o jornalismo visto dessa maneira tem um caráter instrumentalista (FEENBERG, 1995, P. 22), ignorando o contexto no qual a tecnologia é utilizada, concebendo-a como neutra. Nesse caso, o contexto incluiria a subjetividade, não admitida pela Teoria do Espelho, da escrita produzida por pessoas, com suas influências e repertórios particulares, usando uma linguagem cujos códigos técnicos não permitem que seja completamente neutra, como afirmado pela Teoria do Espelho.

A impossibilidade da meta da objetividade total fez com que, aos poucos, os ideais da Teoria do Espelho fossem se abrandando:

Se, nos anos 1890, os jornalistas raramente duvidavam da possibilidade de escrever realisticamente, nos anos 1930 mesmo os jornalistas dedicados à objetividade reconheciam que a reportagem objetiva era, no fim das contas, uma meta fora do seu alcance – os perigos da subjetividade já eram bem reconhecidos. (TRAQUINA, 2005, P. 138)

Essa subjetividade se manifesta, na linguagem jornalística, através de “escolhas de sintaxe, de léxico e de temática” (COSTA, 2010, P. 48) e “posições pessoais, hábitos e emoções” (MANUAL..., 2010, P. 46). Com isso, seria possível pensar que a linguagem é um dos principais problemas da objetividade, já que

a própria linguagem não pode funcionar como transmissora direta do significado inerente aos acontecimentos, porque a linguagem neutra é impossível (TRAQUINA, 2005, P. 169)

Atualmente, a falta de objetividade é admitida pelo jornalismo. O Manual de Redação do jornal Folha de S. Paulo afirma que “não existe objetividade em jornalismo” (MANUAL..., 2010, P. 46), mas

que isso não exime “da obrigação de ser o mais objetivo possível”. Para isso, seria possível pensar não na linguagem, mas nos métodos de apuração, como defendia Walter Lippmann. Ele apontava o uso de métodos com caráter científico como uma possível solução para a imprensa:

A receita, segundo Lippmann, para os males do jornalismo era a ciência. Lippmann defendeu a posição de que a busca do método científico no jornalismo tornaria a imprensa mais profissional. (TRAQUINA, 2005, P. 138)

Esse caráter científico, assim como a tentativa de se aproximar da verdade, está presente também nos primeiros trabalhos com dados estatísticos combinados ao uso de computadores e técnicas vindas das ciências sociais, no chamado “Jornalismo de Precisão” (TRÄSEL, 2013, P. 3).

### 3.1.1 O ideal de objetividade dos grandes conjuntos de dados

Ao assumir que um dado representa um elemento da realidade, é possível traçar um paralelo com uma notícia jornalística, que também é uma representação de um elemento da realidade. Assim como as notícias, os dados estão sujeitos a vários tipos de fatores que podem alterá-los, como imprecisões na metodologia ou inclusive modificações deliberadas, como ilustrado no seguinte diálogo, extraído do longa-metragem de ficção *Tropa de Elite* (PADILHA, 2007, 00:46)<sup>65</sup>:

—Aspira, você vai refazer esse relatório. Veja aqui. Corpo encontrado na praia. Morte na praia, meu filho, é afogamento.  
 —Mas, comandante, havia uma perfuração no corpo...  
 —Você é legista, por acaso?  
 —Não, senhor.  
 —Segunda-feira quero um novo relatório na minha mesa. Porque este aqui não existiu.

A imperfeição na obtenção dos dados e na representação deles também foi advertida por William Playfair, considerado por Edward Tufte (2001, P. 32) um dos “inventores” dos gráficos estatísticos: ele falava na imperfeição tanto na obtenção da

<sup>65</sup> Foi preferido um exemplo ficcional, já que o jornalista entrevistado pelo autor, que estava preparando uma matéria sobre esse tipo de práticas, preferiu não ser identificado até a matéria ser publicada.

informação como em sua apreensão (PLAYFAIR, 1786 *apud* TUFTE, 2001, P. 32).

Essa imperfeição, relacionada aos códigos técnicos das diversas tecnologias e artefatos utilizados, implica em perda de objetividade, seja na obtenção dos dados ou ao escolher um tipo de gráfico que privilegie certos aspectos deles, ou ainda por limitações do meio escolhido para divulgar a visualização, entre outros possíveis fatores.

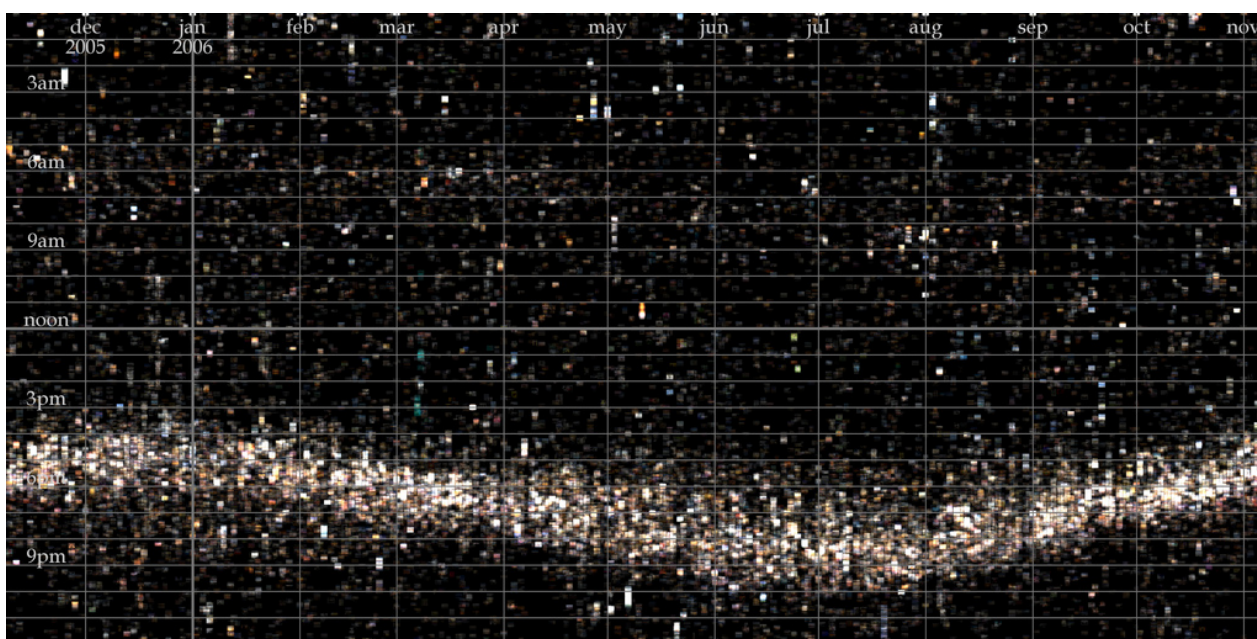


Fig. 4: Visualização feita a partir da tag "sunset". Fonte: BUMGARDNER, 2006.

Um exemplo dessas imperfeições e falta de objetividade são as inconsistências que aparecem em trabalhos da série *Time Graphs*<sup>66</sup> do artista e programador estadunidense Jim Bumgardner. Cada visualização da série consta de dois eixos nos quais se organizam miniaturas de fotografias – o eixo horizontal representa os meses do ano, e o vertical representa as horas do dia – de acordo com as informações presentes nos metadados do arquivo. Para cada visualização, Bumgardner filtrou fotos com diferentes tags<sup>67</sup>; no exemplo da Fig. 4, a tag filtrada foi “sunset” (pôr-do-sol).

<sup>66</sup> Disponível em <<http://www.flickr.com/photos/krazydad/sets/140323/>>. Acesso em 29/10/2014.

<sup>67</sup> Tag ou etiqueta se refere a palavras-chave associadas a um conteúdo eletrônico, como um arquivo, uma fotografia, uma postagem em um blog, ou materiais similares.

A maioria das miniaturas de fotografias se agrupa entre as 17 horas e as 21 horas, inclusive formando uma curva que indica que o sol se põe mais tarde no verão – especialmente entre os meses de junho a agosto, o que também indicaria que a maioria de fotografias foi tirada no hemisfério Norte.

Porém, a presença de miniaturas de fotografias durante o período da manhã pode indicar situações como câmeras com data e hora erradas, pessoas que não ajustaram as horas em suas câmeras quando mudaram de fuso horário, ou que não inseriram os metadados manualmente, entre outras opiniões presentes na seção de comentários da visualização (BUMGARDNER, 2006). Outra possibilidade é que algumas fotos não sejam de um pôr-do-sol, mas tenham sido etiquetadas como “sunset” por outros motivos.

### 3.2 JORNALISMO GUIADO POR DADOS E GÊNEROS JORNALÍSTICOS

Os dados, em qualquer quantidade, são a matéria-prima do jornalismo – segundo Lia Seixas (2009, p. 318), a matéria-prima do jornalismo é a realidade, parte da qual está presente nos dados. Para mapear o papel dos grandes conjuntos de dados no jornalismo, será considerada a relação deles com os gêneros jornalísticos, que podem ser considerados como diferentes maneiras de relacionar a realidade e os produtos jornalísticos. A intenção não é encaixar o JGD em algum gênero, mas sim entender a relação dos métodos do JGD com os diferentes gêneros. Como visto adiante, dentro de cada gênero há, principalmente, formatos já conhecidos, como reportagem, cronologia ou nota; essa divisão sugere que esses formatos possam ser considerados como artefatos que já passaram por um processo de estabilização (FEENBERG, 1995, p. 154), o que ajudaria a entender que tão estabilizados estão o JGD e os formatos que possam dele derivar.

É interessante considerar também o contraponto de Gonzalo Peltzer sobre os gêneros jornalísticos e sobre o termo “jornalismo de dados” – similar ao aqui adotado Jornalismo Guiado por Dados:

Não gosto da expressão *jornalismo de dados* e menos ainda comparada com outros *jornalisms* como o *jornalismo de opinião*. Jornalismo só tem um e os dados são a realidade, seja onde for que se os procure e seja onde for que se os encontre e apesar da



máxima do *Manchester Guardian* que os inspira: *Comment is free, but facts are sacred*. (PELTZER, 2013, GRIFOS NO ORIGINAL)<sup>68</sup>

É considerada, então, uma abordagem menos rígida e mais prática para a compreensão dos gêneros: Seixas (2009, p. 1) começa sua tese dizendo que “Aprender a fazer jornalismo é aprender a produzir gêneros jornalísticos”, mostrando a importância dos gêneros e também dos formatos.

O critério das classificações brasileiras (SEIXAS, 2009, p. 56) se baseia, assim como outras classificações, na primeira separação de gêneros, que data do começo do século XVIII, quando os editores de jornais começaram a separar informação de opinião, como apontam Ana Regina Rêgo e Maria Isabel Amphilo (2010, p. 96). Esse processo se relaciona também ao abrandamento das premissas da Teoria do Espelho (TRAQUINA, 2005, p. 138). Ficam estabelecidos então dois gêneros principais: o jornalismo informativo e o jornalismo opinativo.

Jornalismo informativo é, segundo Marques de Melo (*apud* COSTA, 2010, p. 50)<sup>69</sup> a articulação entre acontecimentos reais e sua expressão jornalística, ou seja, o relato de um fato de maneira objetiva. O jornalismo opinativo é aquele no qual as pessoas “informad[as] de ideias, fatos ou situações conflitantes, exprime[m] a respeito seu juízo” (BELTRÃO *apud* COSTA, 2010, p. 56)<sup>70</sup>. É importante considerar que a já explicada subjetividade do jornalismo dá a todo produto um certo caráter opinativo (RÊGO; AMPHILO, 2010, p. 95–96).

Dentro dessas duas categorias principais – opinativo e informativo –, os gêneros se dividem em formatos. Considerando algumas classificações de jornalismo enumeradas por vários autores (COSTA, 2010; PENA, 2008, p. 65–70; PEÑARANDA, 2000; SEIXAS, 2009, p. 47–59; TRESKA, 2010, p. 89–91), é notável a predominância da divisão informação e opinião, e também que os gêneros dentro dessa divisão são determinados pelos formatos e, em menor

<sup>68</sup> Tradução livre. No original: “No me gusta la expresión *periodismo de datos* (*data journalism*) y menos comparado con otros *periodismos* como el *periodismo de opinión*. Periodismo hay uno solo y los datos son la realidad, se los busque donde se los busque y se los encuentre como se los encuentre y a pesar de la máxima del *Manchester Guardian* que los inspira: *Comment is free, but facts are sacred*.” *Manchester Guardian* era o antigo nome do atual jornal *The Guardian*.

<sup>69</sup> MARQUES DE MELO, José. *Jornalismo opinativo: gêneros opinativos no jornalismo brasileiro*. Campos do Jordão, Mantiqueira, 2003, p. 63–65.

<sup>70</sup> BELTRÃO, Luiz. *Jornalismo opinativo*. Porto Alegre: Sulina, 1980, p. 14.

medida, pela profundidade de cada formato – como por exemplo a diferença entre nota, notícia e reportagem dentro do jornalismo informativo. Referência entre os teóricos brasileiros, a classificação de Marques de Melo, baseada na de Luiz Beltrão (SEIXAS, 2009, P. 56–57), também subdivide os gêneros em formatos. Por exemplo, no gênero informativo, Marques de Melo considera quatro formatos (COSTA, 2010, P. 55):

- *Nota*: relato de acontecimentos que ainda estão se desenvolvendo. Carece de alguns elementos da notícia (ação, agente, tempo, lugar, modo, motivo).
- *Notícia*: relato completo de um fato que já aconteceu. Responde às perguntas que, quem, quando, como, onde e por que. Tem estrutura de “pirâmide invertida” – os fatos de maior importância narrados no começo, no chamado “lide”, e os de menor importância ou que ampliam a informação mais importante, no corpo do texto.
- *Reportagem*: relato ampliado de um fato, incluindo antecedentes, desdobramentos, e outros elementos que possam facilitar a compreensão do fenômeno. Costumava ter autoria individual, mas reportagens feitas por equipes têm se tornado mais comuns.
- *Entrevista* (como formato, não como técnica de apuração): relato que privilegia o ponto de vista de um ou mais protagonistas dos acontecimentos.

Atualmente, Marques de Melo considera cinco gêneros jornalísticos: além do informativo e do opinativo, ele considera os gêneros utilitário – de serviços ao público, como previsão do tempo, cotações etc. –, diversional – histórias de interesse humano e histórias coloridas – e interpretativo (SEIXAS, 2009, P. 56).

### 3.2.1 O caráter interpretativo do JGD

O termo “interpretativo” foi utilizado por primeira vez em 1960, no livro *Interpretative Reporting*, de Curtis MacDougall (SEIXAS, 2009, P. 50). É importante destacar o significado do termo “interpretar”, que é aqui usado como sinônimo de “explicar”. O jornalismo interpretativo é caracterizado por um “esforço analítico e documental que procur[a] situar mais precisamente o cidadão

diante dos acontecimentos” (MARQUES DE MELO *apud* COSTA, 2010, P. 66)<sup>71</sup>. Esse “esforço analítico” poderia se encaixar nas atribuições do jornalismo informativo, e por isso Marques de Melo coloca em dúvida se o interpretativo é um gênero por si só (COSTA, 2010, P. 65–66).

Essa necessidade de um jornalismo que não se limitasse a transmitir os fatos, mas que também usasse da subjetividade para explicá-los, tem origem na 1ª Guerra Mundial, apesar dessa necessidade ter sido identificada mais tarde:

A primeira guerra mundial pegou os americanos de surpresa. Nada os havia anunciado que algo assim iria ocorrer e muito menos os havia explicado porque. A culpa se lançou aos periódicos e especialmente às agências. Haviam-se limitado a transmitir fatos, mas isso não bastava. Acusou-se a *Associated Press* de não querer interpretações, mas só a pura transmissão factual do óbvio. (GOMIS *apud* SEIXAS, 2009, P. 67)<sup>72</sup>

Outros termos que ajudam a entender o que é jornalismo interpretativo são citados por Mário Erbolato: jornalismo em profundidade, jornalismo explicativo ou jornalismo motivacional (ERBOLATO *apud* COSTA; LUCHT, 2010, P. 109)<sup>73</sup>. Paulo Roberto Leandro e Cremilda Medina, apesar de não o considerarem como um gênero por si só, dizem que o jornalismo interpretativo é “o esforço de determinar o sentido de um fato, através da rede de forças que atuam nele” (LEANDRO; MEDINA *apud* COSTA; LUCHT, 2010, P. 110)<sup>74</sup>. Eles também dão as características que fazem dele merecedor da denominação “jornalismo em profundidade”, em oposição ao imediatismo da notícia:

[No jornalismo interpretativo] as linhas de **tempo** e **espaço** se enriquecem: enquanto a notícia registra o **aqui**, o **já**, o **acontecer**, a reportagem interpretativa determina um **sentido** desse aqui num círculo mais amplo, **reconstitui** o já no antes e no depois, deixa os limites do acontecer para um **estar acontecendo** atemporal ou menos presente. Através da complementação de fatos que situem ou interpretem o fato nuclear, através da pesquisa histórica de antecedentes, ou através da busca do humano permanente no acontecimento imediato, a grande reportagem é interpretação do

<sup>71</sup> MARQUES DE MELO, José. Jornalismo opinativo: gêneros opinativos no jornalismo brasileiro. Campos do Jordão, Mantiqueira, 2003, p. 47.

<sup>72</sup> GOMIS, Lorenzo. Teoría del periodismo. Cómo se forma el presente. Barcelona: Paidós Comunicación, 1991, p. 109.

<sup>73</sup> ERBOLATO, Mário. Técnicas de codificação em jornalismo: redação, captação e edição no jornal diário. São Paulo: Ática, 1991, p. 31.

<sup>74</sup> LEANDRO, Paulo Roberto; MEDINA, Cremilda. A arte de tecer o presente: o jornalismo interpretativo. São Paulo: Media, 1973, p. 16.

fato jornalístico (LEANDRO; MEDINA *apud* COSTA; LUCHT, 2010, P. 111, GRIFOS NO ORIGINAL)<sup>75</sup>.

Luiz Beltrão fala na interpretação – ou explicação – como a seleção crítica, feita pelo jornalista, dos dados recolhidos; e também como a disponibilização da “informação em toda a sua integridade, captada, analisada e selecionada pelo jornalista, ao qual não cabe o diagnóstico” (BELTRÃO *apud* COSTA; LUCHT, 2010, P. 112–113)<sup>76</sup>. Essa informação na íntegra, por sua vez, pode ser re-interpretada e analisada pelo público (BELTRÃO *apud* COSTA; LUCHT, 2010, P. 113)<sup>77</sup>. De modo geral, mais especificamente na visão de Beltrão, o Jornalismo Interpretativo se relaciona com a ideia de “completude” que Groth aponta como fundamental para o jornalismo (GROTH *apud* FIDALGO, 2007, P. 108)<sup>78</sup>. A ideia da re-interpretção por parte do público é compartilhada por Lorenzo Gomis:

O que o chamado jornalismo “interpretativo” faz com as informações é ampliar a margem de interpretação. Captar uma notícia no mais típico jornalismo de “fatos” era já, e segue sendo, interpretar a realidade social, mas o “interpretative reporting” amplia a margem de interpretação para servir melhor ao leitor e dar-lhe mais elementos de juízo para que, por sua vez, ele interprete a atualidade que o serve. (GOMIS *apud* SEIXAS, 2009, P. 68)<sup>79</sup>

No que se refere a formatos do jornalismo interpretativo, Marques de Melo identifica quatro (COSTA, 2010, P. 68):

- *Dossiê*: dados que complementam e facilitam a compreensão dos fatos noticiosos, geralmente em forma de “boxes” que podem conter tabelas, gráficos ou mapas.
- *Perfil*: relato que identifica os agentes da notícia, focando nos protagonistas mais frequentes das notícias.
- *Enquete*: descrição de pontos de vista, opinião ou fatos de pessoas escolhidas por amostragem ou por serem parte de um grupo de quem interessa a opinião ou os fatos.

<sup>75</sup> LEANDRO, Paulo Roberto; MEDINA, Cremilda. A arte de tecer o presente: o jornalismo interpretativo. São Paulo: Media, 1973, p. 23.

<sup>76</sup> BELTRÃO, Luiz. Jornalismo interpretativo: filosofia e técnica. Porto Alegre: Sulina, 1976, p. 47 e 52.

<sup>77</sup> BELTRÃO, Luiz. Jornalismo interpretativo: filosofia e técnica. Porto Alegre: Sulina, 1976, p. 42.

<sup>78</sup> GROTH, Otto Groth. Vermittelte Mitteilung - Ein journalistisches Modell der Massenkommunikation, Munique: Verlag Reinhard Fischer, 1998.

<sup>79</sup> GOMIS, Lorenzo. Teoría del periodismo. Cómo se forma el presente. Barcelona: Paidós Comunicación, 1991, p. 111.

- *Cronologia*: relato de fatos na ordem em que aconteceram, com o objetivo de melhorar a compreensão.

Esses formatos – e outros que também possam ser considerados como jornalismo interpretativo –, segundo Lia Seixas,

poderiam ser narrativos, descritivos e explicativos em grau muito maior do que para o jornalismo informativo. A afirmação mais clássica desta compreensão dizia: a reportagem não segue a lógica do lead e pode criar imagens, impressões e invocar sentimentos. No foco, portanto, o grau de subjetividade do enunciador-jornalista, pois sua ação de interpretar permite comparar, explicar, transmitir segundo sentimentos e exige aprofundar e investigar. (SEIXAS, 2009, P. 67)

Cabe questionar o conhecimento dos e das jornalistas de fazer essas comparações e explicações, principalmente quando se trata de grandes conjuntos de dados e de visualizações. Ao falar de Jornalismo de Precisão, conceito explicado na próxima seção, Kunczik (2002, P. 104) conclui que “a investigação sócio-científica não pode ser tarefa do jornalismo” e que é necessário “melhorar de forma geral o nível de conhecimentos sócio-científicos básicos dos jornalistas”. Expandindo essa ideia, é necessário atentar para a capacidade de jornalistas de interpretar e explicar os grandes conjuntos de dados ou os resultados de operações feitas com esses dados – como visualizações. Apesar que essa capacidade de interpretação não é foco deste trabalho, essas habilidades podem ser consideradas como mais uma habilidade relacionada ao JGD e talvez até discutidas da mesma maneira, através dos modos de uso no cotidiano.

Os gêneros e formatos previamente apresentados podem servir para compreender de quais maneiras o JGD e produtos dele derivados se situam dentro do jornalismo, mas sem tentar encaixá-los em categorias que foram concebidas a partir de formatos mais tradicionais. É possível situar o JGD como parte do jornalismo informativo, considerando a objetividade – mesmo que parcial – dos dados, mas também é necessário considerar o caráter interpretativo dos seus métodos de pesquisa e suas intenções de profundidade, seja na obtenção dos dados ou nas explicações necessárias depois do processamento deles.

Os produtos jornalísticos relacionados ao JGD também têm lugar nos formatos aqui apresentados do jornalismo informativo e

interpretativo. Um gráfico pequeno pode ser uma espécie de nota ou notícia, uma reportagem pode ser apurada usando técnicas de JGD, uma visualização organizada em ordem cronológica pode se relacionar ao formato de cronologia incluso no jornalismo interpretativo, visualizações e tabelas podem aparecer em dossiês, uma visualização mais complexa pode ser considerada uma reportagem, entre outras possibilidades.

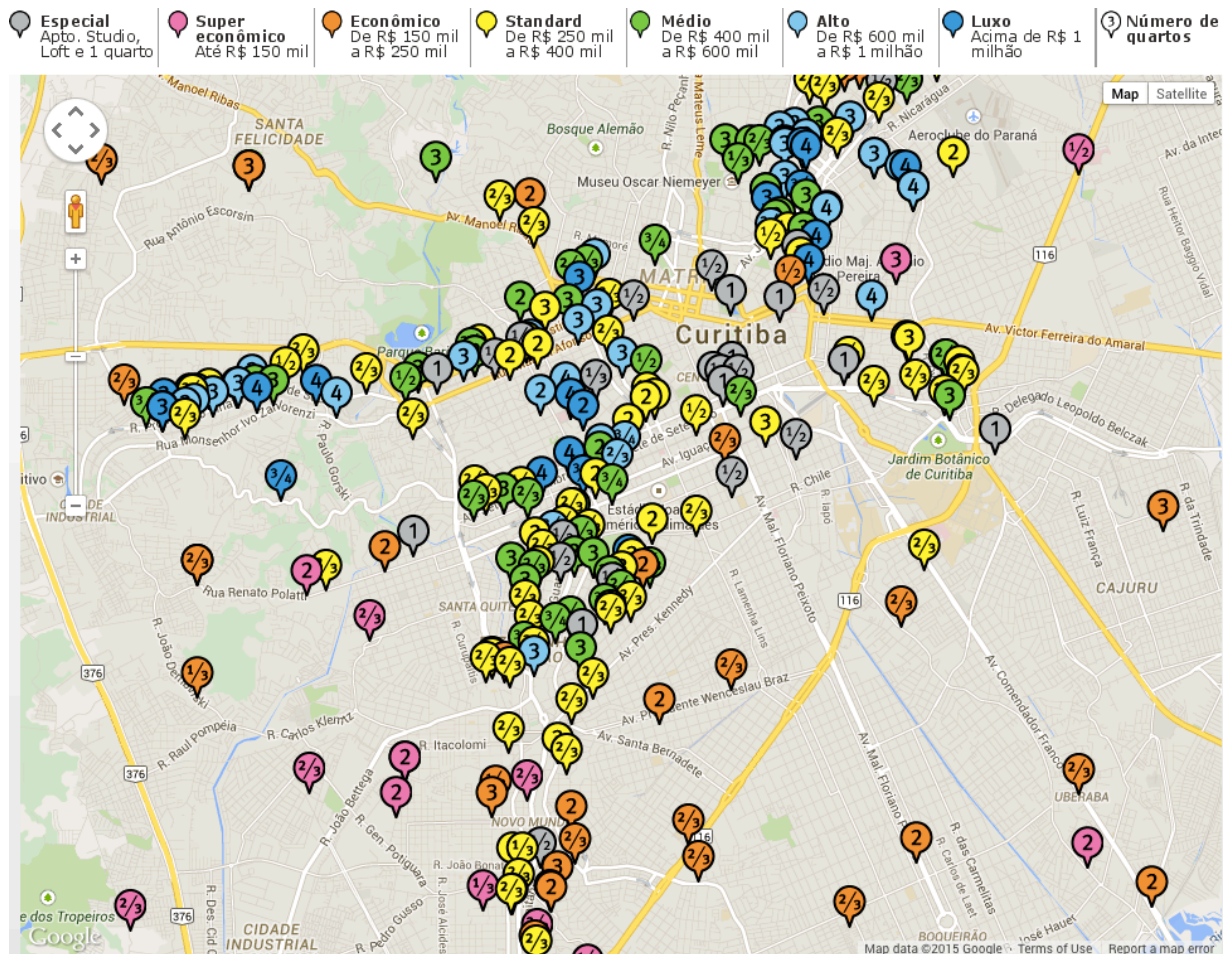


Fig. 5: Mapa dos lançamentos imobiliários de Curitiba. Fonte: GAZETA DO POVO, 2011.

Não se deve, porém, descartar relações com os outros gêneros e/ou formatos jornalísticos, como o jornalismo utilitário, também chamado “de serviço”, que é aquele cujo objetivo é “contribuir para melhores condições de vida do receptor” (TEMER *apud* COSTA, 2010, p. 71)<sup>80</sup>. Um exemplo é o “Mapa dos lançamentos imobiliários de Curitiba” (GAZETA DO POVO, 2011), uma visualização que ajuda a

<sup>80</sup> TEMER, Ana Carolina Rocha Pessoa. **Notícias e serviços:** um estudo sobre o conteúdo dos telejornais da Rede Globo. 2001. 339 f. Tese (Doutorado em Comunicação Social) – Universidade Metodista de São Paulo, São Bernardo do Campo, 2001.

revelar informações latentes em um conjunto de dados – como a predominância de lançamentos econômicos em bairros mais afastados – mas tem a intenção primordial de ajudar na decisão de compra de um imóvel (Fig. 5). Outro exemplo é o mapeamento (Fig. 6) de cruzamentos perigosos para ciclistas (GAZETA DO POVO, 2013), feito pela equipe do jornal Gazeta do Povo a partir de dados levantados pela Associação de Ciclistas do Alto Iguaçu e por leitores do blog Ir e Vir de Bike<sup>81</sup>, que além de levantar discussões sobre planejamento urbano pode orientar a quem anda de bicicleta na busca de percursos menos arriscados.

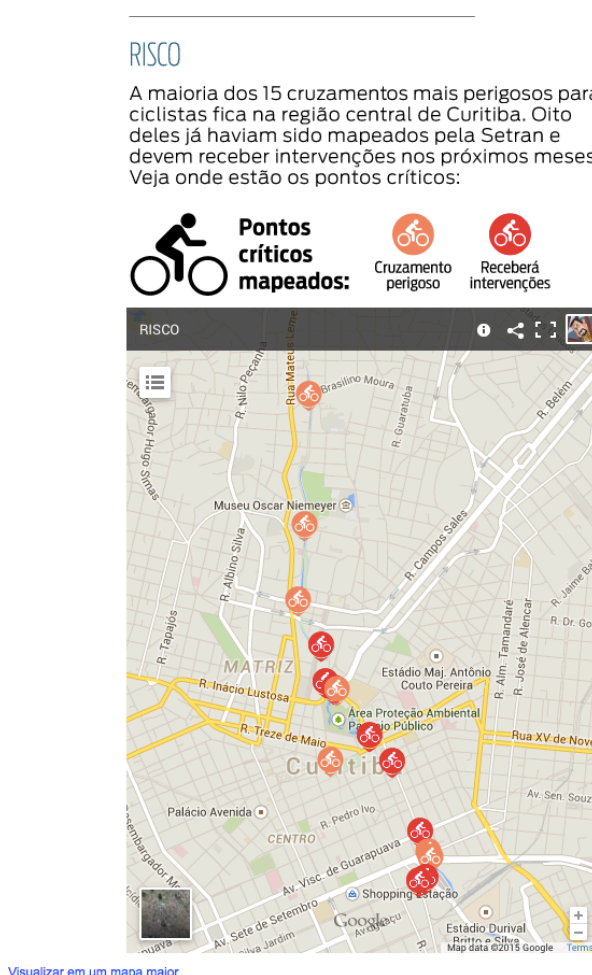


Fig. 6: Mapeamento de cruzamentos perigosos para ciclistas. Fonte: GAZETA DO POVO, 2013.

Sylvain Parasié inclui produtos jornalísticos relacionados ao JGD no que ele chama de “novas relações” entre jornalismo e público (2011, p. 10), o que sugere que produtos jornalísticos como

<sup>81</sup> Blog sobre cicloativismo e mobilidade urbana, antigamente publicado no site do jornal Gazeta do Povo. Disponível em <<http://irevirdebike.com.br>>. Acesso em 13/1/2015.

o mapa de imóveis da Fig. 5 ou o aplicativo do jornal *The New York Times* para calcular se compensa alugar ou comprar um imóvel da Fig. 8 ainda são vistos como novidade e, por isso, não teriam alcançado a estabilização (FEENBERG, 1995, P. 154; PINCH; BIJKER, 1997, P. 44). Produtos jornalísticos mais “tradicionais”, como notas, notícias, dossiês e reportagens, já passaram por esse processo, e têm suas formas mais definidas – estabilizadas – o que permite inclusive o agrupamento em gêneros.

Faltaria então aos produtos do JGD – especificamente àqueles que não se encaixam em nenhum gênero já estabelecido – serem apropriados por jornalistas e leitores, em um processo de estabilização que também colaboraria na estabilização da produção desses “novos” produtos. Estima-se que, como parte desse processo de estabilização, os produtos do JGD ganhem presença dentro dos gêneros jornalísticos, seja se encaixando em gêneros já existentes ou ganhando seus próprios gêneros.

### **3.3 HISTÓRICO DO COMPUTAR NO JORNALISMO**

Vários autores citados por Lailton Costa e Janine Lucht (COSTA; LUCHT, 2010, P. 109) consideram que a fundação do Departamento de Pesquisa e Documentação do *Jornal do Brasil* é um marco do jornalismo interpretativo brasileiro. Entre as funções desse departamento estava a manutenção de bases de dados, o que também o tornaria marco do JGD no Brasil.

O departamento foi fundado pelo jornalista Alberto Dines em meados da década de 1960, período no qual teve começo, no Brasil e em outros países, a informatização das redações, com computadores sendo usados para armazenamento de dados e informações. Os computadores eram usados como uma espécie de “biblioteca digital”, com enciclopédias, resumos de artigos e agendas, mas já era considerado que pudessem servir como auxílio à reportagem. Um caso pioneiro foi no canal estadunidense CBS, ao qual foi proposto usar um computador Univac para a análise de pesquisas e prever o ganhador da eleição daquele ano (BOUNGRU, 2012):



Remington Rand (atual Unisys) procurou a CBS News no verão de 1952 com a ideia de usar o Univac para projetar o resultado da eleição. O chefe de redação Sig Mickelson e o âncora Walter Cronkite foram céticos, mas pensaram que o computador poderia apressar um pouco a análise e que, pelo menos, seria divertido usar um “cérebro eletrônico”. (ALFRED, 2008)<sup>82</sup>

No meio impresso, se destacam como pioneiros a revista *Newsweek*, com a reportagem *The Negro in America*<sup>83</sup>, publicada em 1963, e a reportagem *The People Beyond 12th Street*, publicada pelo jornal *Detroit Free Press*, que analisava pesquisas em uma comunidade afro-americana na qual houve confrontos com a polícia, como aponta Scott Maier (2000, p. 77). Essa matéria ajudou o jornal a ganhar o prêmio Pulitzer na categoria reportagem geral. Em 1968, saiu no jornal *Miami Herald* a série de reportagens *A Scientific Look at Dade Crime*, para a qual foram processadas informações contidas em 13 mil cartões perfurados usando computadores IBM 360<sup>84</sup> (MAIER, 2000, p. 76).

O autor dessas duas últimas matérias é o jornalista Philip Meyer, autor do livro *Precision Journalism*, escrito entre 1969 e 1970 (TRÄSEL, 2013, p. 3) e publicado em 1973 (MEYER, 2002, p. VII). O título do livro deu origem ao termo “jornalismo de precisão” e defende, mais do que o uso do computador, o uso de técnicas das ciências sociais, talvez em oposição às técnicas de escrita vindas da literatura aplicadas à reportagem pelo *New Journalism*<sup>85</sup> (BOUNGRU, 2012). Porém, o uso do computador acabou sendo mais chamativo, como conta o próprio Meyer:

A primeira descrição disto como reportagem por computador que eu lembro foi quando um perfil da [revista] *Newsweek*, em 1968, me chamou de “um repórter de computador”. A aplicação de métodos de pesquisa das ciências sociais era muito complicado de abordar. Computadores eram familiares, se não no uso geral ao menos

<sup>82</sup> Tradução livre. No original: “Remington Rand (now Unisys) approached CBS News in the summer of 1952 with the idea of using Univac to project the election returns. News chief Sig Mickelson and anchor Walter Cronkite were skeptical, but thought it might speed up the analysis somewhat and at least be entertaining to use an ‘electronic brain.’”

<sup>83</sup> Cf. NEWSWEEK. *The Negro in America*. Estados Unidos, 1963. Disponível em: <<http://files.eric.ed.gov/fulltext/ED018488.pdf>>. Acesso em 6/9/2014.

<sup>84</sup> Computador do tipo *mainframe* – de grande porte –, fabricado entre 1965 e 1978.

<sup>85</sup> Estilo de jornalismo que surgiu nos Estados Unidos entre as décadas de 1960 e 1970, caracterizado por reportagens longas, geralmente publicadas em revistas. A escrita era influenciada por elementos da literatura. São expoentes desse estilo jornalistas como Truman Capote, Tom Wolfe e Gay Talese.

conceitualmente, então isso se tornou a definição popular do que eu fazia. (MEYER, 1999, P. 5)<sup>86</sup>

Essa familiarização levou ao protagonismo do computador inclusive na adoção da prática do uso do computador no jornalismo:

Com a adoção dos microcomputadores nas redações nas décadas de 1980 e 1990, as práticas do Jornalismo de Precisão se tornaram acessíveis e tiveram uma disseminação modesta. O termo “Jornalismo de Precisão” acabou sendo abandonado em favor de “Reportagem Assistida por Computador” (RAC), do inglês Computer-Assisted Reporting (CAR). (TRÄSEL, 2013, P. 5)

Por ser menos ligado ao computador, e mais às atividades feitas através dele, nesta dissertação será preferido o termo Jornalismo Guiado por Dados (Data-Driven Journalism), cuja primeira referência (HOLOVATY, 2006) data de 2006 (HOWARDS, 2014). O termo engloba as técnicas da RAC e também outros aspectos dos dados no jornalismo, como descreve Liliana Bounegru:

Alguns defendem que há diferença entre RAC e jornalismo de dados. Defendem que RAC é uma técnica para apurar e analisar dados de forma a aprimorar uma reportagem (normalmente investigativa), enquanto o jornalismo de dados se concentra na maneira como os dados permeiam todo o processo de produção jornalístico. (BOUNEGRU, 2012)

Definições mais recentes e mais simplificadas se concentram na relação entre dados e matérias ou outros produtos jornalísticos. Para a professora Meredith Broussard, por exemplo, o JGD é “encontrar histórias em números e usar números para contar histórias” (BROUSSARD *apud* HOWARD, 2014, P. 5)<sup>87</sup>. Essa definição prioriza o que é feito e por quem é feito, e contraposta ao termo RAC, estabelece uma relação menos determinista entre a tecnologia e quem faz uso dela, tirando o protagonismo do computador próprio do termo RAC.

Considerando as diferentes maneiras em que conjuntos de dados permeiam a produção jornalística, no presente trabalho será

<sup>86</sup> Tradução livre. No original: “The first description of this as computer reporting that I can remember was when a *Newsweek* profile in 1968 called me ‘a computer reporter’. The application of social science research methods was too complicated a frame. Computers were familiar, in concept if not in general use, and so that became the popular definition of what I was doing.”

<sup>87</sup> Tradução livre. No original: “finding stories in numbers and using numbers to tell stories”. O termo “stories” é usado como sinônimo de “matéria” ou de “reportagem”.

priorizado o uso do termo Jornalismo Guiado por Dados como um “modo de fazer”, tentando evitar visões deterministas e sensacionalistas que vejam o JGD como “o futuro do jornalismo”, mas também considerando outros aspectos necessários à discussão, como as novas relações entre jornalismo e público, abordadas na seção 4.1. O termo “encontrar”, usado pela professora Meredith Broussard, sugere que os dados podem ser pesquisados de várias maneiras, fazendo uso de várias técnicas e ferramentas, através dos modos de uso que diferentes jornalistas fazem delas.

### 3.4 POSSIBILIDADES DO JGD

O uso de tecnologias computacionais por parte de jornalistas implica no acesso a bases de dados, que por sua parte também podem estar disponíveis de maneiras que propiciem diversos modos de uso, como pode ser uma visualização ou outro tipo de exploração desses dados. Essas possibilidades são consideradas por Dominique Wolton como um “progresso”:

Também para o acesso às bases de dados que o progresso é real. Acessar, escolher, circular por si mesmo e fazer a sua própria informação permitem não somente ganhar tempo, mas também acessar “reservas” de conhecimento totalmente imprevisíveis. (WOLTON, 2007, p. 87)

Se por um lado o computador é uma forma de “acesso às bases de dados” e também uma forma de “fazer a sua própria informação” e acessar as “reservas” de conhecimento, por outro, segundo Feenberg, o computador é a tecnologia que domina as discussões sobre as mudanças na sociedade, sejam positivas ou negativas (FEENBERG, 1995, p. 123), e aquela que tem funções de transformação cultural e re-codificação de realidades (FEENBERG, 1995, p. 128). Nesse contexto, o uso do computador é questionado:

Como a população fragmentada e desqualificada da nova sociedade usará as tecnologias que dão valor ao trabalho e mediam suas interações sociais? Alcançamos o ponto que os especialistas da computação chamam o problema de “design de

interface” se tornou um problema social, e ainda não sabemos a solução. (FEENBERG, 1995, P. 138-139)<sup>88</sup>

Transpondo essas questões ao jornalismo, é possível problematizar que tão “fragmentado e desqualificado” está o jornalismo dessa “nova sociedade”. Uma possibilidade é a ideia de que repórteres têm que ser multifuncionais – ou seja, acumular funções para atender melhor as necessidades do caráter multimídia das novas redações:

A noção de que os jornalistas do futuro serão uma espécie de McGyver, homens ou mulheres de mil e um recursos, trabalhando sozinhos, equipados com uma câmara de vídeo digital, telefone satélite, [computador] portátil com software de edição vídeo e html, e ligação sem fios à Internet [...] é uma visão futurista do backpack journalist, criticável, e que não esgota, nem preenche, de modo algum, as potencialidades dos novos média<sup>89</sup>. (FIDALGO, 2007, P. 166-167)

Fidalgo critica essa visão “futurista”, bastante comum em discussões sobre o futuro do jornalismo, por estar centrada mais na operação de artefatos tecnológicos do que no resultado final. Para Fidalgo, as possibilidades da tecnologia no jornalismo passam pelo processamento de grandes bases de dados, em processos relacionados ao JGD:

O que a possibilidade e a exequibilidade da mineração de dados nos mostram é a necessidade de formar jornalistas que integrados em equipas multidisciplinares de informáticos, especialistas de estatística e de hermenêutica de dados, saibam detectar o valor noticioso das relações e dos padrões extraídos da massa imensa de dados, coligidos e acumulados a uma rapidez estonteante. (FIDALGO, 2007, P. 166-167)

Além da infraestrutura necessária para a formação de equipas, o que supõe a existência de uma entidade que organize essa formação, é necessário que as entidades – como jornais, por exemplo – apostem nas possibilidades dessas equipas, o que nem sempre acontece, já que

os jornais sempre foram relutantes em investir recursos para a contratação de profissionais em áreas especializadas como

<sup>88</sup> Tradução livre. No original: “How will the fractured and disqualified population of the new society use the technologies that give its work value and mediate its social interactions? We have reached the point where what computer specialists call the problem of ‘interface design’ has become a social problem, and we do not yet know the solution.”

<sup>89</sup> Por “novos média”, Fidalgo se refere principalmente aos sites de notícias, mas sem excluir outros meios similares, como por exemplo um aplicativo de notícias para telefones celulares.

cartografia e estatística, às vezes por ignorância do que o binômio jornalismo de precisão-infografia pode conseguir. (CAIRO, 2008, p. 59)<sup>90</sup>

A formação de equipes se mostra como uma possível solução para conseguir a multidisciplinariedade necessária às práticas do JGD<sup>91</sup>. Porém, pode haver empecilhos para a formação dessas equipes, como a relutância dos jornais em fazer investimentos, narrada por Cairo, ou a dificuldade em encontrar parceiros para fundar organizações independentes. Sendo assim, o interesse desta pesquisa se concentra no indivíduo, estando em uma organização ou não, como pode ser o caso de jornalistas *freelancers* ou outras pessoas que queiram trabalhar com grandes conjuntos de dados, como pessoas que desenvolvam ferramentas para facilitar o acesso a dados públicos, ou ainda jornalistas que trabalhem individualmente ou em equipes pequenas dentro de organizações não jornalísticas. Mesmo em organizações nas quais haja equipes de apoio para executar tarefas – programadores, infografistas, auxiliares, estagiários etc. – pode acontecer das pessoas não conseguirem atender a todas as demandas da redação. Nesse aspecto, ter um conhecimento multidisciplinar permitiria uma menor dependência de outras pessoas, tendo assim uma autonomia maior.

Exemplo disso é a organização Livre.jor, situada em Curitiba e fundada pelos jornalistas José Lazaro Barros Junior e João Guilherme Frey. Eles definem o Livre.jor como

a materialização de uma forma inédita de praticar jornalismo no Paraná, no sentido que é a única iniciativa exclusivamente nas redes sociais e a única totalmente focada em dados oficiais, de qualquer natureza (diários oficiais, bancos de dados, repositórios de documentos etc). A condição exigida é que sejam dados de interesse público e que não tenham sido abordados pelas mídias convencionais. (BARROS JUNIOR; FREY, 2014, P. 1)

A criação do Livre.jor, segundo João e José, aconteceu graças à percepção de lacunas no cenário jornalístico paranaense. Além dos

<sup>90</sup> Tradução livre. No original: “los diarios han sido siempre relucientes a invertir recursos en contratar profesionales en áreas especializadas como cartografía y estadística, a veces por siempre cicatería, a veces por ignorancia de lo que el binomio periodismo de precisión-infografía puede conseguir.” Aqui se entende por infografia a representação gráfica de dados.

<sup>91</sup> Para alguns estudos de casos sobre o assunto, Cf. ZANCHELLI, Michael; CRUCIANELLI, Sandra. Integrating Data Journalism Into Newsrooms. ICFJ; Knight International Journalism Fellowships: Estados Unidos, 2012. Disponível em < [http://www.icfj.org/sites/default/files/integrating%20data%20journalism-english\\_0.pdf](http://www.icfj.org/sites/default/files/integrating%20data%20journalism-english_0.pdf)>. Acesso em 18/6/2015.

métodos e do uso das redes sociais da internet, eles perceberam a carência de jornalistas que soubessem encontrar informações em fontes oficiais, o que teria impedido que as redações fizessem esse tipo de cobertura de maneira sistemática, apesar do sucesso de reportagens como a série “Diários Secretos”<sup>92</sup> (BARROS JUNIOR; FREY, 2014, P. 2). Os interesses do Livre.jor incluem a defesa do direito à informação como instrumento de ativismo social e empoderamento da população. Para pôr isso em prática, outro objetivo deles é

incentivar os colegas jornalistas a usarem dados oficiais em suas reportagens, “fugindo” do declaracionismo e adentrando no mundo das reportagens assistidas por computador e novas técnicas de apuração. (BARROS JUNIOR; FREY, 2014, P. 2)

É possível dizer que João e José se valem de métodos relacionados ao JGD para processar informações disponíveis em portais da transparência e diários oficiais das diferentes esferas de governo, priorizando as informações do estado do Paraná [BARROS JUNIOR; FREY, 2014]. As técnicas do JGD aliadas à divulgação na internet de documentos públicos permitem a eles exercer atividades jornalísticas desde suas casas, conciliando assim o Livre.jor com suas atividades remuneradas. José se dedica mais à leitura de diários oficiais. Quando trabalhou com jornalismo político em um jornal paranaense, percebeu que poderia encontrar furos de reportagem nessas leituras, e desenvolveu técnicas de leitura e busca de informações que agora aproveita no Livre.jor. João se dedica mais à busca de informações orçamentárias e cruzamento de informações de bancos de dados. Eles trabalham com “a pressuposição de que esses documentos sejam verídicos. Eu espero que sejam, porque senão tem que questionar a própria República”, diz João. Como eles não conseguem checar essas informações, tomam cuidado para que o Livre.jor não noticie um fato diretamente, mas deixe claro que é um dado divulgado pelo Estado.

---

<sup>92</sup> Diários Secretos é uma série de reportagens publicadas pelo jornal Gazeta do Povo. Feita a partir de técnicas relacionadas ao JGD combinadas com outros tipos de investigação, revelou um esquema milionário de desvio de recursos e contratação de funcionários fantasmas na Assembleia Legislativa do Paraná. A série ganhou vários prêmios, incluindo o prêmio Esso de Jornalismo. Disponível em: < <http://www.gazetadopovo.com.br/vidapublica/diariossecretos/>>. Acesso em 26/1/2015.

Os resultados encontrados durante suas investigações são publicados diariamente na página do Livre.jor no Facebook<sup>93</sup>, e muitas vezes pautam jornalistas de outros meios, que dão continuidade ao processo de produção jornalística, checando os dados, fazendo entrevistas, e redigindo matérias.

Um exemplo das práticas do Livre.jor é a matéria “No Paraná, média de gastos com rodovias dobra em anos eleitorais” (FREY, 2014), feita a partir de dados pesquisados no Portal da Transparência do Governo do Paraná<sup>94</sup> e também no Sistema de Coleta de Dados Contábeis da Caixa Econômica Federal<sup>95</sup>. A partir das informações coletadas, Frey concluiu – e mostrou em uma visualização feita no Microsoft Excel (Fig. 7) – que os investimentos no transporte rodoviário aumentavam nos anos em que havia eleições. A matéria, publicada na página do Livre.jor do Facebook, foi o ponto de partida de uma matéria televisiva do jornal estadual do canal Band<sup>96</sup> (FREY, 2014).

