



UNIVERSITETI I EVROPËS JUGLINDORE
УНИВЕРЗИТЕТ НА ЈУГОИСТОЧНА ЕВРОПА
SOUTH EAST EUROPEAN UNIVERSITY

STUDIMET E CIKLIT TË DYTË

TEZA:

Vizualizimi interaktiv i të dhënave
tekstuale

Kandidati:
Argjend Bexheti, BSc

Tetovë,
2017

Mentori:
Doc. Dr. Visar Shehu



РЕПУБЛИКА МАКЕДОНИЈА
УНИВЕРЗИТЕТ „СВ. КИРИЛ И МЕТОДИЈ“ - СКОПЈЕ
ФИЛОЛОШКИ ФАКУЛТЕТ „БЛАЖЕ КОНЕСКИ“

УНИВЕРЗИТЕТСКА ДИПЛОМА

РЕКТОРОТ НА УНИВЕРЗИТЕТОТ „СВ. КИРИЛ И МЕТОДИЈ“ ВО СКОПЈЕ
И ДЕКАНОТ НА ФИЛОЛОШКИОТ ФАКУЛТЕТ „БЛАЖЕ КОНЕСКИ“ ВО СКОПЈЕ,
СО СВОИТЕ ПОТПИСИ И СО ПЕЧАТОТ НА УНИВЕРЗИТЕТОТ, ПОТВРДУВААТ ДЕКА

Љејза Насер Аделина

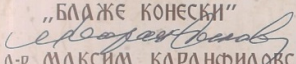
РОДЕН-А НА 29-XII 1984 ГОДИНА, ВО КИЧЕВО Р. МАКЕДОНИЈА
НА ДЕН 26-VI 2008 ГОДИНА ГО ЗАВРШИ ПОЛАГАЊЕТО НА ПРОПИШАНИТЕ ИСПИТИ
НА ДОДИПЛОМСКИТЕ СТУДИИ НА ФИЛОЛОШКИОТ ФАКУЛТЕТ „БЛАЖЕ КОНЕСКИ“,
НА ГРУПАТА АЛБАНСКИ ЈАЗИК И КНИЖЕВНОСТ

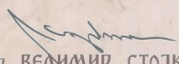
СО СРЕДНА ОЦЕНКА 8.93 И СТЕКНА ПРАВО НА СТРУЧЕН НАЗИВ

ДИПЛОМИРАН ПРОФЕСОР ПО АЛБАНСКИ ЈАЗИК И КНИЖЕВНОСТ

ВРЗ ОСНОВА НА ТОА, УНИВЕРЗИТЕТОТ „СВ. КИРИЛ И МЕТОДИЈ“ ВО СКОПЈЕ
МУ-И ИЗДАВА ДИПЛОМА ЗА ЗАВРШЕНИ ДОДИПЛОМСКИ СТУДИИ,
СО КОЈА КАНДИДАТОТ-КАТА ГИ СТЕКНУВА СИТЕ ПРАВА ШТО МУ-И ПРИПАГААТ
СПОРЕД ЗАКОНИТЕ НА РЕПУБЛИКА МАКЕДОНИЈА

ВО СКОПЈЕ, 15-XII 2008 ГОДИНА, БР. 9331

ДЕКАН
НА ФИЛОЛОШКИОТ ФАКУЛТЕТ
„БЛАЖЕ КОНЕСКИ“

Проф. д-р МАКСИМ КАРАНФИЛОВСКИ

РЕКТОР
НА УНИВЕРЗИТЕТОТ

Проф. д-р ВЕЛИМИР СТОЈКОВСКИ

DEKLARATA E LEKTORIT

Unë Adelina Fejza, studente e diplomuar në fakultetin e Gjuhës shqipe dhe letërsisë, me përgjegjësi të plotë deklaroj se kjo tezë magjistrature është lekturuar nga ana ime.

