

ŁUKASZ ŻYŁA

DZIENNIKARSTWO DANYCH I DATA STORYTELLING

Wszelkie prawa zastrzeżone. Nieautoryzowane rozpowszechnianie całości lub fragmentu niniejszej publikacji w jakiekolwiek postaci jest zabronione. Wykonywanie kopii metodą kserograficzną, fotograficzną, a także kopiowanie książki na nośniku filmowym, magnetycznym lub innym powoduje naruszenie praw autorskich niniejszej publikacji.

Wszystkie znaki występujące w tekście są zastrzeżonymi znakami firmowymi bądź towarowymi ich właścicielami.

Autor oraz wydawca dołożyli wszelkich starań, by zawarte w tej książce informacje były kompletne i rzetelne. Nie biorą jednak żadnej odpowiedzialności ani za ich wykorzystanie, ani za związane z tym ewentualne naruszenie praw patentowych lub autorskich. Autor oraz wydawca nie ponoszą również żadnej odpowiedzialności za ewentualne szkody wynikłe z wykorzystania informacji zawartych w książce.

Redaktor prowadzący: Barbara Gancarz-Wójcicka

Helion S.A.
ul. Kościuszki 1c, 44-100 Gliwice
tel. 32 231 22 19, 32 230 98 63
e-mail: onepress@onepress.pl
WWW: <http://onepress.pl> (księgarnia internetowa, katalog książek)

Drogi Czytelniku!
Jeżeli chcesz ocenić tę książkę, zajrzyj pod adres
<http://onepress.pl/user/opinie?dzidan>
Możesz tam wpisać swoje uwagi, spostrzeżenia, recenzję.

ISBN: 978-83-283-8312-8

Copyright © Łukasz Żyła 2022

Printed in Poland.

Spis treści

Wstęp	7
Rozdział 1. Dane i dziennikarstwo — wprowadzenie	10
1.1. Czym jest dziennikarstwo danych	10
1.2. Dziennikarstwo danych jako metoda	11
1.3. Dane	14
1.4. Dziennikarstwo a świat cyfrowy	15
1.5. Otwarte dane	16
1.6. Warsztat dziennikarza pracującego z danymi	17
1.7. Dokąd zmierza dziennikarstwo cyfrowe?	18
Rozdział 2. Dziennikarstwo danych — przykłady	20
2.1. Odkrywamy	20
2.2. Pokazujemy	25
2.3. Angażujemy	27
Rozdział 3. Internet, czyli dziennikarstwo online	29
3.1. Internet	29
3.2. Jak działają strony internetowe	30
3.3. Język HTML i CSS	32
Rozdział 4. Dane — formaty, źródła, zdobywanie	35
4.1. Przetwarzanie danych	35
4.2. Źródła informacji	41
4.3. Dane na wniosek	42
4.3.1. Bądź przygotowany na opóźnienie	43
4.3.2. Kiedy informacja jest publiczna?	44
4.3.3. Kto ma obowiązek udzielenia informacji?	44

4.3.4. Wnioskuj o dane przetwarzalne	45
4.3.5. Precyza i cierpliwość	46
4.3.6. Pytaj w różnych miejscach i testuj	47
4.3.7. Odmowa i co dalej?	47
4.4. A może zautomatyzować?	48
4.4.1. Wykorzystanie Google Sheets	51
4.4.2. Gdzie przechowywać dane?	52
4.4.3. Jakość i źródła ogólnodostępne	53
Rozdział 5. Czyszczenie i agregacja danych	55
5.1. Pozbądź się śmieci	55
5.1.1. Sortowanie	57
5.1.2. Wyszukiwanie fasetowe	57
5.1.3. Wykrywanie duplikatów	59
5.1.4. Zastosowanie filtra tekstowego	59
5.1.5. Transformacje komórek	59
5.1.6. Usuwanie niewłaściwych danych	61
5.1.7. Grupowanie danych, czyli klastrowanie	61
5.2. Szukanie wzorców, czyli analiza	63
5.2.1. Odrobina statystyki	63
5.2.2. Powiązania i korelacje	66
5.3. Metody analizy w praktyce	67
5.3.1. Przygotowanie tabeli w Google Sheets	67
5.3.2. Analizy bazodanowe	73
Rozdział 6. Jak tworzyć zrozumiałe i przyciągające uwagę wizualizacje	78
6.1. Po co i jak wizualizować?	80
6.2. Jak wygląda dobra wizualizacja?	87
6.2.1. Bądź uczciwy	90
6.2.2. Dobierz odpowiedni wykres	90
6.2.3. Wykorzystaj właściwe skale	90
6.2.4. Zastosuj odpowiednie kolory	91
6.2.5. Spraw, aby wizualizacja była czytelna	91
6.3. Najczęstsze błędy	92
6.3.1. Wykres 3D	92
6.3.2. Brak odwzorowania danych	92
6.3.3. Zbyt duże opisy i legenda	93
6.3.4. Brak różnic	93
6.3.5. Zmiana skali	93

