

Adam Szalach

Uniwersytet Mikołaja Kopernika w Toruniu
ORCID 0000-0001-8040-001X

Veslava Osińska

Uniwersytet Mikołaja Kopernika w Toruniu
ORCID 0000-0002-1306-7832

Adela Jankowska

Uniwersytet Mikołaja Kopernika

Wykorzystanie eye trackingu do analizy wizualnych informacji statystycznych

Wstęp

W społeczeństwie danych, w które ewoluuje społeczeństwo informacji i wiedzy, istotną umiejętnością staje się odczyt zakodowanych treści liczbowych w wykresach i mapach – wszelkiego rodzaju wizualizacjach. Jedne wizualizacje mogą się wydawać odbiorcom nader prostymi reprezentacjami zaobserwowanych zjawisk (np. wykresy słupkowe, kołowe i liniowe), nie generując problemów w ich interpretacji. Inne, złożone, wykorzystujące nowoczesne techniki architektury graficznej, mogą wymagać wiedzy specjalistycznej – matematycznej czy statystycznej oraz wprawy w percepcji i przetwarzaniu przedstawionych na nich informacji. Tym niemniej, niezależnie od formy, architektury i stopnia kompleksowości wykresy i wizualizacje są komunikowane z zamiarem uwidocznienia pewnych zależności pomiędzy elementami, umożliwienia ich porównania, dostrzeżenia zmian lub odwrotnie – ich braku oraz ujęcia całościowego właściwości zbioru danych.

Przedstawienie danych statystycznych w postaci wykresów ma wzmacniać i utrwalać przekaz w dowolnej dziedzinie. Komunikacja danych z założenia przekłada się na interpretację, która może zależeć od pojedynczego odbiorcy lub grup społecznych. W przypadku nieskomplikowanych wykresów przedstawiających dwie, trzy a nawet cztery zmienne, gdzie trzeba odczytać wartość na osi lub oszacować wzrost lub spadek, poradzi sobie każdy – już w szkole podstawowej uczniowie nabywają takie

umiejętności¹. Problem staje się widoczny wraz z ilością zmiennych oraz danych, które mają być zaprezentowane. Duże dane wymagają zarówno innego podejścia, opartego na wykorzystaniu różnych funkcji agregujących, jak i nowych metod wizualizacji informacji (WI). Do usystematyzowania licznych znanych metod WI przyjęło się posługiwanie formalizmem wynikającym z pięciu perspektyw, określanych podczas rozważania problemu badawczego². Pytania, które stawia przed sobą badacz, to: 1) kiedy wydarzyło się dane zjawisko przedstawiane za pomocą danych; 2) gdzie miało miejsce; 3) jak było/jest intensywne; 4) jakiego tematu dotyczy; 5) kto lub co się do tego przyczynił(o) i jakie są relacje pomiędzy podmiotami. W przełożeniu na perspektywy analityczne, a następnie – wizualizacyjne, nazywa się je odpowiednio: czasowa (temporary), geograficzna, ilościowa (statystyczna), tematyczna i sieciowa.

Im więcej wymienionych aspektów analizy chcemy zawrzeć na jednym wykresie, tym bardziej skomplikowany się staje, co z kolei może negatywnie wpłynąć na jego czytelność (*legibility*) lub prowadzić nawet do błędnej interpretacji. Fakt powiązania wielomodalności wizualizacji z trudnością w interpretacji może prowadzić do tego, że użytkownik informacji umieszczonej w mediach może być nimi zmanipulowany³. Ten nasilający się na świecie i w Polsce problem rozwiązywany jest za pomocą praktycznych działań na rzecz alfabetyzacji w zakresie rozumienia języka wizualnego na wykresach i mapach (*graphic literacy*)⁴.

Autorzy postanowili przeprowadzić badania czytelności wizualnych reprezentacji danych statystycznych wśród użytkowników zróżnicowanych wiekowo. W wykresach zawarto intencjonalne błędy projektowe, a różnice w percepcji wizualnej analizowano za pomocą coraz popularniejszej w naukach społecznych techniki eye tracking (ET). Celem niniejszych badań było ustalenie, czy osoby z podstawową wiedzą w zakresie wizualizacji danych zwrócą uwagę na ukryte w wykresach błędy oraz czy mają świadomość błędów w niewłaściwie zaprojektowanych wykresach. Dodatkowe pytanie w badaniu czytelności wizualizacji dotyczyło zależności *legibility* od rodzaju wykresu i/lub aspektu analizy. Równolegle w ramach badań pilotażowych wykorzystana została możliwość analizy podstawowych emocji na podstawie kodowania FACS/FAP oferowanej przez dostawcę narzędzia ET.

Przegląd literatury

Umiejętność odczytywania danych zawartych na wykresach i wizualizacjach uznawana jest za przydatną cechę w dzisiejszym społeczeństwie, pobierającym informacje z różnych mediów. Biorąc pod uwagę skuteczność przekazywania wiedzy za pomocą

1 A. Stolińska, M. Bartkowski, *Analiza procesu percepcji danych przedstawionych na wykresach*, „Edukacja – Technika – Informatyka” 2018, nr 2 (24), s. 56–57.