Adelina Fejza

DEKLARATË

Unë Argjend Bexheti, me përgjegjësi të plotë deklaroj se jam autor origjinal i këtij disertacioni dhe puna e paraqitur këtu ka përbërje origjinale. Informacionet e marra nga puna e të tjerëve janë cituar dhe referencat janë dhënë në listën e burimeve.

Argjend Bexheti

FALËNDERIME

Realizimi i kësaj teme do të kishte qenë i pamundur po të mos ishte përkrahja dhe ndihma e më shumë njerëzve, para së gjithash, mentorit tim Dr. Visar Shehut, këshillat, idetë dhe kritikë konstruktive të së cilit ndihmuan në ngritjen e cilësisë së kësaj teme. Dr. Visar Shehu ishte personi tek i cili gjithmonë e gjeja përgjigjen e duhur të çdo vështirësie që hasja gjatë procesit të krijimit të temës. Nuk do të kisha mundur të imagjinoj mentor më të mirë dhe më të përsosur.

Gjithashtu do të falënderoja Dr. Bujar Raufin dhe Dr. Florije Ismailin për komentet dhe sugjerimet profesionale të cilat ndihmonin në sofistikimin e këtij punimi. Ishte kënaqësi dhe nder për mua që ju kisha anëtar të stafit dhe që pata mundësi të punojë nën observimin tuaj.

Do falënderoja për zemërsisht familjen time. Është i pamundur të përshkruhet roli i secilit nga to, duke filluar nga prindërit e mi, të cilët më ofruan besimin e plotë dhe përkrahjen e vazhdueshme, e deri te shoqja dhe fëmijët të cilët i pata si inspiruesit kyç gjatë të gjithë kësaj periudhe.

ABSTRAKTI

"Një fotografi vlen sa një mijë fjalë" - është idioma e cila në mënyrë implicite përshkruan rolin dhe rëndësinë e vizualizimit. Tendenca e rritjes enorme të numrit të dokumenteve elektronike (uebfaqet, blogjet, artikujt, si dhe publikimet e ndryshme) mundëson sasi të mëdha informacionesh. Një situatë e këtillë sigurisht që hap shumë sfida të reja, ku mes tjerash është gjetja efektive e informacioneve të nevojshme nga një koleksion i madh dokumentesh, duke mos shqyrtuar detajet e çdo dokumenti të koleksionit. Poashtu, sfida tjetër shumë e rëndësishme është edhe përdorimi i mirëfilltë i teknikave të vizualizimit, të cilat bëjnë "shkrirjen" e të dhënave në një përmbledhje grafike, e cila shumë më lehtë do të kuptohet, krahasohet dhe mësohet nga përdoruesi.

Vizualizimi i të dhënave paraqet sintezë të artit dhe shkencës. Këtë besoj që do e dëshmoj nëpërmjet kësaj teme, e cila si qëllim ka analizën e medimeve të Maqedonisë si dhe vizualizimin e të dhënave tekstuale nëpërmjet të së cilave do të arrihen njohuri të reja. Për këtë analizë kemi shfrytëzuar dy aplikacione të ndryshme, me qëllim që të përdorim më tepër teknika vizualizuese.

ABSTRACT

"A picture is worth a thousand words" – is an idiom which implicitly describes the role and importance of visualization. The enormous growing tendency of the electronic documents (websites, blogs, articles and various publications) provides a huge amount of information. Such situation introduce a new challenges, among others discovery of useful knowledge from large document collections effectively without completely going through the details of each document in the collection. Also, another very important challenge is the authentic use of the visualization techniques, which make “melting” of the data into a visual form, for providing a better understanding, comparing and learning.

Data Visualization is synthesis of art and science. I believe that I can prove it in this thesis, which objectives are the media analysis in Macedonia and also the textual data visualization from which we could get new information. For this purpose, we have used two different applications, for providing various visualization techniques.