Dos exemplos de matérias feitas a partir da leitura de diários oficiais, uma que se destacou por ter repercutido em rádio e jornais foi *Defensoria revoga avanço funcional e suspende auxílios transporte e alimentação dos defensores* (BARROS JUNIOR, 2014). Outro exemplo é a matéria *A temporada de 2014 continua*<sup>97</sup>, que trata de como a Orquestra Sinfônica do Paraná superou cortes de verbas originalmente destinados a ela, mas que foram remanejados para outras áreas, informação levantada pelo Livre.jor a partir da leitura de diários oficiais.

---

<sup>93</sup> As publicações se concentram durante os dias úteis, mas há algumas matérias publicadas aos sábados e domingos. A página está disponível em <[www.facebook.com/livre.jor](http://www.facebook.com/livre.jor)>. Acesso em 13/1/2015. A partir de maio de 2015 o site <[www.livre.jor.br](http://www.livre.jor.br)> também passou a ser atualizado com frequência.

<sup>94</sup> Disponível em <<http://www.portaldatransparencia.pr.gov.br/modules/conteudo/conteudo.php?conteudo=61>>. Acesso em 19/1/2015.

<sup>95</sup> Disponível em: <[https://www.contaspublicas.caixa.gov.br/sistncon\\_internet/index.jsp](https://www.contaspublicas.caixa.gov.br/sistncon_internet/index.jsp)>. Acesso em 19/1/2015.

<sup>96</sup> Matéria disponível aos 7 minutos e 30 segundos do vídeo disponível em: <<https://youtu.be/Z6hwhDjmkp8?t=7m30s>>. Acesso em 12/3/2015.

<sup>97</sup> Disponível em: <<http://www.gazetadopovo.com.br/caderno-g/a-temporada-de-2014-continua-eb2x0kqtnke4161190ii49iry>>. Acesso em 23/4/2015.

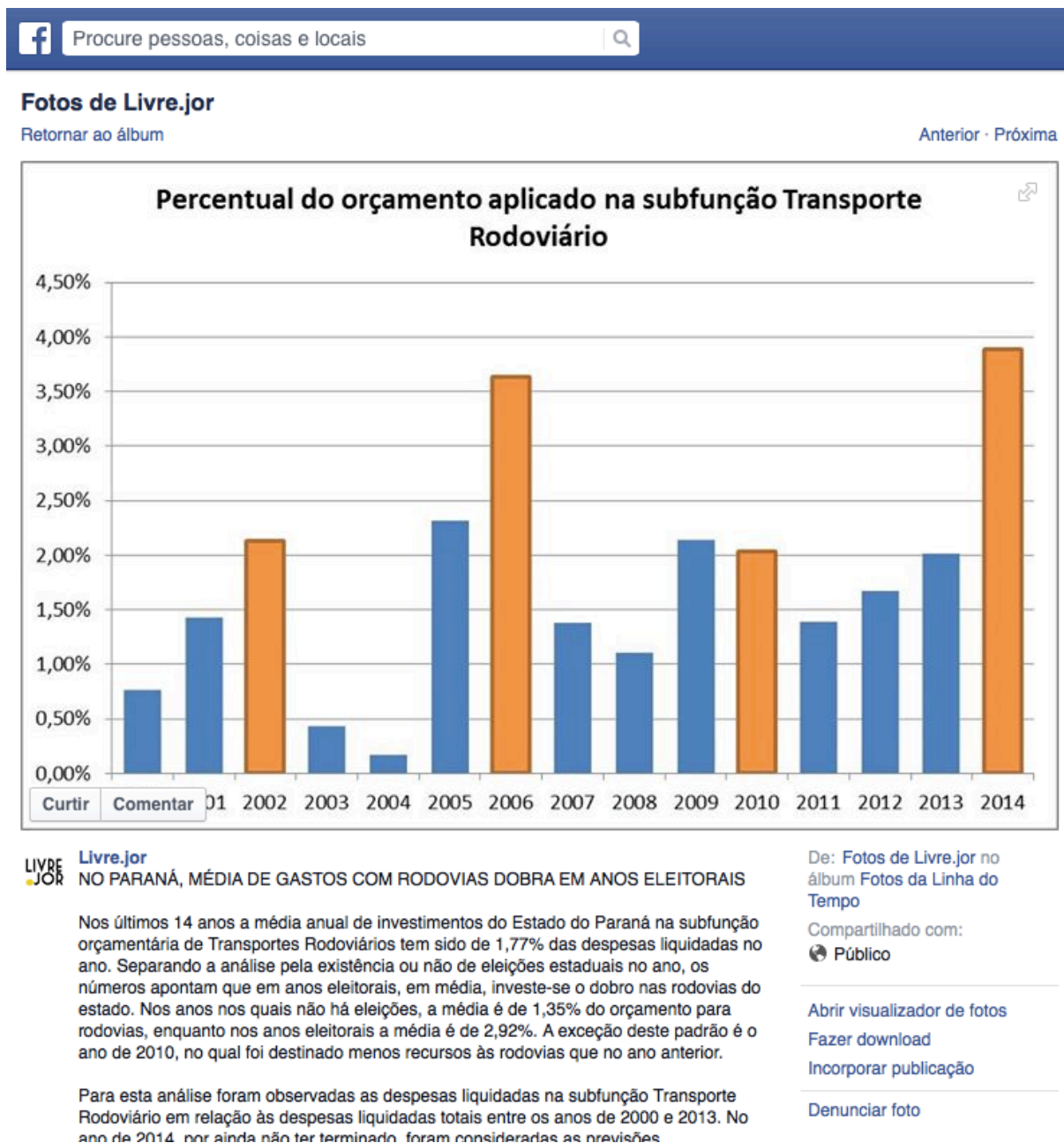


Fig. 7: Matéria e visualização publicadas na página do Facebook do Livre.jor. Fonte: FREY, 2014B.

### 3.5 JGD, ARTEFATOS E IDENTIDADES JORNALÍSTICAS

Como visto na seção 2.1, na página 35, os Estudos Culturais entendem o consumo de tecnologias como uma maneira de construir identidades (MACKAY, 1997B, p. 2). As identidades são construídas pelos artefatos mesmo quando não sejam percebidos como tais, ou seja, quando já se encontram estabilizados (FEENBERG, 1995, p. 156). O que pode ser considerado “tradicional” para algumas pessoas pode ser percebido como “estranho” para outras.



O jornalista estadunidense Truman Capote, autor de *A Sangue Frio*, livro-reportagem considerado marco do jornalismo literário<sup>98</sup>, dispensava tecnologias como gravadores de fita ou blocos de notas, por considerar que eles atrapalhavam a relação com as pessoas que ele entrevistava:

Eu realmente acredito que usar gravador, ou até mesmo fazer anotações, cria artificios e distorções e até mesmo destrói qualquer naturalidade que possa existir entre o observador e o observado, o beija-flor nervoso e seu futuro captor. (CAPOTE APUD KREBS, 1984)<sup>99</sup>

Para Capote, tecnologias como o gravador e o bloco de notas não estariam estabilizadas, ou pelo menos ele considerava – possivelmente em outros termos – que não estariam estabilizadas para quem ele entrevistava, já que seriam elementos estranhos que causariam efeitos não desejados nas entrevistas.

As tecnologias computacionais também causariam efeitos parecidos. Segundo pesquisa etnográfica feita por Isabel Travancas no começo da década de 1990, a presença do computador em redações era percebida como a principal mudança no jornalismo:

Em sua pesquisa etnográfica sobre a identidade dos profissionais das redações cariocas, Travancas (1993, p.57–77) cinco dos sete jornalistas veteranos entrevistados pela autora apontaram o computador como principal mudança em curso na época – mudança recebida com uma mistura de fascínio e resignação pelos informantes, que contrapunham em seu discurso o computador ao principal elemento de sua compreensão da identidade jornalística: a paixão. A informática era vista como um agente de esterilização das redações, um símbolo da transição de um jornalismo com veleidades literárias para a racionalidade capitalista. (TRÄSEL, 2014, p. 34)<sup>100</sup>

Essa percepção do computador como ameaça à identidade dos jornalistas entrevistados por Travancas é exemplo da disrupção causada pela chegada de qualquer tecnologia a uma situação específica, da mesma maneira que acontece quando novas tecnologias são introduzidas no cotidiano de um lar, de uma família:

<sup>98</sup> Gênero de jornalismo caracterizado por uso de recursos literários, cujo marco foi o movimento do *New Journalism* explicado em nota na página 72

<sup>99</sup> Tradução livre. No original: “I devoutly believe that the taking of notes, much less the use of a tape recorder, creates artifice and distorts or even destroys any naturalness that might exist between the observer and the observed, the nervous hummingbird and its would-be captor.”

<sup>100</sup> TRAVANCAS, Isabel. **O mundo dos jornalistas**. São Paulo: Summus, 1993.

quando as tecnologias chegam elas são disruptivas na vida familiar. Seus usos e sucessos estão longe de serem garantidos: mais do que se encaixar naturalmente, elas têm que ser domesticadas, forçadas a preencher os ritmos espaciais e temporais dos lares; e, nesses processos, elas são adaptadas e modificadas. (MACKAY, 1997A, P. 282)<sup>101</sup>

Mackay também aponta que artefatos tecnológicos têm um papel importante em processos de construção de identidades. Eles são “o próprio material com o qual construímos nossas identidades; nós nos transformamos naquilo que consumimos” (MACKAY, 1997B, P. 2)<sup>102</sup>. Sendo assim, é possível que radialistas construam suas identidades a partir do uso das tecnologias de transmissão radiofônica, e que jornalistas que façam uso de “artefatos” como recursos de escrita oriundos da literatura possam se considerar parte do chamado jornalismo literário. Essas identidades também podem ser construídas por outras pessoas, como é o título de “repórter de computador” dado a Philip Meyer pelo uso que ele fez de computadores para a produção de matérias (MEYER, 1999, P. 5), ou a alcunha de “garota do Excel” narrada por Renata [PINTO, 2014].

A construção de identidades através de artefatos tecnológicos pode supor a existência de um modelo a ser seguido. Elliot (APUD TRAQUINA, 2005, P. 101)<sup>103</sup> afirma que a identidade de uma profissão é orientada por um referencial que se aproxima de um estereótipo alcançado pelas práticas. Segundo Kathryn Woodward (2000, P. 8–11), a identidade pode ser construída de várias maneiras: a partir de identidades já existentes, em oposição a elas – na chamada identidade relacional – ou ainda a partir de características comuns entre várias pessoas do mesmo grupo social. A construção de identidades pode ainda ter um caráter essencialista, que sugere a existência de “um conjunto cristalino, autêntico, de características” compartilhadas por um grupo de pessoas, que não se alteraria ao

<sup>101</sup> Tradução livre. No original: “at the time of their arrival new technologies are disruptive to family life. Their uses and success are far from assured: rather than fitting in naturally, they have to be domesticated, made to fit the spatial and temporal rhythms of households; and, in these processes, they become adapted and modified.”

<sup>102</sup> Tradução livre. No original: “the very material out of which we construct our identities; we become what we consume.”

<sup>103</sup> ELLIOT, Philip. **The Sociology of the Professions**. Londres e Basingstoke: The Macmillan Press, 1972, p. 91.

longo do tempo (WOODWARD, 2000, P. 12), o que se aproximaria da ideia de Elliot da existência de um referencial.

Seria possível então partir da categorização de Woodward para pensar que haveria jornalistas que se identificam e agrupam pelas práticas, o que implica no uso de artefatos, e também aqueles que se identificam por não usar algum tipo de artefato. Cláudia Lago fala em um jornalismo romantizado, no qual jornalistas se caracterizavam por serem contra a burocratização das rotinas produtivas, por valorizarem a “missão” do jornalismo, por idealizarem os “bons tempos”, por acreditarem no talento individual, e por se posicionarem como transformadores da sociedade (LAGO APUD TRÄSEL, 2014, P. 89–90)<sup>104</sup>. Nesse estereótipo, de certo caráter essencialista, não haveria lugar para coisas como matemática, cálculo ou tecnologias – ao menos as que ainda não haviam culturalmente aceitas pelos jornalistas:

Nesta cultura profissional, as técnicas jornalísticas são vistas como habilidades inerentes ao espírito [...]. Pode-se até mesmo argumentar que, para a cultura profissional jornalística, estas habilidades estão ligadas biologicamente aos repórteres, uma vez que a capacidade de identificar valor-notícia em acontecimentos do cotidiano é comumente referida como “faro jornalístico”, no jargão das redações. Além disso, os jornalistas são denominados pelos próprios membros deste grupo social como “cães de guarda” da democracia. Neste contexto, o computador é muitas vezes considerado um intruso, um elemento que prejudica o trabalho do repórter ao se intrometer entre sua capacidade de percepção especial dos acontecimentos e do mundo. (TRÄSEL, 2013, P. 6)

Com isso, seria possível pensar que haveria uma identidade jornalística construída a partir de um referencial “romantizado”<sup>105</sup> do que é ser jornalista, e que essa identidade também seria construída em oposição ao jornalismo que faz uso de tecnologias percebidas como “novas”, como é o caso das tecnologias do JGD.

Além dessas tecnologias, o JGD também causaria disrupções pela sua relação com ciências exatas. Nas entrevistas realizadas com docentes, a professora Rosiane Correia de Freitas conta que “a maioria [dos alunos e alunas] não conhece o termo nem pensou que precisaria usar recursos básicos de JGD no trabalho de

<sup>104</sup> LAGO, Cláudia. **O Romantismo Morreu? Viva o romantismo!** Ethos romântico no jornalismo. 2003. 227 f. Tese (Doutorado em Ciências da Comunicação), Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2003, p.169-188.

<sup>105</sup> Cf. TRÄSEL (2014) p. 89–93.

reportagem” [FREITAS, 2014]. Em entrevista, o professor Ricardo Fotios nota que há

uma certa “glamourização” da profissão ainda nos dias de hoje. É comum o aluno chegar ao curso com exemplos de âncoras, apresentadores, correspondentes e colunistas. [FOTIOS, 2014]

Esses exemplos com que discentes chegam ao curso podem ser considerados referenciais na construção de uma identidade buscada por quem começa um curso de jornalismo. Fotios também narra decepções quando veem que o curso não é o que imaginam:

Praticamente ignora[m] a massa profissional que trabalha para que essas poucas pessoas tenham suporte noticioso. Quando descobrem que o grosso da profissão é trabalho duro de pesquisa e apuração, muitos desistem. [FOTIOS, 2014]

Fotios também narra “uma aversão equivocada [por parte de alunas e alunos] em relação ao tema [JGD] por conta da sua proximidade com a tecnologia, com exatas.” Essa aversão ao que estudantes consideram como atividades não relacionadas ao jornalismo se manifesta também na reação dos alunos à apresentação de conteúdos de JGD, como narrado em entrevista ao professor Träsel [2014]: “não houve reações muito negativas, mas poucos alunos se destacaram. Classificaria como apática a recepção.”

Em entrevista, Träsel [2014] ainda nota que os alunos e alunas não estão bem preparados tecnicamente: “conhecimento prévio de Excel ou, ao menos, matemática do segundo grau poderia ser útil.” Rosiane, além da matemática, cita que seria necessário que os alunos e alunas tivessem também conhecimentos básicos sobre apuração jornalística – saber localizar a informação. Nas oficinas de *Data Scraping*, na Universidade Positivo, também houve situações que mostravam ou pouca intenção de se aproximar do assunto – como uma professora de jornalismo que foi convidada a fazer o curso e declinou polidamente o convite afirmando que não tinha “capacidade” para mexer com computadores. Nos e nas jornalistas que fizeram o curso, era notável pouca familiaridade com computadores. Uma aluna afirmou não saber a diferença entre memória RAM e um disco rígido. Quando Álvaro, ministrante

da oficina, explicava o uso de condicionais, perguntou se alguém tinha ouvido falar em lógica booleana – ninguém conhecia o termo. Enquanto explicava a utilidade da condicional, valores verdadeiros ou falsos, uma aluna reclamou, em tom de brincadeira:

—Você não falou que tinha que saber matemática!

—Não é matemática, é lógica.

—Pior ainda!

(NOTAS DO DIÁRIO DE CAMPO, 26/3/2014, 2º ENCONTRO DA OFICINA DE DATA SCRAPING, UNIVERSIDADE POSITIVO)

Porém, o professor Fotios acredita que “o perfil do profissional está mudando.” Essa mudança também é percebida por Rosiane, que conta que

num primeiro momento há muita reclamação. Muitas das reclamações dos alunos a respeito das minhas aulas tem como origem o uso do Excel e de matemática em sala. Já ouvi muito que eles ‘não fizeram jornalismo para ter que fazer conta’. [FREITAS, 2014]

Mas que depois, com os resultados finais do trabalho, essa percepção muda, ao menos para alguns:

é incrível ver os resultados finais do trabalho. Já emplacamos (quase) manchete da Gazeta (matéria sobre a vigilância sanitária). Já furamos a imprensa local (com a matéria dos santinhos de candidatos em 2012). Há um grupo de 15 alunos mobilizados no Núcleo que foram para [o Núcleo] depois de ter contato com JGD em sala de aula. [FREITAS, 2014]<sup>106</sup>

Para Träsel, mais do que conhecimentos prévios, “útil mesmo seria mais boa-vontade com esse conteúdo.” Rosiane também fala que o “mais difícil é romper essa resistência natural contra o tema.” Fotios relata que “No começo, todos reclamam: ‘isso é jornalismo?’, escuto muito. Depois, todos gostam e alguns avançam.”

Alberto Cairo aponta que o jornalista, particularmente o brasileiro, é “um escritor-humanista, não um pesquisador-cientista” (CAIRO, 2012) e que esse perfil é notório nas ocasiões em que jornalistas se deparam com conceitos matemáticos ou de outras áreas:

<sup>106</sup> Exemplos das reportagens citadas pela professora podem ser encontrados em <<http://www.teianoticias.com/2012/10/07/cabos-eleitorais-jogam-santinhos-pelas-ruas-de-curitiba>> e em <<http://brasil.blogfolha.uol.com.br/2014/10/08/estudantes-flagram-arremessos-de-santinhos-no-dia-da-eleicao-veja-video>>.

é muito difícil achar jornalistas diplomados que, ao mesmo tempo, tenham conhecimentos científicos ou técnicos profundos. Não é só que o jornalista médio não saiba mexer com dados; é que não sabe nem ler uma tabela de números, colocar eles em contexto, e extrair histórias, o que é muito mais importante. (CAIRO, 2012)

Kayser-Bril também fala desse tipo de dificuldade, mas em um nível mais básico, narrando uma situação irônica:

A Poynter's News University oferece aulas de matemática on line para jornalistas, nas quais os repórteres recebem ajuda com conceitos como variações percentuais e médias. Curiosamente, esses conceitos são ensinados também perto das salas da Poynter, nas escolas da Flórida a alunos entre 10 e 11 anos, como mostra o currículo. (KAYSER-BRIL, 2012)

Segundo Cairo, esse tipo de dificuldade seria causada “em grande parte, [por] um sistema de educação [universitário] que não tem se adaptado às necessidades dos jornalistas de hoje” (CAIRO, 2012)<sup>107</sup>, aliada à obrigatoriedade do diploma de jornalismo, que “pode dificultar o emprego de gente com perfil diverso para as redações a não ser em posições de segunda categoria” (CAIRO, 2012). Cairo defende que profissionais do jornalismo sejam considerados jornalistas mesmo que não tenham o diploma:

Um hacker que desenvolve ferramentas para que os cidadãos acessem dados públicos, e que segue as regras éticas próprias da profissão, é tão jornalista quanto o repórter que escreve sobre o último escândalo do Governo, gostem os partidários do diploma obrigatório ou não. Se for contratado por um meio de comunicação, deve ser na posição de jornalista ou, pelo menos, com salário e poder de decisão equivalentes aos de um repórter ou editor no mesmo nível. (CAIRO, 2012)

É esse tipo de concepção da profissão que, segundo Cairo, tiraria dos jornalistas a tarefa de mexer com dados e, conseqüentemente, o domínio da profissão:

Quem sabe, talvez sejam estes hackers, desenvolvedores, designers, jornalistas independentes, organizações não governamentais, e fundações os que ocupem um espaço hoje quase vazio, e os que cumpram uma parte importante da tarefa de informação pública que, em tempos anteriores, correspondeu à mídia tradicional. O futuro promete, em qualquer caso. (CAIRO, 2012)

---

<sup>107</sup> Acredita-se que as experiências dos professores Rosiane, Marcelo e Ricardo, anteriormente narradas, estejam em sintonia com as “necessidades” que Cairo aponta.

Essa concepção de ser jornalista atenta para que além de identidades mais tradicionais, cuja construção inclui a rejeição de tecnologias percebidas como “novas”, haveria identidades jornalísticas relacionadas ao JGD, construídas por elementos em comum entre pessoas de um mesmo grupo social (WOODWARD, 2000, P. 11). Por exemplo, as e os integrantes do ProPublica se denominam “nerds”, como aparece nos endereços do seu blog – [www.propublica.org/nerds](http://www.propublica.org/nerds) – e da conta no Twitter – @propubnerds. Em um exemplo observado por Träsel (2014), ele narra que a repórter Amanda Rossi colou adesivos no computador portátil que ela levaria a um evento. Ela questionou “que tipo de pessoa que trabalha com dados sou eu, sem nenhum adesivo?” (TRÄSEL, 2014, P. 159), no que Träsel considera uma preocupação em “manifestar sua participação na subcultura profissional jornalística do JGD” (TRÄSEL, 2014, P. 160). Em entrevista, Renata [PINTO, 2014] fala em se especializar em JGD como diferencial para entrar no mercado de trabalho. João e José [FREY, 2014; BARROS JUNIOR, 2014] constroem suas identidades através das atividades do Livre.jor, se definindo como jornalistas “de dados públicos”. João estabelece uma comparação da sua maneira de trabalhar com bases de dados de orçamentos com a busca de notícias na rua feita por um “jornalista tradicionalzão”, criando um paralelo pela semelhança e ao mesmo tempo situando o fazer jornalismo com base em dados públicos como parte de sua identidade [FREY, 2014; BARROS JUNIOR, 2014].

Esses exemplos atentam para a existência de uma ou várias identidades ligadas ao JGD e a suas tecnologias e artefatos, que configurariam identidades não completamente opostas ao jornalismo tradicional, e sim compartilhando elementos comuns ao que seria ser jornalista.

### **3.6 CONSIDERAÇÕES**

Como visto, o uso de tecnologias computacionais no jornalismo não é algo novo, e mesmo que mais pessoas tenham acesso a bases de dados e a tecnologias para trabalhar com elas, permitindo fazer jornalismo fora do contexto de organizações

maiores, uma análise sócio-construtivista evitaria as consequências de uma visão determinista que apresente o JGD como “um novo jornalismo” ou que retome velhas noções de uma aparente objetividade.

Analisar o JGD a partir da Teoria do Espelho e da concepção construtivista da tecnologia de Feenberg, Pinch e Bijker ajuda a refletir sobre aspectos como a noção de que os dados e os métodos do JGD têm códigos técnicos que dizem respeito à neutralidade dos dados, das maneiras como foram coletados e das tecnologias implicadas nesses processos. Sendo assim, é necessária uma negociação desses códigos técnicos, e não simplesmente aceitá-los, para se precaver da ideia de neutralidade das tecnologias usadas e de um retorno da suposta neutralidade do jornalismo. Atenta-se aqui para uma possível estabilização do JGD que legitime essas noções.

À divisão entre jornalismo informativo e opinativo seguiu a ideia de que a própria informação não poderia ser, como desejado, neutra, já que é enviesada, entre outros fatores, pela própria linguagem usada. Essa divisão informação/opinião também implicou na subdivisão em gêneros que, como visto, surgem principalmente a partir de semelhanças na forma dos produtos jornalísticos.

Isso leva a considerar duas instâncias de estabilização do JGD. A primeira é o uso de técnicas e métodos na apuração e elaboração de produtos jornalísticos já estabelecidos como formatos próprios de um gênero. A segunda é considerar outros formatos jornalísticos concebidos em função dos dados e seu processamento.

Os formatos já típicos do jornalismo, como reportagens, charges, matérias televisivas, colunas de opinião, entre outros presentes em classificações de gêneros jornalísticos, podem ser considerados artefatos que já passaram por um processo de estabilização. Para produzi-los, são usadas várias técnicas de apuração também já estabilizadas, como realizar entrevistas, pesquisar documentos, ou até retirar informação de *releases*<sup>108</sup>.

---

<sup>108</sup> Textos produzidos e repassados a meios de comunicação por assessorias de imprensa, com formato de matéria jornalística para propiciar a sua publicação, preferentemente sem modificações.



Outras técnicas de apuração não estariam estabilizadas, como a análise de bases de dados e outras técnicas do JGD. Essas técnicas ainda precisariam ser apropriadas e negociadas para alcançarem a estabilização e talvez fazer parte da identidade do jornalismo como um todo ou de uma identidade de quem faz JGD. Essa estabilização das tecnologias acontece principalmente no cotidiano, nos níveis meso e micro.

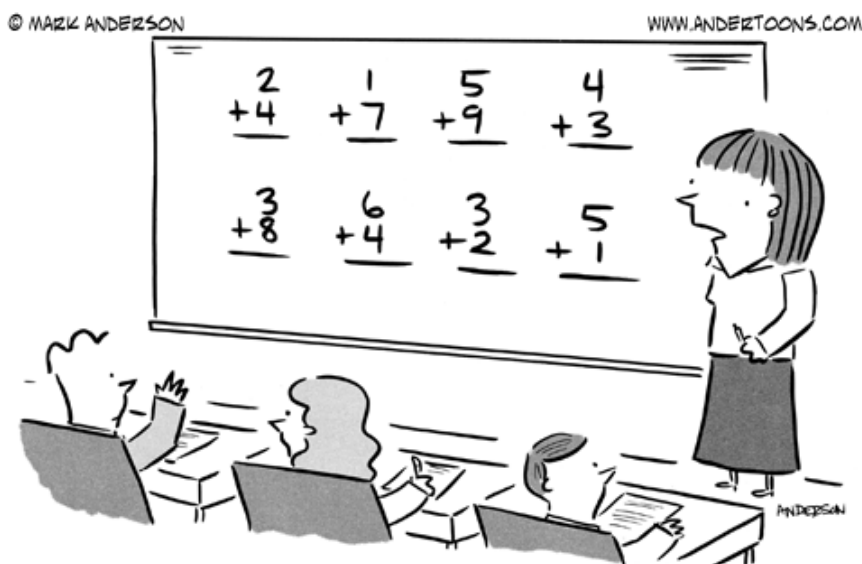
Já a estabilização de formatos jornalísticos que ainda não fazem parte das classificações abordadas, como alguns exemplos citados na seção 4.1, depende não só do grupo social de jornalistas, mas do grupo social de leitoras e leitores, que vão negociar e se apropriar desses “novos” formatos, considerados na seção 4.1.

Questionar a suposta objetividade dos dados e a presença deles em gêneros e formatos jornalísticos ajuda a situar o JGD em geral e as visualizações em particular como mais uma maneira de fazer jornalismo, relacionada à profundidade do jornalismo interpretativo, mas que também pode estar presente em outros gêneros ou formatos. Dessa maneira, o JGD é situado como mais um modo de fazer do cotidiano, através do qual cada pessoa constrói suas identidades.



Visualização 4: localização das referências bibliográficas usadas na presente dissertação. FONTE: elaboração própria.

# 4 FORMATOS E PRÁTICAS DO JGD



*“Primeiro vamos resolver estas contas. Podemos nos preocupar com mineração de dados depois.” (Mark Anderson)*

Neste capítulo é discutida a relação entre JGD e visualização como métodos que se sobrepõem e complementam. Conjuntos de dados podem ser interessantes para encontrar informações específicas que deem origem a matérias ou outros produtos jornalísticos, mas também podem ser interessantes como um todo, desde que disponibilizados de maneira adequada. Nesse aspecto, o JGD propõe outras relações entre o jornalismo e o público, repensando o papel do jornalismo.

Visualizações, em tanto que abordagem gráfica de conjuntos de dados, podem ter um papel importante nessas relações. Por isso, são apresentadas várias definições e considerações que ora confirmam e ora servem para discutir aspectos dessa definição.

A relação entre JGD e visualizações também é considerada a partir das semelhanças entre diferentes propostas de métodos de trabalho que consideram tanto a possibilidade de descobrir informações latentes em grandes conjuntos de dados como o uso de visualizações – ou outros tipos de interface – para disponibilizar

grandes conjuntos de dados ao público. Os métodos de trabalho incluem descrições breves de algumas ferramentas que podem ser utilizadas, assim como uma breve consideração sobre a suposta urgência que jornalistas têm de aprender a programar. Como complemento, são apresentados e descritos alguns exemplos de visualizações além dos que apareceram ao longo deste trabalho.

#### **4.1 O PAPEL DOS GRANDES CONJUNTOS DE DADOS NO JORNALISMO**

Os dados obtidos através de técnicas de JGD podem ser aproveitados de diversas maneiras. É possível usá-los como pista para decidir os rumos de uma investigação jornalística, escrever uma matéria e apresentar os dados ao longo do texto, apresentar os dados em uma tabela ou infográfico anexo, mostrar infográficos durante um jornal televisivo, fazer uma visualização complexa ou pegar um único dado para ser destacado em uma manchete. Mas também é possível divulgar o conjunto de dados obtidos, como aponta Liliana Bounegru:

o jornalismo de dados dedica tanta – às vezes, até mais – atenção aos dados propriamente ditos em vez de apenas empregá-los como forma de descobrir ou melhorar uma reportagem. Por isso, vemos o Datablog do The Guardian e o jornal Texas Tribune publicando conjunto de dados lado a lado com as notícias - ou até mesmo apenas os dados sozinhos – para as pessoas analisarem ou explorá-los. (BOUNEGRU, 2012)

Liliana cita dois jornais fazendo dos “dados sozinhos” um produto jornalístico. É possível que essa divulgação seja feita somente pela facilidade em fazê-lo, mas será considerada aqui qual a função dos “dados sozinhos” e dos grandes conjuntos de dados em tanto que produto jornalístico. Jason Grotto, repórter do Chicago Tribune, em entrevista (PARASIE; DAGIRAL, 2012, P. 30), questiona a utilidade dos dados em si, porque

o que [os dados] te dizem? Ok, aconteceram alguns crimes na sua vizinhança, mas o que isso significa? Jornalistas interpretam, filtram, fazem todo tipo de coisas que acho importantes. É essa

nossa argumentação. Então dados em si não são necessariamente valiosos.<sup>109</sup>

Toni Piqué compara a função do jornalismo com a de um cozinheiro:

Os dados crus não informam. Para chegar aí [a informarem] precisam de cozinha: reportagem e edição, que é verificação, valoração, seleção, contexto, fundo e formato. A carne crua não é comida. Os dados crus não são jornalismo – por muito [jornal *Texas*] *Tribune*, *nonprofit* [possivelmente a agência ProPublica] e Tim Berners-Lee que os amparem. É necessário cozinhar. E nem todos são bons cozinheiros. (PIQUÉ, 2009)<sup>110</sup>

Para Piqué, Grotto, e Little, os dados precisam passar por algum tipo de edição para que sejam considerados um produto jornalístico. Esse “cozinhar” pode ser o “comunicar” que Bradshaw (2011) coloca como último passo do processo do jornalismo de dados e trata de várias práticas, como editar textos que contenham as informações, realizar visualizações, divulgar informações nas redes sociais da internet, considerar os fatores humanos da reportagem, entre outras. Esses e outros processos são aqueles que vão transformar os dados crus em um produto jornalístico adequado a um público mais amplo.

Dentro dessas funções de “comunicar”, parte do trabalho de quem edita pode ser transferido a quem lê, através de recursos interativos que permitem a escolha de uma ordem de leitura, ou de selecionar somente a parte que interesse ao leitor:

Deixar ferramentas de tomada de decisões nas mãos de indivíduos representa uma quebra com as normas do jornalismo, e as reservas expressadas por muitos jornalistas estadunidenses são testemunha disso. Jornalistas construíram sua credibilidade através de poder analisar e oferecer interpretações a um público que não necessariamente tem tempo ou habilidades para eles mesmos extraírem os significados dos eventos. Com a sugestão que as próprias pessoas podem entender os dados aos quais

<sup>109</sup> Tradução livre. No original: “what does it tell you? Okay, you had some crimes in your neighborhood, but what does it mean? Journalists interpret, filter, do all kinds of things that I think are important. That’s what we bring to bear. So data in and of itself isn’t necessarily valuable.”

<sup>110</sup> Tradução livre. No original: “Los datos crudos no informan. Para llegar ahí necesitan cocina: reporte y edición, que es verificación, valoración, selección, contexto, fondo y formato. La carne cruda no es comida. Los datos crudos no son periodismo –por mucho *Tribune*, *nonprofit* y Tim Berners-Lee que los amparen. Hay que cocinar. Y no todos son buenos cocineros.”

conseguiram acesso, esses jornalistas exploram um novo tipo de relação com o público. (PARASIE, 2011, p. 10)<sup>111</sup>

Além de um “novo” tipo de relação, as possibilidades de interação atentam para novos produtos jornalísticos, como newsgames, questionários, testes, entre outros. No que se refere ao JGD, um exemplo é *Is It Better to Rent or Buy?* (BOSTOCK; CARTER; TSE, 2014), ferramenta do *The New York Times* para analisar se compensa comprar ou alugar um imóvel. Se trata de um produto jornalístico que faz uso de uma base de dados e de cálculos para, a partir das informações fornecidas – 21 dados, informados através de gráficos que se modificam de acordo com o controle do leitor – informa o valor máximo pelo qual compensa alugar um imóvel semelhante (Fig. 8).

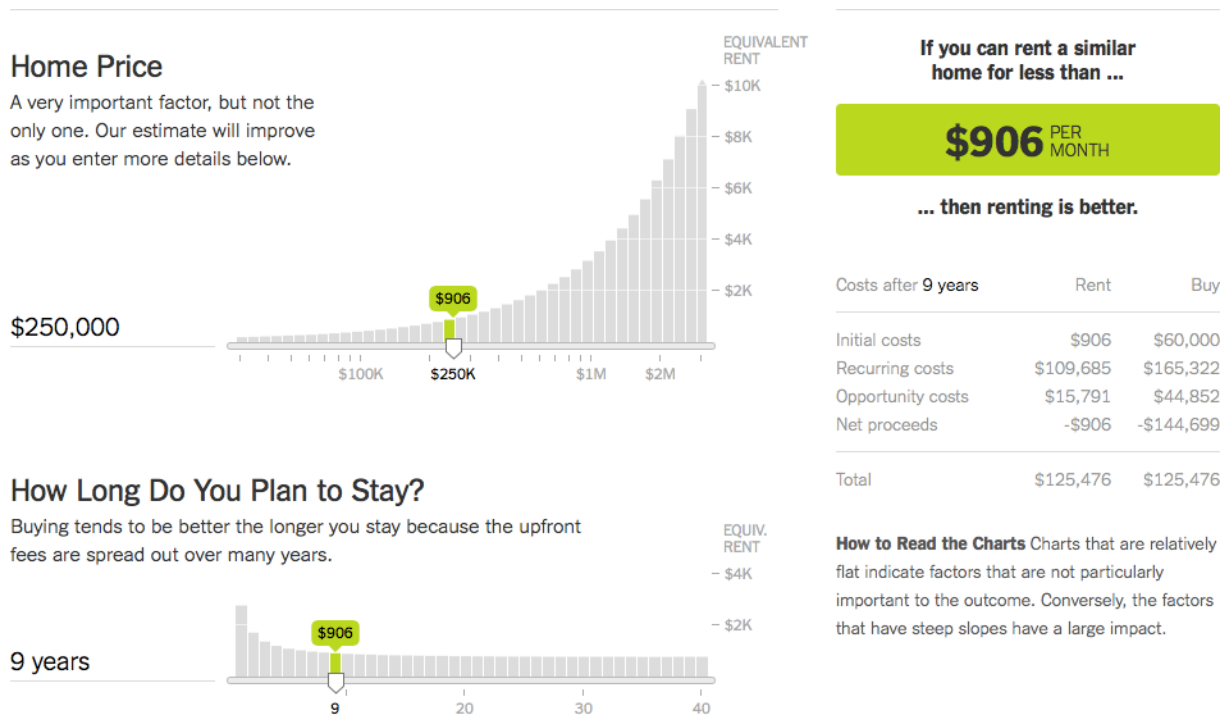


Fig. 8: Tela do aplicativo “Is It Better to Rent or Buy?” com duas das 21 variáveis que podem ser escolhidas para obter um resultado personalizado. Fonte: BOSTOCK; CARTER; TSE, 2014.

O Basômetro (ver seção 4.4.4) também fornece a quem lê informação personalizada. A função do jornalismo de

<sup>111</sup> Tradução livre. No original: “Placing decision-making tools in the hands of individuals represents a break with the previously-held norms of journalism, and the reservations expressed by many American journalists are testament to this. Journalists have built their credibility on being able to analyse and offer interpretations to a public that does not necessarily have the time or the skills to itself extract the significance of events. By suggesting that individuals can themselves make sense of the data to which they gain access, these new journalists are exploring a new relationship with the public.”

interpretação de dados crus adquire mais uma possibilidade, que é a de ajudar a navegar por um conjunto de dados em busca dos dados que o público precise ou queira. É feita a divulgação de conjuntos de dados completos, mas também é proporcionada uma maneira de aproveitá-los.

Outra questão sobre a divulgação de dados em sua totalidade, influenciada pelo movimento *Open Data*<sup>112</sup>, é que a disponibilização da informação crua, junto aos métodos e códigos de programação usados durante a análise, se tornem parte dos métodos e resultados do JGD (HOWARD, 2014, P. 4-5). Entre os exemplos aqui expostos, o que se caracteriza por disponibilizar não só a visualização mas também o código da programação é o Basômetro. Essa seria uma contribuição do jornalismo contemporâneo para o bem público “deixando as coisas visíveis e apontando novas questões” (PARASIE; DAGIRAL, 2012, P. 32)<sup>113</sup>, e também para a democracia:

várias figuras e acadêmicos afirmam que o “JGD” contemporâneo melhora a maneira em que o jornalismo pode contribuir para a democracia – especialmente em um momento no qual cada vez mais conjuntos de dados são disponibilizados por governos. (PARASIE; DAGIRAL, 2012, P. 3)<sup>114</sup>

Isso é feito de três maneiras (PARASIE; DAGIRAL, 2012, P. 3-4): reforçando a objetividade da informação, oferecendo novas ferramentas para manter a responsabilização<sup>115</sup>, e aumentando a participação política dos cidadãos através da sua própria produção e análise dos dados. Nessa situação também é criado um novo tipo de relação entre o meio e quem o lê, criando novas formas de mediação: a ou o jornalista obtém os dados – talvez contornando sistemas que tentem impedir o acesso à informação – e também projeta um sistema que pode ser apropriado pelo público.

Ainda pensando em novas relações entre leitores e jornalismo, há situações nas quais é solicitada a ajuda do público para

<sup>112</sup> Movimento que defende a disponibilização livre de informação, sem impedimentos de uso ou redistribuição como os causados pelos direitos autorais.

<sup>113</sup> Tradução livre. No original: “making things visible and pointing out new issues.”

<sup>114</sup> Tradução livre. No original: “several figures and scholars have claimed that contemporary “data-driven journalism” improves the way journalism can contribute to democracy – especially at a time when a growing number of data sets are released by governments.”

<sup>115</sup> Tradução do termo *accountability*, se refere à relação de fiscalização, por parte de cidadãos, de pessoas que ocupam cargos públicos.

trabalhos que não possam ser automatizados<sup>116</sup>. Um exemplo foi a análise de quase 500 mil documentos de gastos dos parlamentares britânicos proposta pelo jornal *The Guardian*. Quem quisesse colaborar deveria entrar na página do jornal e avaliar se a imagem do documento que lhe era apresentada tinha gastos irrelevantes ou se mereceria ser analisada com mais cuidado. A partir dessa marcação, o jornal saberia quais documentos serviriam para investigações mais aprofundadas (ANDERSEN, 2009). O sistema incluía estímulos como o número de documentos que ainda faltava investigar, além de rankings com os nomes de quem revisou mais documentos. Outro exemplo é a maneira que a organização *Nerds for Nature* encontrou para levantar dados sobre como uma floresta se recupera de um incêndio. Nas trilhas dos parques onde ocorreram incêndios, foram instaladas placas informativas com suportes para telefones celulares (Fig. 9). Quem passasse pelo lugar poderia colocar o telefone no lugar indicado, fotografar a paisagem e enviá-la pela internet. O suporte para telefone fazia com que as fotografias tivessem enquadramentos similares, o que junto às informações de data presentes no arquivo da foto e na postagem na internet permitiam uma análise da recuperação da vegetação (NERDS FOR NATURE, 2013B)<sup>117</sup>.

#### 4.1.1 Relação entre público e jornalismo através da visualização

A visualização de dados pode assumir aspectos de uma ou várias das “novas relações” (PARASIE, 2011, p. 10) entre público e jornalismo aqui apresentadas. Por exemplo, a publicação de um mapeamento extensivo usado para produzir uma reportagem pode permitir a busca por dados que não tenham sido especificamente abordados na matéria. Assim, é possível que uma pessoa queira procurar informações sobre o local onde mora. Essa situação não é nova, mas atenta para o papel do jornalismo como facilitador, como produtor de interfaces que permitam que mais pessoas

---

<sup>116</sup> Se por um lado essa “terceirização” de serviços repensa as maneiras que o público tem de participar na atividade jornalística, seria importante pensar em questões referentes ao trabalho não remunerado feito pelo público, à confiabilidade das informações fornecidas e às maneiras como o público é persuadido a colaborar.

<sup>117</sup> Esse exemplo não é de uma organização jornalística, mas foi incluído aqui pela proximidade com as práticas do JGD.



apreendam a informação de outra maneira, ou que possam escolher quais informações queiram ver.

Aqui o papel dos e das jornalistas como comunicadores pode ser estendido ao papel de projetista de tecnologias que serão apropriadas por quem as lê ou interage com elas da mesma maneira que quem produz visualizações se apropria das ferramentas que usa, em uma outra instância de consumo e de produção (MACKAY, 1997B, P. 7).



Fig. 9: Placa informativa com suporte para fotografar. Fonte: NERDS FOR NATURE, 2013A.

A produção jornalística pode ser analisada como uma tecnologia que pode ser apropriada de várias maneiras e que pode ter passado por um “desenvolvimento privilegiado” (DE CERTEAU,

2007, p. 114). Assim como a escolha de uma palavra menos complicada pode facilitar a leitura de uma matéria, a escolha de certos aspectos de uma visualização pode facilitar ou dificultar a leitura. A questão de que tanto uma visualização é acessível a pessoas com diferentes capacidades e repertórios segue uma lógica similar à do texto, em cuja construção são considerados vários fatores, mas predomina uma ideia subjetiva sobre o conhecimento de quem vai ler. Nota-se também que a capacidades que o público tem de ler uma visualização – ou inclusive outras linguagens visuais, como uma fotografia – são subestimadas, em comparação com a capacidade de ler um texto escrito. Visualizações simples são consideradas “difíceis de entender” por jornalistas que usam palavras complicadas em seus textos. Por exemplo, durante o trabalho com infografia no jornal Gazeta do Povo, um repórter, conhecido pela linguagem complexa de seus textos, questionou ao autor da presente pesquisa se não seria necessário pintar o oceano de azul, já que o leitor supostamente não reconheceria as formas de um mapa-múndi e seria necessário diferenciar o que é terra e o que é água.

#### **4.2 VISUALIZAÇÃO: CONCEITOS E POSSIBILIDADES**

Lev Manovich (2007) define visualização como “um mapeamento entre dados discretos e uma representação visual” (p. 148) e “mapeamento de algumas propriedades dos dados em uma representação visual” (p. 151). Por mapeamento entende-se uma transposição da linguagem numérica para a linguagem visual, levando em consideração a distribuição espacial dos elementos da visualização segundo regras estabelecidas por quem a realiza. Por exemplo, a regra de um gráfico de barras é que a altura das barras sejam proporcionais às quantidades que elas representam, ou que um mapa tenha as diferentes regiões pintadas em tons mais claros ou mais escuros representando diferentes intensidades do valor que está sendo representado. Na atualidade, os aparelhos que produzem esse tipo de imagens técnicas são os computadores, mas mesmo considerando a hipótese de que o gráfico de barras

seja feito em papel, usando canetas e régua, a visualização implica na subordinação da produção da imagem às regras estabelecidas.

Retomando a apresentação de conceitos, Michael Friendly fala na visualização de informação como

representação visual de coleções de grande escala de informação não numérica, como arquivos e linhas de código em programas, bases de dados bibliográficas, redes de relações na internet, e assim por diante. (FRIENDLY, 2009, P. 2)<sup>118</sup>

A visualização de informação abarcaria também a visualização de dados, que seria a abstração dos dados “de alguma maneira esquemática, incluindo atributos ou variáveis para as unidades de informação” (FRIENDLY, 2009, P. 2)<sup>119</sup>. Unwin; Theus e Hofmann (2009, P. 6) também diferenciam os dois tipos de visualização, citando que a visualização de dados “é para exploração, para descobrir informação e também para apresentá-la” enquanto se refere à visualização de informação como tendo o objetivo de “apresentar os dados brutos em si, revelando a variabilidade e as incertezas inerentes” (UNWIN; THEUS; HOFMANN, 2006, P. 6)<sup>120</sup>. Eles ainda especificam algumas das vantagens da visualização como método:

Por que visualizar dados é importante? A visualização de dados é boa para limpar dados, para explorar dados [...], para identificar tendências e agrupamentos, para identificar padrões locais, para avaliar saídas de modelagens e para apresentar resultados. A visualização é essencial para a Análise Exploratória de Dados. (UNWIN; THEUS; HOFMANN, 2006, P. 1, 2006)<sup>121</sup>

A primeira defesa das possibilidades da representação gráfica de números data do final do século 18 e foi feita por William Playfair – considerado um dos “inventores” dos gráficos (TUFTE, 2001, P. 32):

<sup>118</sup> Tradução livre. No original: “visual representation of large-scale collections of non-numerical information, such as files and lines of code in software systems, library and bibliographic databases, networks of relations on the internet, and so forth.”

<sup>119</sup> Tradução livre. No original: “in some schematic form, including attributes or variables for the units of information.”

<sup>120</sup> Tradução livre. No original: “Data Visualization is related to Information Visualization, but there are important differences. Data Visualization is for exploration, for uncovering information, as well as for presenting information. It is certainly a goal of Data Visualization to present any information in the data, but another goal is to display the raw data themselves, revealing the inherent variability and uncertainty.”

<sup>121</sup> Tradução livre. No original: “Why is visualizing data important? Data visualization is good for data cleaning, for exploring data [...], for identifying trends and clusters, for spotting local patterns, for evaluating modelling output and for presenting results. Visualization is essential for Exploratory Data Analysis.”

uma pessoa que investigou cuidadosamente uma tabela impressa, encontra, quando termina, que ela tinha somente uma ideia muito vaga e parcial do que ele leu; e que como uma figura escrita na areia, é em seguida completamente apagada e desfigurada. A quantidade de transações comerciais em dinheiro, e de perdas ou ganhos, é possível de ser facilmente representada em desenhos, como qualquer parte do espaço, ou a figura de um país; mas, até agora, não foi tentado fazer isso. Foi partindo desse princípio que estes Gráficos foram feitos; e, enquanto eles dão uma ideia simples e distinta, eles são tão perfeitamente precisos como úteis. Inspeccionando qualquer um destes Gráficos com atenção, uma impressão suficientemente distinta será obtida, que não será danificada por um tempo considerável, e a ideia que se mantiver será simples e completa, incluindo a duração e a quantidade. (PLAYFAIR *apud* TUFTE, 2001, P. 32)<sup>122</sup>

Tufte dá continuidade ao trabalho de Playfair, falando nas vantagens dos gráficos não como meros substitutos de tabelas, mas como o melhor método para lidar com informação quantitativa:

de todos os métodos para analisar e comunicar informação estatística, gráficos bem desenhados são geralmente a maneira mais simples e ao mesmo tempo a mais poderosa. (TUFTE, 2001, P. 10)<sup>123</sup>

Esse “poder” está na capacidade da visualização de pôr

milhões de bits de informação em uma única página, diante de nossos olhos. Nenhum outro método de mostrar informação estatística é tão poderoso. (TUFTE, 2001, P. 26)<sup>124</sup>

Essa densidade das visualizações permite que um conjunto de dados seja representado em sua totalidade, sem que haja cortes nem amostragens, o que se tornou uma premissa da visualização (MANOVICH, 2007, P. 157). A vantagem de se trabalhar com todo o conjunto e não com amostragens ou outro tipo de redução é que “amostras não vão mostrar as exceções, estruturas locais, ou erros

<sup>122</sup> Tradução livre. No original: “a man who has carefully investigated a printed table, finds, when done, that he has only a very faint and partial idea of what he has read; and that like a figure imprinted on sand, is soon totally erased and defaced. The amount of mercantile transactions in money, and of profit or loss, are capable of being as easily represented in drawing, as any part of space, or as the face of a country; though, till now, it has not been attempted. Upon that principle these Charts were made; and, while they give a simple and distinct idea, they are as near perfect accuracy as is any way useful. On inspecting any one of these Charts attentively, a sufficiently distinct impression will be made, to remain unimpaired for a considerable time, and the idea which does remain will be simple and complete, at once including the duration and the amount.” (PLAYFAIR, William. *The Commercial and Political Atlas*. Londres, 1786).