2 V. Osińska, *Mapy nauki. Badania, Mapy nauki. Badania, potencjał oraz wyzwania w przykładach*, Toruń 2021.

3 V. Osińska, *Wizualizować potrafi każdy: czy jednak robi to poprawnie?*, „Biuletyn EBIB” 2021, nr 2 (197).

4 E. Fry, *Graphical literacy*, „Journal of Reading” 1981, nr 24 (5), s. 383–389.

wykresów, odczytywanie zawartych na nich informacji uważa się za kluczową kompetencję zarówno w szkolnictwie podstawowym, jak i wyższym. Na poziomie szkoły podstawowej uczniowie zdobywają takie umiejętności na lekcjach matematyki, rozwiązując zadania dotyczące odczytu wartości funkcji⁵. Jak dowodzą badania, nie tylko uczniowie, lecz również studenci wykazują braki w zrozumieniu wykorzystywania wizualnych form prezentacji danych⁶. Do nich zaliczają się między innymi: poprawna identyfikacja zmiennych, wykorzystanie odpowiedniego typu wykresu, skalowanie oraz wyciąganie wniosków⁷. Ważnym, powtarzającym się wnioskiem z tak ukierunkowanych badań jest to, iż zdolności poprawnego konstruowania i interpretowania wykresów w dużym stopniu zależą od wiedzy wstępnej i doświadczenia w zakresie wizualizacji.

W ocenie procesów kognitywnych zachodzących podczas percepcji wykresów i grafik informacyjnych w ostatnim dziesięcioleciu przyjęło się korzystać z technologii śledzenia wzroku – eye trackingu. Przesłanką jest to, że ET może się okazać skutecznym sposobem na „skompletowanie dowodów, jak ludzie się uczą, używając materiałów wizualnych”⁸. Jak dowodzą badania, ET może pomóc w odkryciu subtelnych etapów przetwarzania poznawczego, które w przypadku stosowania innych metod trudno byłoby zaobserwować. W studiach porównawczych odczytu tradycyjnych wykresów liniowych z układem promienistym zaobserwowano, iż zadania dotyczące znalezienia punktu danych i odczytu wartości były wykonywane szybciej w pierwszym przypadku⁹. Według autorów przyczyna przypuszczalnie kryje się w tym, że oczy w poszukiwaniu informacji muszą podążać po torze kołowym, a nie liniowym. W innym badaniu z wykorzystaniem ET próbowano ustalić, który typ wykresu najefektywniej zapewnia relatywne porównywanie wartości danych¹⁰. Różnice i podobieństwa w strategii interpretacji informacji na wykresach badano w odniesieniu do kierunku studiów respondentów: przyrodniczych vs społecznych, humanistycznych¹¹. Zaobserwowano znaczące zróżnicowanie w liczbie fiksacji oraz wizualnych wzorcach poszukiwania odpowiedzi, zamieszczonych w treści wykresu.

5 A. Stolińska, M. Bartkowski, ...dz. cyt.

6 N. Glazer, *Challenges with graph interpretation: A review of the literature*, „Studies in Science Education” 2011, nr 47 (2), s. 183–210.

7 A. V. Maltese, J. A. Harsh, D. Svetina, *Data visualization literacy: Investigating data interpretation along the novice–expert continuum*, „Journal of College Science Teaching” 2015, nr 45 (1), s. 84–91.

8 K. Ooms, P. De Maeyer, V. Fack, *Study of the attentive behavior of novice and expert map users using eye tracking*, „Cartography and Geographic Information Science” 2014, nr 41 (1), s. 37–54.

9 J. H. Goldberg, J. I. Helfman, *Eye tracking for visualization evaluation: Reading values on linear versus radial graphs*, „Information Visualization” 2011, nr 10 (3), s. 182–195.

10 J.H. Goldberg, J.I. Helfman, *Comparing information graphics: a critical look at eye tracking. Proceedings BELIV'10 Workshop, Beyond Time and Errors in the Evaluation of Visualizations (Atlanta, USA)*, New York 2010, s. 71–78.

11 J. A. Harsh et al., „Seeing” Data Like an Expert: An Eye-Tracking Study Using Graphical Data Representations, „CBE–Life Sciences Education” 2019, vol. 18, no. 3, <https://doi.org/10.1187/cbe.18-06-0102>.

Autorską propozycją było wykorzystanie otrzymanych wyników w projektowaniu sylabusów w obszarze wspomnianych dyscyplin.

Śledzenie wzroku wykorzystano także do klasyfikacji błędnych strategii odczytywania wykresów¹². W wyniku tych analiz zaproponowano wytyczne do projektowania wykresów wybranych typów. Pomimo zastosowania nowoczesnej metody eksperymentalnej badania ET w kierunku zrozumienia informacji wizualnych nie są pozbawione wad, do których zalicza się między innymi: określenie obszarów zainteresowania, odpowiedni dobór miary do pytania badawczego, urywanie ścieżek wzroku¹³.

Badania ET w warunkach pandemii

Z uwagi, iż badanie prowadzone było w okresie trwania ogólnopolskiego lockdownu wywołanego przez pandemię COVID-19, eksperyment zrealizowano w trybie zdalnym przy wykorzystaniu narzędzia polskiej firmy REALEYE¹⁴. Oferuje ona możliwość wykorzystania kamer internetowych oraz sprzętu komputerowego respondentów bez potrzeby bezpośredniego kontaktu. Rozwiązanie takie, w porównaniu z badaniami bezpośrednimi na ściśle określonej aparaturze badawczej, ma szereg zalet:

- możliwość prowadzenia badań w dowolnym miejscu oraz terminie;
- brak ograniczeń terytorialnych;
- brak stresu wywołanego obecnością w pracowni/laboratorium;
- pomiary odbywają się w warunkach naturalnych dla respondentów, dzięki czemu mają charakter badań etnograficznych;
- badanie nie wymaga bezpośredniego kontaktu i złożonych działań logistycznych.

Dokonując analizy porównawczej z badaniami laboratoryjnymi¹⁵, wykazano, że nie są one jednak pozbawione wad, do których przede wszystkim należy zaliczyć:

- brak możliwości sprawowania bezpośredniej kontroli nad warunkami środowiskowymi (możliwość występowania dystraktorów w postaci zróżnicowanych dźwięków, oświetlenia);
- brak możliwości zapewnienia jednakowych parametrów sprzętu (prędkość łącza, zgodność wersji przeglądarek internetowych);
- brak możliwości stałego i szybkiego monitorowania procesu badawczego;
- wyższy koszt realizacji badań.