АБСТРАКТ

"Една слика е еднаква на илјада зборови" – е идиома која на имплицитен начин ја опишува улогата и значањето на визуализацијата. Енормното зголемување на бројот на електронските документи (вебстраниците, блоговите, артиклите, како и други различни објавувања) овозможува голема количина на информации. Една ваква ситуација отвара нови предизвици, меѓу кои и ефективното наоѓање на вредни информации од голема колекција на документи, без да се разгледуваат деталите на секој документ од колекцијата. Исто така, друг важен предизвик претставува и автентичното употребување на техниките на визуализацијата, кои праваат "топење" на податоците на визуална форма, која би и помогнала на корисникот за полесно да ги разбира, спореди и учи.

Визуализацијата на податоците претставува синтеза на уметноста и науката. Ова верувам дека ќе го докажам со ова теза, која како цел ја има анализата на медиумите на Македонија како и визуализацијата на текстуалните податоци со помош на кои ќе се добиваат нови знаења. За ова анализа сме користеле две различни апликации, со цел да се користат што повеќе техники на визуализација.

PËRMBAJTJA

FALËNDERIME	V
ABSTRAKTI.....	VI
ABSTRACT.....	VII
АБСТРАКТ	VIII
LISTA E FIGURAVE.....	XI
LISTA E SHKURTESAVE	XII
1. HYRJE.....	13
1. 1 Lënda e hulumtimit	13
1. 2 Qëllimet e hulumtimit	14
1. 3 Hipotezat	15
1. 4 Rëndësia e punimit.....	15
1. 5 Struktura e temës.....	16
2. VIZUALIZIMI I TË DHËNAVE	17
2. 1 Faktorët e procesit të vizualizimit	17
2. 2 Shfrytëzuesi si faktor i procesit të vizualizimit	19
2. 2. 1 Rregulla bazuar në perceptim	20
2. 2. 2 Atributet vizuale	22
2. 2. 3 Objektet vizuale bazike	22
2. 3 E dhëna si faktor në procesin e vizualizimit	24
2. 3. 1 Struktura e të dhënave.....	24
2. 3. 2 Tipet e variablave.....	25
2. 4 Objektivi si faktor në procesin e vizualizimit	26
2. 4. 1 Klasifikimi i qëllimeve.....	26
2. 4. 2 Vizualizimet e rekomanduara.....	27
2. 4. 3 Grafikone më të avansuara	29
2. 5 Efekti i interaktivitetit në vizualizim.....	36
2. 6 Vizualizimi efektiv.....	38
2. 6. 1 Këndvështrime në sferën e vizualizimit efektiv.....	38

2. 6. 2 Principet e vizualizimit efektiv.....	39
3. VALIDIMI DHE EVALUIMI I VIZUALIZIMIT.....	41
3. 1 Sfidat për vlerësimin e efikasitetit të vizualizimit.....	41
3. 2 Modeli i ndërthurur për dizajnimin dhe validimin e vizualizimit.....	42
3. 3 Standardi DIN EN ISO 9241-12 "Information Presentation"	44
4. VIZUALIZIMET E ORIENTUARA NË INTERNET	46
4. 1 Pse vizualizim të orientuar në internet?	46
4. 2 Mjetet e vizualizimit të orientuar në internet	47
4. 2. 1 Java Applet.....	48
4. 2. 2 Flash	48
4. 2. 3 Scalable Vector Graphics (SVG)	49
4. 2. 4 HTML 5 Canvas.....	50
4. 2. 5 JavaScript.....	51
4. 3 JavaScript dhe internet shfletuesit si platformë vizuale	51
4. 3. 1 Modeli referues dhe JavaScript	52
4. 3. 2 Interaksioni dhe animacioni	53
4. 3. 3 Kompatibiliteti dhe aksesibiliteti	53
4. 3. 4 Client-side Frameworks.....	54
5. PLATFORMA D3.js.....	56
5. 1 Qëllimet kryesore.....	56
5. 2 Krijimi i dizajnit.....	57
5. 2. 1 Selektorët	57
5. 2. 2 Operacionet me të dhënat	58
5. 2. 3 Interaksioni dhe Animacioni.....	58
6. VIZUALIZIMI I TË DHËNAVE TEKSTUALE.....	60
6. 1 Rezultatet e fituara.....	66
7. PËRFUNDIMI.....	75
7. 1 Udhëzim për zhvillim të mëtutjeshëm	76
LITERATURA.....	77