<sup>123</sup> Tradução livre. No original: “of all methods for analyzing and communicating statistical information, well-designed data graphics are usually the simplest and at the same time the most powerful.”

<sup>124</sup> Tradução livre. No original: “millions of bits of information on a single page before our eyes. No other method for the display of statistical information is so powerful.”

sistemáticos nos dados.” (UNWIN; THEUS; HOFMANN, 2006, P. 1)<sup>125</sup>. Essa premissa, porém, pode ocasionar dificuldades de representação e leitura dos dados, mas a solução não está em diminuir a quantidade de informação e sim em conseguir clareza através da forma em que os dados são visualizados:

a quantidade de detalhe é um problema completamente separado da dificuldade de leitura. *Entralhamento e confusão são falhas do design, não atributos da informação.*[...] E o ponto é encontrar estratégias de design que revelem detalhe e complexidade – mais do que culpar os dados por um excesso de complicação. Ou pior, de culpar os leitores por não entenderem. (TUFTTE, 1990, P. 51 E 53, GRIFOS NO ORIGINAL)<sup>126</sup>

Jacques Bertin (1986, P. 183–184) especifica duas dificuldades: a representação de muitas camadas de informação e a representação minuciosa de redes. Essas dificuldades apontadas por Bertin são inerentes a visualizações impressas, com as quais Bertin trabalhava; meios interativos permitem outras possibilidades para contorná-las. A interatividade permite superar dificuldades inerentes à quantidade de dados e ao meio pelo qual são representados, alterando o gráfico através de opções como mudar o tamanho, estabelecer relações, ampliar detalhes ou classificar (UNWIN; THEUS; HOFMANN, 2006, P. 73).

Fry também fala sobre a interatividade como mudança na leitura da visualização, com foco em que o leitor possa escolher como quer visualizar a informação (FRY, 2008, P. 11), mas também como algo que permite que a visualização seja alterada não só pelo leitor, mas pelos dados em si, atualizados constantemente. Para Fry, os dados são móveis, e uma das propostas do seu livro é construir visualizações dinâmicas com fluxos de dados constantemente atualizados (FRY, 2008, P. 3).

<sup>125</sup> Tradução livre. No original: “You might ask, if a dataset is so large, why not just take a big sample? But samples will not pick out outliers, local structures, or systematic errors in the data.”

<sup>126</sup> Tradução livre. No original: “the quantity of detail is an issue completely separate from the difficulty of reading. *Clutter and confusion are failures of design, not attributes of information.* [...] And so the point is to find design strategies that reveal detail and complexity – rather than to fault the data for an excess of complication. Or, worse, to fault viewers for a lack of understanding.”

#### 4.2.1 Observações sobre os conceitos de visualização

Em uma tentativa de consolidar as definições dos autores aqui apresentados, será considerada, para o presente trabalho, que o termo visualização de dados se refere a **uma representação gráfica de uma grande quantidade de dados não necessariamente numéricos, mas que possam ser quantificados ou classificados**. Porém, essa tentativa de definição tem algumas ambiguidades, discutidas a seguir.

A primeira é considerar o quanto um conjunto de dados é grande. Tufte (2001, p. 56) define que uma tabela é melhor do que um gráfico quando o conjunto de dados tem 20 números ou menos. Bertin observa que para comparar duas tabelas de cem linhas por cem colunas, seriam necessários 20 mil “instantes de percepção”, enquanto em uma visualização a comparação pode até ser instantânea (BERTIN, 2011, p. 3). O ponto em que um volume de dados se torna incompreensível varia segundo quem o visualiza e segundo a forma de visualização escolhida. É possível afirmar que à medida que a quantidade de dados a serem lidos aumenta, a capacidade de apreender seu significado e poder fazer relações entre eles diminui, o que pode justificar a apresentação dos dados de maneira gráfica.

O tamanho do conjunto de dados também pode ser medido em termos de processamento necessário. Unwin se refere a vários autores que problematizam o tamanho dos conjuntos de dados como um problema computacional. Em 1959, 1.000 informações eram consideradas um volume grande. Em 1999, o volume considerado grande alcançava os milhões de números (UNWIN; THEUS; HOFMANN, 2006, p. 9–22). Em 2010, Jer Thorp realizou uma visualização de dados de ligações telefônicas, na qual os dados de 10 milhões de pessoas ocupavam mais de 5,5 gigabytes em arquivos de texto. Como ele estava usando um computador pessoal, usou trechos dos dados para fazer testes antes de visualizá-los em sua totalidade (THORP, 2010).

Diferentes tipos de dados, diferentes formas de visualizá-los e diferentes ferramentas para fazer essas operações exigem mais ou menos processamento e, em consequência, combinações de

hardware e software adequadas a essa exigência. Dependendo do conjunto de dados e dos recursos computacionais disponíveis, pode ser necessário algum tipo de redução. Fry (2008, p. 30) recomenda, caso o volume de dados for muito grande, começar usando uma amostra, o que requer menos infraestrutura e permite avaliar a viabilidade do projeto. Outra maneira de contornar o grande volume de dados é eliminar dados duplicados, como feito por Eric Fischer em *Six Billion Tweets* (FISCHER, 2014B). Dos seis bilhões de *tweets*, apenas 9% seriam visíveis no mapeamento, já que muitas localizações tinham vários *tweets*, seja por terem acontecido no mesmo lugar ou por imprecisões dos aparelhos usados. Segundo Fischer, era inútil que o programa desenhasse o mesmo ponto uma e outra vez (FISCHER, 2014A). Essa redução foi adequada para o objetivo de Fischer, que era realizar um mapeamento (Fig. 10); talvez para outra abordagem fosse importante manter esses dados, mesmo duplicados.

Outra questão é qual tipo de dados pode ser representado. As definições apresentadas anteriormente dão ênfase a dados numéricos, o que está de acordo com a classificação de Jacques Bertin: os dados podem ser quantitativos, números ordinais (que podem representar categorizações) e/ou respostas afirmativas ou positivas (BERTIN, 1986, p. 184). Porém, há dados que não são numéricos e que podem ser visualizados, como é o caso da representação de relações – “sociograma”, nos termos de Yves Deforge (1991, p. 137). A Fig. 12 é um fragmento do infográfico sobre o jogo Brasil–Chile da Copa do Mundo de 2010 (JANNUZZI; MIRANDA, 2010). Há várias informações numéricas, como o número de passes trocados entre os jogadores, representados pela espessura da linha. A relação entre os jogadores não é numérica, mas foi transformada em números referentes a cada jogador para poder escrever o programa – na linguagem Processing – usado para gerar a visualização (PERUYERA, 2010B).

Para evitar uma tentativa de classificação exaustiva e abarcar possíveis exceções, é possível afirmar que os dados que podem ser visualizados são aqueles que compartilham entre si características que permitem encaixá-los em uma estrutura organizada e que



possibilitam que sejam processados, analisados e visualizados da mesma maneira.

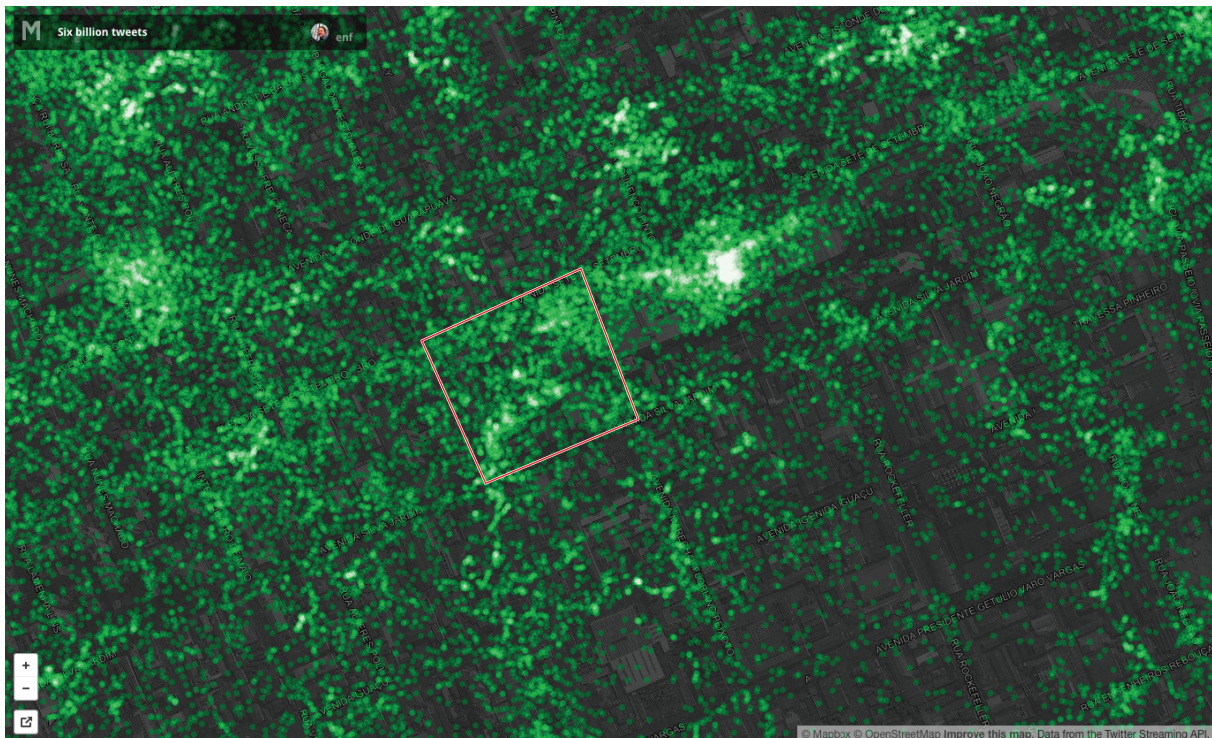


Fig. 10: Trecho da visualização "Six Billion Tweets". Em destaque, a sede da UTFPR do centro de Curitiba.  
Fonte: FISCHER, 2014B, DESTAQUES NOSSOS.

Essa organização e/ou transformação prévia dos dados pode implicar na transformação de características qualitativas em quantitativas, como por exemplo transformar a opinião sobre um filme em uma nota de 0 a 10, ou revelar valores numéricos que estão por trás da informação, como por exemplo obter a latitude e longitude de um lugar do qual se sabe a rua e a numeração. O diagrama de dispersão da Fig. 11 mostra uma classificação de diferentes personagens da atriz Meryl Streep. A escolha do lugar de cada personagem é uma transformação da subjetividade da percepção do autor do diagrama em um valor numérico – representado pela distância do ponto que representa cada personagem em relação aos eixos estabelecidos.

Por último, é necessário considerar que pode haver representações que estão em sintonia com as possibilidades da visualização, mas se valem de outros meios que não o visual. Por exemplo, o infográfico em vídeo *Usain Bolt's Gold in the 2012 Olympics 100 Meter Sprint – All the medalists* simula uma hipotética



competição entre todos os recordistas mundiais dos 100 metros rasos, e representa o momento em que cada corredor cruzaria a linha de chegada com um “bip”, ajudando o leitor a “visualizar” as diferenças de centésimos de segundo entre os tempos (QUEALY; ROBERTS, 2012).

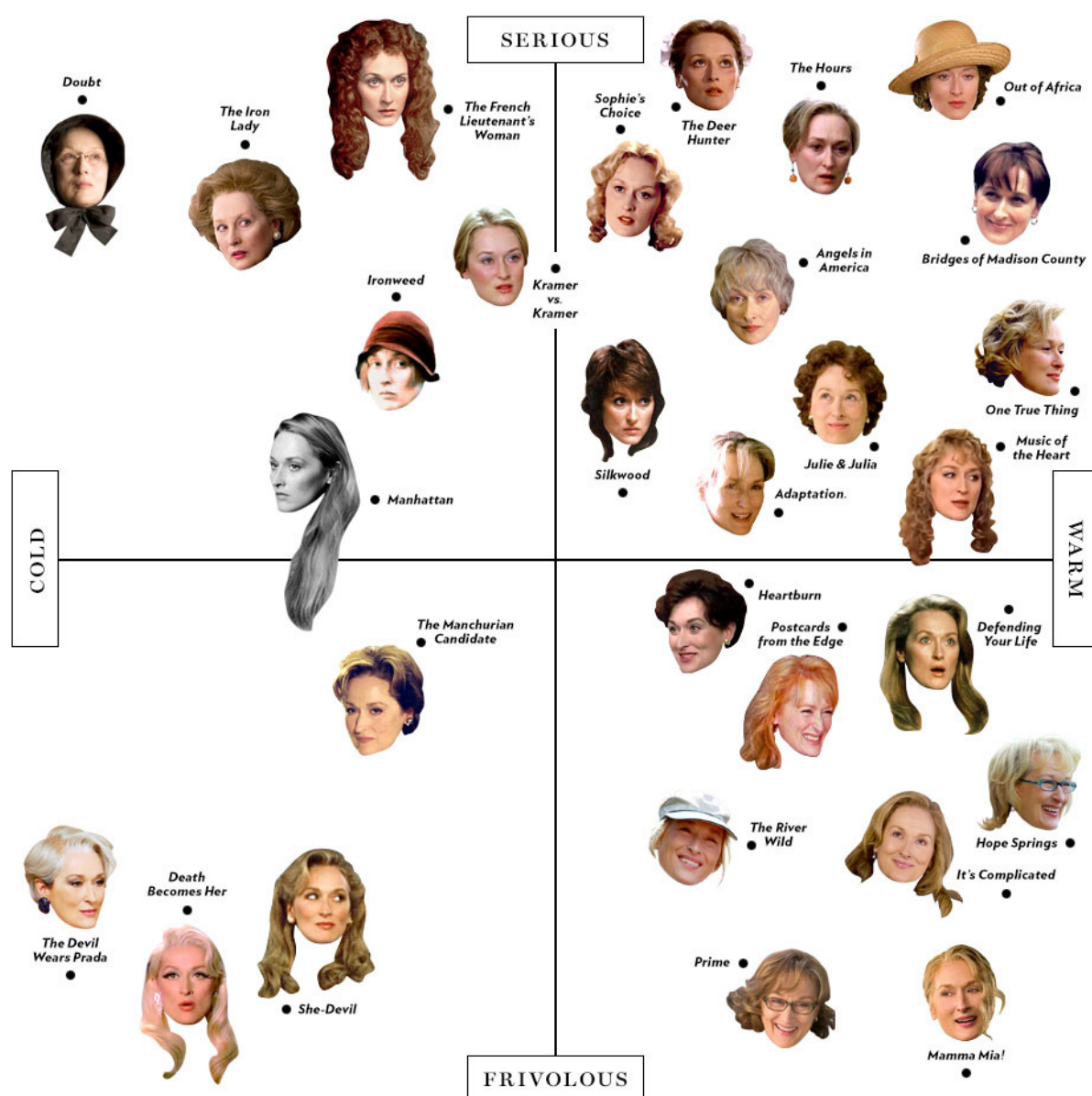


Fig. 11: Diagrama de dispersão com diversos personagens da atriz Meryl Streep. Fonte: GLAZER, 2012.

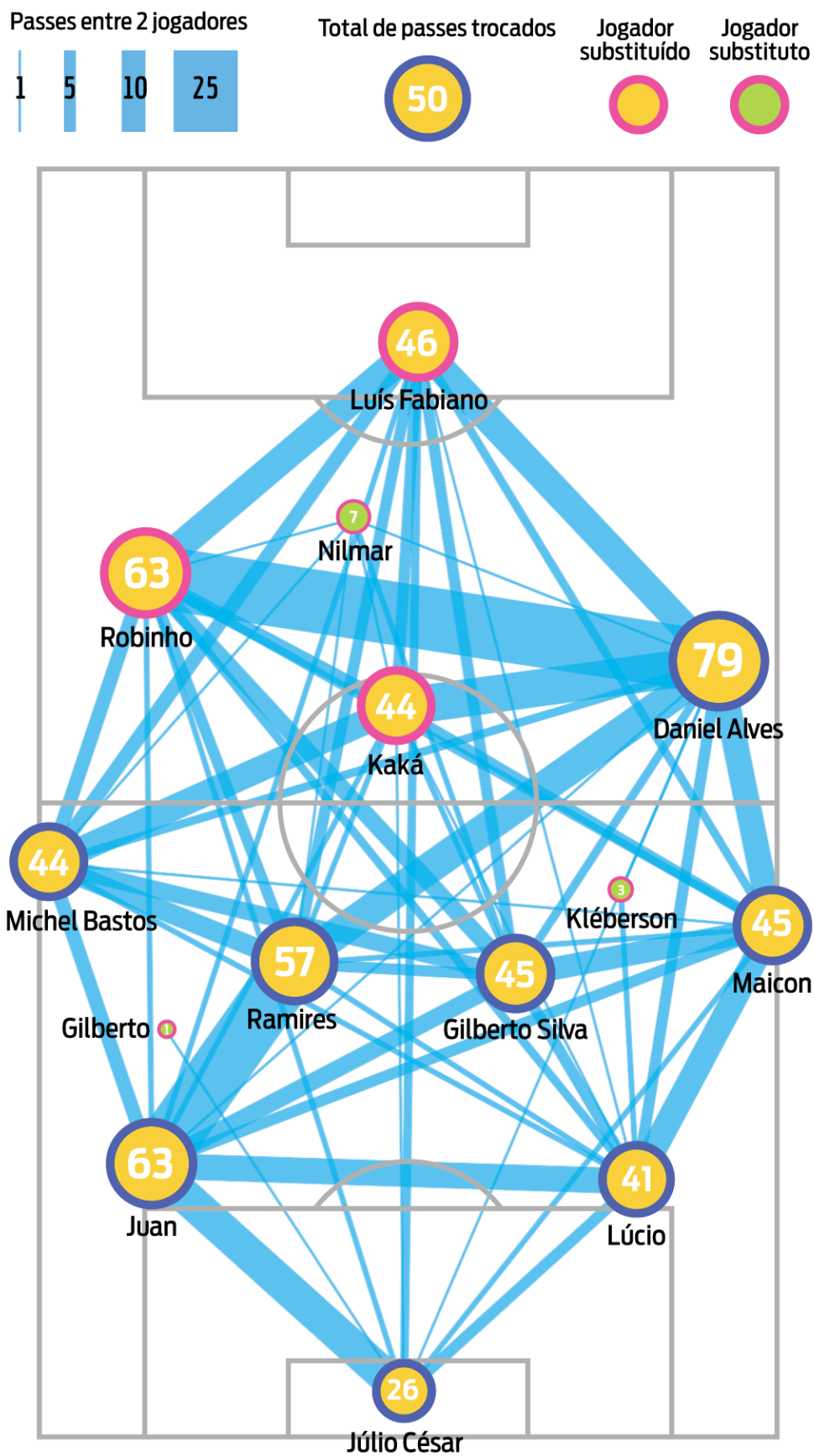


Fig. 12: Passes trocados entre jogadores da Seleção Brasileira Fonte: JANNUZZI; MIRANDA, 2010.

O projeto Tasty Tweets (ZORINA; VAN DER VLEUTEN; FRANTZIS, 2012), consta de um dispositivo que mistura sucos de diferentes frutas em quantidades proporcionais às suas menções no Twitter, transpondo informação numérica para uma informação de sabor. Essas “visualizações não visuais” atentam para que o termo “visual” não seja encarado de maneira tão literal, aumentando assim as possibilidades da linguagem da visualização para além do gráfico, explorando as possibilidades multimídia e multimodal, ou inclusive prescindindo da linguagem gráfica, o que permitiria pensar em visualizações acessíveis a deficientes visuais, por exemplo.

#### 4.2.2 Processos de leitura de visualizações

A leitura de uma visualização propicia que quem a lê não se limite a ver a figura como um todo, e sim que processe a informação de diferentes maneiras:

Nós prosperamos em mundos cheios de informação graças às nossas maravilhosas e cotidianas capacidades para selecionar, editar, isolar, estruturar, destacar, agrupar, formar pares, mesclar, harmonizar, sintetizar, focar, organizar, condensar, reduzir, escolher, categorizar, catalogar, classificar, listar, abstrair, perpassar, observar, isolar, discriminar, distinguir, separar, rotular, ordenar, integrar, misturar, verificar, filtrar, ignorar, fazer médias, aproximar, agregar, resumir, dividir em itens, rever, mergulhar, folhear, olhar, ler rapidamente, refinar, enumerar, extrair informações, fazer sinopses e separar o joio do trigo (TUFTÉ, 1990, p. 50)<sup>127</sup>.

Dentro dessas possibilidades de lidar com informações, podem ser considerados dois extremos na atitude de ler, no que Tufte chama de leituras macro e micro<sup>128</sup>. A leitura macro é a leitura da visualização como um todo. Bertin (1986, p. 13) a chama de “nível de conjunto”. Já a leitura micro é a leitura de perto, dos detalhes, de informação específica. Bertin (1986, p. 12) se refere a essa especificidade da informação como nível elementar. Bertin (1986,

<sup>127</sup> Tradução livre. No original: “We thrive in information-thick worlds because of our marvelous and everyday capacity to select, edit, single out, structure, high-light, group, pair, merge, harmonize, synthesize, focus, organize, condense, reduce, boil down, choose, categorize, catalog, classify, list, abstract, scan, look into, idealize, isolate, discriminate, distinguish, screen, pigeonhole, pick over, sort, integrate, blend, inspect, filter, lump, skip, smooth, chunk, average, approximate, cluster, aggregate, outline, summarize, itemize, review, dip into, flip through, browse, glance into, leaf through, skim, refine, enumerate, glean, synopsisize, winnow the wheat from the chaff and separate the sheep from the goats.”

<sup>128</sup> Não confundir com os níveis macro, meso e microscópico de Spinuzzi (2003, p. 30).

p. 12) ainda se refere a “níveis médios”, para designar as instâncias intermediárias entre os níveis de conjunto e elementares.

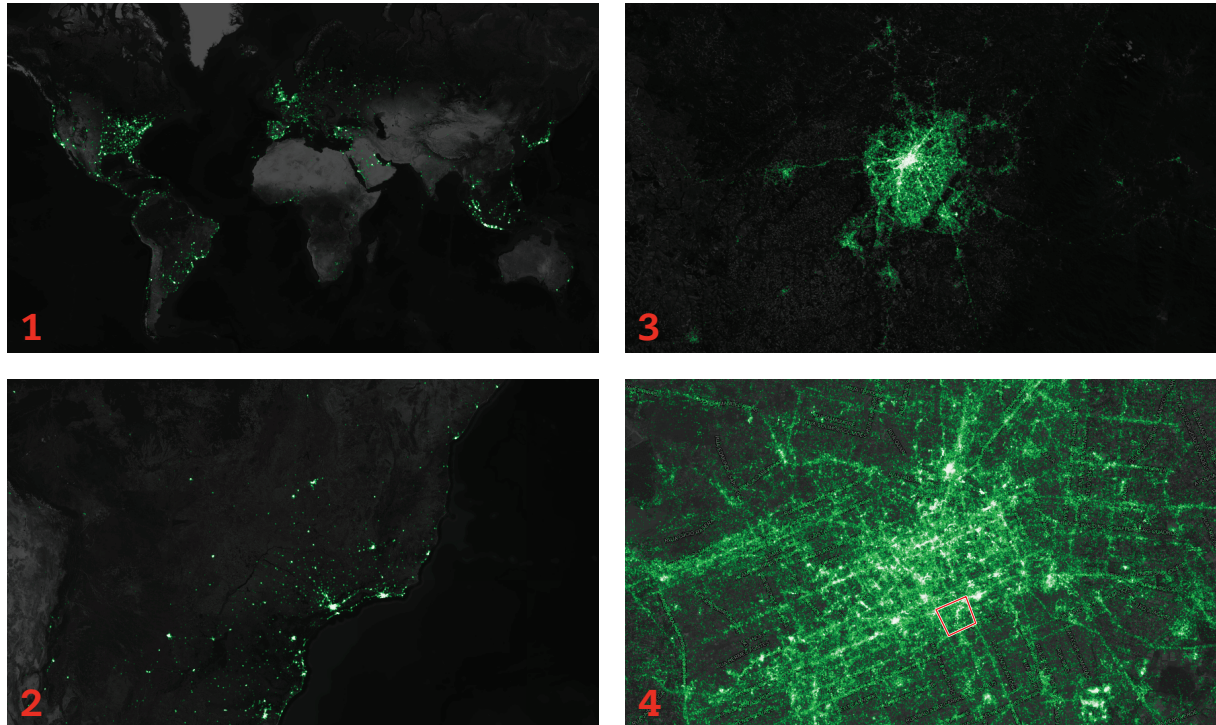


Fig. 13: Quatro níveis da visualização “Six Billion Tweets”: mundial, cidades brasileiras, Curitiba e arredores e centro de Curitiba, com a UTFPR em destaque. Fonte: FISCHER, 2014B.

O nível elementar de uma visualização equivale ao dado que ele representa (BERTIN, 1986, p. 12). É a partir das diferentes relações entre os elementos mínimos, como agrupamentos e distanciamentos, que surgem novas informações, de nível médio e de conjunto. No exemplo da Fig. 13 são apresentados diferentes níveis da visualização “Six Billion Tweets” (FISCHER, 2014B). Na visão de conjunto aparece o mundo todo. Nos dois níveis médios apresentados é possível ver várias cidades e estados em um e a cidade de Curitiba em outro. Na última imagem é visível o nível elementar – no caso, cada ponto que representa um tweet.

Uma visualização que não aporte novas informações nas leituras macro é ineficaz (BERTIN, 1986, p. 13). Essa ineficácia está presente no que Bertin (1986, p. 147) chama de “mapas para ler”: são aqueles nos quais somente é possível a leitura micro, de valores individuais. Em oposição a essa ineficácia temos os “mapas para ver” (1986, p. 147), que são as visualizações que permitem leituras macro e intermediárias, trazendo novas informações. As

visualizações apresentadas nesta dissertação foram escolhidas, entre outros fatores, por serem visualizações “para ver”. A Fig. 14 é o exemplo dado por Bertin de como uma visualização pode ser “para ler”, ou seja, que não acrescenta informações quando lida no nível macro, como acontece na outra versão do mapeamento, na mesma figura.

Considerando a premissa de que a totalidade de um conjunto de dados deve ser representada em uma visualização, temos que a informação visível em uma leitura micro também está presente em uma leitura macro – mas é apreendida de uma outra maneira. Isso leva a pensar na produção de uma visualização, que tem o que Tufte chama de “uma estratégia de desenho bem fora do comum” que é o ato de “*para clarificar, acrescenta detalhes.*” (TUFTE, 1990, P. 37, GRIFOS NO ORIGINAL)<sup>129</sup>. E são esses detalhes, cada um dos elementos, que

nos revela o nível de conjunto da informação quando tudo se torna comparável. Uma vez atingido esse nível, toda informação elementar torna-se interessante, seja como testemunho da tendência geral, seja como exceção. (BERTIN, 1986, P. 14)

A combinação das duas maneiras de leitura permite diferentes comparações entre os dados:

Desenhos [que considerem escalas] micro/macro forçam comparações locais e globais e, ao mesmo tempo, impedem a interrupção da mudança de contexto. Ou seja, exatamente o que é preciso para pensar sobre a informação. (TUFTE, 1990, P. 50)<sup>130</sup>

Assim como fazem com que dados representados uma única vez adquiram diferentes funções e significados quando lidos de uma ou outra maneira, já que “a mesma tinta serve mais de um propósito informacional; os elementos gráficos são multifuncionais” (TUFTE, 1990, P. 47)<sup>131</sup>. Também uma ausência de elementos gráficos são informação, mostrando elementos que representam exceções, por exemplo.

---

<sup>129</sup> Tradução livre. Textos originais: “most unconventional design strategy” e “to clarify, add detail”.

<sup>130</sup> Tradução livre. No original: “Micro/macro designs enforce both local and global comparisons and, at the same time, avoid the disruption of context switching. All told, exactly what is needed for reasoning about information.”

<sup>131</sup> Tradução livre. No original: “the same ink serves more than one informational purpose; graphical elements are multifunctioning”.

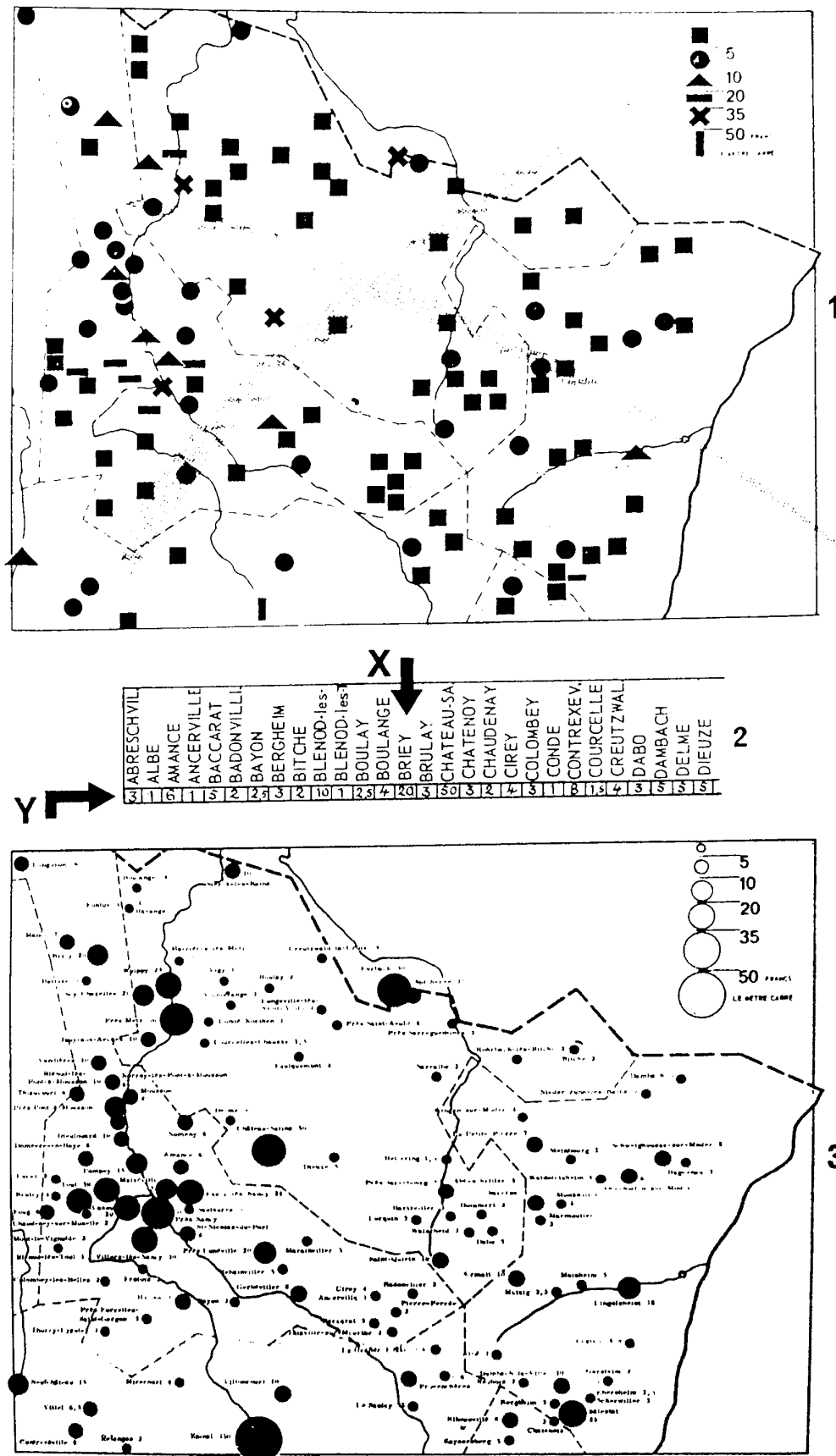


Fig. 14: Exemplos de mapas para "ler" (acima) e para "ver" (embaixo). Fonte: BERTIN, 1986, p. 146.

A alternância entre as leituras micro e macro e suas instâncias intermediárias é o que faz com que seja o leitor que determine o ritmo da leitura, no que Tufte chama de “liberdade de escolha derivada de uma vista geral” (TUFTE, 1990, p. 38)<sup>132</sup>. Quem lê uma visualização estabelece suas próprias relações entre os elementos, vai, volta, olha o conjunto, retorna, faz novas comparações. Essas atitudes são exemplos do que é possível fazer com a “liberdade de escolha” que o leitor tem ao ler uma visualização. Além de “liberdade”, Tufte também fala em “controle”.

Desenhos de alta resolução também permitem que os leitores selecionem, narrem, deem nova forma e personalizem os dados para seus próprios usos. Então, o controle da informação é dada aos leitores, não aos editores, designers ou decoradores. (TUFTE, 1990, p. 50, GRIFOS NO ORIGINAL)<sup>133</sup>

Apesar que Tufte (1990) não aborda visualizações interativas, é possível pensar na interação como mais uma forma de dar ao leitor esse “controle”. A interação também amplia o conceito de “designs de alta resolução”. A alta resolução proposta por Tufte exige suportes adequados, cujas limitações não impeçam a exibição por completo dos dados. Negativos de filme, por exemplo, têm a limitação do tamanho dos sais de prata, que são maiores a medida que a sensibilidade do filme à luz aumenta; monitores de cristal líquido têm a limitação do tamanho do pixel; o processo de impressão offset encontra seu limite na resolução máxima permitida pelas chapas. A interação se mostra como uma possibilidade para contornar essas limitações técnicas, compensando a baixa resolução de uma tela de cristal líquido em comparação com a resolução de uma página impressa. Também se faz necessário ampliar o conceito de leitura macro para abarcar situações como visualizações em vídeo ou interativas, nas quais os tempos são diferentes, talvez determinados por quem fez a visualização, mas também com a ideia de facilitar o acesso e talvez usar o tempo para transmitir noções numéricas.

---

<sup>132</sup> Tradução livre. No original: “freedom of choices that derives from an overview.”

<sup>133</sup> Tradução livre. No original: “High-density designs also allow viewers to select, to narrate, to recast and personalize data for their own uses. Thus control of information is given over to viewers, not to editors, designers, or decorators.”



### 4.2.3 Diferenças entre visualização de dados e infografia

Como o presente trabalho aborda o uso de visualizações de dados no jornalismo, se faz necessária uma diferenciação entre a visualização de dados e a infografia, considerando que o uso de visualizações no jornalismo se dá, na maioria dos casos, dentro do que é comumente chamado de infográfico.

De Pablos (1999, p. 28) define a infografia como “a apresentação impressa de um binômio imagem + texto [...] qualquer que seja o suporte onde se apresente essa união informativa: monitor, papel, plástico, barro, pergaminho, papiro, pedra<sup>134</sup>. A visualização está incluída no que De Pablos chama de imagem, acompanhada de outros tipos de imagem, como os apontados por Gonzalo Peltzer: mapas, símbolos, ilustrações, histórias em quadrinhos, etc<sup>135</sup> (PELTZER, 1991, p. 129). O infográfico “La ballena Franca”, (Fig. 15) de autoria de Jaime Serra, publicado pelo jornal argentino Clarín em 1995, foi considerado o infográfico mais influente dos últimos 20 anos pelos prêmios Malofiej<sup>136</sup> (CLARÍN, 2012). Pelas suas características e influência, pode ser considerado um exemplo de infografia jornalística “canônica”, que representa bem o conceito de De Pablos. Nele foram destacados os elementos imagéticos que se enquadrariam nos conceitos aqui apresentados de visualização de dados. Os demais elementos imagéticos cabem na definição de “infográfico” de De Pablos, mas como não representam dados numéricos, não se encaixam na definição de “visualização de dados” aqui proposta.

Ainda dentro da definição de De Pablos, as visualizações que incluam informação textual – como uso de textos explicativos, anotações (TUFTTE, 1990, p. 53–65) ou apenas legendas para identificar os elementos – também seriam consideradas infográficos. Nesta dissertação, o conceito de infográfico de De Pablos não é considerado de maneira tão estrita; sendo assim, é

---

<sup>134</sup> Tradução livre. No original: “es la presentación impresa de un binomio imagen + texto [...], cualquiera que sea el soporte donde se presente esa unión informativa: pantalla, papel, plástico, barro, pergamino, papiro, piedra.”

<sup>135</sup> Peltzer não usa o termo infográfico ou infografia com o mesmo sentido que De Pablos, mas aqui foram usados os gêneros visuais por ele enumerados por serem bons exemplos das possibilidades de imagens a que De Pablos se refere.

<sup>136</sup> Prêmios internacionais de infografia, entregues anualmente na Espanha.



usado o termo visualização mesmo quando a obra inclua elementos textuais ou legendas.

#### 4.2.4 Outros tipos de visualização

O termo visualização também é utilizado para se referir a várias atividades relacionadas à visualização de dados, mas que não serão considerados nesta dissertação. Além do termo já citado visualização de informação, Friendly (2009, p. 2) cita a visualização científica, que enfatiza a representação tridimensional de fenômenos meteorológicos, biológicos e arquitetônicos, entre outros. Outro termo utilizado é geovisualização ou visualização cartográfica, similar à visualização de dados, mas com ênfase na informação espacial (MENDONÇA, 2013, p. 73-74).

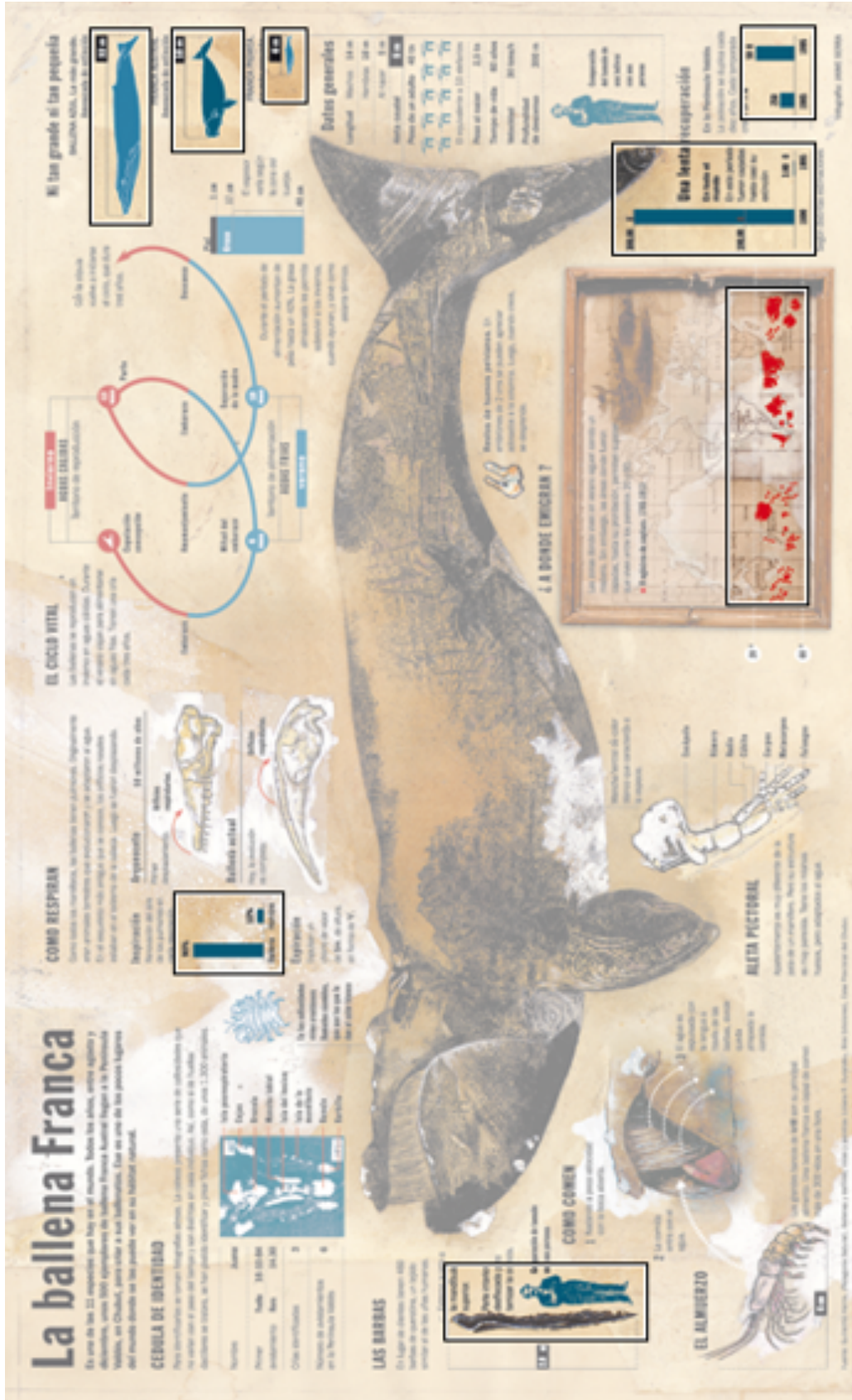


Fig. 15: Utilização de visualizações no infográfico “La ballena Franca” (destaques nossos).  
Fonte: CLARÍN, 2012.

### 4.3 MÉTODOS DO JGD E DA VISUALIZAÇÃO

Consideradas as possibilidades da visualização, e com o objetivo de ajudar na reflexão sobre os modos de uso de tecnologias por parte do jornalismo, serão apresentados alguns métodos e ferramentas usadas no JGD e nas visualizações. Quanto à parte prática, serão consideradas duas abordagens, sem intenção de julgar qual delas é a correta nem de propôr uma divisão rígida de métodos. Uma é a de Fry (2008), que propõe o uso do ambiente de programação Processing<sup>137</sup>. Fry enfatiza a flexibilidade da programação para conseguir ferramentas específicas para um objetivo:

Este livro não está preocupado em prover um punhado de “visualizações” prontas que podem ser plugadas em qualquer base de dados. Visualizações prontas podem ajudar a dar uma visão rápida da sua base de dados, mas são commodities inflexíveis que podem ser implementadas em pacotes de software. Qualquer gráfico de barras ou de dispersão feitos com Excel parecerão um gráfico de barras ou de dispersão feito no Excel. Soluções prontas somente podem dar respostas prontas, como um brinquedo que fala um punhado de frases prontas, como “as vendas tiveram um ligeiro aumento em cada um dos últimos cinco anos!” Cada problema é único, então se aproveite dessa exclusividade para solucioná-lo. (FRY, 2008, p. 16)<sup>138</sup>

Já Yau (2011) apresenta várias ferramentas, oriundas da estatística, da programação, da ilustração vetorial, da programação e da cartografia, destacando os pontos positivos e negativos de cada uma:

...você tem um monte de opções. Algumas são prontas para usar. Outras requerem um pouco de programação, já algumas ferramentas não foram projetadas especificamente para gráficos mas também são úteis.[...] Quanto mais ferramentas de visualização você souber usar e tirar vantagem, é menos provável que você fique sem saber o que fazer

<sup>137</sup> As ferramentas citadas ao longo da dissertação são detalhadas no Apêndice A, “Glossário de Ferramentas”.

<sup>138</sup> Tradução livre. No original: “This book is not concerned with providing a handful of ready-made ‘visualizations’ that can be plugged into any data set. Ready-made visualizations can help produce a quick view of your data set, but they’re inflexible commodity items that can be implemented in packaged software. Any bar chart or scatter plot made with Excel will look like a bar chart or scatter plot made with Excel. Packaged solutions can provide only packaged answers, like a pull-string toy that is limited to a handful of canned phrases, such as ‘Sales show a slight increase in each of the last five years!’ Every problem is unique, so capitalize on that uniqueness to solve the problem.”

com um conjunto de dados e mais provável que você possa fazer um gráfico adequado a sua visão. (YAU, 2011, CAPÍTULO 3)<sup>139</sup>

As duas abordagens se caracterizam pela flexibilidade, seja pela variedade de recursos da programação em Processing como pela variedade de ferramentas existentes, cada uma com suas vantagens. Um resultado desejado dessa flexibilidade é que cada tarefa, seja uma visualização ou uma pesquisa dos dados, tenha uma solução adequada. Essa solução pode incluir as soluções prontas apontadas por Fry (2008, p. 16), desde que sirvam aos objetivos do jornalista. Muitas vezes o resultado evidencia a ferramenta que foi usada:

se quisermos entender como o software já deu nova forma à mídia conceitualmente e praticamente, temos que olhar de perto as ferramentas cotidianas usadas pela grande maioria dos usuários profissionais e não-profissionais. (MANOVICH, 2013, P. 31)<sup>140</sup>

A ideia de o software dar “nova forma à mídia” será aqui relacionada à necessidade, apontada por Fry (2008, p. 16) de não usar ferramentas genéricas que resultem em resultados genéricos. Porém, há algumas vantagens nas ferramentas genéricas, como a facilidade no uso, apontada por Yau: “em troca pelo [facilidade do] clicar e arrastar, você deixa de ter certa flexibilidade no que você quer fazer” (YAU, 2011, CAPÍTULO 3)<sup>141</sup>.

Conseguir resultados mais adequados a cada problema enfrentado pelo jornalista também é algo que deve ser considerado ao refletir sobre a apropriação que eles fazem do computador. Por exemplo, Renata [PINTO, 2014] apesar de sua familiaridade com planilhas eletrônicas, prefere usar o site Infogr.am para produzir gráficos, porque “[no Excel] não é tão bonito”. Nesse caso nota-se o desejo de não apresentar um resultado genérico como o criticado por Fry, mas fazendo uso de várias ferramentas para isso.

<sup>139</sup> Tradução livre. No original: “you have a lot of options. Some are out-of-the-box and click-and-drag. Others require a little bit of programming, whereas some tools weren’t designed specifically for data graphics but are useful nevertheless. [...] The more visualization tools you know how to use and take advantage of, the less likely you’ll get stuck not knowing what to do with a dataset and the more likely you can make a graphic that matches your vision.”

<sup>140</sup> Tradução livre. No original: “if we want to understand how software has already re-shaped media both conceptually and practically, we have to take a close look at the everyday tools used by the great majority of both professional and non-professional users”.

<sup>141</sup> Tradução livre. No original: “In exchange for click-and-drag, you give up some flexibility in what you can do.”

### 4.3.1 Etapas para produzir uma visualização

Ben Fry (2008) propõe sete etapas para realizar uma visualização: obter os dados, processamento dos dados, filtrar, minerar<sup>142</sup>, representar, refinar e adicionar interação (FRY, 2008, p. 5–15)<sup>143</sup>. Como Fry considera esses passos para a realização de visualizações interativas em meio eletrônico, as descrições dos passos serão expandidas para incluírem visualizações em outros meios. Com esse fim, as sete etapas de Fry serão consideradas como um detalhamento das três etapas definidas por Bertin: “análise matricial (dos dados), tratamento gráfico da informação (descobrir as respostas) e gráfico de comunicação (comunicar as respostas, se existem)” (BERTIN, 1986, p. 16 e 17–23). Também serão consideradas a “pirâmide invertida do JGD” (BRADSHAW, 2011)<sup>144</sup> – cujas etapas são compilar, limpar, contextualizar, combinar, comunicar<sup>145</sup> –, e os quatro passos descritos no “manifesto” da organização jornalística FiveThirtyEight: coleção, organização, explicação e generalização (SILVER, 2014)<sup>146</sup>. A Fig. 16 representa os pontos em comum dessas quatro propostas.

É importante não pensar nesses passos como um caminho restrito: “não se pode ser escravo desses passos”<sup>147</sup> (FRY, 2008, p. 6). Por exemplo, pode ser necessário voltar a filtrar os dados por problemas que aparecerem durante a etapa de visualização, ou fazer tentativas de visualização enquanto os dados são analisados.

---

<sup>142</sup> Na computação, o termo minerar pode ter significados diferentes das do jornalismo guiado por dados e do termo usado por Fry.

<sup>143</sup> Os termos originais usados por Fry são: *acquire, parse, filter, mine, represent, refine* e *interact* (FRY, 2008, p. 5).