Metodologicznie badanie prowadzone w formie zdalnej, poza warunkami środowiskowymi, nie odbiega od standardu przyjętego w eksperymentach wykorzystujących eye trackery stacjonarne. Zarówno w przypadku procesu realizowanego bezpośrednio w laboratorium, jak i na odległość istotny jest nadzór prowadzącego nad procesem przygotowania respondentów, rozumiany jako odpowiednie dostosowanie

12 J. H. Goldberg, J. I. Helfman, *Eye tracking for visualization evaluation...*, dz. cyt.

13 J. H. Goldberg, J. I. Helfman, *Comparing information graphics...*, dz. cyt.

14 RealEye – firma oferująca badania eyetrackingowe przy wykorzystaniu kamer internetowych. Zob. Realeye, [on-line:] <https://www.realeye.io/> -19.07 2022.

15 A. T. Duchowski, *Eyetracking methodology*, New York 2017.

odległości od monitora, środowiska pracy oraz oświetlenia. W tym celu pomocne jest narzędzie oferowane przez usługodawcę, w formie wideo-komunikatora, umożliwia bowiem kontrolę zachowania użytkownika i nadzorowanie przebiegu badania, pozostając w bezpośrednim kontakcie. Respondent biorący udział w eksperymentie otrzymuje link kierujący do serwisu i dedykowanego eksperymentu. Po uruchomieniu pierwszym krokiem jest wyrażenie zgód dotyczących ochrony danych osobowych oraz gromadzenia danych w celach badawczych. W wybranych konfiguracjach systemu użytkownika niezbędne jest udzielenie zezwolenia przeglądarce na dostęp do kamery internetowej. Kolejnym elementem jest proces kalibracji, który – podobnie jak w przypadku eye trackerów stacjonarnych – polega na śledzeniu obiektu poruszającego się po ekranie i zatrzymującego się w losowych punktach. W przypadku aplikacji przeznaczonych dla eye trackerów stacjonarnych, zależnie od ustawień wprowadzonych przez badacza, jest to od 9 do około 32 pozycji (np. OGAMA¹⁶). REALEYE dokonuje kalibracji na 44 punktach zarówno na jasnym, jak i ciemnym ekranie w celu analizy zachowania źrenicy. W przypadku nieprawidłowego wyniku proces może być wielokrotnie powtarzany. Jeżeli kalibracja zakończy się pozytywnie, następuje przejście do badania właściwego.

Zależnie od konfiguracji przeprowadzonej przez badacza w procesie tworzenia eksperymentu badania mogą odbywać się na obrazach statycznych, materiałach wideo oraz stronach internetowych. Pytania kierowane do respondenta są wyświetlane na ekranie lub zadawane bezpośrednio przez prowadzącego. Po zakończonym badaniu zależnie od konfiguracji badany proszony jest o podanie danych statystycznych, jak imię, wiek, płeć oraz ewentualnych uwag. Narzędzie oferuje pełną możliwość analizy danych zarówno pojedynczych respondentów, jak i zbiorczo dla całego eksperymentu. Odpowiednio przygotowany dashboard pozwala na: określenie obszarów zainteresowania (AOI), tworzenie map cieplnych, analizę fiksacji, sakkad oraz gaze plotów. Dodatkowo wszystkie niezbędne dane mogą zostać wyeksportowane do formatu .CSV i poddane analizie na zewnętrznym oprogramowaniu.

Badania

Materiał do badań

Badanie odbyło się w styczniu 2022 roku. Zostało podzielone na dwie części o nazwie roboczej: „obrazy statyczne” i „strona internetowa”. Przygotowano cztery plansze i na każdej z nich umieszczono konkretny typ wykresów: słupkowe, ukazujące części całości (kołowy, radarowy, pierścieniowy i mapę drzewiastą) oraz liniowe.

W pierwszej części nazwanej „obrazy statyczne” badani mieli za zadanie odpowiedzieć na pytania dotyczące informacji przedstawionych na wykresach na planszy.

¹⁶ OGAMA – strona główna otwartego oprogramowania do analizy ruchu gałek ocznych oraz urządzeń wskazujących. Zob. Realeye, [on-line] <http://www.ogama.net/> – 29.07.2022.

Mieli 20 sekund na udzielenie odpowiedzi, o czym uczestnicy byli powiadomieni wcześniej. Prezentacja składała się z pięciu slajdów. Pierwszy wprowadzał do stałej topologii rozmieszczenia analizowanych wykresów i przedstawiał przypisane do nich numery w celu ułatwienia komunikacji pomiędzy badaczem a badanym. Drugi slajd przedstawiał wykresy słupkowe i kolumnowe o różnorodnej kolorystyce i stylistyce (tab. 1A), ich tematyka dotyczyła COVID-19.

Tabela 1. Przygotowane plansze z wykresami – słupkowymi (A), słupkowymi ze zmienioną skalą osi Y (B), o promienistej architekturze (C) oraz liniowymi (D)



Źródło: opracowanie własne.

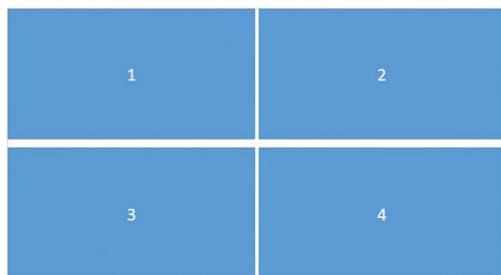
Trzecia plansza przedstawia wykresy słupkowe ze zmienioną skalą osi Y nawiązujące do statystyk szczepień (tab. 1B). Na następnej planszy umieszczono wykresy kołowe lub o architekturze promienistej (tab. 1C), przedstawiające wyniki wyborów na prezydenta z 2020 roku. Na ostatniej planszy zamieszczono wykresy liniowe, które często wykorzystuje się do zademonstrowania wzrostu sprzedaży produktów. Dane pobrane zostały ze strony internetowej gov.pl. Zwizualizowane dane dotyczyły zagadnienia przeciętnego trwania życia noworodka w Polsce (miasto/wieś) w latach 1990–2020.