6.4. Rodzaje wizualizacji	94
6.4.1. Wizualizacja proporcji	94
6.4.2. Porównanie kilku wartości	98
6.4.3. Śledzenie zmian w czasie	100
6.4.4. Obserwacja zależności między danymi	102
6.4.5. Zależności przestrzenne	107
6.4.6. Rozkład cechy statystycznej	107
Rozdział 7. Data storytelling	111
7.1. Składowe	112
7.2. Wzorce, czyli szkielet data story	115
7.2.1. Wzorce argumentacji	116
7.2.2. Wzorce przepływu	117
7.2.3. Wzorce kadrowania narracji	117
7.2.4. Wzorce empatii i emocji	119
7.2.5. Wzorce zaangażowania	121
7.3. Elementy GUI	123
7.4. Data storytelling w praktyce	124
Rozdział 8. Jak stworzyć wizualizacje	136
8.1. Na początek coś prostego — Datawrapper	137
8.2. Flourish	139
8.3. Sieci społecznościowe	141
8.4. Coś bardziej zaawansowanego, czyli Tableau	143
8.5. Ekosystem, czyli Google Charts i Google Maps	149
8.6. Infogram	151
8.7. Łączenie wizualizacji	153
Rozdział 9. Wstęp do programowania	159
9.1. Zmienne	161
9.2. Obiekty	162
9.3. Tablice	162
9.4. Tablice i obiekty	164
9.5. Funkcje	165
9.6. Instrukcje warunkowe	166

Rozdział 10. Wizualizacja zaprogramowana	169
10.1. Chart.js	170
10.2. Highcharts i AmCharts	174
10.3. p5.js	175
10.4. ECharts	180
10.5. Vega-Lite	182
Rozdział 11. Data story w sieci	189
Dodatek A. Wywiad z Dominikiem Uhligiem, redaktorem naczelnym serwisu BIQdata „Gazety Wyborczej”	201
Zakończenie	205
Bibliografia	206

16. [https://nocodewebscraping.com/data-journalism-examples/.](https://nocodewebscraping.com/data-journalism-examples/)
17. [https://observablehq.com/.](https://observablehq.com/)
18. [https://shorthand.com/the-craft/how-data-journalists-create-beautiful-feature-stories/index.html.](https://shorthand.com/the-craft/how-data-journalists-create-beautiful-feature-stories/index.html)
19. [https://www.fractl.com/work/10x-data-journalism-examples/.](https://www.fractl.com/work/10x-data-journalism-examples/)
20. [https://www.google.pl/amp/s/www.pinterest.com/amp/hunterwhitney/data-journalism/.](https://www.google.pl/amp/s/www.pinterest.com/amp/hunterwhitney/data-journalism/)
21. [https://www.internetlivestats.com/.](https://www.internetlivestats.com/)
22. [https://www.jerthorp.com/.](https://www.jerthorp.com/)
23. [https://www.nytimes.com/interactive/2020/03/22/world/coronavirus-spread.html.](https://www.nytimes.com/interactive/2020/03/22/world/coronavirus-spread.html)
24. [https://www.theguardian.com/world/interactive/2012/may/08/gay-rights-united-states.](https://www.theguardian.com/world/interactive/2012/may/08/gay-rights-united-states)