<sup>144</sup> Pirâmide invertida é a estrutura clássica de textos jornalísticos, na qual as informações são organizadas em ordem de importância, começando pela considerada mais importante.

<sup>145</sup> Tradução livre. No original: *compile, clean, context, combine* e *communicate*.

<sup>146</sup> Tradução livre. No original: *collection, organization, explanation* e *generalization*.

<sup>147</sup> Tradução livre. No original: “these steps can’t be followed slavishly”.

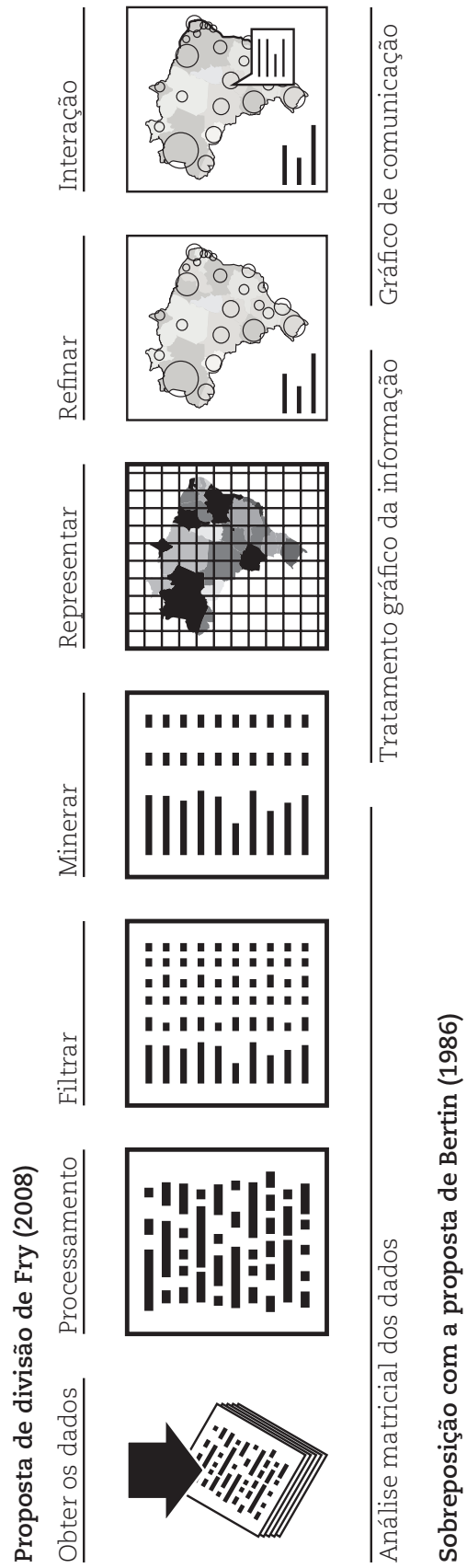


Fig. 16: Esquema das etapas para produzir uma visualização, segundo Bertin (1986), Bradshaw (2011), Fry (2008) e Silver (2014). Fonte: elaboração própria.

### 4.3.2 Obter e processar os dados

Os primeiros quatro passos de Fry estão ligados ao que Bertin (1986, p. 16) chama de “construir o quadro de dados”. O primeiro passo é obter os dados, ou compilá-los. Os dados podem ser fornecidos por uma organização, encontrados na web, raspados de páginas, obtidos através de conversão de documentos – como escaneamento –, puxados através de APIs, obtidos através de observação, de pesquisas, de enquetes online ou enviados pelo público (BRADSWAW, 2011).

Em segundo lugar, processá-los: “dê uma estrutura ao significado dos dados, e organize-os em categorias” (FRY, 2008, p. 5)<sup>148</sup>. O terceiro passo é filtrá-los: “exclua tudo exceto as informações de interesse”<sup>149</sup> (FRY, 2008, p. 5). Bradshaw (2011) fala em limpar como o processo de remover os erros e inconsistências do conjunto de dados, como duplicatas, lacunas, formatações incorretas, despadronização em grafias, uso de abreviaturas, mudanças de nome, entre outras. Limpar também pode se referir a preparar os dados para juntá-los a outros que estejam sendo usados. No que Silver (2014) chama de organização, jornalistas podem organizar os dados submetendo-os a várias análises estatísticas, encontrando relações ou ainda fazendo uma visualização.

Bradshaw também inclui um passo que ele chama de “contextualizar”. Assim como qualquer dado usado na produção jornalística, tem que passar por uma checagem, verificar sua procedência e métodos de coleta para identificar possíveis enviesamentos. Contextualizar pode levar a levantar mais dados. Por exemplo, dados sobre crime podem precisar de dados de população ou renda para serem analisados em contexto, além de cálculos. O contexto pode chegar a outro passo na pirâmide invertida do JGD, que é combinar: semelhante a contextualizar, se trata de combinar mais bancos de dados para obter novas informações. Exemplo típico disso é pôr os dados em um mapa

---

<sup>148</sup> Tradução livre. No original: “Provide some structure for the data’s meaning, and order it into categories.”

<sup>149</sup> Tradução livre. No original: “Remove all but the data of interest.”

para encontrar relações geográficas – o que também está relacionado com a seguinte etapa do processo aqui proposto.

A contextualização dos dados pode ser relacionada com o quarto passo de Fry, que é minerar os dados: “aplique métodos oriundos das estatísticas ou mineração de dados como uma forma de identificar padrões ou colocar os dados em um contexto matemático”<sup>150</sup> (FRY, 2008, P. 5), o que pode exigir cálculos de médias ou transformação de números absolutos em relativos através de processos como cálculos de variação. Bertin (1986, P. 16) usa o termo “simplificar” como parte da descrição dessa etapa, mas essa simplificação não deve ser interpretada como uma redução dos dados através de corte, dentro da premissa de que uma visualização deve exibir os dados em sua totalidade (MANOVICH, 2007, P. 157; UNWIN; THEUS; HOFMANN, 2006, P. 1). Fry fala na seleção de trechos para fazer uma visualização, mas no sentido de facilitar o trabalho e economizar tempo e recursos computacionais até decidir se o projeto vale a pena e então usar a totalidade dos dados (FRY, 2008, P. 30).

Quanto a questões práticas, a maneira como os dados estão disponibilizados é o primeiro problema a ser enfrentado. Pode ser necessário digitá-los manualmente, como na série Crime sem castigo (ver seção 4.4.5), convertê-los a partir de arquivos que não respeitam as recomendações da Lei da Transparência, como é o caso da Assembleia Legislativa do Paraná, que divulga dados em arquivos PDF (OLINDA, 2011), ou “raspá-los” do código-fonte de um site. É neste momento em que são mais evidentes as relações de poder entre quem disponibiliza os dados e quem quer obtê-los. As dificuldades, propositais ou não, podem ser consideradas parte do código técnico das tecnologias usadas para a divulgação dessas informações, o que dá a essas tecnologias um caráter estratégico.

Para contornar essas dificuldades, é possível fazer uso de outras tecnologias. Por exemplo, existem programas que destravam arquivos PDF protegidos, como o site PDF Unlock. Também há programas que convertem tabelas presentes em arquivos PDF em planilhas eletrônicas, como o Tabula. No caso de

---

<sup>150</sup> Tradução livre. No original: “Apply methods from statistics or data mining as a way to discern patterns or place the data in mathematical context.”



arquivos em papel, programas de reconhecimento ótico de caracteres também podem ser usados.

Outra possibilidade é a “raspagem” de sites – extração de dados a partir do código-fonte de páginas da internet. Para isso, existem softwares específicos, como Scaperwiki e Helium Scraper – ambos comerciais. Há várias alternativas gratuitas, como a Scraper, uma extensão para o navegador Chrome. Também há bibliotecas para trabalhar com programação, como BeautifulSoup, uma biblioteca para a linguagem de programação Python.

Uma vez obtidos os dados, é possível que seja necessário organizá-los antes de consolidá-los em uma base de dados ou planilha. Para essa tarefa pode ser útil trabalhar os dados como texto e usar ferramentas como o OpenRefine – antigo Google Refine –, que separa as linhas de texto em colunas de acordo com padrões encontrados pelo programa ou indicados pelo usuário. Editores de texto que suportem expressões regulares<sup>151</sup>, como Notepad++, TextWrangler, Vim, Gedit, entre outros, também podem ser úteis nesta etapa.

Já organizados, os dados podem ser transferidos para um lugar que os contenha, como uma planilha de cálculo – Microsoft® Excel, Google Drive, LibreOffice Calc – ou um software de base de dados – MySQL, Microsoft® Access. Nas planilhas é possível fazer operações e cálculos, e os softwares de bases de dados se destacam por poder operar com maiores quantidades de dados. Por exemplo, a planilha Calc, do pacote LibreOffice, tinha um limite de 65.536 linhas, expandido em versões posteriores para 1.048.576 linhas (LIBREOFFICE, 2010, P. 5). Em algumas situações, um milhão de linhas ainda pode não ser suficiente, sendo necessárias outras soluções, como os aplicativos de banco de dados citados, projetados para lidar com volumes de dados maiores, permitindo outros tipos de consultas e operações. De maneira geral, a apropriação dessas ferramentas é facilitada pela compatibilidade entre os formatos de arquivo, o que permite transpor dados entre aplicativos. O Google Refine, por exemplo, gera arquivos de dados

---

<sup>151</sup> Expressão regular, também conhecida como *regex*, são conjuntos de “coringas” para localizar sequências específicas de caracteres em um texto.

separados por vírgulas, que as planilhas eletrônicas interpretam como separações de colunas.

Nesta etapa já é possível que os dados exponham novas informações a partir de operações como organizá-los por quantidade de valores, fazer buscas específicas, calcular totais e médias, fazer comparações, verificar lacunas, entre outros. Dependendo do objetivo da ou do jornalista, pode ser suficiente. Mas, dependendo do tipo de dados, ainda pode ser necessário visualizar a informação para chegar a novas conclusões, ou a visualização pode resultar em informações sobre o conjunto como um todo. No exemplo da Fig. 17, as planilhas usadas pela reportagem mostravam que a maioria dos deputados estaduais do Paraná gastavam a totalidade da verba indenizatória<sup>152</sup> ao longo do ano, mas a visualização mostra que enquanto alguns o faziam gastando quantias similares mês a mês, outros gastavam no final do ano o que não haviam gastado durante o ano (FÉLIX, 2011).

#### 4.3.3 Visualizar os dados

O quinto passo de Fry, representar, se relaciona com o que Bertin chama de “tratamento gráfico da informação” (BERTIN, 1986, p. 16), que é o momento de trazer à tona, por meio de técnicas de visualização, as informações latentes no conjunto de dados. É nesta instância que alguns aspectos da informação se tornam evidentes, como no caso do vídeo “How Mariano Rivera Dominates Hitters” (seção 4.4.1), das visualizações feitas por estudantes de jornalismo (seção 4.4.7) e da maneira como as verbas indenizatórias dos deputados estaduais do Paraná (Fig. 17, na página 65).

Para esse passo, Fry sugere tentar algum modelo de representação conhecido, como um gráfico de barras, de linhas, ou um mapa, e partir dele para modificá-lo. Para mostrar mais dimensões dos dados, pode ser necessário apelar a outras variáveis visuais (BERTIN, 1986, p. 187), como cores, tons de uma mesma cor, formas geométricas etc:

---

<sup>152</sup> Verba repassada aos deputados para custeio de gastos relacionados à atividade parlamentar, como combustível, aluguel de imóveis, contratação de consultoria, entre outros.

Acredito que a maioria das práticas de visualização de informação da segunda metade do século XVIII até hoje seguem o mesmo princípio – reservando arranjo espacial (podemos chamá-lo de “layout”) para as dimensões mais importantes dos dados, e usando outras variáveis visuais para as dimensões remanescentes. (MANOVICH, 2007, P. 154)

Para esse processo, as planilhas são novamente úteis, já que permitem a elaboração de gráficos a partir de vários modelos disponíveis. Softwares de ilustração vetorial, como o Adobe Illustrator, também têm recursos para a elaboração de gráficos. Há várias ferramentas que permitem fazer visualizações online, como Many Eyes, RAW, infogr.am e Chartbuilder.

Especificamente para mapas, é possível usar ferramentas como ArcGIS e sua alternativa open source, QGIS<sup>153</sup>. Elas trabalham com dados geográficos, como arquivos shape<sup>154</sup>, e permitem combiná-los com dados estatísticos para gerar diferentes tipos de mapas. Também há recursos com foco no online, como Google Maps, MapBox e OpenStreetMap, este último de licença livre. Essas ferramentas podem ser usadas para consultas online, mas também, através de APIs, permitem o uso por parte de desenvolvedores, permitindo assim a criação de outras ferramentas, como o BatchGeo, usado no exercício com estudantes de jornalismo descrito na seção 4.4.7.

Para desenvolver visualizações online, é possível combinar técnicas de HTML e CSS com Javascript, através de bibliotecas como Raphaël e D3, de licença livre, ou HighCharts, de uso gratuito para atividades não comerciais, ou inclusive Google Charts, usados no Atlas do Desenvolvimento Humano no Brasil (PNUD, 2013).

---

<sup>153</sup> GIS é a sigla em inglês para Sistemas de Informação Geográfica (SIG).

<sup>154</sup> Shapefiles são arquivos que contêm informação com referências geográficas que permitem desenhar mapas sobrepondo camadas com diferentes tipos de informação: pontos, linhas, planos ou imagens.

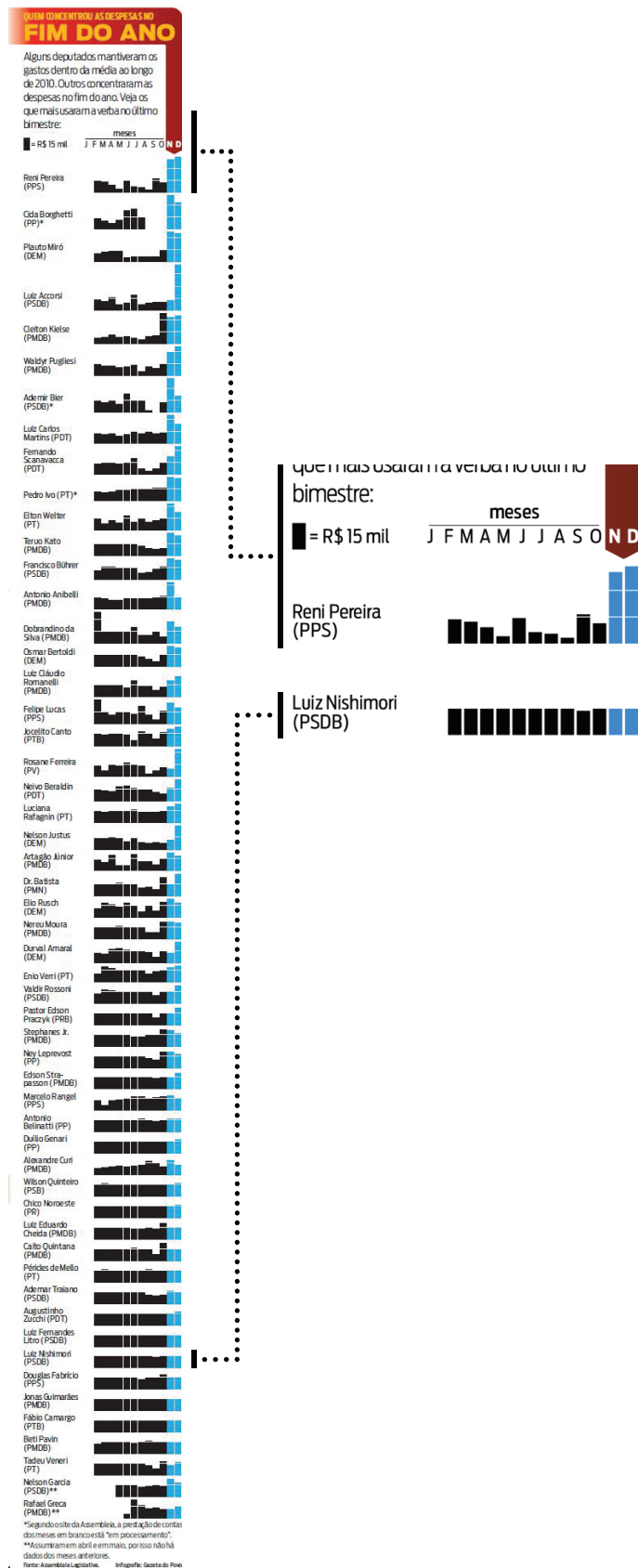


Fig. 17: Infográfico dos gastos dos 34 deputados estaduais do Paraná. No destaque, um deputado que gastou mais no final do ano e outro que gastava regularmente. Fonte: FÉLIX, 2011. DESTAQUES NOSSOS.

Por último, há ferramentas que se propõem a reunir vários recursos para realizar visualizações, como Tableau Public e Qlikview. Esses softwares têm facilidade de uso do tipo arrastar e soltar para elaborar visualizações que podem incluir mapas, gráficos estatísticos, tabelas dinâmicas, entre outros. No caso do Tableau Public, há a exigência, na versão gratuita, de disponibilizar as bases de dados para o público. Uma alternativa Open Source para essas ferramentas é o software Datawrapper, que permite que o usuário instale seu próprio servidor de informações. O software é desenvolvido pela ABZV (*Akademie Berufliche Bildung der deutschen Zeitungsverlage*, Academia para Formação Profissional dos Editores de Jornais da Alemanha, uma instituição alemã para treinamento de jornalistas).

Mesmo considerando que nesta etapa as visualizações podem não estar prontas para serem divulgadas, elas já podem servir para evidenciar aspectos dos dados que não seriam visíveis na etapa anterior. O processo de produção do infográfico *How Mariano Rivera Dominates Hitters* (ver seção 4.4.1) inclui uma exploração prévia dos dados através de uma visualização feita em Processing (Fig. 28, na página 141), a partir da qual foram tiradas as conclusões apresentadas no vídeo e foi decidido quais aspectos da visualização seriam evidenciados. Os mapeamentos da seção 4.4.7 também tiveram uma intenção “exploratória” mais do que apresentar os dados ao público.

Essa produção de visualizações para explorar conjuntos de dados atenta para que a visualização seja considerada como um método do JGD e não somente como um produto final. A entrevistada Renata, por exemplo, tem a percepção que no campo do JGD se fala pouco em visualizações, e ela mesma não as vê como algo que seja parte do JGD e sim como “algo pra mais tarde. Porque acho que a maioria das redações ainda tem infografistas; eu acho que ainda dá pra deixar com esse pessoal” [PINTO, 2014]. Essa percepção atenta para que as visualizações, dentro do JGD, podem ser consideradas como algo separado, ou como um produto jornalístico, ignorando assim o potencial delas na etapa de exploração dos dados.

#### 4.3.4 Comunicar os dados

O sexto e sétimo passos de Fry – refinar e adicionar interação – serão considerados aqui como o que Bertin chama de “comunicar as respostas” (1986, p. 16). O ato de “comunicar” implica fazer com que uma visualização seja entendida por mais pessoas e, se usada em um contexto jornalístico, que tenha cuidado com as questões éticas que a informação comunicada exige, em um processo análogo ao de edição de uma matéria.

Dentro dessa “edição”, há vários fatores que devem ser considerados. A visualização deve ser adequada ao suporte final. Em determinadas condições de impressão, visualizações em papel podem suportar maiores resoluções do que monitores de computador, mas estes podem exibir visualizações dinâmicas e permitir interatividade, por exemplo.

O passo de “adicionar interação” pode ser adequado para algumas visualizações, mas para outras pode atrapalhar a compreensão. Em “How Mariano Rivera Dominates Hitters” (item 4.4.1), foi decidido que o formato final seria um vídeo, já que os autores consideraram que, nesse caso, interações – como selecionar diferentes maneiras de visualizar os dados – deixariam a visualização mais confusa, e preferiram ficar com o controle do que seria visualizado (GONZÁLEZ VEIRAS, 2010).

Também devem ser consideradas as possibilidades de interação em suportes não-digitais, como a visualização em forma de bola de futebol do jornal omanense *Times of Oman* (FARACH, 2014), que propõe ao leitor a construção, usando tesoura e cola, de vários poliedros que representam as seleções que disputaram a Copa do Mundo de Futebol de 2014. Cada elemento tem várias visualizações de dados, que podem ser comparadas entre si (Fig. 18) e, quando agrupadas, formam uma bola de futebol (Fig. 19).

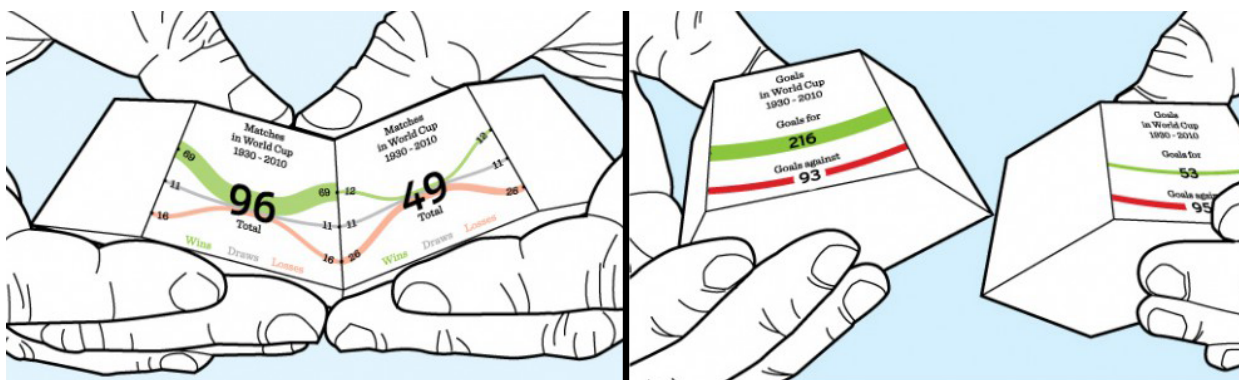


Fig. 18: Exemplos de manuseio dos elementos da bola de futebol para comparar dados e visualizações das faces de diferentes peças. Fonte: FARACH, 2014.



Fig. 19: Exemplar da visualização em forma de bola de futebol do jornal Times of Oman. Fonte: FARACH, 2014.





Fig. 20: Primeira e última páginas do primeiro caderno do jornal argentino “La Voz del Interior”. Fonte: COLOMBATO, 2010.

Há outras possibilidades de interação que não exigem tanto esforço por parte de quem lê, como a representação em tamanho real (Fig. 20) do diâmetro do túnel pelo qual seriam resgatados os mineiros presos em uma mina no Chile (COLOMBATO, 2010), que permitia, por exemplo, pôr a folha de jornal no chão e pisar nela para ter uma ideia do “aperto” pelo qual os mineiros passariam durante o resgate.

É fundamental não pensar em que comunicar implica em um embelezamento da informação através de “decorações”, no que Tufte chama de *chartjunk*:

Se esconder por trás de *chartjunk* é menosprezar tanto a informação como a audiência. Os defensores do *chartjunk* imaginam que números e detalhes são chatos, desinteressantes e tediosos, e precisam de ornamentos para ter vida. A decoração

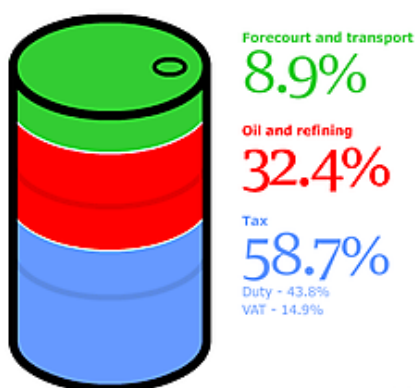


cosmética, que frequentemente distorce os dados, nunca vai salvar a falta de conteúdo. (TUFTTE, 1990, P. 34)<sup>155</sup>

Tufte cunhou o termo *chartjunk* para se referir a gráficos nos quais a tentativa de deixá-los “visualmente atraentes” distorce a informação e dificulta a leitura, como acontece no gráfico da Fig. 21, no qual a tampa do barril está na mesma cor da faixa superior, deixando a área que deveria ser proporcional ao valor de 8,9% de um tamanho similar à área que representa o valor de 32,4%.

### Black gold at an all time high

Every \$1 increase in the price of a barrel of crude equates to roughly 0.42p on a litre of petrol at the pump. At \$150 a barrel, a litre of unleaded would cost £1.25 on the forecourt; at \$200 that rises to £1.46 - or £80 to fill up the family runabout.



Start

SOURCE: petrolprices.com

Fig. 21: Gráfico em formato de barril publicado online pelo jornal The Guardian. Fonte: FUNG, 2014A.

Excessos de informação e preponderância de elementos que deveriam ser somente de apoio à informação, como quadrículas ou cores dos elementos, também são considerados *chartjunk*. São necessárias então estratégias de desenho para que a informação seja destacada. A Fig. 22 é um gráfico retirado de um artigo científico sobre como cães demonstram ciúme. Esse gráfico foi retrabalhado pela equipe do The New York Times (Fig. 23), corrigindo os elementos desnecessariamente destacados e acrescentando elementos que o tornam mais fácil de ler, mas sem distorcer a informação. Os padrões usados no gráfico original – preto, branco e hachurado – foram substituídos por cores, equilibrando mais as duas primeiras barras e “acalmando” o efeito das hachuras; os grupos de barras foram reorganizados; a escala

<sup>155</sup> Tradução livre. No original: “Lurking behind chartjunk is contempt both for information and for the audience. Chartjunk promoters imagine that numbers and details are boring, dull, and tedious, requiring ornament to enliven. Cosmetic decoration, which frequently distorts the data, will never salvage an underlying lack of content.”

lateral foi simplificada; textos foram acrescentados ou modificados para explicar melhor os dados. Também foram acrescentadas ilustrações, mas estas não interferem na visualização dos dados, tornando o infográfico mais divertido sem dificultar sua leitura (FUNG, 2014B).

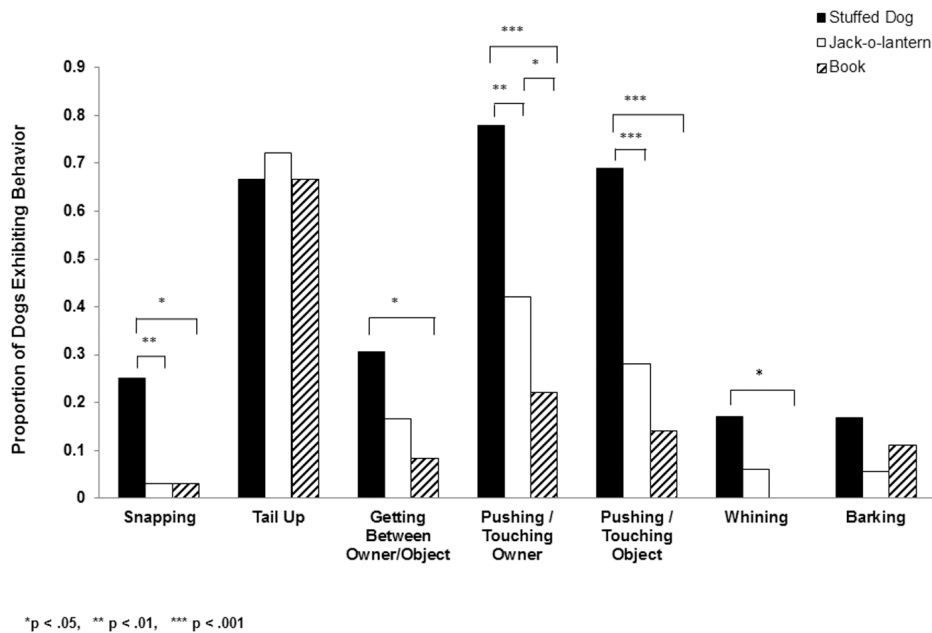
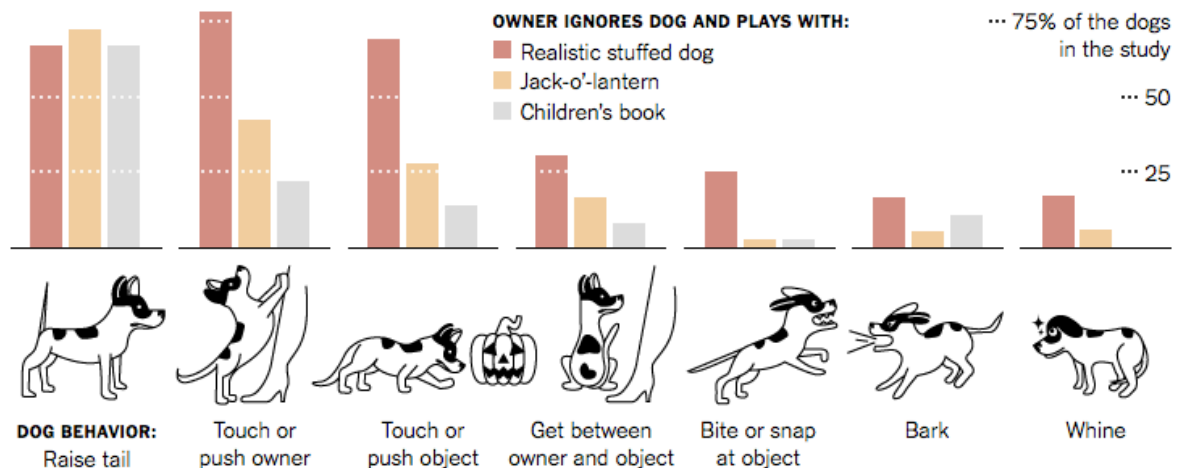


Fig. 22: Gráfico sobre comportamento de cachorros, retirado de um artigo científico. Fonte: FUNG, 2014B.

### Your Jealous Dog

In a study, dogs reacted more strongly when their owners paid attention to stuffed dogs than to more generic objects.



By The New York Times; Illustrations by Jennifer Daniel  
 Source: PLoS One

Fig. 23: Infográfico redesenhado pela equipe do The New York Times. Fonte: FUNG, 2014B.

Com base nas práticas sugeridas por Tufte (1990, 52–65), a Fig. 24 apresenta uma proposta de redesenho de um dos mapas da

seção 4.4.7, já que o mapeamento original foi produzido através da apropriação de uma ferramenta que talvez não fosse adequada para os objetivos da visualização. O mapa feito por Mariana Delamuta, durante oficina ministrada pelo autor na Universidade Positivo em março de 2013, mostra quais municípios do Paraná têm menos pessoas empregadas com carteira assinada.

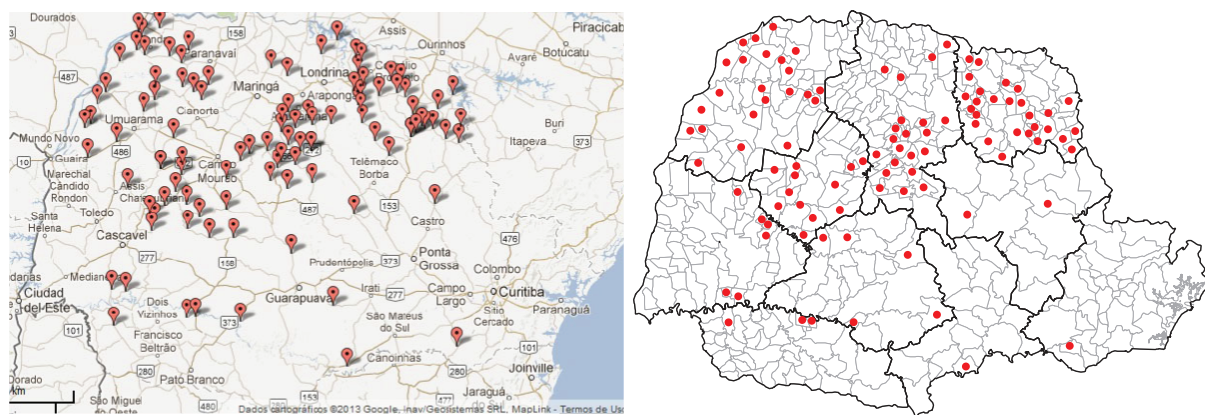


Fig. 24: Mapeamento (DELAMUTA, 2013) e proposta de redesenho (elaboração própria).

Para conseguir uma separação melhor entre as camadas de informação, foram eliminadas as estradas e referências geográficas, por serem aqui consideradas “ruído” e também por considerar que o mapa não seria comunicado a um grande público e sim usado por jornalistas em uma instância de exploração dos dados, como era o exercício durante o qual foi produzido o mapa original. As estradas e nomes foram substituídos pelos contornos dos municípios e a divisão em regiões do IBGE, que talvez sejam melhores para localizar os municípios. O fundo da visualização foi eliminado, já que o contorno do estado se mostrou suficiente. Foi preferida a representação dos municípios com pontos, para separar melhor a informação e para substituir o ruído causado pelo marcador padrão do Google Maps. O redesenho foi feito no Adobe Illustrator, programa de ilustração vetorial, mas resultados semelhantes podem ser alcançados com outras ferramentas, como QGIS. As cores e espessura das linhas foram escolhidas considerando as possibilidades da reprodução em papel.

Quanto aos softwares usados na etapa de comunicação dos dados, todos os citados na seção anterior têm possibilidades de gerar visualizações que podem ser usadas para “comunicar”, caso

isso seja considerado necessário pelo jornalista. Por exemplo, é possível fazer um mapa no aplicativo QGIS que sirva para o jornalista visualizar os dados, mas caso ele queira disponibilizá-lo, é possível escolher cores adequadas, colocar legenda e títulos, entre outros. Também é possível exportar o mapa para ser editado em um software de ilustração vetorial, ou ainda usá-lo em um aplicativo online. Da mesma maneira que na seção anterior, é possível se apropriar de diferentes ferramentas e combiná-las da maneira que for possível.

É importante notar as limitações e diferentes possibilidades de apropriação desses aplicativos e sistemas. Sistemas concebidos para funcionarem online podem apresentar dificuldades para quem quiser transpor as visualizações para outros aplicativos, como os mapas eleitorais da seção 4.4.3, que foram gerados no Tableau, e precisaram ser modificados para gerar arquivos PDF que pudessem ser editados em um editor de desenho vetorial. Outra dificuldade pode aparecer quando um conjunto de dados exige um tipo de visualização muito específica e incomum, não prevista por nenhum aplicativo, como é o caso de “How Mariano Rivera Dominates Hitters” (CARTER; GRAHAM; WARD, 2010), que exigiu o uso de programação combinada com modelagem em 3D (ver seção 4.4.1).

#### 4.3.5 Considerações sobre programação

Em vez de usar um aplicativo pronto, cada pessoa pode também programar seus próprios aplicativos. Isso leva à questão de se os jornalistas devem saber programar ou não, com profissionais defendendo que sim, é necessário aprender programação – várias opiniões podem ser encontradas no artigo de Moreno Osório (OSÓRIO, 2014) e no fluxograma feito por David Homes (Fig. 25).

Em entrevista, Renata [PINTO, 2014], como usuária do Excel, consegue programar algumas coisas e “brincar um pouquinho com macro”, mas ela não sabe se algum dia poderá dispensar as planilhas:

Eu não acho que tenha que deixar de usar Excel porque ainda não sei o suficiente de programação. Mas eu não sei também se [...] existe uma etapa que eu não vou precisar mais usar Excel, se só a programação vai ser suficiente. [PINTO, 2014]

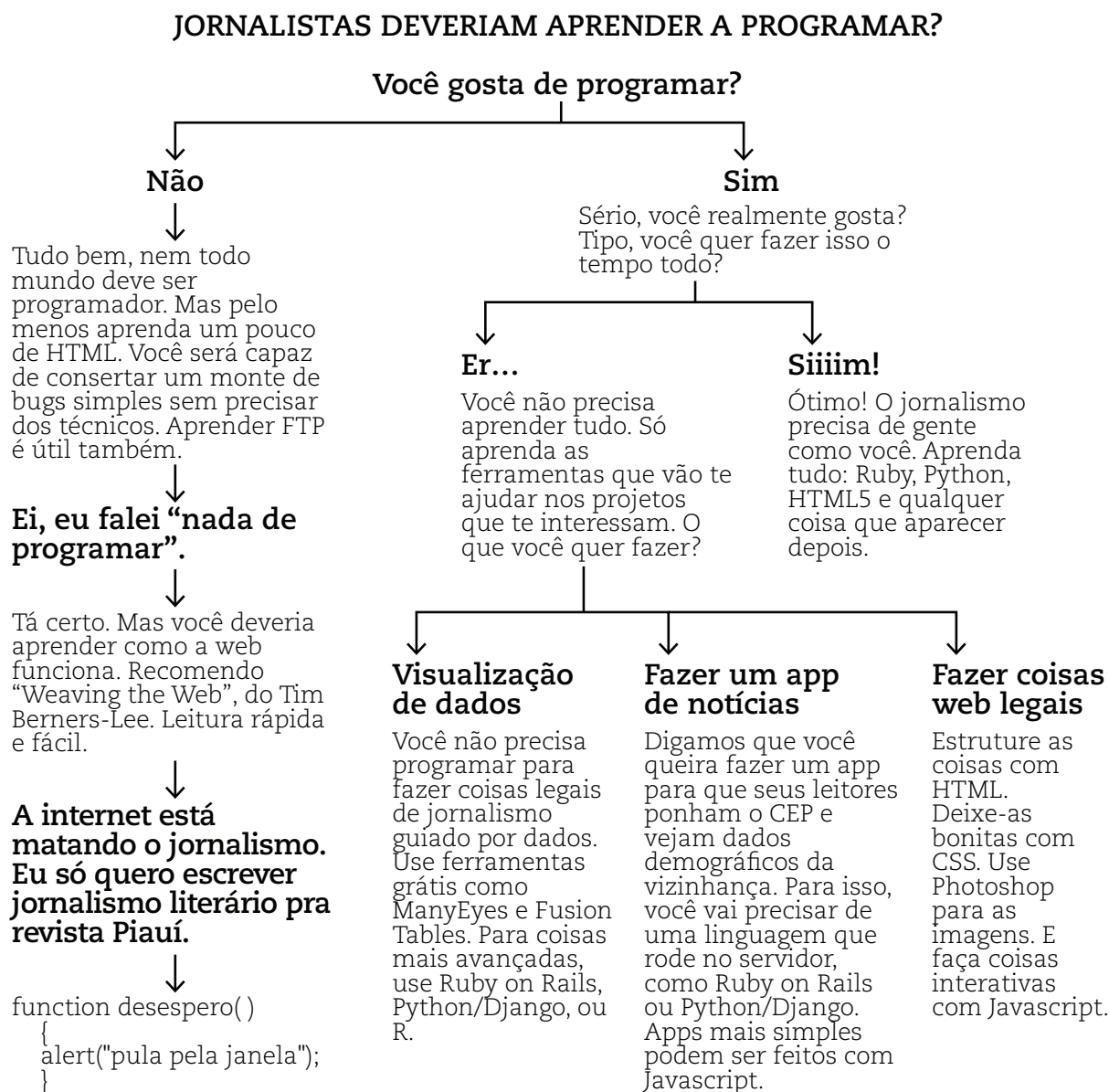


Fig. 25: Fluxograma para jornalistas que queiram aprender – ou não – programação. Fonte: adaptado e traduzido de HOLMES, 2013.

Ela aposta no uso de recursos de programação combinados com outras ferramentas, “um em auxílio do outro, sabe? Pegar a informação, e daí passar pro Excel.” Ou ainda, recursos da linguagem R: “acho que muita coisa que eu fico brincando de *pivotable*<sup>156</sup> eu ia colocar dois comandos e *páh!*” Na ocasião da entrevista, ela ainda estava tentando entrar “de vez com a

<sup>156</sup> Termo em inglês que, no contexto, pode ser traduzido como “tabela dinâmica”.

programação.” Ela já aprendeu “um pouquinho de Python, um pouquinho de Ruby...”. Ela cita que aprendeu Python durante a conferência da Necir<sup>157</sup>, ocasião que aproveitou para entrar em contato com pessoas do jornal *The New York Times* e do *ProPublica*<sup>158</sup>. Segundo ela, havia cursos introdutórios e avançados, que ela considerou “avançados demais”. Ela cita que nos cursos introdutórios os instrutores “passam bastante programas legais.”

No começo, Renata pensou que “tinha que aprender na hora e fazer cinquenta mil coisas”, mas agora ela não acredita que precise de um conhecimento profundo para realizar tarefas. Nos congressos, ela viu que

o pessoal gosta bastante de programação, mas eu vi que você não precisa aprender tudo de programação, porque tem opções. Tem como você aprender programação pra você fazer uma raspagem, mineração de dados, e tem como você usar programação pra você fazer aplicativos de notícias. Meu objetivo mais próximo é fazer um scraping legal.

Já Guilherme [STORCK, 2014] não vê o aprendizado de programação como algo muito necessário:

Desde que eu comecei a mexer com dados abertos, visualização de dados e jornalismo de dados meio que virou moda falar sobre programação. E eu pesquisei algumas coisas, quis aprender Javascript, Python, algumas outras coisas. Mas basicamente todos os problemas que eu tive até hoje, e que foram muitos, para resolver com dados, eu consegui resolver de outras formas. Eu acho que se você sabe uma linguagem, se você sabe Javascript, sabe Python, você realiza as tarefas mais rápido, mas nenhuma delas resolve todos seus problemas. Você vai ter, que nem eu comentei no processo, que usar outras ferramentas. [...] Em três ou quatro anos que eu trabalho com isso, não tive tempo de aprender a mexer com programação e eu consegui fazer de outro jeito, tipo, não é uma coisa que eu digo nossa, alguém precisa muito saber fazer isso. É legal se você souber, mas não considero essencial. [STORCK, 2014]

Guilherme só vê a programação como essencial na etapa de raspagem de dados, pela falta de ferramentas disponíveis [STORCK, 2014]. Quanto ao uso de programação na visualização, ele acha que é mais prático o uso de ferramentas prontas:

a parte de programação que é mais ligada à visualização, do tipo, ah, vou fazer um mapa de bolhas e vou fazer em Javascript... até você aprender Javascript e conseguir fazer um negócio decente,

<sup>157</sup> New English Center of Investigative Reporting, da Universidade de Boston.

<sup>158</sup> Agência de notícias sem fins lucrativos, com sede em Nova Iorque.

you lost much more time than someone who already knows and who would do it in very little time. You would lose a lot of time to do a job. And for this there are some solutions that are ready for you that you will not need to lose your time to do something that will be worse. [STORCK, 2014]

José [BARROS JUNIOR; FREY, 2014] relata que tentou aprender programação, mas não conseguiu. “Eu cheguei a fazer alguns cursos online do Codecademy e... não serviu pra nada. Eu não consegui nada.” José percebeu que saber programação poderia ser útil em um evento no qual teve contato com bases de dados obtidas por raspagens usando scripts em Python. “Aí falei: tenho que aprender essa porra!” Ele tentou então fazer um curso no Codecademy, mas acabou desistindo, talvez pela demora para aprender os conceitos básicos:

Eu não consegui perceber como iria conectar aquilo [o aprendizado de noções básicas de programação em Python] mais pra frente naquele modelo de aula. Acho que a didática do Codecademy ainda não me convenceu a persistir nela. Não sei se ela funciona ou não.

[...]

Eu quero fazer robô, me ensina a fazer robô, do zero, depois eu aprendo o resto. Acho que não houve uma conciliação entre as minhas perspectivas muito imediatas com o tempo que eu demoraria pra [cumprí-las]. Sou muito mais a fim de pagar milão prum cara fazer um robô pra mim do que eu aprender.

Yau (2011) aconselha a quem quiser fazer visualizações que aprenda o básico de programação para estar pronto para situações que softwares genéricos não podem resolver, deixando o jornalista pronto para mais situações do que estaria se ele se apegasse a softwares genéricos. A sugestão de Yau é “comece com o básico e vá treinando. Antes de que você perceba, você vai estar escrevendo código” (YAU, 2011, CAPÍTULO 3)<sup>159</sup>. Para iniciantes, ele recomenda começar com uma única linguagem.

Trabalhar com programação tem vários “trade-offs”. Um deles é o tempo que demora dominar uma linguagem de programação, como aconteceu com José, e que tanto trabalho será feito para compensar o tempo e esforço investido no aprendizado, como ilustrado na tira da Fig. 26. Se for uma necessidade única, compensa usar ferramentas genéricas. Mas

<sup>159</sup> Tradução livre. No original: “Start with the basics first and then work your way up. Before you know it, you’ll be coding.”

o tempo gasto aprendendo a programar pode se transformar em tempo ganho em outros projetos, com resultados mais impressionantes. Você vai ficar melhor em programação a cada projeto que você fizer, e vai começar a ficar muito mais fácil. (YAU, 2011, CAPÍTULO 3)<sup>160</sup>

Yau também pondera sobre a facilidade de uso das ferramentas genéricas, que é um outro “trade-off”: em troca da facilidade de uso, as possibilidades do software são limitadas:

Você geralmente pode mudar cores, fontes e títulos, mas você está restrito ao que o software oferece. Se não há um botão pro gráfico que você quer, azar o seu. (YAU, 2011, CAPÍTULO 3)<sup>161</sup>

Além dessa limitação, softwares genéricos não são necessariamente fáceis de serem usados:

alguns softwares podem ter um monte de funções, mas em contrapartida têm um monte de botões que você precisa aprender. Por exemplo, tinha um programa [...] que poderia fazer um monte de coisas se eu dispusesse do tempo. Os processos para fazer as coisas eram tão contraintuitivos que me fizeram não querer aprender mais. Também era difícil de repetir meu trabalho para diferentes conjuntos de dados, porque tinha que lembrar de tudo que eu tinha clicado. Em comparação, quando você escreve código para manipular seus dados, costuma ser fácil reusar código para um outro conjunto de dados. (YAU, 2011, CAPÍTULO 3)<sup>162</sup>

Mas Yau não descarta o uso de ferramentas prontas, como uma planilha de cálculo ou outro software que gere gráficos, e a prática de usar os dois é comum, como mostram inclusive alguns dos entrevistados desta pesquisa. “Só não deixe o software te limitar” (YAU, 2011, CAPÍTULO 3)<sup>163</sup>.

Dificuldades em fazer uso de técnicas de programação narradas em entrevistas (ver seção 5.4) levam a pensar que, apesar de poder ser considerada um sistema aberto (SPINUZZI, 2003, P. 204) que permite o uso por parte de pessoas com diferentes graus de conhecimento, ela ainda parece estar longe da estabilização como

<sup>160</sup> Tradução livre. No original: “the time spent learning how to program now could end up as saved time on other projects, with more impressive results. You’ll get better at programming on each project you go through, and it’ll start to come much easier”

<sup>161</sup> Tradução livre. No original: “You can usually change colors, fonts, and titles, but you’re restricted to what the software offers. If there is no button for the chart you want, you’re out of luck”

<sup>162</sup> Tradução livre. No original: “some software might have a lot of functions, but in turn have a ton of buttons that you need to learn. For example, there was one program [...] that it could do a lot if I put in the time. The processes to get things done though were so counterintuitive that it made me not want to learn anymore. It was also hard to repeat my work for different datasets, because I had to remember everything I clicked. In contrast, when you write code to handle your data, it’s often easy to reuse code and plug in a different dataset.”

<sup>163</sup> Tradução livre. No original: “Just don’t let the software restrict you”.



artefato do jornalismo. É possível que o código técnico da programação implique em dificuldades para ter um primeiro acesso a ela, possivelmente por ela não se assemelhar a outros artefatos já conhecidos, mais próximos da estabilização.

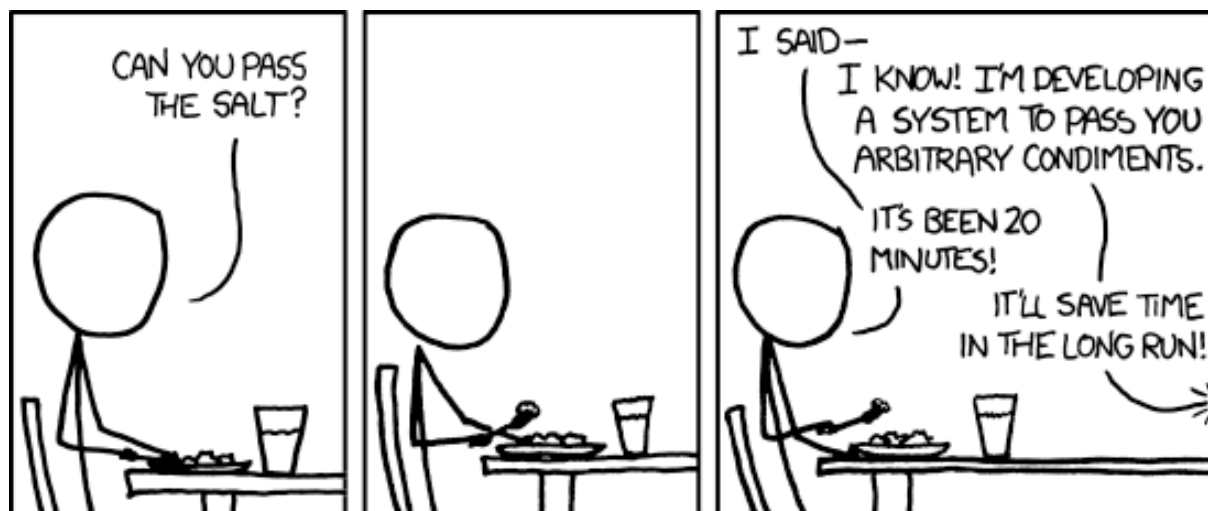


Fig. 26: Tira satirizando o tempo de desenvolvimento de programas. Fonte: MUNROE, 2011.