W drugim etapie badani oglądali identyczne wykresy, lecz opublikowane na stronie WWW (zaprojektowanej przy użyciu platformy wix.com), a czas interakcji został wydłużony do 10 minut, dzięki czemu respondenci mogli dłużej eksplorować wizualizacje, przybliżać, skalować i oddalać. W ten sposób badani sami mogli kontrolować czas i poziom szczegółowości każdej planszy z wykresami.

Przebieg badań

Badanie dla obrazów statycznych przeprowadzone zostało na 13 respondentach, a wspierane było metodą ankietową. Po przeprowadzonej kalibracji użytkownikowi wyświetlano przez 20 sekund kolejny slajd zawierający cztery wykresy w układzie przedstawionym na rys. 1. Do każdego slajdu zadawano pytanie. Po tym czasie zadaniem badanego było udzielenie odpowiedzi. Respondentom łącznie wyświetlono 4 slajdy.

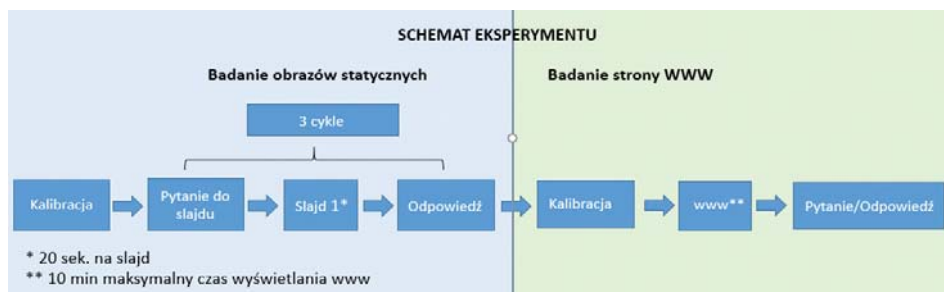
Rys. 1. Układ wykresów przewidziany dla pojedynczego slajdu



Źródło: opracowanie własne.

Każdy z etapów eksperymentu – z wyświetleniem obrazów statycznych i stron internetowych – wymagał ponownej kalibracji. Jak wspomniano wcześniej, respondenci mogli skalować witryny i dokładnie eksplorować każdy wykres. Z uwagi na problemy techniczne dotyczące niewystarczającej przepustowości łącza internetowego w tej części badania udział wzięło 12 uczestników.

Rys. 3. Schemat przeprowadzonego eksperymentu.



Źródło: opracowanie własne.

Wyniki

W pytaniu pierwszym (tab. 1A): **Czy liczba zgonów w wyniku COVID i chorób współistniejących jest większa/mniejsza niż liczba zgonów w wyniku COVID bez chorób współistniejących oraz czy wszystkie wykresy podają ten sam wynik**

poprawna odpowiedź brzmiała „mniejsza”. Odpowiedzi z dwóch badań w większości się pokrywały, ponieważ przy badaniu obrazów statycznych 5 osób odpowiedziało: „większa”, 4 osoby – „mniejsza”, 1 osoba stwierdziła: „ten sam wynik”, a 2 osoby udzieliły odpowiedzi: „nie wiem”. Przy badaniu strony internetowej odpowiedzi rozłożyły się następująco: 5 osób odpowiedziało: „większa”, 6 osób – „mniejsza” i 1 osoba stwierdziła: „ten sam wynik”.

W pytaniu dotyczącym slajdu 2 (tab. 1B) brzmiącym: **Czy w województwie lubelskim liczba wykonanych testów jest równa liczbie wykonanych testów z wynikiem pozytywnym?** – poprawna była odpowiedź „nie”. Przy badaniu obrazów statycznych 4 osoby odpowiedziały „tak”, 9 osób odpowiedziało „nie”. W przypadku drugiego etapu ze stroną internetową wszyscy jednogłośnie zadeklarowali „nie”.

Trzecie pytanie to: **Czy różnica pomiędzy 1 i 2 kandydatem (A. Dudą i R. Trzaskowskim) jest większa/mniejsza od różnicy pomiędzy 2 i 3 (R. Trzaskowski i S. Hołownia)?** (tab. 1C). W przypadku obrazów statycznych po 5 osób wybrało prawidłową odpowiedź „mniejsza” i nieprawidłową „większa”, natomiast „nie wiem” wybrały 3 osoby. Przy wyświetleniu strony internetowej wymienionych wyżej odpowiedzi odpowiednio udzieliło 6 i 7 osób.

Ostatnie pytanie dotyczyło grupowania danych na wykresach liniowych i brzmiało: **Jakie 2 głównie różniące się grupy widać na wykresach miasto-wieś czy kobiety-mężczyźni?** (tab. 1D). Odpowiedź „wieś-miasto” (prawidłowa) wybrało 6 osób, „kobieta-mężczyzna” – 5 osób oraz obie odpowiedzi padły dwukrotnie. Na etapie analizy strony internetowej tylko jedna osoba zmieniła swoją odpowiedź.