#### 4.4 ALGUNS EXEMPLOS DE USO DE VISUALIZAÇÕES NO JORNALISMO

Na escolha dos seguintes exemplos, foi procurada uma amostragem que representasse exemplos que fossem úteis à discussão deste trabalho. Há exemplos de visualizações interativas e estáticas. A maioria têm relação com o jornalismo, enquanto outras foram incluídas por terem alguma relação com as questões aqui discutidas.

##### 4.4.1 “How Mariano Rivera Dominates Hitters”, the new York Times

A visualização em vídeo “How Mariano Rivera Dominates Hitters” (CARTER; GRAHAM; WARD, 2010) foi publicada em 29 de junho de 2010 no site do jornal nova-iorquino *The New York Times*. Foram visualizados os dados de 1,3 mil lançamentos feitos em 2009 pelo jogador de beisebol Mariano Rivera. Esses dados foram fornecidos em forma de tabela (Fig. 27) por uma empresa especializada em coleta de dados de jogos de beisebol usando

câmeras para a captura dessas informações, que incluem dados como distância e aceleração dos arremessos (FAST, 2010).

<b>X0</b>	<b>Y0</b>	<b>Z0</b>	<b>VX0</b>	<b>VY0</b>	<b>VZ0</b>	<b>AX0</b>	<b>AY0</b>	<b>AZ0</b>
-2.486	50	5.903	3.895	-133.673	-6.276	-3.489	30.531	-29.546
-2.196	50	5.996	5.998	-135.705	-6.762	4.278	32.082	-18.036

Fig. 27: Trecho da base de dados dos lançamentos de Rivera. Fonte: GONZÁLEZ VEIRAS, 2010.

Em primeiro lugar, a equipe responsável pela produção da visualização, composta por Shan Carter, Graham Roberts e Joe Ward, desenvolveu um programa na linguagem Processing para fazer uma visualização prévia dos dados (Fig. 28). Segundo Xaquín González Veiras, então chefe do departamento de gráficos do *The New York Times*, essa instância “os ajudou a decidir quais dados tinham que ser mostrados e como deveriam mostrá-los” (GONZÁLEZ VEIRAS, 2010)<sup>164</sup>. A visualização também fez com que decidissem fazer um material que não fosse interativo, já que o objetivo do gráfico era “mostrar o quê fazia de Rivera o melhor – e deixar o usuário explorar todos seus lançamentos não era a maneira mais eficiente de consegui-lo” (CARTER *apud* GONZÁLEZ VEIRAS, 2010).<sup>165</sup>

O vídeo inclui outras instâncias de visualização, explicando outros fatos, até que aparece a animação de todos os lançamentos ao mesmo tempo (Fig. 29). A animação final foi feita no software Maya, com a ajuda de um script na linguagem Python (GONZÁLEZ VEIRAS, 2010).

A partir dessa animação, são mostradas novas animações, desde o ponto de vista do rebatedor (Fig. 30), mostrando onde se concentram os arremessos. Para uma melhor compreensão, a representação dos lançamentos em um plano localizado na frente do rebatedor foi feita na forma de mapa de calor (Fig. 31) que mostra que Rivera evitava lançar a bola na região onde seria fácil de rebatê-la.

<sup>164</sup> Tradução livre. No original: “les ayudó a decidir qué datos había que mostrar y cómo debían mostrarlos”.

<sup>165</sup> Tradução livre. No original: “mostrar qué era lo que hacía a Rivera el mejor – y dejar al usuario explorar todos sus lanzamientos no era la manera más efectiva de conseguirlo.”

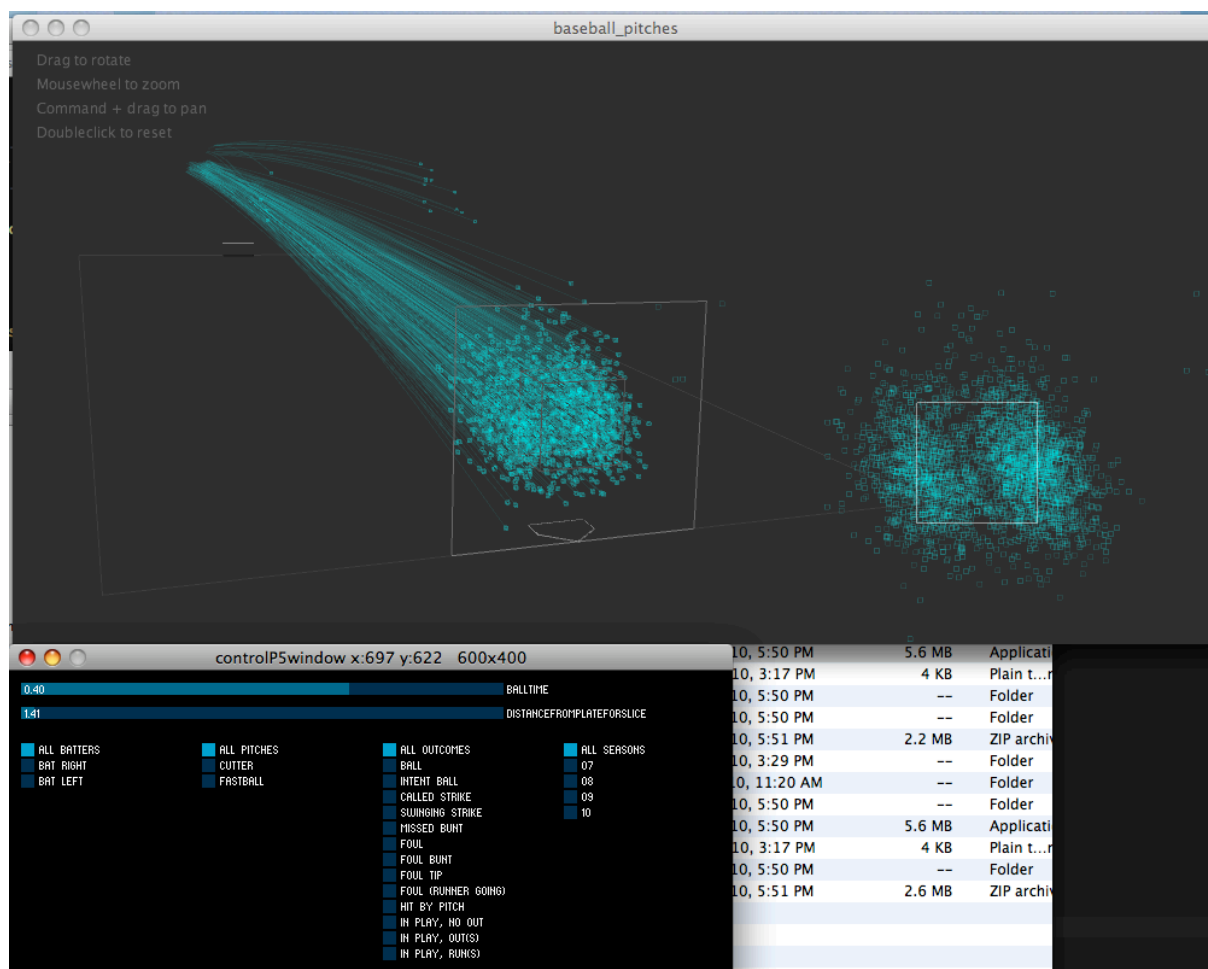


Fig. 28: Tela da visualização feita para explorar os dados dos arremessos. Fonte: GONZÁLES VEIRAS, 2010.

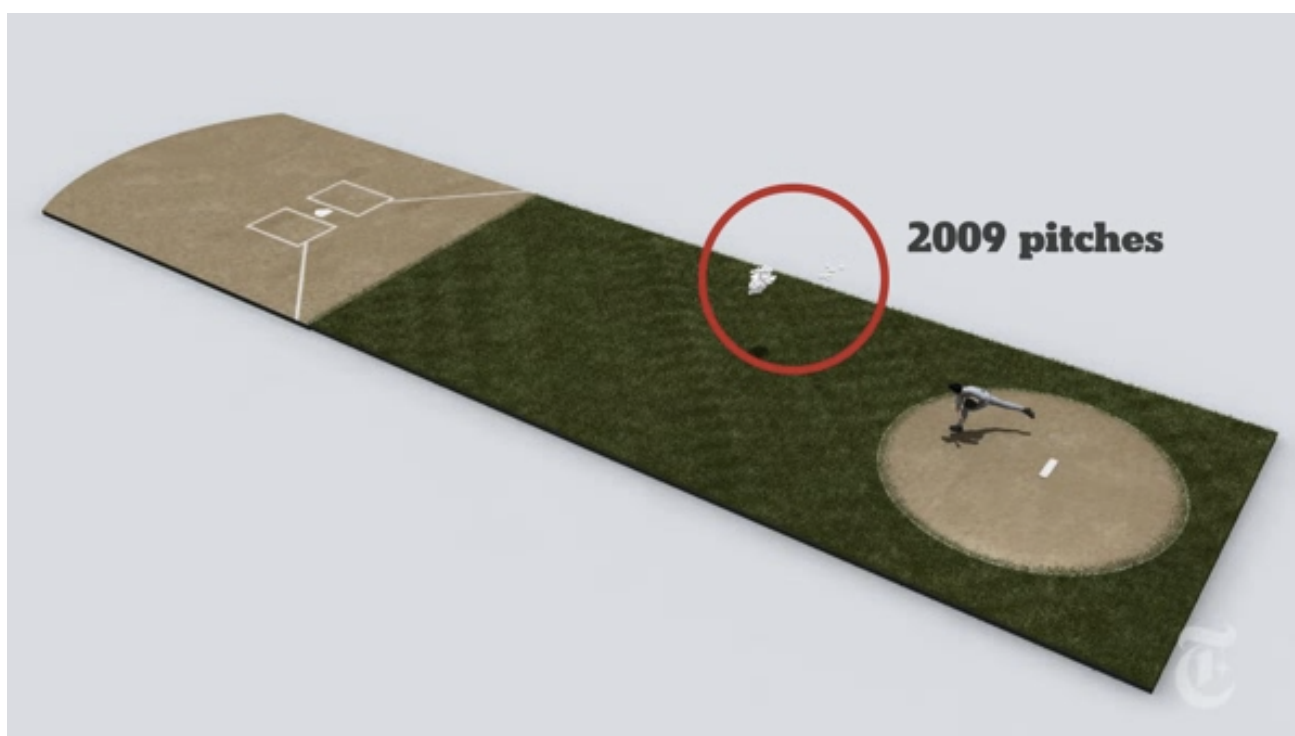


Fig. 29: Momento do vídeo no qual aparecem representados ao mesmo tempo todos os lançamentos feitos em 2009. Fonte: CARTER; GRAHAM; WARD, 2010.

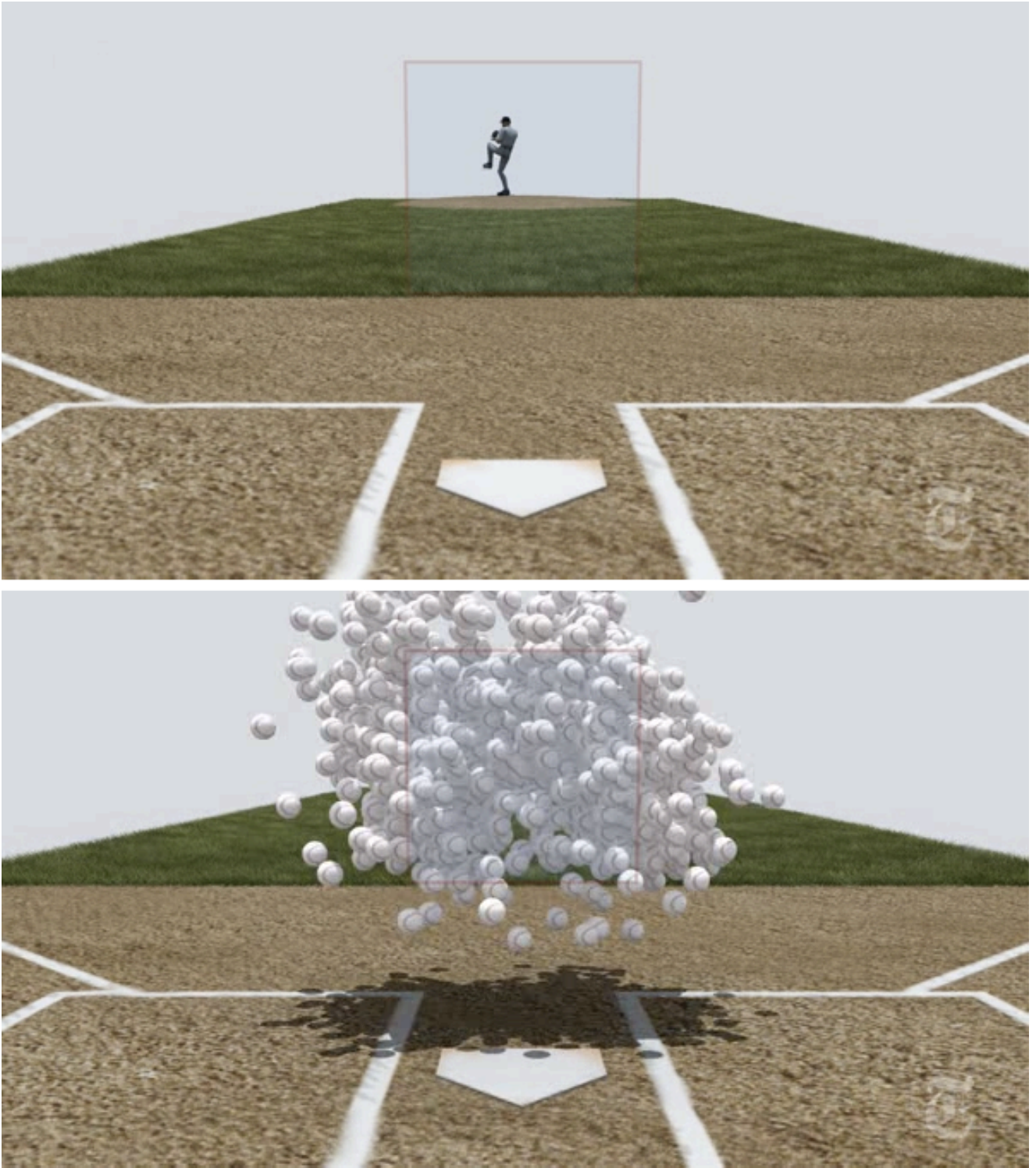


Fig. 30: Dois frames da animação na qual são mostrados todos os lançamentos. Fonte: CARTER; GRAHAM; WARD, 2010.

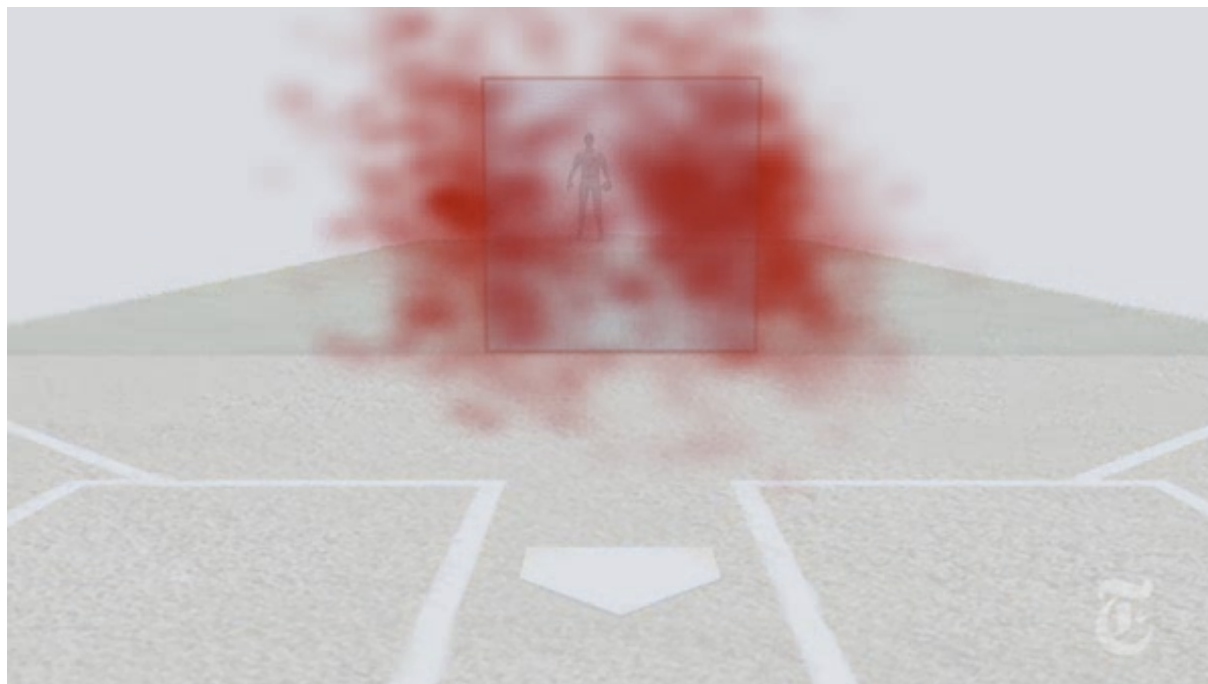


Fig. 31: Mapa de calor representando a concentração de lançamentos em duas regiões específicas. Fonte: CARTER; GRAHAM; WARD, 2010.

O modelo tridimensional do arremessador foi animado usando técnicas de captura de movimento a partir de vídeos nos quais ele aparecia (GONZÁLEZ VEIRAS, 2010). Esse fato não influencia na visualização em si, mas o uso dessa tecnologia é um fato curioso para o jornalismo.

#### 4.4.2 “See Something or Say Something”, de Eric Fischer

A série “See Something or Say Something”, elaborada por Eric Fischer, que se intitula “artista de dados e desenvolvedor de software” e tem vários projetos similares, faz uso de informações de latitude, longitude e altitude presentes nas postagens do Flickr e do Twitter. O autor usou ferramentas de programação para baixar essas informações e representá-las em um plano de coordenadas: cada tweet foi representado com um ponto azul, e cada foto com um ponto vermelho. Não foram usadas outras informações nem o conteúdo das postagens; foi marcado apenas o lugar no qual foi gerada cada postagem. Também não foram usados mapas ou qualquer outra informação cartográfica: os pontos formam, por si sós, os contornos das cidades, avenidas, mares. Fischer fez mapeamentos do mundo todo, de continentes e



de algumas cidades; aqui é apresentada a visualização da cidade do Rio de Janeiro (Fig. 32).

Na página da visualização (FISCHER, 2011), os usuários podem incluir anotações; foram marcados o Cristo Redentor, que tem muitos pontos indicando fotografias, é possível identificar bairros e avenidas mais importantes, assim como os morros, cuja ausência de informação denota uma menor atividade nesses sites.

É importante ressaltar que o trabalho de Fischer não está em um contexto jornalístico, mas corresponde aos conceitos de visualização aqui abordados.

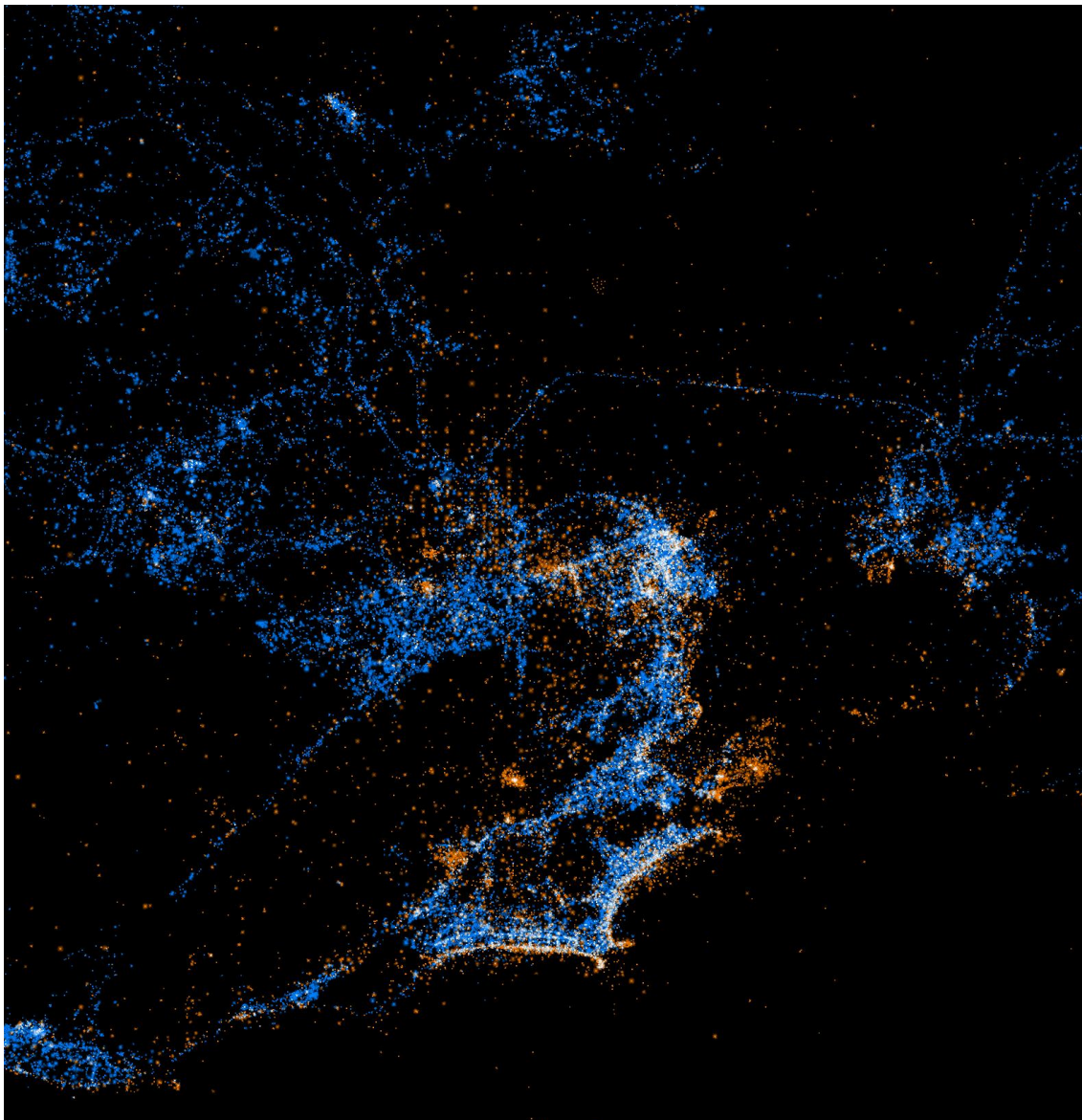


Fig. 32: Visualização da série "See Something or Say Something", de Eric Fischer, mostrando a cidade do Rio de Janeiro. Fonte: FISCHER, 2011.

#### 4.4.3 “Bases eleitorais”, Gazeta do Povo

Publicada em 3 de novembro de 2012 no jornal curitibano Gazeta do Povo (STORCK, 2012), a visualização (Fig. 33) mostra um pequeno mapa da cidade de Curitiba para cada um dos 38 vereadores eleitos em 2012, mostrando quantos votos ele recebeu em cada bairro. O TSE libera dados de cada lugar de votação uma semana depois da eleição. Esses dados foram agrupados por bairro usando planilhas eletrônicas, e foi feita uma visualização interativa no software Tableau. Essa visualização, por sua vez, foi usada para gerar arquivos PDF com o mapa de cada um dos candidatos eleitos. A visualização foi preparada para impressão usando softwares de desenho vetorial. Uma cópia foi fornecida ao repórter responsável pela matéria, que preparou textos para os casos que considerasse mais interessantes. Feitos os textos, foram aplicados novamente em softwares de desenho vetorial.

#### 4.4.4 “Basômetro” – O Estado de São Paulo

O Basômetro<sup>166</sup>, inspirado pelo projeto “Camaraws”, do Grupo de Estudos de Software Livre da Poli-USP (ESTADÃO DADOS, 2013), é uma visualização interativa (Fig. 34) que mede que tanto cada parlamentar apoia ou não o governo. Cada parlamentar é representado por um círculo, e quanto mais próximo do topo o círculo estiver, mais governista ele é. Clicando em cada círculo, aparecem informações sobre o parlamentar em questão. É possível escolher quais partidos e quais estados comparar. Também é possível alternar para uma visualização que prioriza as votações (Fig. 35). Os dados foram obtidos nos sites da Câmara dos Deputados e do Senado, e o próprio Estadão os redistribuiu através do Google Docs e do Github<sup>167</sup>. Neste último site também é disponibilizado o código da ferramenta.

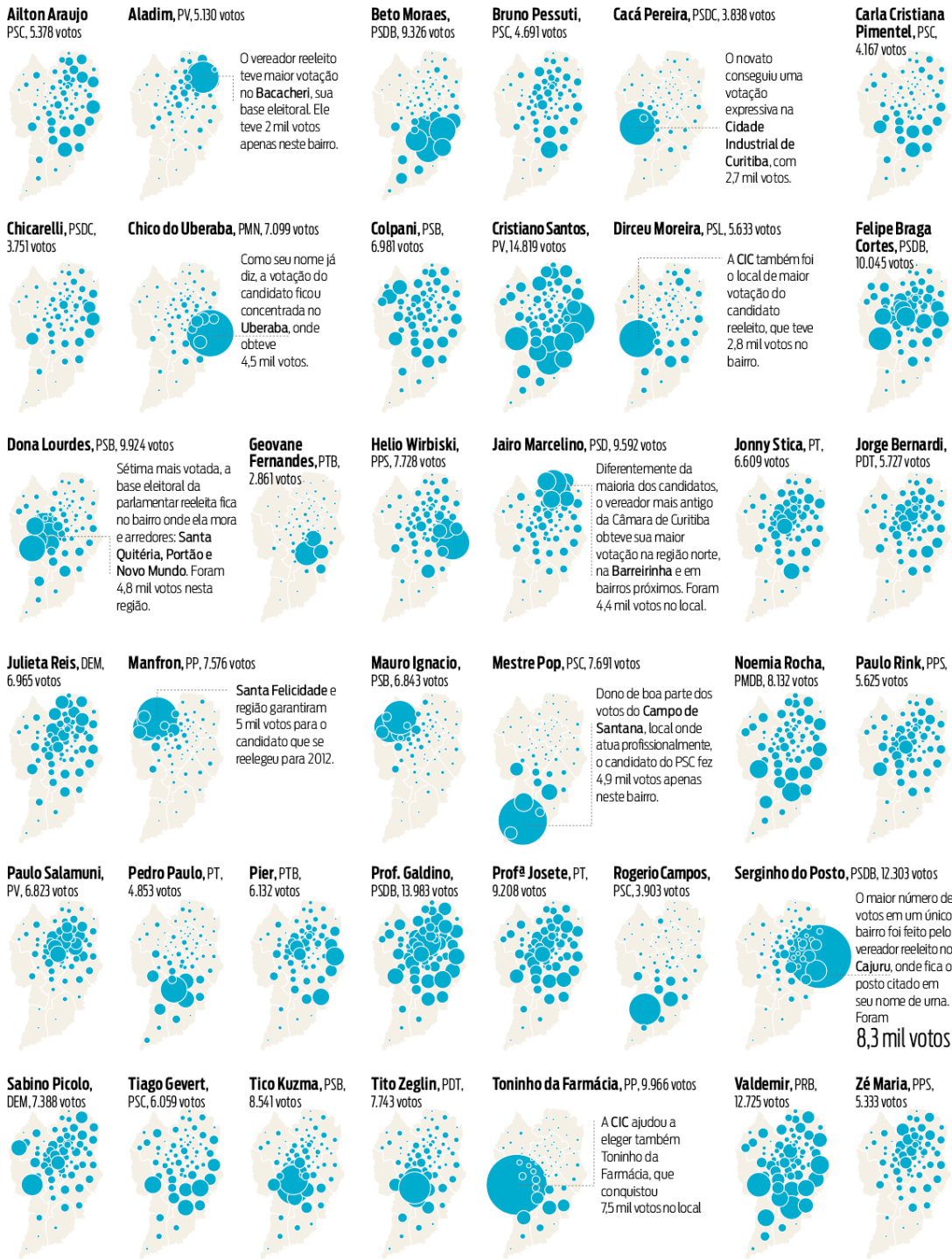
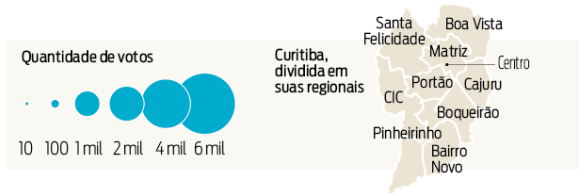
---

<sup>166</sup> Cf. PINTO, Renata Martins de Oliveira Silva. Programação de Alto Nível no Jornalismo de Dados. 2014. 95 páginas. TCC. Universidade Positivo, Curitiba, 2014.

<sup>167</sup> Plataforma online para gerenciamento de projetos que envolvam código. Disponível em <<http://github.com>>.

## BASES ELEITORAIS

Veja como foi a votação dos 38 vereadores eleitos nos bairros da capital. Quanto maior a área azul, mais votos o candidato recebeu.



Fonte: TSE. Infografia: Guilherme R. Storck/GP.

Fig. 33: Infográfico com a quantidade de votos nos bairros de Curitiba para cada vereador eleito. Fonte: STORCK, 2012.





Fig. 34: Basômetro. Fonte: ESTADÃO DADOS, 2013

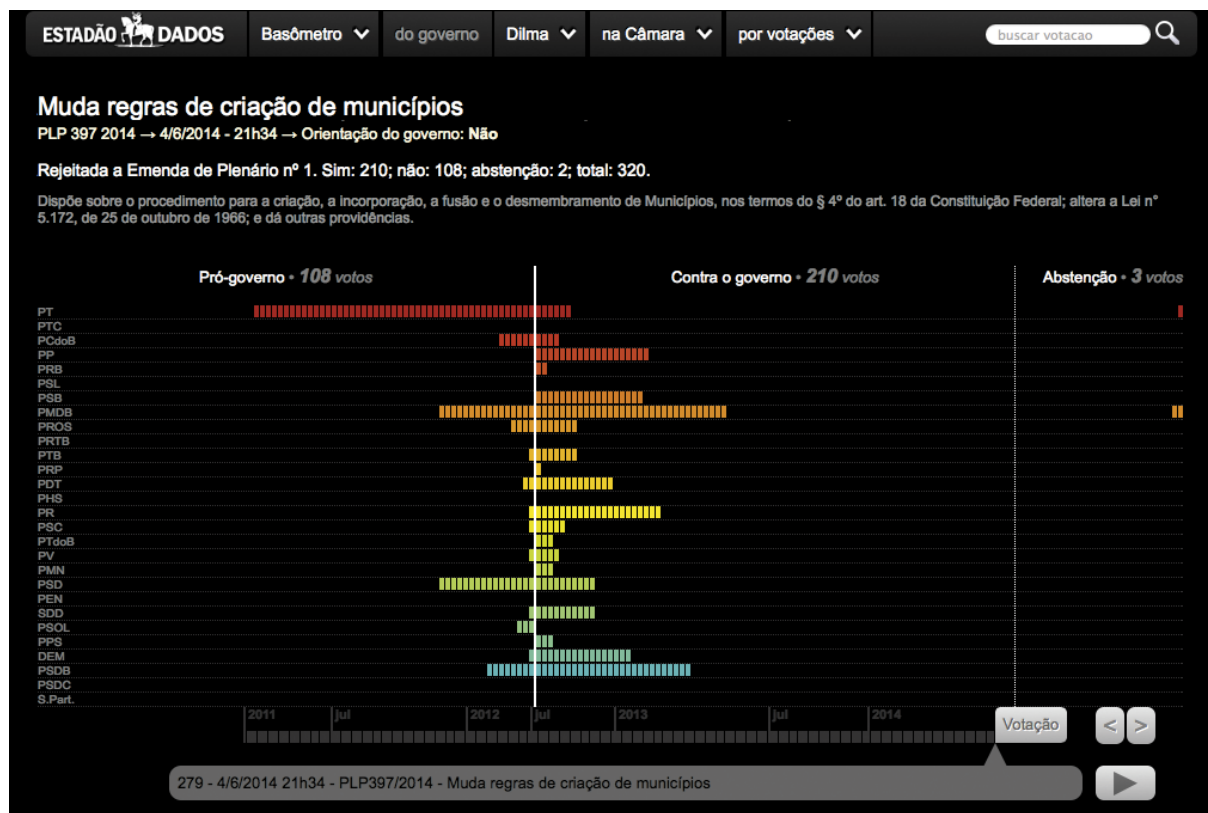


Fig. 35: Outra visão do basômetro. Fonte: ESTADÃO DADOS, 2013.

#### 4.4.5 “World War II in Europe: Every Day” – Emperor Tigerstar

A visualização “World War II in Europe: Every Day” (EMPEROR TIGERSTAR, 2013), produzida pelo estudante estadunidense Bob Smith (ver entrevista no capítulo 5) é um vídeo no qual são apresentados mapas das regiões envolvidas na Segunda Guerra Mundial, a um ritmo de aproximadamente 4 dias por segundo (Fig. 36). Os mapas foram coloridos usando o Microsoft® Paint e editados usando o Windows Movie Maker.

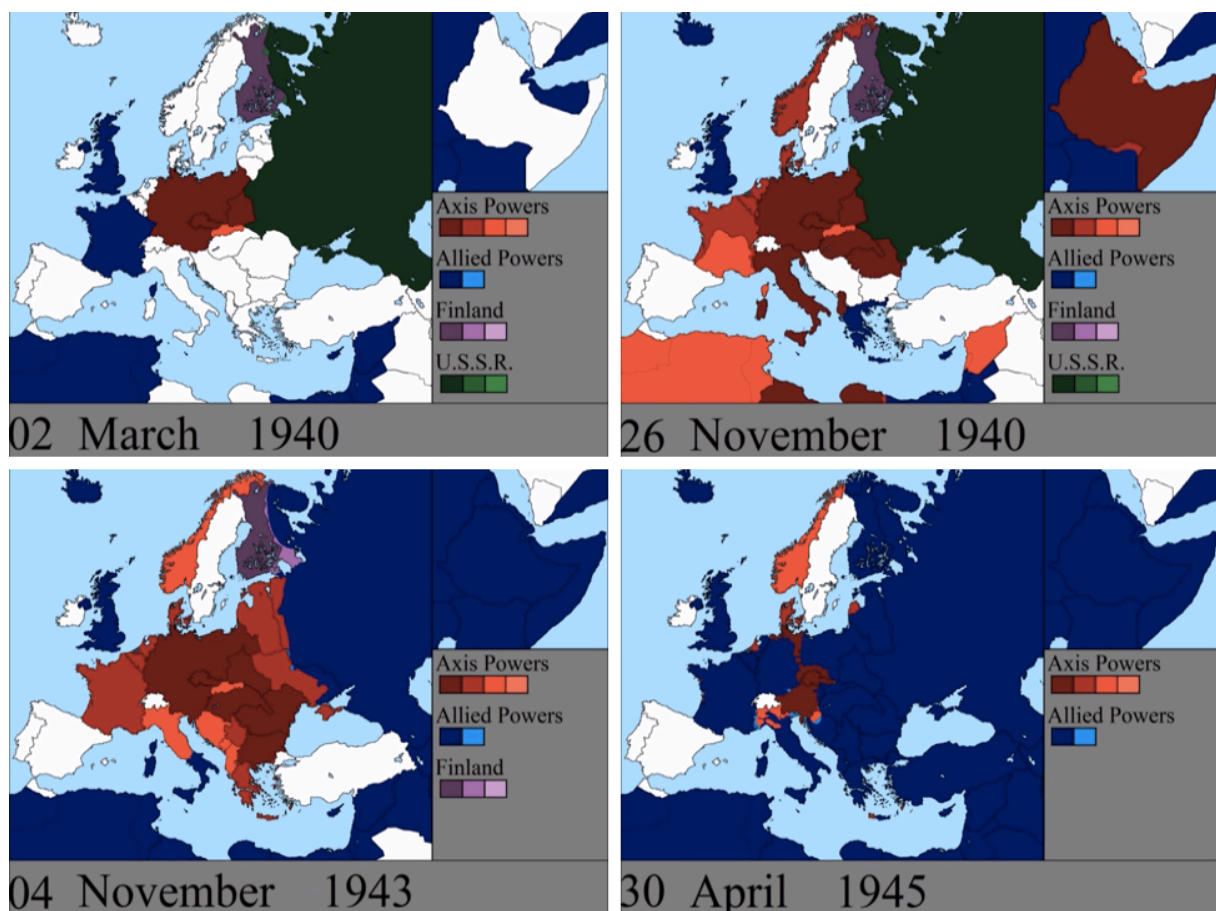


Fig. 36: Quatro quadros da visualização em vídeo "World War II in Europe: Every Day". Fonte: EMPEROR TIGERSTAR, 2013.

#### 4.4.6 “Crime sem castigo” – Gazeta do Povo

A visualização, realizada e publicada no software Tableau, consta de três abas. A primeira é um mapa interativo, a segunda reúne dados em forma de gráficos de barras, e a terceira é um mapa que pode ser animado (Fig. 37).

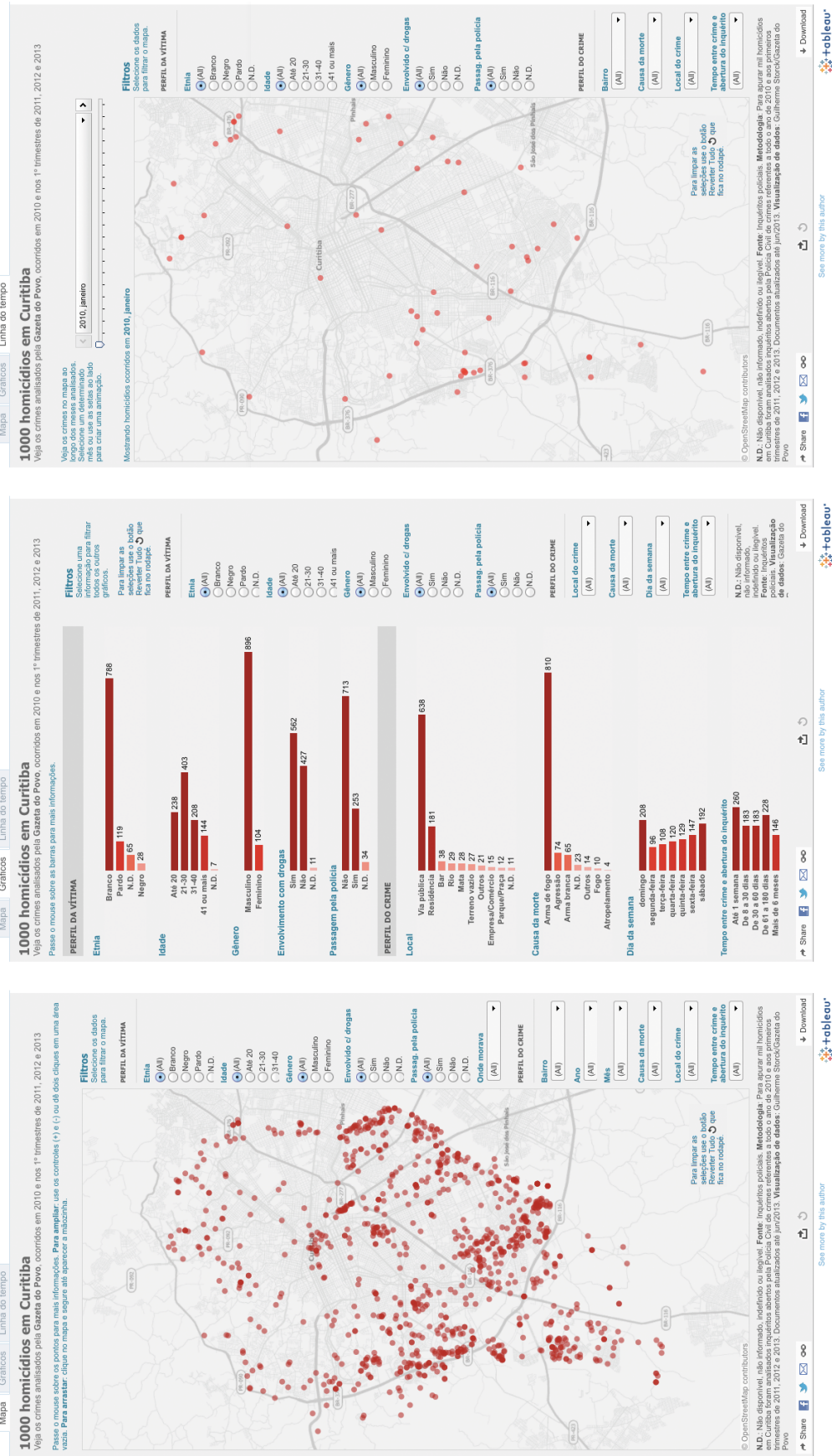


Fig. 37: Três abas da visualização "Crime sem castigo". Fonte: STORCK, 2013.

O conjunto de dados foi montado a partir da análise de inquéritos da Polícia Civil e de informações do Poder Judiciário. Foram analisadas mais de 145 mil páginas de inquéritos referentes a mil homicídios ocorridos nos primeiros trimestres entre 2010 e 2013 (FÉLIX ET AL., 2013). Na base de dados que foi montada em planilha eletrônica, constavam o endereço, a arma usada, se havia alguém preso pelo crime, quantas testemunhas foram ouvidas, entre outros. As informações foram obtidas a partir de cópias xerográficas dos processos, realizadas no Fórum.

#### 4.4.7 Visualizações feitas por estudantes de jornalismo

Os três mapeamentos de diferentes estatísticas do estado do Paraná (Fig. 38, Fig. 39 e Fig. 40) foram feitos por estudantes do curso de jornalismo, em uma atividade de introdução à visualização de dados, ministrada pelo autor desta dissertação, realizada em março de 2013. Os dados foram extraídos do sistema SIDRA<sup>168</sup>, do IBGE, processados usando o pacote Microsoft® Office, e mapeados usando a ferramenta online BatchGeo. Depois do exemplo trabalhado em sala de aula, os alunos e alunas repetiram o processo em suas casas, escolhendo outras informações. A professora da disciplina separou alguns exemplos que se destacaram, três dos quais são os aqui apresentados.

A então aluna Larissa Nichele fez um mapeamento com dados de emigrantes que saíram do estado em 2010, e identificou que o norte concentrou os municípios que tiveram mais de cem pessoas indo para outro país (Fig. 38). Mariana Delamuta notou uma concentração em Curitiba e Região Metropolitana de empregados com carteira assinada, e o norte e noroeste têm mais pessoas sem carteira (Fig. 39). E Fernanda Cercal Odppes mapeou quais os municípios paranaenses com mais pessoas solteiras (Fig. 40).

É possível que tenham acontecido erros durante os mapeamentos, no procedimento ou na escolha da informação. Apesar disso, é interessante observar como a visualização permitiu

---

<sup>168</sup> Sistema IBGE de Recuperação Automática, interface para gerar tabelas com dados selecionados dentre os disponíveis nas bases de dados do IBGE. Disponível em: <<http://www.sidra.ibge.gov.br>>. Acesso em 19/8/2014.



que fossem encontrados padrões e, graças a eles, observações que as próprias alunas consideraram que seriam bons pontos de partida para pautas jornalísticas, em sintonia com a ideia defendida aqui de utilizar a visualização no processo de exploração de dados, para encontrar informações novas.

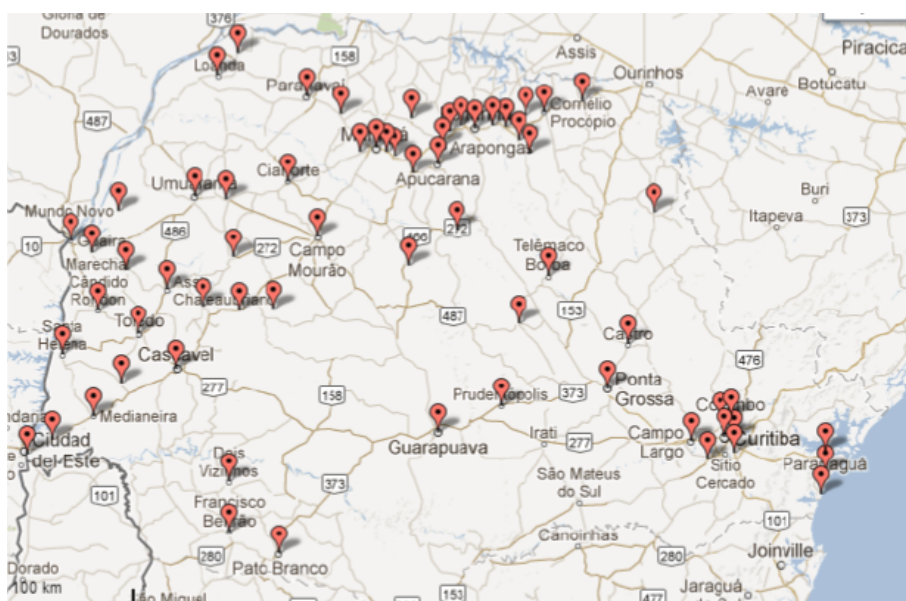


Fig. 38: Mapeamento de emigração. Fonte: NICHELE, 2013.

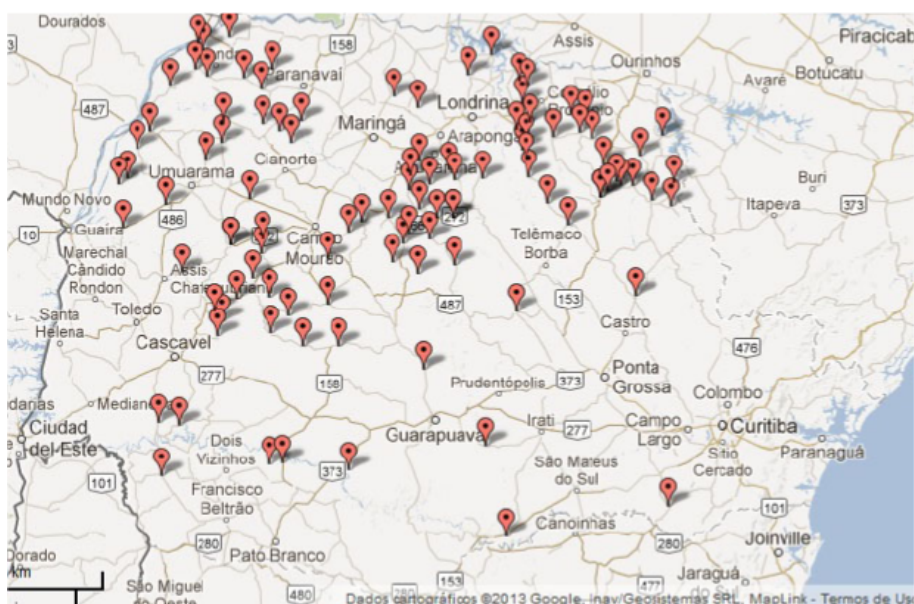


Fig. 39: Mapeamento de empregados com carteira assinada. Fonte: DELAMUTA, 2013.

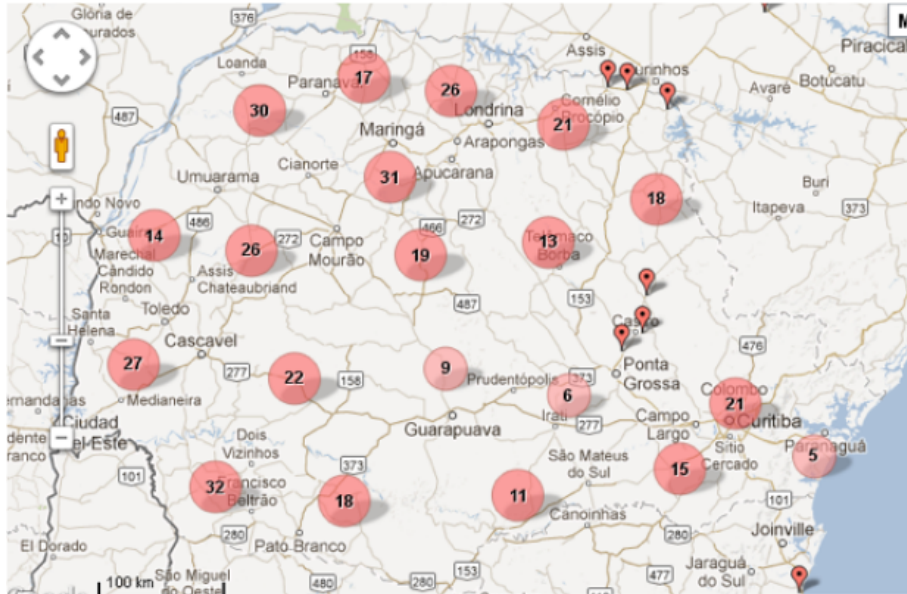


Fig. 40: Mapeamento de concentração de pessoas solteiras. Fonte: ODPPEs, 2013.

#### 4.5 CONSIDERAÇÕES

Produtos jornalísticos concebidos a partir de grandes conjuntos de dados resultam em um jornalismo que, além de interpretar, possibilita outras maneiras de acesso aos dados, como por exemplo através de visualizações. Isso implica em uma reabertura do código técnico do jornalismo que propicia outro tipo de relação entre quem produz produtos jornalísticos e quem os consome. Com o código técnico novamente aberto, há outras possibilidades de apropriação desses produtos/artefatos.

Esses novos artefatos são questionados em tanto que pertencentes ao jornalismo, que tradicionalmente tem um papel de mediar a relação entre a realidade – como matéria prima – e o público. Apresentar grandes conjuntos de dados implica em outro tipo de relações e produtos jornalísticos, situando quem os produz em funções que ainda não foram estabilizadas, como sim o são funções mais tradicionais: escrever textos, fotografar, ou editar matérias televisivas. Em alguns dos casos apresentados, a função da ou do jornalista é produzir uma interface para consulta ou apresentação dos dados – o que, de certa maneira, também é escrever o texto de uma notícia. A diferença é que o texto de uma notícia, como artefato, já está estabilizado, e produtos como

visualizações ainda não estariam estabilizados nem para o grupo social de quem os produz nem para o grupo de quem os consome.

Para quem produz jornalismo, as possibilidades da visualização de dados apresentadas atentam para a possibilidade de facilitar a exploração deles durante a produção jornalística – como acontece com os mapas das alunas de jornalismo e a visualização feita em Processing pela equipe do *The New York Times* para explorar os dados de *How Mariano Rivera Dominates Hitters*. Esse uso das visualizações para exploração de dados está em sintonia com o proposto por Gert Nielsen:

A visualização de dados é para pesquisar. Não conta a história por si só. [...] Se o pesquisador trabalha como jornalista, ele pode usar visualização de dados para desenterrar uma história. Essa história tem que ser contada com outros meios que não somente publicar seu material de pesquisa e chamá-lo de infográfico. [...] Visualização de dados é uma maneira de alcançar um objetivo – não o objetivo em si.[...] Quando você trabalha nas notícias com narrativas jornalísticas, você tem que começar a contar essas histórias, mais do que apresentar um conjunto enorme de dados ao leitor e pedir pra ele fazer o trabalho do jornalista. (NIELSEN, 2011)<sup>169</sup>

Nielsen critica, ainda que com certo exagero no título – *A morte da visualização de dados nas notícias* –, o uso sem sentido da visualização de dados, e defende que sejam usadas com mais parcimônia e fazendo uma edição dos dados antes de apresentá-las como parte de uma reportagem ou infográfico. É esse tipo de tarefa que daria continuidade à função do jornalismo como ponte entre a realidade e o público, mas de outras maneiras, ainda não estabilizadas. Nielsen também cita o vídeo *How Mariano Rivera Dominates Hitters* como exemplo de bom uso de visualizações na mídia, por combinar visualizações com uma narrativa linear que facilita o entendimento do leitor; qualidades presentes em várias das visualizações aqui apresentadas, ou na combinação da visualização com as matérias que a contextualizavam, ou nos recursos de desenho usados para facilitar a leitura. É esse tipo de edição que situa o trabalho de jornalistas – incluindo profissionais

---

<sup>169</sup> Tradução livre. No original: “Dataviz is for research. It doesn’t tell the story on its own.[...] If the researcher works as a journalist, he can use dataviz to dig up a story. That story has to be told with other means than just publishing your research-material and call it an infographic.[...] Dataviz is a means to an end – not an end in itself.[...] When you work in the news with journalistic storytelling you bloody well start tell these stories, rather than presenting a huge dataset to the reader and ask them to do your work.”

da computação, designers, infografistas, *hackers* – como mais próximo de projetar artefatos.

As visualizações de dados, enquanto artefatos, têm seus próprios códigos técnicos, alguns dos quais foram aqui apresentados como possibilidades. É necessário pensar que cada visualização terá seus próprios códigos técnicos decorrentes dos códigos técnicos dos dados, das ferramentas usadas, das decisões de desenho, entre outros. Como já problematizado com os grandes conjuntos de dados (ver seção 3.1), é necessário considerar possíveis vieses tanto dos dados usados como da maneira com que foram visualizados.

As diferentes maneiras de ler uma visualização podem ser consideradas negociações do código técnico, parte de uma negociação maior que levaria o uso de visualizações – ou algum aspecto ou tipo específico – à estabilização. Assim como no JGD de maneira geral, é necessário considerar duas instâncias de estabilização, que é o uso por jornalistas, para explorar informação, e pelo público, para consumi-la. No próximo capítulo, são apresentadas entrevistas que sugerem que a visualização não é considerada pelas pessoas entrevistadas como uma maneira de fazer JGD e sim como uma maneira de levar os dados a público, o que indicaria que, em tanto que técnica para interpretar grandes conjuntos de dados, está longe da estabilização.

O método de visualização aqui apresentado, composto de quatro métodos similares, pode ser apropriado parcialmente para a exploração de dados, mesmo que o produto final não seja uma visualização. Os códigos técnicos envolvidos dependem da apropriação que cada pessoa possa fazer do método e das ferramentas que ela escolha. Algumas podem ser mais abertas, e outras mais fechadas – nos conceitos de Spinuzzi (2003, p. 204); outras são *open source* e outras de licença fechada; algumas são de uso bem específico, e outras teriam sido projetadas para usos mais amplos. Em cada sistema podem conviver ferramentas de códigos técnicos bastante diferentes e até contraditórios. Acredita-se que cada pessoa vai se apropriar de um conjunto de ferramentas e vai trabalhar com elas até se aproximarem da estabilização, para a pessoa ou para o grupo que usa essa ferramenta. É importante notar que para que a estabilização aconteça,

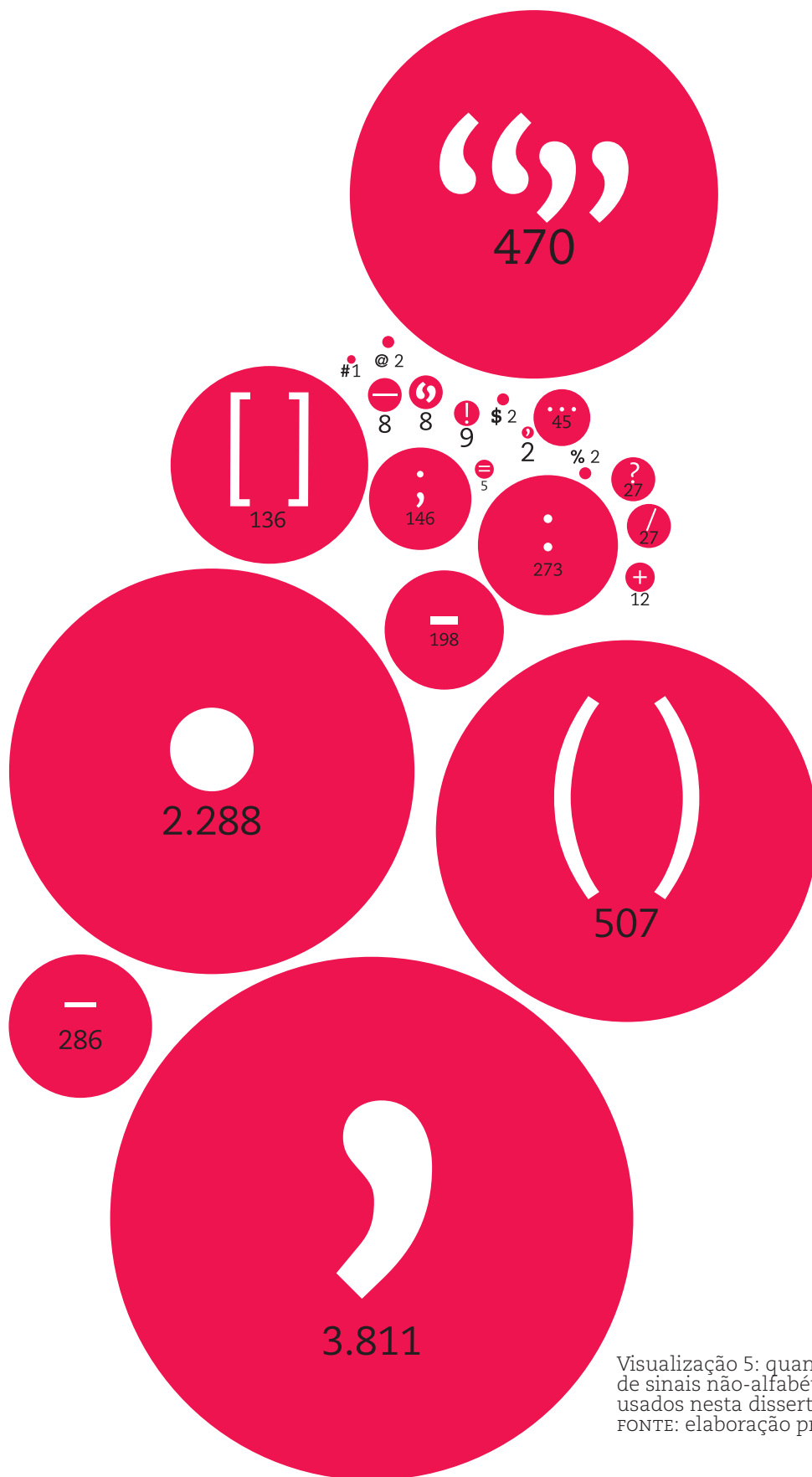


as pessoas não precisam dominar as formas de uso: é possível que aconteça a estabilização retórica (PINCH; BIJKER, 1997, p. 44), que é quanto o problema não é mais percebido como tal, o que pode acontecer por assumirem que a ferramenta cumpre a função, ou por decidirem não procurar outras alternativas.

Para certos grupos, os artefatos necessários para as etapas de obter e processar os dados podem estar mais próximos da estabilização do que os artefatos para fazer visualizações, ou dos artefatos para transformá-las em um produto que seja considerado pronto para ser publicado. Essas situações podem levar a uma divisão de trabalho, possivelmente entre quem produz a parte considerada jornalística e a parte de deixar o produto pronto para consumo – programar um aplicativo ou preparar um infográfico ou visualização para a impressão – o que pode também estabelecer uma hierarquia entre quem se vê como jornalista e quem tem funções consideradas cosméticas. Essa hierarquia também pode se estabelecer de outras maneiras, priorizando pessoas que trabalhem com outras partes do processo. Isso sugere que, para um aproveitamento completo dos recursos da visualização – e de outras maneiras de trabalhar com grandes conjuntos de dados –, é necessário não só estabilizar os artefatos usados, mas também estabilizá-los como parte da identidade de jornalista, ou pelo menos não excluir os artefatos dessa identidade.

Como último aspecto, é necessário pensar nos códigos técnicos dos produtos jornalísticos, no que se refere ao material no qual os produtos são apresentados, como a baixa resolução de um monitor ou o pouco espaço de um telefone celular, que podem ser compensados com recursos interativos como ampliação de setores de imagens, por exemplo, ou a falta de interatividade de publicações em papel, ou ter que privilegiar um ou outro sistema operacional para lançar um aplicativo, ou ainda que os produtos jornalísticos exijam um poder de processamento tal que exclua pessoas com computadores mais antigos, por exemplo.

Em resumo, acredita-se que uma visão construtivista colabore para que artefatos – visualizações e ferramentas do JGD – não sejam vistos como alheios ao trabalho jornalístico, permitindo que mais pessoas aproveitem suas possibilidades, além de reconhecer que outras pessoas fazem parte da atividade jornalística.



Visualização 5: quantidade de sinais não-alfabéticos usados nesta dissertação. FONTE: elaboração própria.

# 5 ENTREVISTAS E OBSERVAÇÕES

Neste capítulo constam as observações coletadas através de entrevistas e observações em campo, assim como considerações que as relacionam com os autores previamente apresentados. Foram incluídas citações literais das pessoas entrevistadas, mantendo a linguagem coloquial, e também citações indiretas, com o objetivo de tornar o texto mais claro. Algumas das citações, já incluídas em exemplos ao longo do texto, são aqui repetidas.

A seção 5.1 apresenta as pessoas entrevistadas, e a seção 5.2 descreve as situações nas quais foram feitas observações. Nas seções seguintes (5.3 a 5.6) são apresentados exemplos de modos de uso coletados e/ou observados, primeiro de maneira geral e depois especificando práticas relacionadas ao uso de programação, à maneira como são buscadas maneiras de aprender a realizar tarefas, e ao uso de visualizações.

## 5.1 ENTREVISTADOS

### *Guilherme Storck*

Formado em Design Gráfico pela Universidade Federal do Paraná, tem 25 anos, mora com a esposa. Depois de estagiar e ocupar o cargo de infografista, Guilherme ocupa o cargo de Editor de Infografia do jornal Gazeta do Povo.

Segundo ele, o ambiente jornalístico o favorece para perceber a importância do Jornalismo Guiado por Dados, como por exemplo a participação nos congressos da Abraji<sup>170</sup>, nos quais viu “a parte final dos trabalhos, como eles são publicados, que impacto têm na sociedade” e teve a percepção de que “vai meio longe essa coisa aí [o Jornalismo Guiado por Dados], tem uma importância legal.”

Ele é o principal encarregado das tarefas operacionais relacionadas ao Jornalismo Guiado por Dados, como é o caso da

---

<sup>170</sup> Associação Brasileira de Jornalismo Investigativo.

extração dos dados de gastos da Assembleia Legislativa do Paraná (GAZETA DO POVO, 2014). Foi a partir desse trabalho que a redação começou a encomendar tarefas de raspagem e visualização de dados. Esse trabalho sobre a Assembleia foi seu primeiro trabalho relacionado ao JGD, em 2010.

### *Renata Silva Pinto*

Curitibana, estudante do último ano do curso de jornalismo da Universidade Positivo, tem 21 anos, solteira, mora com os pais em Curitiba. Renata entrou em contato com os conceitos de RAC e Jornalismo Guiado por Dados nas aulas da disciplina de Jornalismo Digital, nas quais não houve um foco nessas atividades, e sim algumas atividades dentro da disciplina. Começou, em suas palavras, “a meter a mão na massa” quando entrou no Núcleo de Jornalismo Investigativo, que, junto à professora Rosiane Correia de Freitas, da própria Universidade Positivo, começaram a experimentar com dados: “foi um experimento, tanto as aulas como o núcleo em si.” Também estagiou por um mês no jornal Gazeta do Povo, ocasião na qual ajudou na preparação da base de dados usada para a série de reportagens “Crime sem castigo” (ver seção 4.4.6).

### *Bob Smith*

“Emperor Tigerstar” é o pseudônimo do estudante de ensino médio estadunidense Bob Smith. Ele pretende, algum dia, obter “um doutorado ou coisa parecida” em história. Em seu canal no YouTube, ele têm alguns mapas que mostram a situação durante as guerras, como o do exemplo da seção 4.4.5.

### *Ricardo Ruas Silva*

Ricardo tem 25 anos e é de Minas Gerais, onde estudou Sistemas de Informação. Solteiro, mora em Curitiba há um ano, e atualmente mora com dois colegas. Ele participa da elaboração do Atlas do Desenvolvimento Humano no Brasil (PNUD, 2013). Nesse trabalho, ele e a sua equipe fizeram uso de ferramentas prontas para a realização de gráficos – Google Charts – e para a realização de mapas – MapBox.

*José Lazaro Junior e João Guilherme Frey, do Livre.jor*

José nasceu em Curitiba, tem 31 anos e é jornalista concursado da Câmara Municipal de Curitiba. Estudou jornalismo na Universidade Federal do Paraná, instituição na qual também fez especialização e mestrado em Sociologia. Mora com a namorada.

João tem 25 anos, é de Umuarama, oeste do Paraná, e atualmente mora em Curitiba com a namorada e uma filha. Formado em jornalismo, trabalha com assessoria de imprensa na Câmara Municipal de Curitiba. Está concluindo uma pós-graduação em Direito à Cidade e Gestão Urbana, na Universidade Positivo.

Eles fundaram o Livre.jor no início de 2014, com a proposta de fazer jornalismo com base em dados públicos – diários oficiais, bancos de dados, repositórios de documentos, entre outros. “A gente trabalha com outro conceito de notícia, que não é só o extraordinário, interessa mais o interesse público”, disse João. Sendo assim, não têm as preocupações de conseguir matérias surpreendentes ou chamativas: “a nossa leitura de que aquilo é um furo [de reportagem] é diferente”, diz José.

Para realizar seu trabalho, eles aplicam técnicas próximas do JGD, que usam para pesquisar informações disponíveis em portais da transparência e diários oficiais, com foco no estado do Paraná. Os métodos do Livre.jor permitem que José e João trabalhem em suas casas, em horários compatíveis com suas outras atividades.

Até o fechamento desta dissertação, o Livre.jor tinha em sua página no Facebook seu principal meio de publicação. As matérias, que tentam ser “um texto gerado por uma máquina”, nos termos de José, são publicadas acompanhadas de imagens, geralmente capturas de telas de diários oficiais ou gráficos estatísticos. Muitas vezes essas publicações servem de ponto de partida para que jornalistas de outros meios desenvolvam matérias, complementando com entrevistas e checagem de dados, atividades que o Livre.jor não desenvolve. Eles acabam servindo de apoio a reportagens, segundo José:

O pessoal faz [as matérias] e conversa com a gente por mensagem [no Facebook]: ó, tamo pegando a pauta lá, beleza? Vocês têm mais alguma coisa? A gente acaba dando um suporte pra essas pautas, mesmo que a pauta não cite a gente diretamente.

José considera que as redações trabalham com jornalistas cada vez mais jovens, sem experiência, o que dificultaria o uso de dados públicos. O conhecimento necessário para isso “requer um mínimo de tempo de estrada” assim como uma disposição.

A pessoa consegue fazer uma carreira de sucesso em [jornalismo de] política sem isso. Consegue ter fontes, criar canais, ter os bastidores, sem ter acesso aos dados. Mas a gente entende que é uma carreira mais pobre.

João aponta que os métodos do jornalismo de dados dão mais independência, principalmente quando se trata de jornalismo político: “se você depende de fonte [pessoa que fornece informações a jornalistas], uma hora você vai ter que pagar esse favor. Isso [o uso de dados públicos] te liberta.”

### *Docentes que lecionam JGD*

Como complemento às entrevistas com jornalistas, serão apresentadas algumas percepções da professora Rosiane Correia de Freitas, da Universidade Positivo, de Curitiba; do professor Marcelo Träsel, da PUCRS; e do professor Ricardo Fotios, que dá aulas nas instituições paulistas Escola Superior de Propaganda e Marketing e na Universidade Metodista. Todos já trabalharam com jornalismo impresso e online e todos abordam o ensino de JGD com alunos e alunas de cursos de graduação em jornalismo. O objetivo é conhecer de que maneira o assunto é abordado em sala de aula, e também como os alunos e alunas recebem esse tipo de conteúdo.

Os três docentes afirmam que em suas respectivas universidades o ensino de Jornalismo Guiado por Dados é apoiado pela coordenação do curso, apesar de não existir uma disciplina específica – a exceção é a Universidade Metodista. Na Universidade Positivo, explica Rosiane, o apoio da universidade acontece, em parte, pelos resultados das matérias produzidas com técnicas de JGD. O Jornalismo Guiado por Dados é abordado, geralmente, na disciplina de Jornalismo Online, e talvez seja abordado, no caso da PUCRS, numa futura disciplina de Jornalismo Investigativo. Porém, para Träsel, seria necessário mais tempo de aula para conseguir “desenvolver um trabalho razoável”. Na Universidade Positivo há

um Núcleo de Jornalismo Investigativo – também citado por Renata – no qual são feitas matérias usando técnicas de JGD, além de ter parcerias com alguns meios de comunicação, como RIC e Gazeta do Povo. “A princípio é difícil de envolver os alunos, mas o resultado sempre impressiona tanto quem está envolvido no trabalho quanto quem vê de fora”, diz Rosiane.

## 5.2 OBSERVAÇÕES EM OFICINAS DE VISUALIZAÇÃO

### *Oficinas de visualização com alunos de jornalismo*

A primeira oficina observada foi ministrada em março de 2013, a alunas e alunos de jornalismo da Universidade Positivo. A convite da professora e jornalista Kátia Brembatti, foi feita uma atividade em sala de aula para apresentar conceitos de JGD e visualizações. A aula foi no período noturno e eram aproximadamente trinta alunas e alunos que teriam entre 18 e 22 anos.

Com base nessa experiência, e já com a intenção de levantar informações para a presente dissertação, foi realizada uma nova oficina, também com estudantes de jornalismo, desta vez na UFPR, a convite do professor e jornalista José Carlos Fernandes. A aula foi no período matutino e eram aproximadamente vinte alunos e alunas que teriam entre 18 e 22 anos.

Uma terceira oficina, organizada com apoio do Centro Universitário Uninter, foi ministrada durante dois sábados de manhã. Foi uma atividade fora da aula, para três alunas. No primeiro dia foi mostrado como baixar dados do IBGE, como arrumá-los usando Notepad++, tabulá-los no Excel e visualizá-los no BatchGeo. Houve dificuldades com os computadores, o que demorou um pouco o processo e desestimulou as alunas. Acredita-se que ter começado a atividade pelas planilhas também pode ter tornado a tarefa enfadonha. As que estiveram no primeiro dia não foram no segundo – somente uma aluna que não pode ir no primeiro dia. Depois da oficina, as alunas não enviaram o exercício solicitado – assim como na primeira oficina, foi solicitado que elaborassem uma sugestão de pauta com base nos dados visualizados.

### *Grupo de estudos em Ruby e Hackaton*

Outra situação escolhida para colher percepções sobre como jornalistas se apropriam do uso de computadores foi um grupo de estudos de Data Scraping da Universidade Positivo, organizado pela professora Rosiane Correia de Freitas. Os encontros do grupo aconteciam em média a cada duas semanas, com duração de aproximadamente uma hora. O formato era de uma aula, ministrada por um estudante do curso de Sistemas de Informação, nas quais eram apresentadas noções de programação na linguagem Ruby. O objetivo era fazer programas em Ruby para coletar dados de páginas da internet. As aulas eram assistidas por entre três e nove alunos e alunas, sendo alguns profissionais, docentes e estudantes de jornalismo.

Como exercício para o encerramento do grupo de estudos, a professora Rosiane propôs um Hackaton, que foi realizado no sábado 1º de novembro de 2014, feriado. A data foi escolhida para trabalhar com dados das eleições de 2014. Inicialmente pensado como um evento para os integrantes do grupo, mas, graças a divulgação e convites, teve aproximadamente 30 participantes. Além de jornalistas e estudantes de jornalismo, participaram alunos de cursos de computação e de estatística. Foram organizadas pequenas oficinas, com apresentações sobre JGD, apresentação dos trabalhos do Livre.jor, dicas de uso de expressões regulares – ministrada pelo autor desta dissertação – e de SEO.

## **5.3 MODOS DE USO**

### *Guilherme Storck*

Guilherme relata quais as ferramentas que ele usa para executar uma tarefa de preparação e exploração de dados, em uma ordem que ele considera típica:

Primeiro [a ser usado] é o TextWrangler, editor de texto com funções avançadas. É o que mais uso: entre um programa e outro, sempre passo por ele.

Em segundo vem o Excel, porque quase todos os tipos de informação que a gente tem acabam tendo meio que formato de tabela, ou a tabela como formato intermediário ajuda a resolver um problema ou a conferir os dados.



Depois que os dados estão prontos, uma função principal pra bastante coisa que eu faço aqui é mapear, trabalhar com os dados geográficos, eu uso o QGIS, que é um software livre, para trabalhar com mapeamento, *shapefiles*, camadas de satélite. Ele é sensacional, ele faz tudo o que o ArcGIS, que é super caro, faz, e faz melhor, e roda no Mac, tranquilo.

Tableau: é meio que um *all-in-one*, você tendo os dados prontos e no formato certo, o Tableau é um jeito de visualizar de várias formas e ver como pode ficar. Vou jogando os campos, ele vai me dando umas tabelas, uns gráficos de barras, mapas, e eu consigo ver o que tava escondido na minha tabela, desse jeito. O bom dele é que ele não é a melhor ferramenta que existe, mas gera uma visualização interativa. Eu consigo mandar o link para o reporter visualizar mesmo antes de eu ter finalizado, e consigo também gerar um pdf depois que está publicado, para, por exemplo, editar num programa de vetor para usá-lo no jornal impresso.

O Acrobat acabo tendo que usar, mas só pra tirar as coisas de dentro [de PDFs]. O Google Refine, que é para limpar os dados, para conferir... estou dando uma olhada na minha lista... tem as ferramentas do Google, tem o Fusion Tables, que não é tão frequente, mas acho que dá pra colocar o Fusion Tables e o BatchGeo no mesmo pacote, que eu uso pra mapear bastantes locais. Essas não são de todo dia, mas uma vez por semana vou acabar usando.

O que a gente está usando mais agora é o Raw. É no navegador, você joga uns dados lá, tem vários tipos de visualização. Ele já baixa um arquivo em vetor, você joga no software e usa pra jogar no infográfico.

Às vezes aparecem novas tarefas que exigem o aprendizado de novas ferramentas. Porém, se esse aprendizado não acontecer, não há problemas: “se não der da melhor forma, faz alguma gambiarra parecida que leva alguns passos a mais.” Um exemplo de como às vezes é necessário sair dos passos “típicos” explicados anteriormente é a extração e divulgação dos dados de gastos da Assembleia Legislativa do Paraná (GAZETA DO POVO, 2014). Nesse trabalho, Guilherme relembra que precisou achar a lógica da localização dos arquivos dentro da estrutura da página da internet na qual as informações estavam disponíveis, fazer uma tabela com as URLs de cada documento e usar um plug-in do navegador Firefox para baixar os arquivos. Depois, precisou achar um programa para destravar os arquivos PDF baixados, que como “[o programa] não abre um monte de coisa ao mesmo tempo, [então] tive que aprender a fazer uma *action*<sup>171</sup> para abrir os arquivos,

---

<sup>171</sup> Termo que se refere a um pequeno programa destinado a automatizar uma tarefa específica.

destravá-los e salvá-los”. Essa *action* foi feita no *Automator*, aplicativo de automação de tarefas incluído no Mac OSX, e gera cópias destravadas dos PDFs, que são unidos no Adobe Acrobat Pro, para poder copiar o conteúdo para o processador de texto TextWrangler, no qual Guilherme tem um *script* para preparar o texto para ser colado em uma planilha. Antes da planilha, Guilherme passa o texto pelo Google Refine, para preencher alguns campos vazios. Finalmente, a planilha é repassada ao setor responsável pela publicação em uma página do sítio da Gazeta do Povo na qual podem ser consultados pelo público (GAZETA DO POVO, 2014).

### *Renata Silva Pinto*

Enquanto para Guilherme o Excel é importante para certas tarefas, para Renata é a ferramenta mais usada. O primeiro contato dela com Excel mais avançado foi num curso de jornalismo investigativo nos Estados Unidos, onde aprendeu Tableau, Excel e Access. Segundo ela, esse curso “foi o que me deu o pontapé inicial”. Depois, ela fez curso de Excel avançado no CIEE/PR<sup>172</sup>, já com a ideia de trabalhar com Jornalismo Guiado por Dados. Era a única pessoa do curso que estudava jornalismo. Ela consegue programar algumas coisas, ou “brincar um pouquinho com macro”, nas palavras dela.

Ela usa Excel para uma primeira aproximação aos dados, comparando-os e fazendo cruzamentos: a “primeira coisa que eu faço é brincar com a tabela e ver os dados diferentes, comparar e por aí vai”. Ela também conta que se sente “estranha de usar muito Excel, é engraçado quando eu começo a abrir porque é como um vício, eu me sinto estupidamente idiota e bem quando eu tô usando Excel”. Ela acredita que talvez não encontre uma ferramenta que substitua as planilhas:

Eu não acho que tenha que deixar de usar Excel porque ainda não sei o suficiente de programação. Mas eu não sei também se [...] existe uma etapa que eu não vou precisar mais usar Excel, se só a programação vai ser suficiente.

---

<sup>172</sup> Centro de Integração Empresa-Escola do Paraná.

Quando começou a trabalhar com técnicas de JGD, Renata pensou que “tinha que aprender na hora [rapidamente] e fazer cinquenta mil coisas”, mas agora ela não acredita que precise de um conhecimento profundo para realizar tarefas.

Renata também cita ferramentas como Tabula e PDF2XL Enterprise, que usa para converter arquivos PDF em planilhas.

### *Bob Smith*

Para fazer seus mapas animados, Bob Smith usa os softwares Paint e Movie Maker, que vêm pré-instalados no sistema operacional Windows, “simplesmente porque são gratuitos, já estão instalados, e são fáceis de usar.” Ele cita a dificuldade de uma ferramenta simples como o Paint, mas não sente necessidade de algo mais complexo “como o Photoshop”, inclusive porque “de qualquer maneira, não faço ideia de como usá-lo”. Ele não pretende usar outras ferramentas a menos que realmente precise delas, e, nesse caso, “provavelmente tentaria coisas aleatórias até eu conseguir o que eu quero.”<sup>173</sup>

### *Ricardo Ruas Silva*

Durante a elaboração do Atlas do Desenvolvimento Humano no Brasil (PNUD, 2013), Ricardo e sua equipe decidiram fazer uso de ferramentas prontas – Google Charts e MapBox – para fazer gráficos e mapas. Segundo Ricardo, essa decisão foi tomada por causa do pouco tempo disponível para realizar o projeto. A decisão sobre o formato no qual os dados deveriam ser representados era de uma equipe de designers gráficos, mas houve negociação com relação a algumas ideias difíceis de implementar e que exigiriam soluções muito fora dos padrões do Google Charts.

Houve algumas situações nas quais os padrões do Google Charts tiveram de ser contornados. Por exemplo, para conseguir realizar uma pirâmide populacional, Ricardo precisou, através de programação, transformar em negativo o número que representa a população feminina, para que as barras fossem representadas da

---

<sup>173</sup> Traduções nossas. Textos originais: “a doctorate or something like that”; “simply because they’re free, already installed, and easy to use”; “like Photoshop”; “I have no idea how to use it anyway”; “I’d probably just try random things until I get what I want.”

direita para a esquerda (Fig. 41). Feito isso, ele modificou os números da escala (abaixo do gráfico).

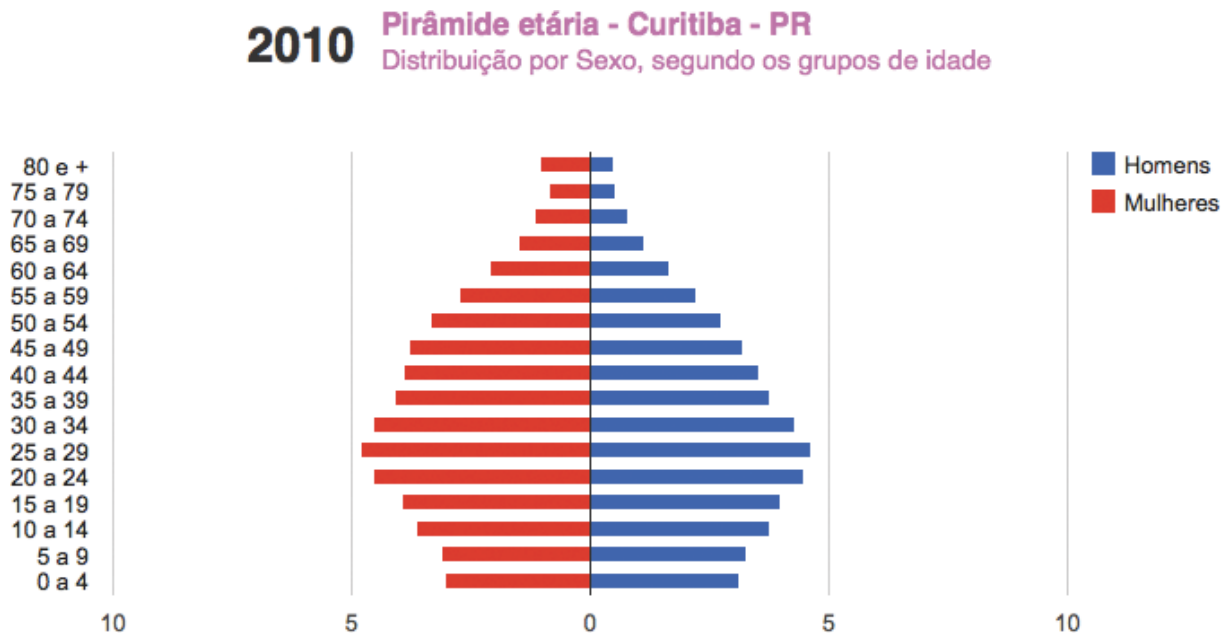


Fig. 41: Pirâmide de população. Fonte: PNUD, 2013.

### *Livre.jor*

Quando questionados sobre ferramentas, a primeira que vem à cabeça de João e José não é um aplicativo de computador e sim o pedido de informação, direito garantido pela Lei da Transparência. Segundo João, “a gente faz muito pedido de acesso porque o próprio pedido de acesso já é uma notícia: se ele vai ser respondido no prazo correto, a resposta... a gente gosta de publicar isso. Publicamos e veio incorreto, publicamos e demorou, publicamos e chegou ok. A gente gosta de fazer isso porque é um teste.”

Eles também citam alguns sistemas, como o Finbra, da Secretaria do Tesouro Nacional, e o Siga Brasil, do Senado. O Finbra, é um banco de dados financeiros do Brasil, é preferido por João:

Por que que eu uso muito mais o Finbra? Além de ser um documento muito bacana, é um documento em Access. É um documento aberto [não na licença, mas na facilidade de uso]. Eu já fiz alguns trabalhos de ficar seis, sete dias seguidos, porque eu sou meio ruim com tecnologia, no ctrl-c, ctrl-v de PDF. Isso é ruim porque eu posso ter errado em algum ctrl-c ctrl-v, e é um trabalho horrível.

Já o Siga Brasil gera arquivos a partir das informações escolhidas. João exemplifica: “eu quero ver gastos do governo federal na subfunção desenvolvimento da educação infantil no estado do Paraná”. É possível baixar os bancos de dados em vários formatos, incluindo Excel, que é o que João prefere. Excel também é o programa que João usa para cuidar dos dados, João usa a planilha Excel:

Do Access pro Excel é copia e cola. Copio do Access o que me interessa e joga no Excel. Faço assim mesmo no manual, nem sei se tem outro jeito. Eu fiz um curso de Excel até pra tentar melhorar, ajudou algumas coisas.

Seja no Access ou no Excel, João faz uma analogia entre a abordagem que ele faz dos dados e o jornalismo tradicional:

Eu vou flinando pelo sistema. É igual pegar pauta andando na rua. Uma hora você vai pegar um acidente. Eu vou mexendo nas coisas: se eu cruzar isso com isso, vou fazer alguma coisa legal? Descobriram [foi noticiado] que aconteceu alguma coisa no Piauí. Como tá [a situação] no Paraná? Meio jornalista tradicionalzão.

Se João trabalha mais com números, José prefere informações textuais. Uma das ferramentas usadas por José é o RSS, que “é fundamental pra estar em dia com o conteúdo. A gente precisa estar em dia, a gente precisa saber o que não foi dado [por outros meios jornalísticos].” Quando acontece de algum site não ter o RSS funcionando, ele solicita que seja disponibilizado, e costuma ser atendido, como aconteceu no caso dos blogs de política do jornal Gazeta do Povo. Ele usava o extinto leitor de RSS online Google Reader, e teve que substituí-lo pelo Feedly – e pagar uma taxa anual para ter recursos de busca similares aos do Google Reader.

na pasta que eu tinha com coisas da Câmara, do Senado e da União, costumava dar busca por “risquinho [hífen] PR [abreviatura do estado do Paraná]”, porque eu sabia que os políticos, por padrão federal, é sempre “partido-PR”, então eu dava essa busca lá e achava tudo que a bancada do Paraná tinha feito no dia, organizado. Era maravilhoso.

Ele também faz buscas por termos geográficos ou palavras-chave relacionadas a algum acontecimento pontual.

José conta que também já fez “controle georreferenciado de posts no Twitter e Facebook”. Ele já fez isso em uma cobertura do Enem para o portal UOL. “Tem como pôr uma marca em um ponto,

pegar um raio de um quilômetro, e ver quem está pondo tweets na rede.” Ele considerou fazer isso durante as eleições, para ver se havia algum funcionário público fazendo propaganda indevida. Para essas atividades, ele usa o aplicativo gratuito Hootsuite. Ele diz também que “tem vários truques que acabo usando uma vez ou outra”.

Outras ferramentas citadas foram o Tabula, para converter PDFs em planilhas, e caneta e papel, especificamente para fazer esboços e mapas mentais para “organizar ideias”.

### *Observação com José*

A primeira observação foi feita com José, em 18 de novembro de 2014, à tarde, em seu apartamento, que é onde e quando ele costuma fazer essas atividades. O trabalho principal dele é a leitura dos Diários Oficiais. Ele acessa a página da Imprensa Oficial do estado do Paraná, na qual há links para três dos diários que ele lê. Ele repete o trabalho que faz no dia a dia, mas com pausas para explicá-lo ao pesquisador.

José começa pela leitura do índice, e já pela quantidade de páginas de cada seção ele decide em qual parte prestar atenção. No exemplo do dia 18 de novembro de 2014, eram 40 páginas de finanças. Ele está acostumado a “passar o olho” por essas páginas em busca de expressões que signifiquem alguma coisa para ele. Passa rapidamente pela seção de decretos, já que ele sabe que provavelmente não haverá nada de interesse. “Se fosse a assembleia [legislativa], eu leria com mais calma”. Ele encontra um ato de mudança de natureza da despesa, cujo código ele desconhece. “Se eu quiser achar isso tenho que olhar a pedra de Rosetta” que é como ele se referiu ao Manual Técnico da Receita Federal. Ele acha o código através do comando CTRL-F. Ele puxa o site da Receita Federal para uma janela separada, “para poder ter o ctrl-tab [atalho de teclado para alternar as janelas]”. Acha o código que procurava e vê que se refere à Justiça.

Continuando com a leitura, lhe chama a atenção uma transferência no valor de R\$ 85 milhões. “Posso supor que é fundo de transferência do governo federal”. Deixa para ver isso depois. “É inviável olhar os diários todos no detalhe”, diz.

“Chegamos onde tem coisa quente: MP [Ministério Público].” José lê os arquivos PDF do diário usando o navegador Google Chrome, porque “me dá opções de busca melhores”. Com o comando CTRL-F, ele localiza as ocorrências do termo “inquerito civil” – em minúsculas e sem acento – e aparecem 774 ocorrências.

Um dos motivos pelos quais José lê os PDFs no Chrome é que, quando feita uma busca por um termo, aparecem pontos amarelos na barra de rolagem, representando cada ocorrência. O acúmulo de pontos amarelos indica que não há ocorrências do termo fora da seção referente ao Ministério Público.

Das 774 ocorrências, ele quer extrair o número de inquéritos que foram abertos. Para isso ele seleciona a primeira ocorrência e usa o atalho CTRL-SHIFT-G para ir para a última, e segura o atalho para ir de trás pra frente até encontrar o ponto em que a listagem se divide, que ele identifica pela diagramação e tipografia. Ele vê que a última ocorrência da primeira parte é a “545 de 774”, e anota esse número num papel. Ele repete o processo com a busca por “procedimento preparatorio”, que retorna 492 resultados, e “procedimento investigatorio criminal”, que retorna 77 resultados. Nesses dois casos, ele não precisa separar, somente visualiza se as ocorrências, marcadas em amarelo, se acumulam na parte da barra de rolagem referente ao Ministério Público. No mesmo papel em que anotou, faz a soma dos três tipos de inquérito (ele calcula 1124, mas o total é 1114) investigações do Ministério Público.

Quando encontra algo que considera que vai ser a notícia do dia do Livre.Jor, José faz uma captura da tela e a edita no Microsoft Paint – programa de desenho incluso com o sistema operacional Windows – e salva o arquivo com um nome que representa a data da publicação, no formato ano-mês-dia (141118), para facilitar possíveis buscas futuras. Ele inclui indicações de que partes da captura foram omitidas, e publica a imagem junto de um texto explicativo na página do Livre.jor no Facebook.

A continuação, ele abre o Diário da Justiça. “Aqui olho com um pouquinho mais de calma, olhando sobrenomes”, diz José, procurando por processos que envolvam parentes de políticos famosos.

No Diário Oficial da União, procura por Paraná e vê os resultados da busca. Um dos resultados abre como uma imagem, mas tem a opção de baixar o PDF, o que ele faz. Ele não usa muito, pela dificuldade que o sistema impõe.

A pedido do pesquisador, José mostra um exemplo de como ele procura informação nos feeds RSS usando o Feedly.

Ele aproveita a entrevista para trabalhar com a tabela do orçamento, que é distribuída em PDF. Ele converte a tabela com o Tabula, e cola numa planilha do LibreOffice. Ele digita as fórmulas manualmente, “=D5/C5”. Ele faz essa operação a cada três meses, que é quando o balanço é divulgado.

### *Observação com João*

A observação com João foi feita em seu apartamento, no final da tarde do dia 4 de dezembro de 2014. Ele começa mostrando os sites de alguns sistemas que ele usa e que foram citados na entrevista anterior, como o SISTN<sup>174</sup>, da Caixa Econômica Federal, usado para baixar o arquivo do RREO – Relatório Resumido da Execução Orçamentária –, que é um documento bimestral.

Ele acessa e baixa o Finbra, da Secretaria do Tesouro Nacional, e abre o arquivo no Microsoft Access. A seguir, ele mostra como ele procura dados. Geralmente ele quer saber informações dos municípios paranaenses, que é o foco do Livre.jor. Ele copia e cola colunas inteiras para uma planilha no Excel, já que “eu nem sei mexer no Access, só copio e colo”. Então João usa o recurso “formato automático de tabela”, que pinta as linhas da planilha alternando branco e cinza claro, para facilitar a leitura, e também põe filtros, para facilitar a reorganização dos dados. Ele organiza os dados, e mostra uma lacuna de dados de municípios que não lançam os dados, mas ressalta que “eu evito dizer que o município não gastou nada” e sim dizer que o município não lançou as informações dos gastos no sistema. “Tento ser bem fiel ao que o documento diz”.

Ele vai agregando dados na tabela, imaginando qual cruzamento de informações pode originar uma informação nova.

---

<sup>174</sup> Disponível em: <<http://www.caixa.gov.br/poder-publico/apoio-poder-publico/SISTN/Paginas/default.aspx>>. Acesso em 28/1/2015.



Na observação feita, João não pensava tanto sobre quais dados poderiam ser interessantes, priorizando mostrar ao pesquisador a maneira de trabalhar. Também mostra que faz cálculos, como médias e proporções per capita – a preferida por ele. Ele não costuma usar recursos como mediana, e define o seu conhecimento de estatística como “estatística pra *newba*<sup>175</sup>”. Ele organiza a tabela por ordem alfabética de município ou por ordem crescente de população para poder colar novas colunas e que as linhas sejam do mesmo município. Vai acrescentando informações à tabela, procurando correlações. “Tem muitas [vezes] que eu desisto”, afirma.

Outro recurso usado por João é um email que recebe do Portal da Transparência do Governo Federal. Como o email é enviado de madrugada, João lê no celular, de manhã, as informações que solicitou.

João narra de alguns trabalhos que tentou fazer através de copiar e colar informações para tentar criar uma base de dados que lhe interessasse, mas que acabou desistindo porque demoraria horas ou até dias para chegar em uma base de dados que talvez não tivesse resultado satisfatório. Apesar que esse tipo de trabalho se beneficiaria de processos como conversão automática das bases de dados originais – em PDF – em planilhas eletrônicas, ele tem receio de que aconteça algum erro no processo, o que pode ocasionar em erros na informação. “Não automatizar dá controle”, afirma.

Ele ainda cita algumas questões de interface, que ele chama de “detalhezinhas que acabam atrapalhando”, como o fato de que as planilhas têm a opção de deixar o cabeçalho de uma tabela sempre visível, enquanto o PDF não tem essa possibilidade. Relacionado a esses problemas, João usa seu laptop ligado a um monitor externo, o que facilita as tarefas, já que dessa maneira “não tem que ficar trocando de tela o tempo todo”.

Como ele trabalha mais com dados numéricos, de orçamentos por exemplo, as notícias que ele posta no Livre.jor às vezes são ilustradas com um gráfico, que ele faz no Excel e finaliza no editor

---

<sup>175</sup> Derivação do termo em inglês *newbie*, gíria para inexperiente, novato, iniciante.

de imagens Gimp, aplicativo que ele prefere para tarefas como acrescentar títulos. Ele começou a mexer com o Tableau para conseguir gráficos melhores. Ele também cita as limitações de formatação de texto do Facebook, o que exige o uso de quebras de linha para espaçar parágrafos e caixa alta para diferenciar o título do corpo do texto. Ele também mostra gráficos de círculos, feitos no aplicativo de ilustração Inkscape. Para deixar os círculos proporcionais à área, contou com a ajuda de um familiar, formado em engenharia.

### *Experiências narradas pela professora e professores entrevistados*

A professora Rosiane descreve os conteúdos que trabalha com alunas e alunos:

Eu começo com matemática básica, Excel e banco de dados. Daí vou avançando para noções básicas de programação (HTML, PHP), uso de ferramentas de mapeamento, visualização (Tableau Public, Google Fusion Tables) e termino com raspagem de dados.

Não gosto de trabalhar com software, porque acho que limita o aluno. Mas acabo usando alguns, como o Outwit Hub, o Tableau Public. E uso aplicativos web como o BatchGeo.

Essa abordagem de não focar no software é compartilhada por Fotios, que tenta

focar no uso e tirar o peso do meio [ferramentas], já que os alunos são resistentes a falar sobre softwares, tecnologia e contas. Mas usam bem se for na prática.

Procuro não relacionar muito a este ou aquele programa. Me interessa mais o “o quê” do que o “como”. Eles são melhores que eu para descobrir e operar softwares.

Tanto Fotios como Träsel apontam que os alunos conseguem usar as ferramentas propostas sem problemas. Porém, Träsel aponta que o principal problema é que alunas e alunos entendam o conceito básico de encontrar novas informações utilizando esses recursos:

[os alunos e alunas] têm dificuldade em compreender a diferença entre cruzar dados e simplesmente pegar uma tabela e transformar num gráfico. Acabam entendendo que o objetivo da atividade de Jornalismo Guiado por Dados é copiar dados de algum portal de transparência e apresentar num texto ou tabela. Poucos

conseguem dar o passo adiante, que é combinar duas ou mais bases para chegar a algum conhecimento novo.

### *Observações das oficinas ministradas pelo autor*

Na oficina realizada na Universidade Positivo foi usado o pacote Microsoft® Office em combinação com ferramentas disponíveis na internet. A escolha dessas ferramentas foi feita considerando que não haveria tempo de solicitar a instalação de outros programas. Também foi considerado que a maioria das alunas e dos alunos teria familiaridade com esses programas. As atividades eram realizadas pelo ministrante em um computador ligado a um projetor, e replicadas pelos estudantes em cada computador.

Para fazer o exercício proposto, foram extraídos dados do IBGE<sup>176</sup> sobre a porcentagem da população atendida por serviços de esgoto nos 399 municípios paranaenses. Foi gerada uma planilha no sistema Sidra<sup>177</sup>, que foi copiada para um editor de texto e preparada para colar os dados em uma planilha, na qual foram organizados em ordem decrescente. Os dados foram copiados e colados no site BatchGeo, ferramenta online que localiza endereços usando mapas online. O mapa mostrava quais os municípios com porcentagem de população atendida por esgoto inferior a 50%, sem fazer distinções quantitativas.