Zgromadzone dane w formacie CSV pozwoliły na szczegółową analizę zachowań użytkowników w procesie obserwowania poszczególnych slajdów. Pod uwagę wzięte zostały odpowiednio dane statystyczne dla ilości fiksacji w uśrednionym czasie ich trwania. Kolumny danych rejestrowanych podczas badania prezentuje rys. 4. Największym wyzwaniem było odkodowanie współrzędnych punktów fiksacji, które zapisywane były w jednym ciągu. Dzięki znajomości zaawansowanych technik przetwarzania danych dane dotyczące położenia i czasu każdej fiksacji przedstawiono w kolumnach. Za pomocą tabel przestawnych można było uzyskać dane liczby fiksacji, uśrednionego, maksymalnego i minimalnego czasu fiksacji, punktów rozrzutu fiksacji zarówno dla poszczególnego slajdu, jak i użytkownika.

Równolegle do badań ET narzędzie REALEYE pozwala na dokonanie pomiaru emocji respondentów na podstawie systemu kodowania FACS. Facial Actions Coding System (lub Facial Animation Parameters, związany z kompresją danych VIDEO w ramach standardu MPEG-4)¹⁷ jest systemem rozpoznawania emocji i detekcji ruchów twarzy. Przyjęty w standardzie MPEG-4 FAP model geometryczny obejmuje blisko 80 reakcji mięśni twarzy uzależnionych od reakcji użytkownika na dany bodziec w chwili jego rzeczywistej obserwacji. Wspomniany model na bazie wiedzy anatomicznej o skurczach mięśni lub ich grup opracowano już w 1970 roku. Celem

17 F. Lavagetto, R. Pockaj, *An efficient use of MPEG-4 FAP interpolation for facial animation at 70 bits/frame*, „IEEE Transactions on Circuits and Systems for Video Technology” 2001, vol. 11, s. 1085–1097, doi: 10.1109/76.954495.

tego eksperymentu było skategoryzowanie i rozróżnienie ruchów mimicznych¹⁸. W przyjętym modelu najistotniejsze grupy mięśniowe zostały skategoryzowane, zaś intensywność natężenia oceniana jest w 5-stopniowej skali, gdzie A oznacza minimalny, a E maksymalny stopień. Przykładowo, E4 w kodowaniu głównym wskazuje maksymalne opuszczenie brwi.

Rys. 4 Przykładowa tabela z danymi CSV

	A	B	D	F	I	K	L	M	O	P	Q	R
1	tester_id	tester_age	tester_display_g	tester_quality_g	test_id	test_duration_m	test_browser_vl	test_browser_he	test_raw_data	notes		
2	0029dc2-1384-	24	female	2	22d211b2-61b3-	26928	1366	768	[802.610.2030.(RealEye v8.13.10 - testerQualityGrader			
3	24d59ae7-258c-	24	female	1	ca1fa1ba-03cd-	25079	1536	864	[701.631.107.0.1332.497](899.795.161.0.1332.497)[1033.			
4	e005cdcd-7fe6-	22	male	5	3a2751dc-6dad-	25022	1536	864	[874.404.40.0.1000.415](100.442.70.0.1000.415)[878.421			
5	088e9d5d-490d-	20	female	4	457a7696-6558-	25056	1280	720	[573.348.74.0.1127.327](573.348.111.0.1127.327)[604.36			
6	35a3965a-530b-	22	female	3	27ec1f58-45ea-	25078	1536	864	[846.487.95.0.1062.540](853.508.140.0.1062.540)[964.49			
7	8118a84f-080d-4	17	female	1	15f842a-545a-	25387	1600	900	[990.418.450.0.802.525](1091.563.511.0.802.525)[953.46			
8	187a7d8a-5ccc-4	28	male	5	36d0f96d-3d06-	25035	1536	864	[808.380.62.0.741.452](895.309.93.0.741.452)[892.447.1;			
9	27b7b2af-744-	28	female	5	4d080eb0-8631-	25018	1536	864	[713.395.34.0.617.394](713.395.66.0.617.394)[695.411.94			
10	5638a1d2-5184-	22	female	5	7a971c2c-0208-	25053	1366	768	[768.881.150.0.1348.407](836.817.203.0.1348.407)[572.7			
11	91d506f5-fa35-4	55	female	4	7ac20444-640b-	25019	1536	864	[1127.635.50.0.1380.763](1127.635.79.0.1380.763)[1127.1			
12	06a0785b-b096-48fc-9ab4-c0c3f	unknown		5	8675fa1-cdc3-	25013	1536	864	[384.913.33.0.1507.830](384.913.65.0.1507.830)[384.913			
13	0d7d05d2-d7e9-	22	female	2	e68cd3a5-710a-	25101	1366	768	[840.131.150.0.1287.323](850.159.211.0.1287.323)[850.1			
14	a8e0dad1-d9e7-	23	female	4	ba48f6d2-90a7-	25005	1280	720	[883.562.71.0.627.863](834.521.106.0.627.863)[818.531-			
15	0029dc2-1384-	24	female	2	a9ad08de-af8a-	25116	1366	768	[1164.392.96.0.1146.345](1204.253.214.0.1365.324)[1156			
16	24d59ae7-258c-	24	female	1	807515dc-0587-	24991	1536	864	[1069.611.174.0.1343.400](1077.508.183.0.1343.400)[101			
17	e005cdcd-7fe6-	22	male	5	d142d27c-d7fd-	25001	1536	864	[905.643.35.0.1000.415](934.663.66.0.1000.415)[778.445			
18	088e9d5d-490d-	20	female	4	1bb2c50-919d-	25051	1280	720	[482.482.124.0.1127.327](505.522.158.0.1127.327)[484.5			
19	35a3965a-530b-	22	female	3	5a5b4c8f-abb5-	25020	1536	864	[329.656.63.0.1066.487](329.656.105.0.1066.487)[331.69			
20	8118a84f-080d-4	17	female	1	5ca8400f-554f-4	25005	1600	900	[73.122.70.0.604.431]-[73.122.146.0.604.431]-[24.116.23			
21	187a7d8a-5ccc-4	28	male	5	38d2014-5ead-	24998	1536	864	[773.621.35.0.763.443](773.621.69.0.769.443)[773.621.11			
22	27b7b2af-744-	28	female	5	15c36e83-9ead-	24998	1536	864	[764.592.31.0.617.394]-[764.592.68.0.617.394]-[770.466.11			
23	5638a1d2-5184-	22	female	5	6c63399e-5d33-	25021	1366	768	[758.670.68.0.1331.367](841.602.122.0.1331.367)[789.65			
24	91d506f5-fa35-4	55	female	4	9cdfa99-d854-	25015	1536	864	[881.589.33.0.1380.763](881.589.65.0.1380.763)[346.533			
25	06a0785b-b096-48fc-9ab4-c0c3f	unknown		5	16b0e0fc-83d7-	24982	1536	864	[964.646.31.0.1522.794](739.691.65.0.1522.794)[739.691			
26	0d7d05d2-d7e9-	22	female	2	b7815315-89ac-	25068	1366	768	[939.710.99.0.930.176](939.710.116.0.930.176)[422.218-			
27	a8e0dad1-d9e7-	23	female	4	41312919-0890-	25042	1280	720	[703.249.61.0.627.863](722.251.112.0.627.863)[690.254.1			
28	0029dc2-1384-	24	female	2	8c1c141f-4243-	25001	1366	768	[837.240.54.0.1365.324](1070.493.194.0.1365.324)[1070.			
29	24d59ae7-258c-	24	female	1	0607201c-c595-	25039	1536	864	[525.553.73.0.1337.482](402.526.168.0.1337.482)[403.51			
30	e005cdcd-7fe6-	22	male	5	1079a284-39de-	25006	1536	864	[799.559.32.0.1000.415](799.559.64.0.1000.415)[810.567			
31	088e9d5d-490d-	20	female	4	24d856f4-df45-4e	25016	1280	720	[640.790.74.0.1127.327](633.787.139.0.1127.327)[633.78			
32	35a3965a-530b-	22	female	3	71509160-78d9-	25058	1536	864	[765.401.58.0.1166.491](827.365.97.0.1166.491)[827.365.			
33	8118a84f-080d-4	17	female	1	54a79117-ff4f-4f	25114	1600	900	[678.571.92.0.1090.681](678.571.140.0.1090.681)[817.647.60			