Os alunos e alunas não tiveram muitas dificuldades, considerando que tudo foi feito ao mesmo tempo, especificando cada recurso que deveria ser utilizado. De modo geral, o que mais surpreendeu foi o uso de recursos do Word que para eles eram inéditos, como recursos avançados de busca e substituição de texto, assim como o próprio uso de planilhas eletrônicas em atividades jornalísticas. Também surpreendeu a possibilidade de encontrar informações em um mapeamento, mais ainda em um mapeamento que eles e elas conseguiram fazer.

Em oficina similar ministrada na UFPR, a estrutura disponível era parecida à da Universidade Positivo, mas como havia mais tempo disponível, foram feitas duas visualizações. Na primeira

---

<sup>176</sup> Instituto Brasileiro de Geografia e Estatística.

<sup>177</sup> Sistema IBGE de Recuperação Automática, interface para gerar tabelas com dados selecionados dentre os disponíveis nas bases de dados do IBGE. Disponível em: <<http://www.sidra.ibge.gov.br>>. Acesso em 19/8/2014.

foram mapeados dados de todos os países do mundo, levantados por alunas e alunos em diversas fontes, entre elas a Organização Mundial da Saúde. Alguns puderam ser baixados já no formato de planilha, e outros tiveram que ser preparados pelos discentes usando editores de texto e planilhas dos pacotes Microsoft Office e LibreOffice. Por fim, os dados foram carregados no Google Docs, cuja planilha eletrônica tem opções de gráficos que incluem mapas mundiais, permitindo que os países apresentassem variações de cores de acordo com os dados da planilha.

Os discentes da tiveram dificuldades, em entender que os dados que eles deveriam obter tinham que ter um formato fixo, ou seja, uma coluna com os nomes de países. Dois alunos quiseram trabalhar com dados divididos por regiões mundiais, e não por países, o que impossibilitaria a visualização de dados no Google Docs. Outra dificuldade foi a falta de padronização entre as fontes escolhidas pelas alunas e alunos. Como estavam disponíveis em vários formatos, cada dupla de alunos teve problemas diferentes, que demoraram a ser resolvidos. Outro problema foi a variedade nas grafias dos nomes dos países, o que fez com que o Google Docs não localizasse corretamente muitos dos países, resultando em mapas incompletos. Mas, mesmo com vários países sem dados, conseguiram visualizar informações importantes, como por exemplo o contraste de situações entre países europeus e africanos.

O exercício seguinte foi realizado com rapidez. Dado que o primeiro exercício consumiu mais tempo do que o esperado e a dificuldade dos alunos e alunas de arrumar os dados, foi escolhido uma única informação, extraída do IBGE através do sistema Sidra, para que não houvesse a necessidade de resolver diferentes problemas. Os dados foram filtrados em planilha eletrônica e mapeados no BatchGeo. O conhecimento geral das regiões do Paraná facilitou tirar conclusões a partir do mapeamento realizado. Os alunos, então, repetiram o exercício com dados escolhidos por eles. A dificuldade na escolha de dados foi, novamente, achar dados que tivessem o nível de informação necessário – no caso, a informação dos municípios. Em um primeiro mapeamento, somente uma das informações mapeadas por uma aluna não

apresentou informações visuais que mostrassem algum padrão ou concentração regional.

### *Observações no grupo de estudos de Data Scraping*

O primeiro encontro do grupo de estudos aconteceu em um laboratório de computadores destinado ao jornalismo, e por esse motivo os alunos e alunas não puderam configurar os computadores para permitir acesso a pastas e arquivos e outras configurações necessárias para rodar o ambiente de programação Ruby. Para os primeiros exemplos, foi usado o site [tryruby.org](http://tryruby.org), já que os responsáveis pela manutenção dos computadores não fizeram as modificações necessárias para rodar os scripts em Ruby; a partir da terceira aula, os encontros foram em um laboratório usado pelos cursos de computação, com computadores com acesso liberado a essas configurações.

O que caracterizou esse primeiro encontro foi a surpresa que os e as jornalistas tiveram com as possibilidades da programação, mesmo com funções básicas. O professor mostrou o comando `puts`, que resultou em um “Hello World”, e também que o site retornava resultados de operações simples, como `2+2`, o que ocasiona surpresa nos alunos e alunas. Ele insere os comandos `idade=19` e depois `puts=idade`. O programa retorna o valor da variável `idade`, o que provocou uma interjeição de surpresa e admiração. “Isso pra dados é muito bom!”, exclamou uma aluna. Os alunos e alunas também tentavam adivinhar a sintaxe da linguagem, sugerindo o uso de aspas e colchetes para escrever comandos, e perguntavam se poderiam fazer operações matemáticas com as variáveis. O professor usou o comando `puts idade + 10`, o que ocasionou uma expressão de surpresa um pouco menos animada.

Ao longo das aulas, foram observadas dificuldades com interfaces de texto, como usar comandos do MS-DOS e do Terminal do Mac OSX. Somente uma aluna conhecia o Terminal, e houve dificuldade para tarefas como a criação de um arquivo com extensão “.bat” – necessário para contornar as configurações do sistema e tentar rodar os scripts em Ruby. Navegar por pastas no MS-DOS também se mostrou uma tarefa difícil, assim como compreender a representação textual da estrutura de pastas do

sistema. Também foi difícil a compreensão de que arquivos têm uma extensão no nome que precisava ser acrescentada para serem reconhecidos como um script de Ruby. Mesmo com o curso avançando, eram usadas expressões como “magia negra” para se referir ao funcionamento da programação.

### *Observações no Hackaton*

As primeiras atividades do Hackaton foram organizadas pela professora Rosiane, que dividiu as tarefas segundo as capacidades dos e das assistentes. Os estudantes de computação se concentraram na raspagem de dados. As e os jornalistas, de maneira geral, trabalharam manipulando dados em planilhas de Excel. O autor ensinou uma jornalista a juntar vários arquivos CSV usando um recurso presente no MS-DOS, processo que ela repetiu e conseguiu seu objetivo, que era reunir em uma única planilha dezenas de arquivos CSV. Guilherme Storck, entrevistado desta pesquisa, realizou visualizações usando a plataforma Tableau, trabalhando junto a dois estudantes de estatística, que usavam a plataforma R e passaram a maior parte do dia montando uma base de dados para ser usada em visualizações que realizariam no R. Um estudante e uma estudante de jornalismo, mais inexperientes, ficaram responsáveis por catalogar santinhos e folhetos de candidatos e fazer uma planilha descrevendo os itens e os CNPJs das gráficas que os imprimiram, para depois procurarem os CNPJs na internet e anotarem os nomes das empresas – eles foram orientados pelo autor nas duas tarefas.

Em sua palestra, João Guilherme Frey – também entrevistado na presente pesquisa – mostrou uma base de dados de reclamações recebidas no telefone 156, da Prefeitura de Curitiba. Jornalistas se surpreenderam com a existência dessas informações, e notaram a possibilidade de encontrar informações interessantes na exploração desses dados.

A oficina de expressões regulares ministrada pelo autor foi acompanhada por aproximadamente dez pessoas, que usaram o editor de texto Notepad++ – disponível nos computadores por ser usado nos cursos de computação – para padronizar uma base de dados deliberadamente bagunçada. De maneira geral,

conseguiram acompanhar os exercícios propostos e entenderam o uso de expressões regulares.

As maneiras de trabalhar das pessoas eram diferentes de acordo com as suas profissões. De maneira geral, jornalistas combinavam várias ferramentas e tinham métodos mais irregulares, e estavam abertos a receber dicas para realizar tarefas específicas ou funções desconhecidas dos aplicativos. Já quem sabia de computação e os estudantes de estatística aparentavam maior familiaridade com os aplicativos que estavam usando, além de seguirem métodos mais rígidos e saberem qual seria a próxima tarefa necessária para alcançar seus objetivos.

O tempo acabou sendo curto para que as tarefas, tanto as planejadas previamente como as que surgiram no evento, fossem terminadas. Porém, a percepção geral, tanto das e dos jornalistas como das pessoas de outras áreas, foi de que a combinação de conhecimentos de áreas diferentes, típica dos métodos do JGD, podem colaborar para chegar a resultados que não seriam alcançados por métodos mais tradicionais.

### 5.3.1 Considerações sobre os modos de uso observados

No geral, as pessoas entrevistadas e/ou observadas mostram uma certa proficiência no uso de ferramentas, fazendo uso delas de várias maneiras, e aproveitando os mesmos recursos para várias situações. É comum que uma pessoa se sinta à vontade com uma ou duas ferramentas, e que essa convivência faça com que a familiaridade com a ferramenta aumente. Guilherme, por exemplo, usa o editor de texto TextWrangler para várias tarefas e para a transferência de dados entre aplicativos [STORCK, 2014]. João e Renata têm familiaridade com a planilha Excel, a ponto de Renata dizer que “é como um vício” [PINTO, 2014]. Esse conforto também é visível no uso que Bob Smith faz do aplicativo Paint, que apesar de se tratar de uma ferramenta que não permite tantas possibilidades de apropriação quanto planilhas ou editores de texto, serve para o que Bob procura. A mesma ferramenta é usada por José para compilar trechos dos diários em uma única imagem para ser publicada no Facebook.

As características de desenho de planilhas e editores de texto podem ser relacionadas à ideia de abertura ou fechamento de um sistema de Spinuzzi (2003, p. 205), sendo um sistema mais aberto quando possibilita que quem o use possa usá-lo de várias maneiras, de acordo com seus modos de uso e com o que a pessoa quer fazer. A planilha eletrônica é justamente citada como exemplo de sistema aberto e modular (NARDI *apud* SPINUZZI, 2003, p. 205). O TextWrangler também pode ser considerado um sistema aberto, já que pode ser usado de maneira mais básica e também aprender, de maneira gradual, funções mais avançadas, como o uso de expressões regulares e *scripts*.

Um exemplo desse tipo de aprendizado foi observado durante o Hackaton. Renata – que também participou – precisou fazer uma contagem de ocorrências em uma planilha, e para isto precisou aprender uma função que ainda lhe era desconhecida. Feito isso, usou o resultado dessa nova função em combinação com *pivotables*, que usou de maneira diferente do que ela costuma usar (NOTAS DO DIÁRIO DE CAMPO, 1º/11/2014, HACKATON NA UNIVERSIDADE POSITIVO).

No caso do TextWrangler, é possível perceber como Guilherme fez uso, segundo ele, primeiro de funções de busca e substituição de texto usando expressões regulares, e depois de automatizações com *scripts*. O TextWrangler trabalha com texto simples, sem formatação, o que permite que ele seja usado para manipular arquivos que também são de texto simples, como bases de dados em formato CSV, XML ou JSON<sup>178</sup>. O texto simples também facilita o intercâmbio entre diferentes programas, permitindo fazer alterações em um arquivo de texto usando expressões regulares antes de colar esses dados em uma planilha, que tem menos recursos de busca e substituição, como Guilherme faz rotineiramente [STORCK, 2014].

Considerando que tão próximos da estabilização (FEENBERG, 1995, p. 154) esses aplicativos estão, ou seja, que tanto falta para que deixem de ser percebidos como problemas (PINCH; BIJKER, 1997, p. 44), é possível que as planilhas, pela sua presença no cotidiano

---

<sup>178</sup> CSV (*Comma Separated Value*), XML (*eXtended Markup Language*) e JSON (*JavaScript Object Notation*) são formatos de arquivos de texto que podem ser usados para intercâmbio e armazenamento de dados.



fora do jornalismo, estejam mais próximas da estabilização do que outras tecnologias usadas no JGD. Em oficina ministrada durante o Hackaton na Universidade Positivo, nenhuma das pessoas que participou do exercício de expressões regulares conhecia o Notepad++ – aplicativo similar ao TextWrangler –, apesar de terem contato com técnicas de JGD. Algumas pessoas estranharam, por exemplo, o fato de não poderem usar negrito ou itálico no texto. Mas as pessoas identificaram o uso de funções de busca e substituição de texto, talvez devido à semelhança com editores de texto já conhecidos por elas.

Outro sistema que permite vários modos de uso é o Tableau, citado por Renata e amplamente utilizado por Guilherme, como nos exemplos das seções 4.4.3 e 4.4.6. O programa disponibiliza uma grande variedade de recursos e permite usar dados de várias fontes, além de imagens e arquivos *shapefiles* – arquivos de informação geográfica. Pelas várias possibilidades de uso que ele oferece, pode ser considerado um sistema mais aberto (SPINUZZI, 2003, p. 205). Outro programa citado – por Renata – como facilitador na publicação de dados online sem a necessidade do aprendizado de HTML e Javascript é o infogr.am. Ele é específico para a realização e publicação de gráficos, sem muitos recursos para exploração dos dados. Talvez por ter sido projetado para um uso mais específico, ele não oferece tantos recursos como outros programas e nem tantas possibilidades para quem quiser usá-lo de outras maneiras, sendo assim um sistema mais fechado (SPINUZZI, 2003, p. 205).

João [FREY, 2014] usa o Tableau com um objetivo similar ao uso que Renata faz do infogr.am, que é conseguir visualizações diferentes das que conseguem produzir no Excel, que são “batidas” [FREY, 2014]. Essa preocupação com o resultado e a visualização como produto final também é visível nos exercícios que os professores Träsel e Fotios propuseram aos alunos.

As manifestações do código técnico do Tableau incluem a exigência de que as visualizações nele feitas fiquem armazenadas nos servidores próprios do Tableau, exigindo conexão à internet. Por um lado, isso exige que quem quiser disponibilizar as informações de outra maneira deverá contornar essa dificuldade –

Guilherme, por exemplo, faz uso do menu de impressão do navegador, que permite salvar a “impressão” em arquivos PDF, como fez no exemplo da seção 4.4.3. Já João, como divulga suas pesquisas no Facebook, fica limitado a usar uma imagem estática da visualização – em manifestações do código técnico presentes no suporte final, da mesma maneira que uma visualização em papel propicia interações e recursos diferentes das da internet, por exemplo.

Aceitar as condições impostas pelo Tableau também implica em não poder esconder as informações utilizadas na visualização, que podem ser acessadas a partir da própria visualização. Somente quem usa a versão paga do Tableau tem direito a não disponibilizar as informações.

É possível pensar que os artefatos, ao estarem inseridos em um contexto de uso por pessoas, têm ao mesmo tempo faces táticas e estratégicas. Quando as ferramentas são usadas de maneiras mais próximas para o que foram projetadas – ou para o que supostamente servem –, como fazer uso das planilhas para fazer cálculos, pode haver uma percepção, por parte do usuário, de estar seguindo regras – o que não quer dizer que não haja apropriação e negociação do código técnico.

A “aceitação” desse modo de uso proposto – ou imposto, no caso de sistemas que tenham sido assim projetados – por quem projetou as ferramentas pode ser contraposta à combinação de várias dessas ferramentas. As pessoas entrevistadas e/ou observadas por um lado fazem um uso maior de ferramentas às quais estão mais acostumadas, e por outro fazem uso de outras ferramentas para uso muito específico. Renata, por exemplo, combina o uso de várias ferramentas para extrair dados de PDFs, já que a maneira como os PDFs foram gerados pode variar e o resultado pode ser melhor em um ou outro aplicativo. Guilherme, quando faz mapas usando o QGIS, geralmente usa programas de ilustração vetorial como Adobe Illustrator ou Macromedia Freehand para prepará-los para serem impressos. João usa o Microsoft Access apenas para retirar os dados que lhe interessam das bases de dados como o Finbra, e também usa aplicativos de ilustração para fazer gráficos que os outros aplicativos não fazem –

como disse Yau (2011, CAPÍTULO 3), “Se não há um botão pro gráfico que você quer, azar o seu.” Mesmo que uma ferramenta tenha um certo recurso, é possível usar outra ferramenta que execute a mesma tarefa de uma maneira mais fácil para quem a usa.

Outro tipo de negociação do código técnico pode ser percebida no uso que José e João fazem das fontes dos seus dados. O Diário Oficial do Estado é divulgado em formato PDF, possivelmente com a intenção de oferecer um produto assemelhado a um diário impresso, mas que tem consequências como a dificuldade de fazer operações com o texto do diário. A negociação de José inclui se aproveitar da paginação característica dos arquivos PDF, já que José escolhe quais partes do diário ele deve ler a partir da quantidade de páginas de cada seção, que ele calcula na leitura do sumário.

A forma que José lê os diários oficiais também é um modo de uso. Para encontrar o tipo de informação que procura, José não faz uma leitura completa do texto mas sim uma leitura rápida, buscando palavras-chave e navegando pelo texto de várias maneiras, como as citadas por Tufte (1990, p. 50), que incluem selecionar, isolar, estruturar, observar e “separar o joio do trigo”. Essas diferentes maneiras de ler podem ser consideradas táticas, se considerado que o texto dos diários realiza o código técnico de diversas maneiras, como através do uso de linguagem jurídica – a qual José conhece – ou pelo desenho e tipografia das páginas. O uso que José faz do navegador Chrome para visualizar em quais partes do documento há ocorrências das palavras que ele procura também pode ser considerado um modo de uso próprio dele, já que ele se aproveita, de maneira pouco comum, de um recurso de um aplicativo para atingir seus objetivos.

O modo de uso feito por João dos dados disponíveis na base de dados Finbra, descrita por ele como “flanar” pelos dados, também se relaciona com as diferentes capacidades de encarar informação enumeradas por Tufte (1990, p. 50). Assim como José navega pelas informações dos diários em busca de palavras que chamem sua atenção, João seleciona dados que lhe parecem interessantes, e os processa de várias maneiras para interpretá-los novamente e descobrir informações de interesse. O Finbra é distribuído em

formato do banco de dados do Microsoft Access, talvez com a intenção de torná-lo disponível a um maior número de pessoas, apesar de ser um formato ligado a um aplicativo comercial, de código fechado. Essa decisão implica em manifestações do código técnico que podem ser negociadas de várias maneiras. No caso de João, a negociação passa por fazer uso do aplicativo Access para uma primeira aproximação aos dados e retirar aqueles que ele considere interessantes para então usar uma ferramenta com a qual ele tem mais familiaridade. Na planilha Excel, João faz uso de recursos que facilitam a exploração dos dados por ele selecionados, usando formatações automáticas para separar melhor as linhas, e filtros que permitem a ele reorganizar a informação para evidenciar tendências e agrupamentos, e fórmulas para calcular médias e obter os valores máximos e mínimos de um conjunto de dados.

#### 5.4 PROGRAMAÇÃO

*Guilherme Storck*

Guilherme não vê o aprendizado de programação como algo muito necessário:

Desde que eu comecei a mexer com dados abertos, visualização de dados e jornalismo de dados meio que virou moda falar sobre programação. E eu pesquisei algumas coisas, quis aprender Javascript, Python, algumas outras coisas. Mas basicamente todos os problemas que eu tive até hoje, e que foram muitos, para resolver com dados, eu consegui resolver de outras formas. Eu acho que se você sabe uma linguagem, se você sabe Javascript, sabe Python, você realiza as tarefas mais rápido, mas nenhuma delas resolve todos seus problemas. Você vai ter, que nem eu comentei no processo, que usar outras ferramentas. [...] Em três ou quatro anos que eu trabalho com isso, não tive tempo de aprender a mexer com programação e eu consegui fazer de outro jeito, tipo, não é uma coisa que eu digo nossa, alguém precisa muito saber fazer isso. É legal se você souber, mas não considero essencial.

Na opinião de Guilherme, o processo que mais se beneficia da programação, pela falta de ferramentas prontas, é a raspagem de dados. Quanto ao uso de programação na etapa de visualização, o uso de ferramentas prontas seria mais prático:

a parte de programação que é mais ligada à visualização, do tipo, ah, vou fazer um mapa de bolhas e vou fazer em Javascript... até

você aprender Javascript e conseguir fazer um negócio decente, você perdeu muito mais tempo do que alguém que já sabe e que faria em muito pouco tempo. Você perderia muito tempo pra fazer uma porcaria. E que pra isso tem algumas soluções mais prontas que você não vai precisar perder seu tempo pra fazer uma coisa que vai ficar pior.

### *Renata Silva Pinto*

Renata tomou contato com as possibilidades do uso de programação no jornalismo junto ao Núcleo de Jornalismo Investigativo da Universidade Positivo. Por ocasião de um evento, ela conheceu um programador, não envolvido diretamente com jornalismo, que afirmou que poderia resolver um problema que o núcleo tinha:

A gente falou da nossa matéria que tava com bastante coisa, que precisava pegar tabela do site do governo só que tinha que entrar em 50 mil sites e o cara falou “não, eu faço isso aqui em uma hora, meia hora e já mando pra vocês” e daí foi... mágico.

Apesar dela conhecer gente que programa, ela “não tinha ideia” que era possível usar esses recursos no jornalismo. “Caiu uma pedra, uma bigorna na minha cabeça”, afirma. Essa “bigorna” fez com que Renata começasse a focar no aprendizado de técnicas, para “aprender bem Jornalismo de Dados e programação pra eu conseguir produzir matérias mais completas e... ter [essas habilidades] como diferencial profissional.”

Em outros eventos e cursos, como na Conferência Global de Jornalismo, no Rio de Janeiro, ela viu “que mais pessoas já perceberam isso, e eu achei muito legal, que é um grupo [...] com pessoal que ensina bastante”, se referindo à comunidade que lida com Jornalismo Guiado por Dados. Nesses eventos, ela viu que

o pessoal gosta bastante de programação, mas eu vi que você não precisa aprender tudo de programação, porque tem opções. Tem como você aprender programação pra você fazer uma raspagem, mineração de dados, e tem como você usar programação pra você fazer aplicativos de notícias. Meu objetivo mais próximo é fazer um scraping legal.

Ela também vê possibilidades na combinação de ferramentas, como o uso de Excel junto à programação, “um em auxílio do outro, sabe? Pegar a informação, e daí passar pro Excel.” Ou ainda,

recursos da linguagem R: “acho que muita coisa que eu fico brincando de *pivotable*<sup>179</sup> eu ia colocar dois comandos e *páh!*” Ou também usar o site *infogr.am* para produzir gráficos, já que no Excel “não é tão bonito”.

### *Livre.jor*

Quando questionados sobre o uso ou intenções de usar técnicas de programação, José fala que “programação é o nosso desafio...”. João diz que “nem sei como começar, não sei nem como funciona a programação. O máximo de programação que eu fiz... [foram] algumas fórmulas mais avançadas de Excel.”

José narra que tentou aprender a programar através de cursos online: “cheguei a fazer alguns cursos online do Codeacademy e... não serviu pra nada. Eu não consegui nada.” José percebeu que a programação poderia ser útil quando tomou contato, durante um evento, com bases de dados obtidas por raspagens usando scripts em Python. “Aí falei: tenho que aprender essa porra!” Ele tentou o curso do Codeacademy, mas acabou desistindo, talvez pela demora para aprender os conceitos básicos:

Eu não consegui perceber como iria conectar aquilo [o aprendizado de noções básicas de programação em Python] mais pra frente naquele modelo de aula. Acho que a didática do Codeacademy ainda não me convenceu a persistir nela. Não sei se ela funciona ou não.

[...]

Eu quero fazer robô, me ensina a fazer robô, do zero, depois eu aprendo o resto. Acho que não houve uma conciliação entre as minhas perspectivas muito imediatas com o tempo que eu demoraria pra [cumprí-las]. Sou muito mais a fim de pagar milão prum cara fazer um robô pra mim do que eu aprender.

A aposta de José estaria mais ligada à criação de equipes. João cita a ideia de programador-jornalista em vez do jornalista que aprendesse a programar, participando de uma equipe com mais integrantes. Já para o *Livre.jor*, José considera que o ideal seria “um programador porra louca. Programador é uma parte política do *Livre.jor*, precisaria de um programador que estivesse sintonizado”.

---

<sup>179</sup> Termo em inglês que, no contexto, pode ser traduzido como “tabela dinâmica”.

#### 5.4.1 Considerações sobre os modos de uso de programação

Guilherme, Renata e José concebem a programação como uma ferramenta poderosa que permite a realização de tarefas específicas. Eles relatam que o aprendizado de programação está sendo discutido em congressos e eventos como uma habilidade necessária para jornalistas que queiram trabalhar com JGD. Estimulados por essas discussões, os três tentaram se aproximar, mas encontraram diferentes dificuldades.

Guilherme não considerou que o tempo que levaria para se familiarizar com uma linguagem de programação valeria a pena, já que ele consegue realizar as tarefas usando ferramentas prontas e combinações entre elas, e no dia-a-dia do seu trabalho não tem tempo para se dedicar a esse aprendizado. Renata também vai usando as ferramentas que domina, mas encara o aprendizado de programação como uma maneira de estar preparada para futuras situações.

No caso de José, ele desistiu de aprender esses conceitos básicos, por não ver como essas primeiras noções ensinadas serviriam a seus objetivos. Talvez o caminho não estivesse explicitado, já que o curso era mais geral, sem foco em jornalismo. Já no grupo de estudos de Data Scraping da Universidade Positivo, notou-se uma predisposição das alunas e dos alunos para aprender conceitos básicos como variáveis e condicionais, já que era explicitado que esses conceitos eram necessários para aprender a usar programação para fazer uma raspagem de dados.

Pensando a programação em tanto que sistema aberto ou fechado, nos termos de Spinuzzi (2003, p. 204), é possível considerá-la um sistema mais aberto, ou seja, que permite o uso por parte de pessoas com diferentes graus de conhecimento. É possível que haja pessoas que façam uso da programação através do aproveitamento de códigos prontos, por exemplo, assim como pessoas com mais conhecimento, que façam uso de recursos mais avançados.

Por outro lado, o deslumbamento descrito pelos entrevistados – principalmente por Renata e José – pode indicar que o uso de programação no jornalismo ainda não alcançou o estágio de

estabilização (FEENBERG, 1995, P. 156), ou seja, seu código técnico ainda vai ser negociado, modificado e/ou aceito. Considera-se que, neste caso, as negociações do código técnico teriam como consequência uma concepção da programação como um sistema que permite apropriações, inclusive por parte de pessoas que não se considerem capazes disso, validando todo tipo de tentativa de apropriação e modos de uso.

É possível que a dificuldade das pessoas de se aproximarem da programação esteja na pouca semelhança com ferramentas mais conhecidas. O Tableau, por exemplo, faz uso de planilhas semelhantes às do Microsoft Excel, no que seria uma manifestação do código técnico com vistas à facilidade de uso. Outro aspecto do código técnico seria a menor familiaridade com outros aplicativos que os ambientes de programação têm, e a impossibilidade de “fuçar” da mesma maneira com que as pessoas entrevistadas fazem com outros aplicativos. Apesar que a programação permite vários modos de uso, as pessoas entrevistadas encontram dificuldades para começar a usá-la com frequência.

Essas dificuldades podem ser classificadas dentro das quatro categorias enumeradas por Cibele Reis, Henrique Sarmiento e Vinicius Zaramella (2014, P. 12): as dificuldades relacionadas ao aprendizado, às metodologias de ensino, à escolha da linguagem de programação e ao chamado “pensamento computacional”. Este último, segundo Jeannette Wing (*apud* REIS; SARMENTO; ZARAMELLA, 2014, P. 12)<sup>180</sup>, não é a habilidade de programar, e sim a capacidade de abstração para decompor e reformular problemas, chegando assim a soluções mais simples. Sendo assim, segundo Wing, o pensamento computacional não deveria ser exclusivo dos cientistas da computação.

O pensamento computacional também considera as maneiras através das quais as pessoas resolvem problemas, envolvendo outras áreas do conhecimento, inclusive para a análise de dados. Se por um lado o conceito de pensamento computacional pode envolver áreas que seriam parte de identidades do jornalismo mais

---

<sup>180</sup> WING, Jeannette M. **Computational Thinking and Thinking About Computing.** *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, The Royal Society, v. 366, n. 1881, p. 3717–3725, 2008.



tradicionais, por outra inclui elementos que não fariam parte dessas identidades, como lógica e matemática. As dificuldades citadas por Reis, Sarmiento e Zaramella serviriam não só para considerar as dificuldades do aprendizado de programação por parte de jornalistas, mas também o aprendizado de outras técnicas e ferramentas do JGD e das visualizações.

As dificuldades encontradas pelas pessoas entrevistadas podem não se restringir à programação, mas também a outras ferramentas que exijam um treinamento prévio ou alguma outra característica que dificulte sua apropriação, como no exemplo citado por Yau (2011, CAPÍTULO 3) de um programa cujos “processos para fazer as coisas eram tão contraintuitivos que me fizeram não querer aprender mais.”

## 5.5 APRENDIZADO

### *Guilherme Storck*

Guilherme tem um perfil “meio autodidata”, que ele define como o “esforço em buscar informação e ir aplicando ela, fazendo tentativa e erro até chegar [a um resultado].” Normalmente ele procura esse aprendizado quando precisa resolver uma tarefa específica. A partir de uma tarefa ele divide em passos e procura a solução para cada passo.

Vou fazendo buscas sobre alguém que explique como fazer isso. Então uma coisa que eu percebo nesse processo inteiro é que é muito bom você aprender a fazer buscas. Porque não é muito fácil “ah, como eu faço isso”, mas “isso” é difícil de explicar às vezes. E às vezes as pessoas têm outros termos, elas dizem de outro jeito, mas até você achar a informação que você quer, demora um pouco... é bom aprender a fazer buscas.

Guilherme cita a documentação dos processos como de importância para poder refazer tarefas que já foram feitas, ou para futura referência:

durante o processo eu vou anotando: joguei no navegador, tentei baixar e não deu certo por causa disso; tentei limpar o negócio e usei esse site. Essa anotação foi sempre importante, tanto pra eu ter noção de onde que eu estou, tanto como para quando eu precisar fazer o negócio de novo; eu preciso voltar lá e fazer a mesma tarefa e que bom que eu encontro uma anotação velha.

Guilherme decidiu compartilhar essas “anotações velhas” no blog “Dados Finos”<sup>181</sup>, no qual publica tutoriais para realizar diversas tarefas, fazendo uso de uma ou várias ferramentas.

### *Renata Silva Pinto*

Depois de se aprofundar no Excel, Renata quer ver se entra “de vez com a programação.” Ela já aprendeu “um pouquinho de Python, um pouquinho de Ruby”. Ela cita que aprendeu Python durante uma conferência da Necir<sup>182</sup>, ocasião que aproveitou para entrar em contato com pessoas do jornal *The New York Times* e do *ProPublica*<sup>183</sup>. Segundo ela, havia cursos introdutórios e avançados, que ela considerou “avançados demais”. Ela cita que nos cursos introdutórios os instrutores “passam bastante programas legais.”

Ela considera que cursos formais, como o que ela fez de Excel, são importantes, “porque o Excel tem muita função, não tem como você aprender assim na pira.” Além desse curso e dos encontros de jornalismo, Renata cita plataformas de aprendizado online. Ela começou vários Moocs<sup>184</sup>, mas acabou desistindo: “não consegui terminar assim. Me cadastrei em quatro cursos de graça e eu não não consegui manter o ritmo.” Para o aprendizado pela internet, ela prefere a plataforma Codecademy, na qual pode fazer exercícios mais curtos:

comecei a brincar lá, é bem didático. Daí você começa, aquela coisa, “Hello World”. Aí você vai um pouquinho, ah, faz string, chama add 3, soma a com b dá 4, aí vai. Não precisa separar uma hora pra fazer, pode entrar, se exercitar um pouco e sair, que o curso estará lá quando voltar.

Para aprender, ela ainda cita recursos online, como a lista de email do Necir – que usa para tirar dúvidas específicas –, um grupo online de usuários de Python, além do “nosso querido amigo Google.” Segundo ela, há um senso comunitário entre quem trabalha com Jornalismo Guiado por Dados: “gosto bastante dessa comunidade jornalística de dados [...] o pessoal é super aberto [...] elas não se importam em passar os dados e dicas e tudo o mais.”

<sup>181</sup> Disponível em: <<http://dadosfinos.blogspot.com>>. Acesso em 1º/4/2015.

<sup>182</sup> New English Center of Investigative Reporting, da Universidade de Boston.

<sup>183</sup> Agência de notícias sem fins lucrativos, com sede em Nova Iorque.

<sup>184</sup> Massive Open Online Course (Curso Online Aberto e Massivo).

### 5.5.1 Considerações sobre maneiras de aprendizado

Também é possível pensar em modos de uso para procurar informação que ajude na execução de tarefas – ou seja, para aprender novos modos de uso a partir de soluções descritas por outras pessoas. As pessoas entrevistadas narraram várias maneiras de aprender novas ferramentas. Algumas preferem aprender de maneiras mais estruturadas, como através de cursos, mas na maioria é perceptível uma preferência por buscar ajuda para solucionar casos mais específicos, talvez pelas diferenças entre as maneiras de trabalhar que diferentes profissionais têm, talvez por se depararem com problemas que podem ser resolvidos de diferentes maneiras.

Aqui se faz presente a figura que Bonnie Nardi (APUD SPINUZZI, 2003, p. 207)<sup>185</sup> apelidou de “jardineiro”: pessoas que (re)distribuem conhecimento “informal” através de conexões também “informais”. Renata, em seu papel de “garota do Excel”, é fonte de informações para colegas. Guilherme, através de seu blog, disponibiliza tutoriais a partir de tarefas que teve que realizar. A internet, especialmente “nosso querido amigo Google”, nas palavras de Renata, também é citada como fonte desses conhecimentos.

Assim como De Certeau (2007, p. 47) fala da euforia de marcar uma presença autoral em uma prática cotidiana, como é conseguir realizar uma tarefa através da combinação de elementos heterogêneos, Bonnie Nardi (APUD SPINUZZI, 2003, p. 206) fala no orgulho de compilar e divulgar essas soluções. Seria possível pensar nessa disseminação de conhecimentos como táticas, como maneiras de negociar – ou até resistir – o código técnico de formas de aprendizado mais estruturadas ou “tradicionais”, como cursos e seminários. Quando José tentou aprender a linguagem de programação Python, não conseguiu negociar a maneira de aprendizado e acabou desistindo por não ver como esse curso ajudaria ele a atingir seus objetivos, talvez por ser um curso na internet, que não propicia o mesmo tipo de negociação que um curso presencial, como foram as oficinas descritas previamente.

---

<sup>185</sup> NARDI, Bonnie; O'DAY, Vicki. *Information ecologies: Using technology with heart*. Cambridge, MA: MIT Press, 1999. P. 139–151.

Sendo assim, distribuições de conhecimento “informais” parecem ser mais adequadas para as atividades do JGD, considerando a heterogeneidade de modos de uso descritos pelas pessoas entrevistadas e as diversas dificuldades encontradas no cotidiano.

## 5.6 VISUALIZAÇÕES

### *Renata*

Quando questionada sobre como faria um mapa com estados brasileiros pintados em duas cores, Renata disse: “não sei se tem um programa pra isso, não conheço pelo menos.” Ela mencionou também que procuraria alguém que trabalhasse com isso, já que “mexo com os dados, mas gráfico eu sei bem pouco.” Sobre visualizações, ela diz que “é algo pra mais tarde. Porque acho que a maioria das redações ainda tem infografistas; eu acho que ainda dá pra deixar com esse pessoal” já que ela preferiu se focar no nicho do Jornalismo Guiado por Dados. Ela tem a percepção que no campo do Jornalismo Guiado por Dados se fala pouco em visualizações, e cita expressões como “dons artísticos” para se referir à parte de visualização.

### *Livre.jor*

Para publicar os dados, João usa, algumas vezes, o Google Fusion Tables ou o Tableau, que começou a usar para evitar os gráficos do Excel, “que são meio batidos”. José usa o Microsoft Paint para compôr as figuras com trechos dos diários oficiais. Quando questionados sobre o uso de visualização, João afirmou que vê “muito potencial, tem que ter. Já que a gente está buscando uma linguagem nova, que seja completa, não dá pra fazer pela metade.”

### *Como atividades acadêmicas dos professores*

Tanto Fotios como Träsel usaram infográficos em seus trabalhos. Os alunos e alunas de Fotios tiveram que usar gráficos para apresentar os resultados de uma pesquisa; já os de Träsel tiveram que usar um infográfico como complemento de uma matéria desenvolvida com técnicas de Jornalismo Guiado por

Dados. Rosiane trabalha “só com gráficos básicos e mapas”, mas vê a visualização como um recurso fundamental para “identificar possíveis lides<sup>186</sup>. Então tenho tentando puxar mais isso em sala de aula.”

### *Nas oficinas*

Na oficina ministrada na Universidade Positivo, quando terminado o mapeamento feito como exemplo, a professora mostrou as conclusões que poderiam ser tiradas a partir do mapa, como a relação entre esses dados e pobreza e população. Encomendou às alunas e alunos que repetissem o procedimento com os dados de sua escolha, enviassem o mapa resultante e elaborassem uma proposta de pauta a partir dos dados que encontrassem. Durante a semana seguinte, a professora separou algumas das propostas que ela recebeu e que considerou muito boas, três das quais foram apresentadas na seção 4.4.7. De maneira geral, o retorno do exercício mostrou que alunos e alunas conseguiram fazer uso das ferramentas propostas e também das visualizações realizadas.

#### 5.6.1 Considerações sobre modos de uso de visualizações

Se por um lado as entrevistas narram certa proficiência nos modos de uso das tecnologias e ferramentas do JGD, o mesmo não acontece com as tecnologias e ferramentas da visualização de dados. As exceções são o trabalho de Guilherme, que começou a trabalhar com JGD depois de trabalhar com visualizações, e o trabalho de Bob Smith, baseado no uso de mapas. Gráficos são usados como uma consequência da exploração de dados ou como produto final, e não para a exploração dos dados em si, como feito pela professora Rosiane e como feito nas oficinas com alunas e alunos de jornalismo em universidades.

As tecnologias da visualização e a visualização em si ainda estariam começando a serem negociadas e apropriadas por

---

<sup>186</sup> Lide – também usado na forma inglesa *lead* – é o primeiro parágrafo de um texto jornalístico. Quando o texto segue a estrutura de pirâmide invertida (ver página 68), o lide concentra as informações consideradas mais importantes.

jornalistas. Segundo Renata, a visualização de dados não é uma questão muito abordada na comunidade de JGD. Ela atribui isso a que redações e organizações jornalísticas contam com departamentos de infografia, que seriam responsáveis pela visualização e outros tipos de apresentação gráfica de informação. Em algumas situações haveria uma divisão de trabalho entre jornalistas e profissionais que dão forma aos produtos jornalísticos, deixando a visualização fora do escopo do que é ser jornalista, e inclusive do JGD.

Isso leva a considerar que a visualização ainda não teria alcançado o estágio de estabilização, mesmo no contexto do JGD. Sendo assim, ela ainda pode ser percebida como um problema, caso seja considerada parte do JGD. Em algumas situações, ela pode não ser considerada um problema, mas não por ter atingido a estabilização, e sim por não ser percebida como uma atividade própria do JGD.

Acredita-se que haja pessoas que trabalhem com JGD e/ou visualizações que esteja se apropriando e negociando as tecnologias de visualização de dados, e que a visualização de dados como tecnologia, graças às negociações dessas pessoas – que podem ser consideradas um grupo social relevante (PINCH; BIJKER, 1997, P. 30) – esteja cada vez mais presente no JGD, primeiro como problema e depois rumo à estabilização, tanto dentro do JGD como do jornalismo de modo geral.

## **5.7 CONSIDERAÇÕES SOBRE AS ENTREVISTAS E OBSERVAÇÕES**

Partir do conceito de táticas, de modos de uso do cotidiano, implica em conceber que as pessoas fazem uso de elementos heterogêneos para conseguir seus objetivos (DE CERTEAU, 2007, P. 47). Nesse processo, cada pessoa perceberá certos aspectos das tecnologias como problema, como manifestações do código técnico. Considerando as várias ferramentas que podem ser usadas no JGD, citadas pelos entrevistados e compiladas em livros, seria necessário pensar que cada uma delas – ou até aspectos de cada uma – é negociada de maneira diferente por diferentes pessoas,

que vão fazendo uso de uma ou várias ferramentas para tentar solucionar o que elas percebem como problemas.

Nesses processos, algumas tecnologias podem terminar sendo priorizadas, sendo vistas como solução do problema e chegando assim à estabilização. Planilhas eletrônicas, por exemplo, parecem ser uma tecnologia mais estabilizada dentro do JGD, talvez por estarem mais estabilizadas em um contexto mais geral. Já outras tecnologias, como as da programação ou da visualização, ainda têm problemas que precisam deixar de serem percebidos como tal para que essas tecnologias comecem a ser apropriadas de outras maneiras e assim comecem a se estabilizar dentro do JGD. É possível que essas tecnologias sejam percebidas como problemas por serem oriundas de outras áreas, ou percebidas como alheias à atividade jornalística.

A observação dos modos de uso de tecnologias, tanto no nível de escolher qual tecnologia será usada – nível meso – como no ato de clicar, apertar teclas ou ler um documento – nível micro –, revela que cada pessoa pode fazer uso de uma tecnologia de várias maneiras, e que são essas maneiras de usar que dão independência às pessoas para conseguir seus objetivos.





# 6 CONSIDERAÇÕES FINAIS E PERSPECTIVAS

O jornalismo, assim como outras profissões, faz uso de tecnologias que vão sendo modificadas ora constantemente, ora de maneira abrupta. Do telégrafo à comunicação pela internet, da máquina de escrever ao computador, do filme às câmeras digitais. Essas mudanças implicam em novos códigos técnicos, novas apropriações, novos modos de uso, novas táticas, e também diferentes instâncias de estabilização das tecnologias.

O estudo do cotidiano, especificamente das formas de apropriação, colaborou para o entendimento do papel do JGD não como conjunto isolado de tecnologias e métodos, e sim em seu papel dentro de um – ou vários – grupos sociais. A primeira questão de pesquisa, sobre de quais maneiras os modos de uso de tecnologias levam à sua estabilização, levou à reflexão sobre negociações de poder entre táticas e estratégia, entre apropriações cotidianas e códigos técnicos, entre pessoas que usam os sistemas e quem os projeta. É através desse processo que os artefatos alcançam sua estabilização, processo que foi considerado um indicativo de que tão pertencente ao jornalismo é o uso de técnicas de JGD.

Nesses termos, o JGD, como um todo, parece estar próximo da estabilização diante do grupo social de jornalistas. O que ainda parece estar longe da estabilização são as técnicas e os aplicativos utilizados. As pessoas entrevistadas usam combinações próprias de ferramentas, com algumas coincidências, como o uso de planilhas de cálculo. O JGD também tem a imprevisibilidade do jornalismo, e talvez por isso não haja muitas ferramentas ou plataformas destinada ao JGD – a única ferramenta encontrada durante a pesquisa que parece ter sido concebida para o jornalismo é o Datawrapper – e sim sejam apropriados aplicativos

trazidos de outras áreas – como aplicativos de geoprocessamento – ou aplicativos concebidos para usos mais gerais – como planilhas e sistemas gerenciadores de bancos de dados. Sendo assim, cada pessoa que precisar usar técnicas de JGD em seu trabalho vai montar seu próprio sistema a partir das ferramentas que ela saiba usar, ou queira saber usar, ou acredite que seja a melhor para as tarefas que queira realizar – de forma similar a quem quer realizar outras atividades. Essa combinação de elementos tem aspectos de tática, mesmo que nas ferramentas em si esse aspecto não predomine.

Sendo assim, defende-se que a estabilização das ferramentas do JGD é importante para cada pessoa que faz uso delas, sendo também importante em um nível mais macro, no qual a estabilização das tecnologias implica na validação do JGD como um todo, seja como conjunto de técnicas, seja como área, seja como identidade possível dentro da identidade jornalística.

Com o segundo objetivo específico como norte, foi situado o JGD com relação aos gêneros jornalísticos. O questionamento da neutralidade da tecnologia foi relacionado à suposta neutralidade do jornalismo, que pode se mostrar como uma meta do JGD, mas pelos próprios códigos técnicos das tecnologias envolvidas, o que inclui as próprias técnicas jornalísticas. Sendo assim, é necessário alertar para possíveis vieses presentes nos métodos e produtos jornalísticos relacionados ao JGD. Acredita-se que uma visão menos determinista da presença do computar no jornalismo colaboraria para que mais jornalistas pudessem se aproveitar dessas técnicas no dia-a-dia. Outro aspecto da visão determinista que deve ser considerado é a suposta neutralidade dos grandes conjuntos de dados. É necessário pensar nos dados como articulações da realidade e que, como tais, podem não representá-la com a exatidão desejada.

Foi apresentado que artefatos fazem parte da construção de identidades, o que também acontece no jornalismo, inclusive com identidades jornalísticas que supostamente não envolvem o uso de tecnologias. As tecnologias computacionais já fazem parte da identidade jornalística quando consideradas substitutas da máquina de escrever ou dos teletipos, mas ainda há resistência

quando a proposta é fazer uso dessas tecnologias para a apuração de informações, ou talvez quando a proposta é incluir essas tecnologias na identidade do jornalismo. O computador que é usado em técnicas de JGD é visto como “tecnologia”, enquanto o computador usado para escrever e ler notícias não. Essa evidência do computar sugere que ele não alcançou ainda a instância de estabilização, pelo menos quando considerados alguns aspectos dele, mas também o aponta como importante na construção de uma identidade relacionada ao com JGD, ou ainda à identidade jornalística em si.

Em tanto que formato, estima-se que o JGD pode ser usado tanto para a produção de formatos jornalísticos tradicionais – já estabilizados – como em outro tipo de produtos – ainda não estabilizados. Esses “novos” produtos, o que inclui as visualizações, propiciam outro tipo de relações entre jornalismo e quem o consome.

Porém, foi observado que, mesmo entre jornalistas que trabalham com JGD, há uma certa resistência ao uso de visualizações – análoga à resistência dos jornalistas “tradicionais” para com os métodos do JGD. Considerando um cenário proposto no qual jornalistas se beneficiariam de certa autonomia, tal atitude implica na perda das possibilidades que a visualização ofereceria. Situação similar acontece com o aprendizado de programação. Assim como é necessário legitimar a apropriação de tecnologias, seria benéfico estabilizar a visualização como mais um método do JGD e não somente como algo a ser feito para divulgar a informação.

Com as tecnologias usadas no JGD mais próximas do momento de estabilização, elas seriam mais facilmente concebidas como parte do que é ser jornalista. É possível afirmar o mesmo das visualizações. Como exposto na seção 4.2, elas têm várias possibilidades que são próprias do formato, e afins às intenções do Jornalismo Guiado por Dados. Espera-se que este trabalho contribua para que elas também façam parte do repertório jornalístico, tanto para a comunicação como para a descoberta de novas informações latentes em grandes conjuntos de dados.

Por conclusão principal, defende-se que é possível conseguir resultados desejados do Jornalismo Guiado por Dados e das visualizações através da apropriação cotidiana, das táticas presentes nas gambiarras, da negociação com as estratégias. Todas essas formas de uso devem ser legitimadas. Essas formas de uso podem ter limitações quando comparadas a pessoas com conhecimento mais amplo ou a equipes de pessoas com diferentes habilidades trabalhando juntas. Porém, a autonomia que resulta dos processos de apropriação é o que pode dar às e aos jornalistas estímulo para aproveitarem as possibilidades do Jornalismo Guiado por Dados e das visualizações por si sós, com os recursos que têm à mão. Essa proficiência resulta de vitórias cotidianas aparentemente pequenas e da conquista de um espaço que parecia alheio à profissão, e essas conquistas fazem repensar, se não a identidade do jornalismo, as maneiras em que essa identidade se constrói e desconstrói.

# REFERÊNCIAS BIBLIOGRÁFICAS

ALFRED, Randy. Nov. 4, 1952: **Univac Gets Election Right, But CBS Balks**. Wired, Estados Unidos, 4 nov. 2010. Disponível em: <[www.wired.com/2010/11/1104cbs-tv-univac-election/](http://www.wired.com/2010/11/1104cbs-tv-univac-election/)>. Acesso em: 14/5/2014.

ANDERSEN, Michael. Four crowdsourcing lessons from the Guardian's (spectacular) expenses-scandal experiment. **Nieman Lab**, Estados Unidos, 2009. Disponível em: <[www.niemanlab.org/2009/06/four-crowdsourcing-lessons-from-the-guardians-spectacular-expenses-scandal-experiment/](http://www.niemanlab.org/2009/06/four-crowdsourcing-lessons-from-the-guardians-spectacular-expenses-scandal-experiment/)>. Acesso em 21/1/2015.