Źródło: zasoby własne.

Z uwagi, iż badanie emocji w ramach oferowanego systemu jest realizowane jeszcze na etapie beta testów, zespół postanowił poddać analizie dane zgromadzone w części eksperymentu odbywającego się na obrazach statycznych. Oprogramowanie rozróżnia 3 główne stany: neutralność, radość oraz zaskoczenie. Dane, zgodnie ze wskazanym powyżej systemem, oznaczone są wartościami od 0 (brak emocji) do 1 (emocja wyraźna) (tab. 2). Łącznie w trakcie obserwacji obrazów statycznych zarejestrowanych zostało 61 tysięcy zdarzeń.

Na potrzeby analizy danych przyjęto, iż wartości powyżej 0,5 uznano za emocje wiarygodne i te poddawane były analizie – takich zdarzeń zarejestrowano ponad 18 tysięcy. Według wskazanego systemu kodowania największą liczbę emocji wzbudził slajd 3 związany z wyborami prezydenckimi i tu dominowało zaskoczenie: zliczono 2936 rekordów o wartości powyżej 0,5. Najmniej emocji towarzyszyły slajdowi 2 – 3742. Najbardziej występowała radość, zarejestrowana od 276 do 347 razy. Dominującym stanem towarzyszącym respondentom podczas obserwacji obrazów na wszystkich slajdach było zaskoczenie – 11320 w stosunku do neutralności na poziomie 5612 zarejestrowanych zdarzeń.

18 R. N. Devendra Kumar, A. Chakrapani, *Facial Expression Recognition System „Sentiment Analysis”*, „Journal of Advanced Research in Dynamical and Control Systems” 2017.

Tabela 2. Wyniki analizy systemu kodowania emocji

Slajd 1	Neutralny	Radość	Zaskoczenie	SUMA
suma emocji	9089	2032	4398	15519
średnia ilość emocji	699,15	156,31	338,31	
suma emocji >0,5	2111	347	2757	5215
średnia ilość emocji >0,5	162,38	26,69	212,08	
Slajd 2	Neutralny	Radość	Zaskoczenie	SUMA
suma emocji	9098	2036	4402	15536
średnia ilość emocji	699,85	156,62	338,62	
suma emocji >0,5	636	347	2759	3742
średnia ilość emocji >0,5	48,92	26,69	212,23	
Slajd 3	Neutralny	Radość	Zaskoczenie	SUMA
suma emocji	9102	2045	4715	15862
średnia ilość emocji	700,15	157,31	362,69	
suma emocji >0,5	2111	347	2936	5394
średnia ilość emocji >0,5	162,38	26,69	225,85	
Slajd 4	Neutralny	Radość	Zaskoczenie	SUMA
suma emocji	8675	1241	4304	14220
średnia ilość emocji	667,31	95,46	331,08	
suma emocji >0,5	754	276	2868	3898
średnia ilość emocji >0,5	58	21,23	220,62	

Źródło: opracowanie własne.

Analiza wyników

Policzenie odpowiedzi poprawnych i błędnych, których udzielił każdy badany, nie wykazało znaczących różnic, z wyjątkiem serii wykresów z pytaniem o liczbę wykonanych testów. Poprawnie odpowiedziało 9 osób wobec 4 odpowiedzi niepoprawnych. Dla reszty slajdów przeprowadzono analizy ilościowe z wykorzystaniem testów statystycznych. Tabela 3 prezentuje wyniki obliczeń podstawowych statystyk śledzenia ruchu gałek ocznych dla slajdu pierwszego w wersji statycznej i interaktywnej.