BARROS JUNIOR, José Lazaro Ferreira. Defensoria revoga avanço funcional e suspende auxílios transporte e alimentação dos defensores. **Livre.jor**, Curitiba, 2014a. Disponível em: <<http://on.fb.me/1ys7gHc>> ou em: <[www.facebook.com/livre.jor/photos/a.426976960770689.1073741828.422619687873083/579702378831479](https://www.facebook.com/livre.jor/photos/a.426976960770689.1073741828.422619687873083/579702378831479)>. Acesso em: 19/1/2015.

BARROS JUNIOR, José Lazaro Ferreira; FREY, João Guilherme Bernardo. **Inscrição 9º prêmio Sangue Bom**. Curitiba, 2014b.

BERTIN, Jacques. **A neográfica e o tratamento gráfico da informação**. Curitiba: Editora da Universidade Federal do Paraná, 1986.

BERTIN, Jacques. **Semiology of Graphics**. Nova Iorque: Esri Press, 2011.

BOSTOCK, Mike; CARTER, Shan; TSE, Archie. Is It Better to Rent or Buy? **The New York Times**, Nova Iorque, 2014. Disponível em: <[www.nytimes.com/interactive/2014/upshot/buy-rent-calculator.html](http://www.nytimes.com/interactive/2014/upshot/buy-rent-calculator.html)>. Acesso em 13/1/2015.

BOUNEGRU, Liliana; CHAMBERS, Lucy; GRAY, Jonathan. **Manual de Jornalismo de Dados**. Disponível em <[datajournalismhandbook.org/pt/index.html](http://datajournalismhandbook.org/pt/index.html)>. Acesso em 17/7/2014.

BOUNEGRU, Liliana. **Jornalismo de dados em perspectiva**. In: BOUNEGRU, Liliana (coord.). Manual de Jornalismo de Dados 1.0. [S.l.]: European Journalism Centre, 2012. Disponível em <[datajournalismhandbook.org/pt/introducao\\_4.html](http://datajournalismhandbook.org/pt/introducao_4.html)>. Acesso em 14/5/2014.

BRADSHAW, Paul. The inverted pyramid of data journalism. **Online Journalism Blog**, Inglaterra, 2011. Disponível em: <<http://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>>. Acesso em 18/6/2015.

BUMGARDNER, Jim. **A year of sunsets** (2006 version). Estados Unidos, 2006. Disponível em: <<https://www.flickr.com/photos/krazydad/292081922/in/set-140323>>. Acesso em 29/10/2014.

CAIRO, Alberto. **Existe jornalismo de dados e visualização no Brasil?** In: BOUNEGRU, Liliana (coord.). Manual de Jornalismo de Dados 1.0. [S.l.]: European Journalism Centre, 2012. Disponível em <[datajournalismhandbook.org/pt/introducao\\_6.html](http://datajournalismhandbook.org/pt/introducao_6.html)>. Acesso em 14/5/2014.

CAIRO, Alberto. **Infografía 2.0: visualización interactiva de información en prensa**. Madrid: Alamut, 2008.

CARTER, S.; GRAHAM R.; WARD, J. **How Mariano Rivera Dominates Hitters**. The New York Times, 29/6/2010. Disponível em <[www.nytimes.com/interactive/2010/06/29/magazine/rivera-pitches.html](http://www.nytimes.com/interactive/2010/06/29/magazine/rivera-pitches.html)>. Acesso em 7/4/2012.

CHURRAS Calculator. Brasil, 28/3/2001. 1 arquivo (20 kilobytes). Microsoft® Excel. Disponível em: <[ftp://unilins.edu.br/formigoni/Utilitarios/churras\\_calculator\\_tabajara.xls](ftp://unilins.edu.br/formigoni/Utilitarios/churras_calculator_tabajara.xls)>. Acesso em: 21/7/2014.

CLARÍN. **Eligieron una infografía de Clarín como la mejor de los últimos 20 años**. Buenos Aires, 2012. Disponível em: <[www.clarin.com/sociedad/Eligieron-infografia-Clarín-mejor-ultimos\\_0\\_677932336.html](http://www.clarin.com/sociedad/Eligieron-infografia-Clarín-mejor-ultimos_0_677932336.html)>. Acesso em 12/11/2012.

COLOMBATO, Juan. La más real de todas las infografías. **Zaveka Infográfica**, Argentina, 2010. Disponível em: <<http://zavekainfografica.blogspot.com.br/2010/10/la-mas-real-de-todas-las-infografias.html/>>. Acesso em 19/1/2015.

COSTA, Lailton; LUCHT, Janine. **Gênero Interpretativo**. In: ASSIS, Francisco de; MARQUES DE MELO, José. Gêneros jornalísticos no Brasil. São Bernardo do Campo: Universidade Metodista de São Paulo, 2010.

COSTA, Lailton. **Gêneros Jornalísticos**. In: ASSIS, Francisco de; MARQUES DE MELO, José. Gêneros jornalísticos no Brasil. São Bernardo do Campo: Universidade Metodista de São Paulo, 2010.

DE CERTEAU, Michel. **L'invention du quotidien 1**. arts de faire. França: Gallimard, 1990.

DE CERTEAU, Michel. **A invenção do cotidiano: 1**. Artes de fazer. Petrópolis: Vozes, 2007.

DE CERTEAU, Michel; GIARD, Luce; MAYOL, Pierre. **A invenção do cotidiano: 2**. Morar, cozinhar. Petrópolis: Vozes, 1996.

DE PABLOS, José Manuel. **Infoperiodismo – el periodista como creador de infografía**. Madri: Editorial Síntesis, 1999.

DEFORGE, Yves *et al.* **Imagen Didáctica**. Barcelona: Ediciones Ceac, 1991.

DELAMUTA, Mariana. **Mapa IBGE**. Mensagem recebida por <[matiasperuyera@gmail.com](mailto:matiasperuyera@gmail.com)> em 12/3/2013.

Du GAY, Paul; HALL, Stuart; MACKAY, Hugh; NEGUS, Keith. **Doing cultural studies: the story of the Sony Walkman**. Londres: Sage Publications; 2003.

EMPEROR TIGERSTAR. **World War II in Europe: Every Day**. Estados Unidos, 2013. Disponível em <<http://youtu.be/WOVEy1tC7nk>>. Acesso em: 11/7/2014.

ESTADÃO DADOS. Basômetro. **Estadão Dados**, 2013. Disponível em: <[estadaodados.com/basometro/](http://estadaodados.com/basometro/)>. Acesso em 21/7/2014.

FARACH, Antonio. **3D World Cup Dataviz Ball, by Times of Oman**. [Omã], 2014. Disponível em: <<http://visualoop.com/22881/3d-world-cup-dataviz-ball-by-times-of-oman>>. Acesso em 15/8/2014.

FAST, Mike. **What the Heck is PITCHf/x?** Estados Unidos, 2010. Disponível em <[baseball.physics.illinois.edu/FastPFXGuide.pdf](http://baseball.physics.illinois.edu/FastPFXGuide.pdf)>. Acesso em 21/7/2014.

FEENBERG, Andrew. **Alternative modernity: the technical turn in philosophy and social theory**. Berkley; Londres: University of California Press, 1995.

FEENBERG, Andrew. **Teoría crítica de la tecnología.** In: Revista CTS, no 5, vol. 2, Junio de 2005 (pág. 109-123) Disponível em <<http://www.scielo.org.ar/pdf/cts/v2n5/v2n5a07.pdf>>. Acesso em 13/1/2015.

FEENBERG, Andrew. **Da informação à comunicação:** a experiência francesa com o videotexto. In: NEDER, Ricardo T. (org.). Andrew Feenberg: racionalização democrática, poder e tecnologia. Brasília: Observatório do Movimento pela Tecnologia Social na América Latina/Centro de Desenvolvimento, 2010.

FÉLIX, Rosana et al. Banco de dados exigiu a leitura de 145 mil páginas. **Gazeta do Povo**, Curitiba, 4/8/2013. Disponível em: <[www.gazetadopovo.com.br/vidaecidadania/crime-sem-castigo/conteudo.phtml?tl=1&id=1396485&tit=Banco-de-dados-exigiu-a-leitura-de-145-mil-paginas](http://www.gazetadopovo.com.br/vidaecidadania/crime-sem-castigo/conteudo.phtml?tl=1&id=1396485&tit=Banco-de-dados-exigiu-a-leitura-de-145-mil-paginas)>. Acesso em 21/7/2014.

FÉLIX, Rosana. Deputados “estouram” gastos em dezembro. **Gazeta do Povo**, Curitiba, 20 de fevereiro de 2011. Caderno Vida Pública.

FIDALGO, António. **A resolução semântica no jornalismo online.** In: BARBOSA, Suzana (org.). Jornalismo Digital de Terceira Geração. Covilhã: Labcom – Universidade da Beira Interior, 2007.

FISCHER, Eric. Making the most detailed tweet map ever. **Mapbox**, Washington DC, 3 de dezembro de 2014. Disponível em: <<https://www.mapbox.com/blog/twitter-map-every-tweet/>>. Acesso em 13/1/2015.

FISCHER, Eric. **See Something or Say Something: Rio de Janeiro.** [San Francisco], 2011. Disponível em <[www.flickr.com/photos/walkingsf/5926353358/](http://www.flickr.com/photos/walkingsf/5926353358/)>. Acesso em 21/7/2014.

FISCHER, Eric. Six billion tweets. **Mapbox**, Washington DC, 3 de dezembro de 2014. Disponível em: <<http://bit.ly/1u2rZl5>>. Acesso em 13/1/2015.

FLUSSER, Vilém. **Ensaio sobre a fotografia para uma filosofia da técnica.** Lisboa: Relógio D'Água Editores, 1998.

FREY, João Guilherme. No Paraná, média de gastos com rodovias dobra em anos eleitorais. **Livre.jor**, Curitiba, 2014b. Disponível em: <<http://on.fb.me/1J8PwnB>> ou em: <[www.facebook.com/livre.jor/photos/pb.422619687873083.-2207520000.1421355370./561922453942805/](http://www.facebook.com/livre.jor/photos/pb.422619687873083.-2207520000.1421355370./561922453942805/)>. Acesso em 19/1/2015.

FRIENDLY, Michael. **Milestones in the history of thematic cartography, statistical graphics, and data visualization.** 2009. Disponível em: <[euclid.psych.yorku.ca/SCS/Gallery/milestone/milestone.pdf](http://euclid.psych.yorku.ca/SCS/Gallery/milestone/milestone.pdf)>. Acesso em 09/06/2012.

FRY, Ben. **Visualizing Data.** Sebastopol: O'Reilly Media, 2008.

FUNG, Kaiser. Return of the barrel. **Junk Charts**, Estados Unidos, 2014. Disponível em: <[http://junkcharts.typepad.com/junk\\_charts/2014/07/the-top-dog-among-jealous-dogs.html](http://junkcharts.typepad.com/junk_charts/2014/07/the-top-dog-among-jealous-dogs.html)>. Acesso em: 20/1/2015.

FUNG, Kaiser. The top dog among jealous dogs. **Junk Charts**, Estados Unidos, 2014. Disponível em: <[http://junkcharts.typepad.com/junk\\_charts/2014/07/return-of-the-barrel.html](http://junkcharts.typepad.com/junk_charts/2014/07/return-of-the-barrel.html)>. Acesso em: 20/1/2015.

GAZETA DO POVO. Ciclistas mapeiam os 15 cruzamentos mais perigosos. **Gazeta do Povo**, Curitiba, 23/10/2013. Disponível em: <[www.gazetadopovo.com.br/vidaecidadania/conteudo.phtml?tl=1&id=1419127](http://www.gazetadopovo.com.br/vidaecidadania/conteudo.phtml?tl=1&id=1419127)>. Acesso em 13/1/2015.

GAZETA DO POVO. Gastos dos deputados estaduais. **Gazeta do Povo**: Curitiba, 2014. Disponível em: <<http://www.gazetadopovo.com.br/dados/vidapublica/gastos>>. Acesso em 16/8/2014.

GAZETA DO POVO. Mapa dos lançamentos imobiliários de Curitiba. **Gazeta do Povo**, Curitiba, 2011. Disponível em: <[www.gazetadopovo.com.br/perfilimobiliario-2011/lançamentos](http://www.gazetadopovo.com.br/perfilimobiliario-2011/lançamentos)>. Acesso em 13/1/2015.

GONZÁLEZ VEIRAS, Xaquín. **Cómo se hizo: los lanzamientos de Mariano Rivera**. Disponível em <[www.xocas.com/blog/?p=758](http://www.xocas.com/blog/?p=758)>. Acesso em 7/4/2012.

HOLMES, David. **Should journalists learn to code?** I made a flowchart. <http://pndo.ly/HfkjEU> cc @webjournalist @jbenton. Estados Unidos, 2013. Twitter: @holmesdm. Disponível em: <<https://twitter.com/holmesdm/status/393090781038125056/photo/1>>. Acesso em: 15/8/2014.

HOLOVATY, Adrian. **A fundamental way newspaper sites need to change**. Adrian Holovaty, Estados Unidos, 6 set. 2006. Disponível em: <[www.holovaty.com/writing/fundamental-change](http://www.holovaty.com/writing/fundamental-change)>. Acesso em: 14/5/2014.

HOWARD, Alexander Benjamin. **Data-Driven Journalism**. Estados Unidos: Columbia Journalism School, 2014.

JANNUZI, Lyn; MIRANDA, Fabiano. Mais agressivo. **Gazeta do Povo**, Curitiba, 29 de junho de 2010. Caderno Esportiva.

JOBIM E SOUZA, Solange. **Dialogismo e alteridade na utilização da imagem técnica em pesquisa acadêmica**: questões éticas e metodológicas. In: FREITAS, Maria Teresa; JOBIM E SOUZA, Solange, KRAMER, Sonia (orgs.). Ciências Humanas e Pesquisa: a leitura de Mikhail Bakhtin. São Paulo: Cortez, 2003.

KAYSER-BRIL, Nicolas. **Familiarizando-se com os dados em três passos**. In: BOUNEGRU, Liliana (coord.). Manual de Jornalismo de Dados 1.0. [S.l.]: European Journalism Centre, 2012. Disponível em <[datajournalismhandbook.org/pt/entendendo\\_os\\_dados\\_0.html](http://datajournalismhandbook.org/pt/entendendo_os_dados_0.html)>. Acesso em 14/5/2014.

KREBS, Albin. **Truman Capote Is Dead at 59; Novelist of Style and Clarity**. The New York Times, Nova Iorque, 28 de agosto de 1984. Disponível em: <[www.nytimes.com/books/97/12/28/home/capote-obit.html](http://www.nytimes.com/books/97/12/28/home/capote-obit.html)>. Acesso em 8/4/2015.

KUNCZIK, Michael. **Conceitos de jornalismo**: Norte e Sul. São Paulo: Edusp, 2002.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Técnicas de pesquisa**: planejamento e execução de pesquisas, amostragens e técnicas de pesquisa, elaboração, análise e interpretação de dados. São Paulo, Atlas, 1985 (edição 2008, reimpressão 2012).

LIBREOFFICE. **Guia Calc**. 2010. Disponível em: <[https://wiki.documentfoundation.org/images/1/11/0301CG3-Introducao\\_ao\\_Calc\\_ptbr.pdf](https://wiki.documentfoundation.org/images/1/11/0301CG3-Introducao_ao_Calc_ptbr.pdf)>. Acesso em 18/6/2015.

MACKAY, Hugh. Consuming Communication Technologies at Home. In: MACKAY, Hugh (ed.). **Consumption and Everyday Life**. Londres: Sage, 1997.

MACKAY, Hugh. Introduction. In: MACKAY, Hugh (ed.). **Consumption and Everyday Life**. Londres: Sage, 1997.



MAIER, Scott R. **The Digital Watchdog's First Byte: Journalisms First Computer Analysis of Public Records.** In: American Journalism Vol. 17 Ed. 4. Estados Unidos: American Journalism Historians Association, 2000. Disponível em: <archive.org/details/americanjournali17amer. Acesso em: 14/5/2014.

MANOVICH, Lev. **O que é visualização?** In: Estudos em Jornalismo e Mídia - Vol. 8 No 1. Florianópolis, 2011. Disponível em <www.periodicos.ufsc.br/index.php/jornalismo/article/view/1984-6924.2010v8n1p146/18947>. Acesso em 11/7/2013.

MANUAL DA REDAÇÃO: Folha de S. Paulo. São Paulo: Publifolha, 2010.

MARTÍN-BARBERO, Jesús. **Dos Meios às Mediações: comunicação, cultura e hegemonia.** Rio de Janeiro: Editora UFRJ, 2009.

MENDONÇA, André Luiz Alencar. **Avaliação de interfaces cartográficas para dispositivos com tela sensível ao multitoque.** 2013. 269 f. Tese (doutorado em cartografia) – Universidade Federal do Paraná, Curitiba, 2013.

MEYER, Philip. **Precision Journalism.** Estados Unidos: Rowman & Littlefield, 2002.

MEYER, Philip. **The Future of CAR: Declare Victory and Get Out!** In: When Nerds and Words Collide: reflections on the development of Computer Assisted Reporting. São Petersburgo: The Poynter Institute for Media Studies, 1999. Disponível em: <www.unc.edu/~pmeyer/carfuture.doc>. Acesso em: 14/5/2014.

MUNROE, Randall. **XKCD: a webcomic.** Disponível em: <http://xkcd.com/974/>. Estados Unidos, 2011. Acesso em: 29/1/2015.

NERDS FOR NATURE. Crowd-sourced Time Lapse Photography. **Instructables**, Estados Unidos, 2013. Disponível em: <http://www.instructables.com/id/Crowd-sourced-Time-Lapse-Photography/>. Acesso em 21/1/2015.

NERDS FOR NATURE. Mt. Diablo Morgan Fire. **Nerds for Nature**, Estados Unidos, 2013. Disponível em: <http://nerdsfornature.org/monitor-change/index.html>. Acesso em 21/1/2015.

NICHELE, Larissa. **Mapa Emigrantes.** Mensagem recebida por <matiasperuyera@gmail.com> em 12/3/2013.

NIELSEN, Gert. **The Death of Datavisualization in the News.** Dinamarca, 2011. Disponível em <visualjournalism.com/the-death-of-datavisualization-in-the-news/2011/03/31/>. Acesso em 21/7/2014.

ODPPER, Fernanda Cercal. **Exercício batchgeo.** Mensagem recebida por <matiasperuyera@gmail.com> em 18/3/2013.

OLINDA, Caroline. Portal da Assembleia não facilita análise de dados. **Gazeta do Povo**, Curitiba, 5/4/2011. Vida Pública, p. 13.

OSÓRIO, Moreno. Jornalistas devem saber programar? **Farol Jornalismo**, Rio Grande do Sul, 2014. Disponível em: <http://faroljornalismo.cc/blog/2014/02/10/jornalistas-devem-saber-programar/>. Acesso em 15/8/2014.

PADILHA, José. **Tropa de Elite.** Brasil: Universal Pictures, 2007. 118 min.

PARASIE, Sylvain; DAGIRAL, Eric. **Data-Driven Journalism and the Public Good**. Paris, 2012. Disponível em: <[https://sympar.files.wordpress.com/2011/05/parasie-dagiral-2012\\_data-driven-journalism\\_nms\\_preprint.pdf](https://sympar.files.wordpress.com/2011/05/parasie-dagiral-2012_data-driven-journalism_nms_preprint.pdf)>. Acesso em 13/1/2015.

PARASIE, Sylvain. **'Hacker' journalism – A new utopia for the press?** Paris, 2011. Disponível em: <[http://sympar.files.wordpress.com/2011/05/20111012\\_hacker-journalism.pdf](http://sympar.files.wordpress.com/2011/05/20111012_hacker-journalism.pdf)>. Acesso em 13/1/2015.

PELTZER, Gonzalo. **Jornalismo Iconográfico**. Lisboa, Portugal: Planeta Editora, 1991.

PELTZER, Gonzalo. Periodismo hay uno solo y los datos son la realidad. **Paper Papers**, [Buenos Aires], 16/4/2013. Disponível em <<http://www.paperpapers.net/2013/04/periodismo-hay-uno-solo-y-los-datos-son.html>>. Acesso em 19/8/2014.

PENA, Felipe. **Teoria do jornalismo**. São Paulo: Contexto, 2008.

PEÑARANDA U., Raúl. **Géneros periodísticos: ¿Qué son y para qué sirven?** 2000. Disponível em: <<http://www.saladeprensa.org/art180.htm>>. Acesso em 29/10/2014.

PERUYERA, Matias. **“Every Picture Tells a Story”: visualizações de dados e seu uso no jornalismo**. 2012. 66 f. Trabalho de Conclusão de Curso (Comunicação Social – Jornalismo) – Centro Universitário Uninter, 2012. Disponível em: <[www.academia.edu/7733857/Visualizacoes\\_de\\_dados\\_e\\_seu\\_uso\\_no\\_jornalismo](http://www.academia.edu/7733857/Visualizacoes_de_dados_e_seu_uso_no_jornalismo)>. Acesso em 21/7/2014.

PERUYERA, Matias. **Visualizações De Dados Disponíveis Na Internet e Novas Possibilidades Na Comunicação**. In: Congresso de Ciências da Comunicação na Região Sul, XI, 2010, Novo Hamburgo. Anais eletrônicos... Novo Hamburgo: Intercom, 2010. Disponível em: <[www.intercom.org.br/papers/regionais/sul2010/resumos/R20-0317-1.pdf](http://www.intercom.org.br/papers/regionais/sul2010/resumos/R20-0317-1.pdf)>. Acesso em 21/7/2014.

PERUYERA, Matias. **Visualizações de Dados na Infografia Jornalística e seu uso na Copa do Mundo pela Gazeta do Povo**. In: Congresso Brasileiro de Ciências da Comunicação, XXXIII, 2010, Caxias do Sul. Anais eletrônicos... Caxias do Sul: Intercom, 2010. Disponível em: <[www.intercom.org.br/sis/2010/resumos/R5-2236-1.pdf](http://www.intercom.org.br/sis/2010/resumos/R5-2236-1.pdf)>. Acesso em 21/7/2014.

PICARD, Robert G. **Measuring and Interpreting Productivity of Journalists**. In: Newspaper Research Journal, Vol. 19, Nº 4. Estados Unidos: Association for Education in Journalism and Mass Communication, 1998. Disponível em: <<https://www.academia.edu/3632528>>. Acesso em 27/4/2015.

PINCH, Trevor; BIJKER, Wiebe. **The social construction of facts and artifacts: or how the Sociology of Science and the Sociology of Technology might benefit each other**. In: BIJKER, Wiebe; HUGHES, Thomas; PINCH, Trevor (editores). *The Social Construction of Technological Systems*. Cambridge, EUA : MIT Press, 1997.

PIQUÉ, Antoni. Esto NO es periodismo. **Paper Papers**, [Barcelona], 18/11/2009. Disponível em <<http://www.paperpapers.net/2009/11/esto-no-es-periodismo.html>>. Acesso em 19/8/2014.

PNUD. **Atlas do Desenvolvimento Humano no Brasil**. Programa das Nações Unidas para o Desenvolvimento: Brasil, 2013. Disponível em: <[www.atlas2013.org.br](http://www.atlas2013.org.br)>. Acesso em 21/8/2014.

QUEALY, Kevin; ROBERTS, Graham. **London Olympics 2012 – Usain Bolt's Gold in the 100 Meter Sprint**. The New York Times, Nova Iorque, 8 ago. 2012. Disponível em <[www.youtube.com/watch?v=\\_L\\_vq5JYQlE](http://www.youtube.com/watch?v=_L_vq5JYQlE)>. Acesso em: 14/5/2014.

RÊGO, Ana Regina; AMPHILO, Maria Isabel. **Gênero opinativo**. In: ASSIS, Francisco de; MARQUES DE MELO, José. Gêneros jornalísticos no Brasil. São Bernardo do Campo: Universidade Metodista de São Paulo, 2010.

REIS, Cibele; SARMENTO, Henrique; ZARAMELLA, Vinicius. **Ferramenta de auxílio ao desenvolvimento do pensamento computacional**: uma plataforma robótica controlada por smartphone. 113 f. Trabalho de Conclusão de Curso – Departamento Acadêmico de Informática, Universidade Tecnológica Federal do Paraná. Curitiba, 2014.

SAMPAIO, Dafne. **“no meio da redação do @Estadao instalaram uma grande lâmpada que acende toda hora que a audiência do portal cai” #meninojornalismotámorto**. Brasil, 2015. Twitter: @dafnesampaio. Disponível em: <<https://twitter.com/dafnesampaio/status/587718633289338880>>. Acesso em: 22/4/2015.

SEIXAS, Lia. **Redefinindo os gêneros jornalísticos**. Portugal, Covilhã: Labcom, 2009.

SILVER, Nate. What the fox knows. **FiveThirtyEight**, Estados Unidos, 2014. Disponível em: <<http://fivethirtyeight.com/features/what-the-fox-knows/>>. Acesso em 19/1/2015.

SPINUZZI, Clay. **Tracing Genres Through Organizations**. Cambridge: MIT Press, 2003.

STORCK, Guilherme. Bases eleitorais. **Gazeta do Povo**, Curitiba, 3/11/2012. Vida Pública, p. 11.

STORCK, Guilherme. Consulte as estatísticas sobre assassinatos. **Gazeta do Povo**, Curitiba, 4/8/2013. Disponível em: <[www.gazetadopovo.com.br/vidaecidadania/crime-sem-castigo/conteudo.phtml?id=1396257](http://www.gazetadopovo.com.br/vidaecidadania/crime-sem-castigo/conteudo.phtml?id=1396257)>. Acesso em 21/7/2014.

THORP, Jer. **Wired UK, Barabási Lab and BIG data**. Blprnt.com, Nova Iorque, 12 jul. 2010. Disponível em: <[blog.blprnt.com/blog/blprnt/wired-uk-barabasi-lab-and-big-data](http://blog.blprnt.com/blog/blprnt/wired-uk-barabasi-lab-and-big-data)>. Acesso em: 14/5/2014.

TRAQUINA, Nelson. **Teorias do Jornalismo – Volume 1**. Florianópolis: Insular, 2005.

TRÄSEL, Marcelo. **Jornalismo Guiado por Dados: relações da cultura hacker com a cultura jornalística**. In: XXII Encontro Anual da Compós. 2013, Salvador. Anais... Salvador, 2013. Disponível em: <[compos.org.br/data/biblioteca\\_2065.pdf](http://compos.org.br/data/biblioteca_2065.pdf)>. Acesso em: 14/5/2014.

TRÄSEL, Marcelo R. **Entrevistando planilhas**: estudo das crenças e do ethos de um grupo de profissionais de jornalismo guiado por dados no Brasil. 2014. 314 f. Tese (Doutorado em Comunicação Social) – Faculdade de Comunicação Social, Pontifícia Universidade Católica do Rio Grande do Sul, PUCRS, Porto Alegre, 2014.

TRESCA, Laura. **Gênero Informativo**. In: ASSIS, Francisco de; MARQUES DE MELO, José. Gêneros jornalísticos no Brasil. São Bernardo do Campo: Universidade Metodista de São Paulo, 2010.

TUFTE, Edward. **Envisioning Information**. Connecticut: Graphics Press, 1990.

TUFTE, Edward. **The Visual Display of Quantitative Information**. Connecticut: Graphic Press, 2001.

UNWIN, A.; THEUS, M.; HOFMANN, H. **Graphics of Large Datasets**. Nova Iorque: Springer, 2006.

WOLTON, Dominique. **Internet, e depois?** Porto Alegre: Sulina, 2007.

WOODWARD, Kathryn. **Identidade e diferença: uma introdução teórica e conceitual**. In: SILVA, Tomaz Tadeu (org.). *Identidade e diferença – a perspectiva dos estudos culturais*. Petrópolis: Vozes, 2000.

YAU, Nathan. **Visualize This**. 2011.

ZORINA, Kat; VAN DER VLEUTEN, Ruben; FRANTZIS, Kostantinos. **Tasty Tweets**. Disponível em: <ciid.dk/education/portfolio/idp12/courses/data-visualisation/projects/tasty-tweets/>. 2012. Acesso em: 14/5/2014.

## ENTREVISTAS

BARROS JUNIOR, José Lazaro Ferreira. **Entrevista**. Arquivo pessoal, 18/11/2014. Entrevista concedida ao autor.

BARROS JUNIOR, José Lazaro Ferreira; FREY, João Guilherme Bernardo. **Entrevista**. Arquivo pessoal, 16/10/2014. Entrevista concedida ao autor.

FOTIOS, Ricardo. **Entrevista**. Mensagem recebida por <matiasperuyera@gmail.com> em 5/8/2014.

FREITAS, Rosiane Correia de. **Entrevista**. Mensagem recebida por <matiasperuyera@gmail.com> em 26/8/2014.

FREY, João Guilherme Bernardo. **Entrevista**. Arquivo pessoal, 4/12/2014. Entrevista concedida ao autor.

PINTO, Renata Silva. **Entrevista**. Arquivo pessoal, 7/7/2014. Entrevista concedida ao autor.

SILVA, Ricardo Ruas. **Entrevista**. Arquivo pessoal, 20/8/2014. Entrevista concedida ao autor.

SMITH, Bob. **Entrevista**. Arquivo pessoal. Mensagem recebida por <matiasperuyera@gmail.com> em 10/2/2014.

STORCK, Guilherme. **Entrevista**. Arquivo pessoal, 24/7/2014. Entrevista concedida ao autor.

TRÄSEL, Marcelo. **Entrevista**. Mensagem recebida por <matiasperuyera@gmail.com> em 27/6/2014.

# APÊNDICE A: GLOSSÁRIO DE FERRAMENTAS

Neste glossário foram incluídas todas as ferramentas citadas na dissertação e mais algumas que foram consideradas importantes ou representavam alternativas livres para programas pagos que foram citados. Em vez de organizá-las alfabeticamente, foram agrupadas pela utilidade. Por esse critério, as suítes LibreOffice e OpenOffice estão presentes em duas categorias.

## FERRAMENTAS PARA EXTRAÇÃO DE DADOS

<b>Adobe Acrobat Pro</b> Pago, para Mac, Windows	Manipulação de arquivos PDF. Permite exportar texto, imagens ou planilhas contidas no arquivo.
<b>BeautifulSoup</b> Livre	Biblioteca para a linguagem Python para interpretação e raspagem de código HTML.
<b>Helium Scraper</b> Pago, para Windows	Ferramenta para raspagem de informação de páginas na internet.
<b>Nitro</b> Pago	Aplicativo para operações com arquivos PDF, incluindo extração de texto. Disponibiliza versão gratuita, com limitações.
<b>Outwit Hub</b> Pago, para Linux, Mac, Windows e para o navegador Firefox	Ferramenta para raspagem de informação de páginas na internet. Disponibiliza uma versão gratuita com limitações, e opção de comprar o aplicativo ou pagar uma taxa anual.
<b>Scrapewiki</b> Pago, para Web	Ferramenta online para transformação de PDFs e HTML em planilhas eletrônicas.
<b>Tabula</b> Livre, para Mac, Windows	Programa que converte arquivos PDF em planilhas eletrônicas.

## FERRAMENTAS PARA PREPARAR, ORGANIZAR, LIMPAR E FAZER OPERAÇÕES COM OS DADOS

<b>Google Refine</b>	Ver OpenRefine.
<b>LibreOffice</b> Livre, para Linux, Mac, Windows	Suíte que inclui editor de texto, planilha eletrônica e banco de dados.
<b>Microsoft® Access</b> Pago, para Windows	Aplicativo para criação, manipulação e consulta de bases de dados.
<b>Microsoft® Excel</b> Pago, para Mac, Windows	Planilha eletrônica.

<b>Microsoft® Word</b> Pago, para Mac, Windows	Editor de texto tradicional, com recursos de paginação e formatação.
<b>Notepad++</b> Livre, para Linux, Windows	Editor de texto sem formatação que suporta grandes quantidades de texto (gigabytes) e recursos como expressões regulares.
<b>OpenOffice</b> Livre, para Linux, Mac, Windows	Suíte que inclui editor de texto, planilha eletrônica e banco de dados.
<b>OpenRefine</b> Livre, para Linux, Mac, Windows	Programa que ajuda a organizar conjuntos de dados bagunçados, encontrando padrões para separar colunas e linhas, preenchimento automático de células etc. Antigamente conhecido como Google Refine.
<b>TextWrangler</b> Gratuito, para Mac	Editor de texto sem formatação que suporta grandes quantidades de texto (gigabytes) e recursos como expressões regulares.

## FERRAMENTAS DE PROGRAMAÇÃO

<b>Processing</b> Livre, para Linux, Mac, Windows	Linguagem e ambiente de programação, com foco no aprendizado de programação e em gerar resultados visuais.
<b>Python</b> Livre	Linguagem de programação.
<b>R</b> Livre, para Linux, Mac, Windows	Linguagem e ambiente de programação com ênfase em estatística e produção de visualizações.
<b>Ruby</b> Livre	Linguagem de programação.

## FERRAMENTAS PARA VISUALIZAÇÃO

<b>D3</b> Livre	Biblioteca JavaScript para trabalhar com dados usando HTML, CSS e SVG.
<b>Raphaël</b> Livre	Biblioteca JavaScript para trabalhar com desenho vetorial.
<b>Chartbuilder</b> Livre, na internet	Ferramenta online para construção de gráficos. Exporta em SVG e imagem raster.
<b>Highcharts</b> Pago	Biblioteca JavaScript para trabalhar com dados usando HTML, CSS e SVG. É gratuita para uso não-comercial.
<b>Hootsuite</b> Grátis	Aplicativo para gerenciamento de sites de redes sociais. Incluso nesta lista pela sua integração com Google Maps para mapear postagens.

<b>Infogr.am</b> Grátis, na internet	Ferramenta online para produzir visualizações.
<b>Many Eyes</b> Grátis, na internet	Ferramenta online para produzir visualizações.
<b>Raw</b> Grátis, na internet	Ferramenta online para produzir visualizações. Exporta as visualizações em formatos vetoriais.
<b>Microsoft® Paint</b> Incluso no Windows	Programa de desenho básico, vem com o sistema operacional Windows
<b>Adobe Illustrator</b> Pago, para Mac, Windows	Programa de desenho vetorial, tem funções de realização de gráficos.
<b>Gimp</b> Livre, para Linux, Mac, Windows	Programa de edição de imagens raster.
<b>Photoshop</b> Pago, para Mac, Windows	Programa de edição de imagens raster. Ver a alternativa livre, Gimp.
<b>Tableau</b> Grátis, com opções pagas; para Mac e Windows	Programa para criação de “painéis” online, permitindo vários tipos de visualização de bases de dados. A versão gratuita exige que os dados fiquem disponíveis ao público.
<b>Qlikview</b> Grátis, com opções pagas, para Windows	Programa para criação de “painéis” online, permitindo vários tipos de visualização de bases de dados.
<b>Google Charts</b> Gratuito, para Web	Serviço que permite a criação de gráficos estáticos ou dinâmicos para serem incluídos em arquivos HTML.
<b>Google Fusion Tables</b> Grátis, para Web	Tabelas online com a possibilidade de se integrar a mapas e a vários tipos de visualizações. Incorporado agora ao Google Drive.

## FERRAMENTAS PARA PRODUÇÃO DE MAPAS CARTOGRÁFICOS

<b>ArcGIS</b> Pago, para Windows	Processamento de dados espaciais, permite, a partir de camadas de arquivos shapefile, gerar mapas com gráficos. Ver a alternativa livre, QGIS.
<b>BatchGeo</b> Grátis, com opções pagas. Disponível na internet.	Automatiza a localização de locais no Google Maps. A versão gratuita permite localizar até 250 locais por mapa.
<b>Google Maps</b> Grátis, para Web	Plataforma que provê mapas para uso em sites e aplicativos.
<b>MapBox</b> Pago	Plataforma que provê mapas para uso em sites e aplicativos.
<b>OpenStreetMap</b> Livre	Plataforma que provê mapas para uso em sites e aplicativos.
<b>QGIS</b> Livre, para Linux, Mac, Windows	Processamento de dados espaciais, permite, a partir de camadas de arquivos shapefile, gerar mapas com gráficos.

# APÊNDICE B: ROTEIRO DE ENTREVISTA SEMI ESTRUTURADA

- **Idade**
- **Gênero**
- **Formação:** estudante de jornalismo, estudante de outro curso, formado em jornalismo, formado em outro curso, pesquisador, professor ou pós-graduando em jornalismo, pesquisador, professor ou pós-graduando em outra área, outros.
- **Ocupação atual:** fazendo estágio em jornalismo, trabalhando em jornalismo, já trabalhou em jornalismo, trabalhando ou estagiando em outra área, outros.
- **Veículos em que trabalha ou trabalhou:** impresso, web (portal, blog, etc), blog jornalístico próprio, televisão, rádio, várias mídias, outros.
- **O que você sabe fazer usando o computador, tablet ou telefone?:** internet, chat, email, mídias sociais em geral, editores de texto (Word e similares), planilhas de cálculo (Excel e similares), apresentação (PowerPoint e similares), editoração e tratamento de imagem (InDesign, Photoshop, Gimp e similares), edição de vídeo (Premiere, Final Cut e similares), modelagem 3D (SketchUp, Maya, Blender e similares), HTML básico, HTML avançado, programação, outros.
- **Quais dessas habilidades você usa no dia a dia como jornalista?:** internet, chat, email, mídias sociais em geral, editores de texto (Word e similares), planilhas de cálculo (Excel e similares), apresentação (PowerPoint e similares), editoração e tratamento de imagem (InDesign, Photoshop, Gimp e similares), edição de vídeo (Premiere, Final Cut e similares), modelagem 3D (SketchUp, Maya, Blender e similares), HTML básico, HTML avançado, programação, outros.
- **Onde ou como você aprendeu a usar computador?:** sozinho, fazendo tutoriais ou similares na internet, fazendo cursos, no trabalho, outros.
- **Pretende aprender alguma nova habilidade com o computador?**
- **Quais destes conceitos você conhece?:** RAC (Reportagem Assistida por Computador, também conhecida como CAR), Jornalismo de dados, Visualização de dados
- **Conhece alguma reportagem ou trabalho que se encaixe nas categorias acima?**
- **Como você procuraria saber como usar uma função de um software que você desconhece?**
- **Se você precisasse aprender a usar um programa que você não conhece, como você aprenderia?**
- **Como você acha que o computador pode ajudar no dia a dia do jornalista?**



- **Pingue-pongue: descreva brevemente como você executaria as ações abaixo, usando ou não o computador. Mencione softwares, sites, fontes, a quem você pediria ajuda, as tarefas que você sabe que é possível fazer mas não lembra exatamente como e o que mais você queira detalhar:**
- Diagramar uma página de revista ou jornal
- Montar um blog pessoal
- Mostrar cinco cidades brasileiras em um mapa a partir de uma lista
- Mostrar duzentas cidades brasileiras em um mapa a partir de uma lista
- Fazer um gráfico com o PIB dos dez últimos anos
- Descobrir a população de todas as capitais brasileiras
- Descobrir o número de homicídios de todas as capitais brasileiras
- A partir dessas duas últimas informações, calcular a taxa de homicídios a cada 100 mil habitantes e fazer um ranking das capitais brasileiras mais violentas
- Pintar um mapa do Brasil com duas cores representando os estados de oposição e de situação
- Fazer um mapa do Paraná usando vários tons de cinza segundo o Índice de Desenvolvimento Humano de cada um dos 399 municípios

# APÊNDICE C: QUESTÕES PARA PROFESSORES

## **Na universidade**

Um breve histórico seu como profissional e/ou professor.

Em quais universidades você dá ou deu aula? São públicas ou particulares? Qual o perfil dos alunos?

Você leciona ou lecionou outras disciplinas relacionadas ao jornalismo?

O conteúdo era abordado dentro de alguma disciplina já existente? Qual?

Você organizou atividades extracurriculares relacionadas ao JGD? Quais? Como foi a recepção?

Qual a postura da universidade com relação a esse tipo de conteúdo?

Como você vê o ensino de JGD atualmente no Brasil, inclusive em atividades para jornalistas já formados?

Como você aprendeu o que ensina de JGD na universidade?

Já trabalhou com JGD fora da sala de aula? Já fez alguma parceria entre a universidade e algum meio de comunicação ou similar?

## **Na sala de aula**

Como foi a apresentação dos conceitos de JGD (ou RAC) aos alunos? Eles conheciam os conceitos?

Quais foram os conteúdos abordados?

O foco das aulas é em algum software específico? Qual ou quais?

De maneira geral, qual a reação dos alunos? Teve algum aluno que se destacou durante o curso? Houve reações muito negativas?

Pretende mudar alguma coisa no seu método de ensino? Quais experiências te levaram a isso?

Você acha que algum tipo de conhecimento prévio por parte dos alunos teria facilitado as aulas?

## **Visualizações de dados**

Você fez alguma atividade com os alunos usando algum tipo de visualização de dados?

Notou alguma dificuldade a mais que os alunos tiveram ou poderiam ter caso fossem necessárias técnicas de visualização de dados?

# APÊNDICE D: DESCRIÇÃO DA PRODUÇÃO DAS VISUALIZAÇÕES ILUSTRATIVAS

## *Capa*

O texto principal da dissertação (capítulos 1 a 6) foi editado no editor de texto TextWrangler para separar cada frase a partir dos pontos finais. Foram eliminadas as frases soltas, como referências, e o texto foi colado no Excel para calcular a extensão de cada uma, e marcar quais tinham as palavras “visualização”, “jornalismo” ou “jornalista”, assim como as posições que cada elemento deveria ter. No ambiente de programação Processing, foi feito um programa para desenhar um quadrado para cada frase – ou um círculo, caso a frase tivesse uma das palavras marcadas. Cores mais claras representam frases mais curtas, e cores mais próximas do preto representam frases mais longas. O programa gerava um arquivo PDF que foi aberto no programa de desenho vetorial Adobe Illustrator, para diagramar a visualização junto aos títulos da capa.

## *Extensão dos capítulos segundo tipo de texto*

Cada tipo de texto dos capítulos da dissertação foi deixado em fonte Courier corpo 12, e teve uma cor de fundo aplicada para cada tipo de texto (texto corrido, citação recuada, títulos, notas de rodapé). Foram retiradas as imagens. Foi gerado um arquivo PDF e aberto no Macromedia Freehand MX, aplicativo de desenho vetorial que abre todas as páginas ao mesmo tempo. Foram selecionados somente os retângulos coloridos que indicam o tempo de texto, e organizados para comparar a quantidade de texto de cada tipo presentes em cada capítulo. A visualização foi finalizada no aplicativo de desenho vetorial Adobe Illustrator.

## *Páginas das citações*

Foram buscadas as ocorrências de páginas da dissertação usando a expressão regular `p. (\d+)` para encontrar todas as vezes que fora indicada uma página. Foi editado o texto para criar duas sequências que foram escritas como sequências numéricas – arrays – no Processing, programa usado para escrever a visualização, já que exigia o espaçamento entre as barras na proporção das páginas. Acrescentou-se ao laço linhas a cada 10 páginas, para facilitar a edição final, no Adobe Illustrator.

```
import processing.pdf.*;
```

```

int[] pag = {
  1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 18, 22, 26, 27, 28, 30, 31, 32, 34, 36, 37, 38,
  39, 41, 42, 44, 45, 46, 47, 48, 50, 51, 52, 53, 55, 56, 57, 59, 61, 64, 65, 66, 67, 68, 69, 70, 71,
  73, 75, 76, 77, 78, 79, 80, 87, 89, 93, 94, 95, 96, 99, 100, 101, 104, 108, 109, 110, 111, 112, 113,
  114, 123, 128, 129, 131, 137, 138, 139, 146, 147, 148, 151, 154, 155, 156, 157, 162, 166, 169, 183,
  184, 187, 202, 204, 205, 206, 207, 255, 261, 265, 269, 271, 274, 282, 294, 318, 327
};

int[] alt = {
  7, 8, 8, 6, 9, 8, 9, 6, 2, 5, 3, 4, 4, 1, 9, 1, 2, 1, 6, 3, 4, 1, 5, 2, 1, 3, 5, 2, 1, 1, 1, 1, 4,
  5, 1, 5, 2, 2, 1, 1, 5, 1, 1, 2, 1, 2, 1, 5, 4, 2, 1, 2, 2, 1, 1, 3, 3, 1, 1, 1, 1, 3, 1, 1, 1, 1,
  1, 1, 1, 2, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 5, 1, 2, 3, 1, 1, 5, 2, 5, 2, 1, 2, 1, 1, 1, 1, 1, 5, 7, 2,
  4, 1, 1, 1, 2, 3, 1, 1, 1, 1, 1
};

void setup() {
  size((10+max(pag)*4), 200);
  noLoop();
  background(255);
  fill(128, 0, 0);
}

void draw() {

  beginRecord(PDF, "citacoes.pdf");

  int largura = (width/max(pag));

  for (int i=0; i < pag.length; i=i+1) {
    rect (largura*pag[i], largura, largura, largura*alt[i]);
    line(5+(largura*10*i), 30, 5+(largura*10*i), 60);
    textAlign(CENTER);
    text("e", 5+(largura*10*i), 40);
  }

  endRecord();
}

```

### *Frequência de letras*

Foi gerado um arquivo da dissertação em formato de texto puro para ser puxado por um script na linguagem Python (abaixo). O resultado do script foi colado no Microsoft Excel, para ser associado à tabela de frequência de letras na língua portuguesa. A tabela então foi usada para fazer um gráfico no aplicativo de ilustração vetorial Adobe Illustrator.

```

#!/usr/bin/env python
# -*- coding: utf-8 -*-

with open ("disset 2.txt") as texto:
    str=texto.read()

#transforma tudo em caixa baixa
str=str.lower()

#elimina acentos
str=str.replace("ã", "a");
str=str.replace("á", "a");
str=str.replace("Á", "a");
str=str.replace("à", "a");
str=str.replace("À", "a");
str=str.replace("â", "a");
str=str.replace("Â", "a");
str=str.replace("ã", "a");

```

```

str=str.replace("Ã", "a");
str=str.replace("ç", "c");
str=str.replace("Ç", "c");
str=str.replace("é", "e");
str=str.replace("É", "e");
str=str.replace("ê", "e");
str=str.replace("Ê", "e");
str=str.replace("í", "i");
str=str.replace("Í", "i");
str=str.replace("ó", "o");
str=str.replace("Ó", "o");
str=str.replace("ô", "o");
str=str.replace("Ô", "o");
str=str.replace("u", "u");
str=str.replace("U", "u");
str=str.replace("ú", "u");
str=str.replace("Ú", "u");

#não substitui caracteres de outros idiomas
#str=str.replace("ä", "a");
#str=str.replace("Ä", "a");
#str=str.replace("ë", "e");
#str=str.replace("Ë", "e");
#str=str.replace("ñ", "n");
#str=str.replace("Ñ", "n");
#str=str.replace("ö", "o");
#str=str.replace("Ö", "o");

from collections import Counter

counter = Counter(str)

total =
counter['a']+counter['b']+counter['c']+counter['d']+counter['e']+counter['f']+counter['g']+counter['
h']+counter['i']+counter['j']+counter['k']+counter['l']+counter['m']+counter['n']+counter['o']+count
er['p']+counter['q']+counter['r']+counter['s']+counter['t']+counter['u']+counter['v']+counter['w']+c
ounter['x']+counter['y']+counter['z']

#apresenta o total de letras e o total de cada uma
print total
print counter['a']
print counter['b']
print counter['c']
print counter['d']
print counter['e']
print counter['f']
print counter['g']
print counter['h']
print counter['i']
print counter['j']
print counter['k']
print counter['l']
print counter['m']
print counter['n']
print counter['o']
print counter['p']
print counter['q']
print counter['r']
print counter['s']
print counter['t']
print counter['u']
print counter['v']
print counter['w']
print counter['x']
print counter['y']
print counter['z']

```

### *Mapeamento dos autores*

A partir das referências bibliográficas, foi gerada uma tabela usando o editor de texto TextWrangler e a planilha Excel. Quando não foi possível descobrir a nacionalidade, foi usado o lugar de publicação da obra. O mapa foi feito no aplicativo Google Fusion Tables, disponível online, e redesenhado no Adobe Illustrator.

### *Contagem de caracteres não-alfanuméricos*

Foram eliminados os caracteres alfanuméricos do texto principal usando o editor de texto TextWrangler. Foram organizados alfabeticamente no mesmo aplicativo. Na planilha Excel, foi feita uma formula para contar os caracteres. A partir desse número foram desenhados os círculos no aplicativo Illustrator, com áreas proporcionais à contagem de caracteres.

### *Nuvem de palavras*

Foi usado o aplicativo online Wordle, disponível em <[www.wordle.net](http://www.wordle.net)>. Foi gerado um arquivo PDF a partir do menu de impressão, e editado no Illustrator.