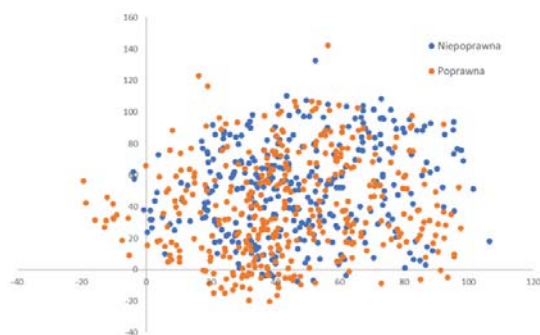
Tabela 3. Statystyki fiksacji dla pierwszego slajdu w wersji statycznej i interaktywnej (grup nie stanowiły te same osoby)

Odpowiedzi	Obrazy statyczne		Strona WWW	
	Liczba fiksacji	Średni czas trwania fiksacji, ms	Liczba fiksacji	Średni czas trwania fiksacji, ms
Niepoprawna	303	294.9	386	258.1
Poprawna	345	238.9	380	268.4

Źródło: opracowanie własne.

Test Studenta wykazał, iż różnica średniego czasu fiksacji w przypadku obrazów statystycznych w grupach z odpowiedziami poprawnymi i błędnymi jest istotna statystycznie: $p = 0,0001$, $\alpha = 0,05$. Odwrotną sytuację wykryto dla interaktywnej wersji slajdu, zatem różnica pomiędzy grupami nie była statystycznie istotna: $p = 0,424$, $\alpha = 0,05$. Takie sprzeczne wyniki zaobserwowano również dla następných slajdów. Natomiast w obu grupach nie wykazano istotnych zmian w położeniu fiksacji, jeśli weźmiemy pod uwagę dane statystyczne wszystkich respondentów. Rysunek 4 ilustruje mapę fiksacji dla jednego z eksperymentów, w którym kolorem oznaczono grupę odpowiedzi. Analiza wizualna potwierdza tę konkluzję dla wszystkich przeprowadzonych testów.

Rys. 4. Mapa fiksacji dla obu grup odpowiedzi w jednym z eksperymentów



Źródło: opracowanie własne.

W przypadku analizy FACS można przyjąć, iż obserwowane emocje w głównej mierze przyjmowały stan neutralny oraz zaskoczenia, przy czym zaskoczenie występowało najczęściej (o ile przyjmujemy wartości 0,5 jako istotniejsze dla przeprowadzonego testu). W przypadku całościowych odczytów bez progu istotności odbiorcy najczęściej zachowywali się neutralnie dla wszystkich zaprezentowanych slajdów. Radość jako najrzadziej występująca emocja potwierdza, iż w przypadku odczytu wykresów statystycznych może ona wynikać z udzielenia prawidłowej odpowiedzi na zadawane pytanie.

Dyskusja i wnioski

Oczywiste jest, iż czytanie wykresów zarówno uruchamia procesy percepcji wzrokowej, jak również aktywuje struktury w mózgu odpowiedzialne za myślenie analityczne, ponieważ błyskawicznie trzeba zestawiać wzajemne relacje wartości liczbowych oraz ich graficznych atrybutów. W badaniach nad przyswajaniem informacji i wyciąganiem wniosków z wizualizacji, jak dowodzi szereg opracowań (p. przegląd literatury), obiecującą metodą może być śledzenie ruchu gałek ocznych. Eksperyment eye trackingowy zaplanowano tak, żeby prześledzić różnice w percepcji wizualnej poprawnie

i błędnie zaprojektowanych wykresów. Poza tym respondentom zaoferowano dwie wersje materiału badawczego: statyczną oraz interaktywną, zamieszczoną na stronie WWW. Dodatkowym elementem badań była ankieta sprawdzająca poprawność odczytu danych na wykresach – uzyskane wyniki dały podstawę grupowania danych według poprawnych i błędnych odpowiedzi. Za pomocą testów statystycznych przeanalizowano różnice zachowań wzrokowych pomiędzy grupami. Różnice istotne statystycznie średniego czasu fiksacji wykryto w części eksperymentów i tylko dla jednej z wersji. Mapa fiksacji nie potwierdziła również pierwotnego założenia, iż osoby właściwie odczytujące informacje z wykresów postrzegają je inaczej niż reszta. Po ponownym przeanalizowaniu warunków eksperymentu autorzy rozważyli następujące przyczyny tego stanu: 1) nieprecyzyjne pytania, wywołujące wahania respondentów, 2) niewystarczająca rozdzielczość obrazów, uniemożliwiająca odczyt liczb skali, podpisów na osi, legend itp., 3) nadmiarowa liczba wariacji (zmiennych), które rozważono przy projektowaniu eksperymentu. Te spostrzeżenia zostaną wzięte pod uwagę w przyszłych badaniach. Z pewnością zbadanie czytelności wielomodalnych (czyli prezentujących wiele typów analiz – p. wstęp) wizualizacji powinno się przeprowadzić w prostszy sposób i z wykorzystaniem dashboardów. Końcowe pytanie dotyczyło zapamiętanego przez badanych wykresu, który sprawił im trudności w odczycie – okazało się, że najwięcej osób (9/13) wskazało diagram radarowy, na którym przedstawiono wyniki sondażu przed wyborami prezydenckimi.

Badanie przy wykorzystaniu kodowania FACS wykazało, że najczęstszą emocją było zaskoczenie. Trudno jednoznacznie wskazać, czy omawiana emocja była wynikiem różnorodności wykresów oraz ich złożoności, czy też była skutkiem interpretacji ich wyników. Radość, jako emocja występująca najrzadziej, może wskazywać na momenty, kiedy użytkownicy udzielali prawidłowej odpowiedzi na zadawane pytanie. Analiza kodowania FACS wymaga zatem dalszych badań oraz dopracowania zarówno metod ich dekodowania, jak również precyzyjnego przetestowania w obszarze interpretacji danych statystycznych. Warto rozważenia jest skorelowanie wyników z innymi metodami badania czytelności i użyteczności.

Bardzo ważnym kryterium okazał się być czas badania, ograniczony do 10 minut – ze względu na wykupioną licencję. Jak zanotowano, podczas badania na stronie internetowej nikt nie wykorzystał całego czasu: najdłuższe badanie z przyczyn technicznych, dotyczących przepustowości łącza, trwało 8 minut. Natomiast w przypadku obrazów statycznych badani skarżyli się na brak czasu na to, by mogli przyswoić i przetworzyć informacje zawarte w wizualizacjach. Z tego też powodu mocno odczuwali stres. Wynikiem powyższych przemyśleń stał się wyklarowany kierunek dalszych badań, w których można byłoby odpowiedzieć na pytanie, ile czasu potrzeba, żeby zrozumieć dane na wykresie. Autorzy zdają też sobie sprawę, że zarejestrowanie momentu, kiedy faza „widzę” przekształca się w „rozumiem”, wymaga zastosowania hybrydowych eksperymentów ET, z użyciem np. EEG i FACS. Innym odnotowanym problemem jest zaprojektowanie stres-testu na wzór testu stosowanego w architekturze informacji i UX, mającego na celu minimalizację negatywnych emocji badanych. Tym samym zastane warunki i okoliczności eksperymentu determinują kolejne ważne pytania badawcze.

Bibliografia

- Devendra Kumar R. N., Chakrapani A., *Facial Expression Recognition System „Sentiment Analysis”*, „Journal of Advanced Research in Dynamical and Control Systems” 2017.
- Duchowski A.T., *Eyetracking methodology*, New York 2017.
- Fry E., *Graphical literacy*, „Journal of Reading” 1981, nr 24 (5), s. 383–389.
- Glazer N., *Challenges with graph interpretation: A review of the literature*, „Studies in Science Education” 2011, nr 47 (2), s. 183–210.
- Goldberg J. H., Helfman J. I., *Comparing information graphics: a critical look at eye tracking. Proceedings BELIV’10 Workshop, Beyond Time and Errors in the Evaluation of Visualizations (Atlanta, USA)*, New York 2010, s. 71–78.
- Goldberg J. H., Helfman J. I., *Eye tracking for visualization evaluation: Reading values on linear versus radial graphs*, „Information Visualization” 2011, nr 10 (3), s. 182–195.
- Harsh J. A. et al., *„Seeing” Data Like an Expert: An Eye-Tracking Study Using Graphical Data Representations*, „CBE–Life Sciences Education” 2019, vol. 18, no. 3, <https://doi.org/10.1187/cbe.18-06-0102>.
- Lavagetto F., Pockaj R., *An efficient use of MPEG-4 FAP interpolation for facial animation at 70 bits/frame*, „IEEE Transactions on Circuits and Systems for Video Technology” 2001, vol. 11, s. 1085–1097, doi:10.1109/76.954495.
- Maltese A. V., Harsh J. A., Svetina D., *Data visualization literacy: Investigating data interpretation along the novice–expert continuum*, „Journal of College Science Teaching” 2015, nr 45 (1), s. 84–91.
- Ooms K., De Maeyer P., Fack V., *Study of the attentive behavior of novice and expert map users using eye tracking*, „Cartography and Geographic Information Science” 2014, nr 41 (1), s. 37–54.
- Osińska V., *Mapy nauki. Badania, potencjał oraz wyzwania w przykładach*, Toruń 2021.
- Osińska V., *Wizualizować potrafi każdy: czy jednak robi to poprawnie?*, „Biuletyn EBIB” 2021, nr 2 (197).
- Pantic M. i Rothkrantz L., *Automatic Analysis of Facial Expressions: The State of the Art*. „IEEE Transactions on Pattern Analysis and Machine Intelligence” 2001, vol. 22, s. 1424–1445, doi:1109/34.895976.
- Stolińska A., Bartkowski M., *Analiza procesu percepcji danych przedstawionych na wykresach*, „Edukacja – Technika – Informatyka” 2018, nr 2 (24), s. 56–57.

Źródła cyfrowe

- Ogama – strona główna otwartego oprogramowania do analizy ruchu gałek ocznych oraz urządzeń wskazujących. Realeye [on-line] <http://www.ogama.net/> – 29.07.2022.
- RealEye – firma oferująca badania eyetrackingowe przy wykorzystaniu kamer internetowych. Realeye, [on-line] <https://www.realeye.io/> – 19.07.2022.

„Niniejsze badania były przeprowadzone w ramach projektu Bitscope, finansowanego przez Narodowe Centrum Nauki w ramach programu CHIST-ERA IV, który otrzymał dofinansowanie na podstawie Umowy Finansowej nr 857925 w ramach Programu finansowania badań naukowych i innowacji Unii Europejskiej Horyzont 2020”.

Using eye tracking to analyze visual statistical information

Abstract

The correct interpretation of statistical data presented using visualization methods is a challenge in relation to modern media and communication. This is mainly due to poorly designed charts and often intentional manipulation of data presented graphically. Another problem may be the limited time that the recipients have at their disposal to familiarize themselves with the visualizations, e.g. via television. The task is additionally more difficult when the presented information concerns socially important matters. Users' perception of such specific visual information can be tested using a currently proven eye-tracking method.

The following article will present the results of a pilot study conducted on a group of 13 respondents.

Developed and thematically diversified surveys presented to the respondents were intended to help determine how the recipients read the charts, how long it takes them to read the data correctly, and whether the respondents are able to locate any errors.

The experiment was carried out using the eye-tracking technique supported by the survey method.

Keywords: eye tracking, information visualization, readability, FACS.