LISTA E FIGURAVE

Fig. 1: Shembull i pipeline-it të vizualizimit.....	19
Fig. 2: Ndërtimi i bloqeve për vizualizim të informacioneve kuantitative	21
Fig. 3: Katër llojet e objekteve vizuale bazike	22
Fig. 4: Hierarkia e strukturës së të dhënave	24
Fig. 5: Hierarkia e tipeve të ndryshme të variablave	25
Fig. 6: Matrica e grafikoneve rekomanduese	28
Fig. 7: Shembull grafikoni me shirite të grupuara	29
Fig. 8: Shembull grafikoni Stacked Bar	30
Fig. 9: Shembull grafikoni Box Plot	30
Fig. 10: Shembull grafikoni histogram	31
Fig. 11: Shembull i grafikonit me boshte paralele	31
Fig. 12: Shembull grafikoni heatmap	32
Fig. 13: Shembull grafikoni Scatter Plot.....	33
Fig. 14: Shembull të diagramit network	33
Fig. 15: Shembull të diagramit Sankey	34
Fig. 16: Diagrami Sankey i Çarls Minard	35
Fig. 17: Shembull të vizualizimit cirkular.....	35
Fig. 18: Definicioni i vizualizimit efektiv	39
Fig. 19: Modeli i ndërthurur për dizajnim dhe validim të vizualizimit	42
Fig. 20: Kërcënimet dhe validimet në modelin e ndërthurur	44
Fig. 21: Statistikë e kategorisë së kërcënimeve në Flash	49
Fig. 22: Modeli referues i adaptuar për JS vizualizim të orientuar në internet.....	52
Fig. 23: Aplikacioni 1	61
Fig. 24: Web aplikacioni wordle.net	62
Fig. 25: Aplikacioni 2	63
Fig. 26: Web aplikacioni mkweb.bcgsc.ca/tableviewer/visualize/	65
Fig. 27: Frekuenca e përdorimit të emrave dhe mbiemrave të politikanëve të R. M. - 1.....	66
Fig. 28: Frekuenca e përdorimit të emrave dhe mbiemrave të politikanëve të R. M. - 2.....	67
Fig. 29: Fjalët më të përdorura në mediumin “TV-Shenja”	68
Fig. 30: Frekuenca e përdorimit të fjalëve në platformën e VMRO-DPMNE.....	69
Fig. 31: Frekuenca e përdorimit të fjalëve në platformën e LSDM	69
Fig. 32: Trendi i përdorimit të fjalëve.....	71
Fig. 33: Frekuenca e përdorimit të termit VMRO-DPMNE në përmbajtjet e katër medimeve	72
Fig. 34: Vizualizimi nëpërmjet teknikës “word cloud”	73
Fig. 35: Vizualizimi nëpërmjet teknikës “circos”	73

LISTA E SHKURTESAVE

VOI: Vizualizimet e orientuara në internet

SVG: Scalable Vector Graphics

API: Application Programming Interface

D3: Data-Driven Document

JS: JavaScript

GUI: Graphical User Interface

HTML: HyperText Markup Language

VML: Vector Markup Language

DOM: Document Object Model

CSS: Cascading Style Sheets

JSON: JavaScript Object Notation

1. HYRJE

Vështirë mund të definohet origjina dhe zanafilla e saktë e vizualizimit, edhe pse besohet se përdorimi i saj lidhet me periudhë shumë të hershme, kur nuk ka ekzistuar as edhe gjuha e shkruar. Edhe pse zbulimi i kësaj llogaritë shumë i hershëm, zhvillimin më esencial duhej ta priste në shekullin XVII, me zbulimin e "Sistemit Kartezian" nga Rene Dekart, i cili mundësonte paraqitjen e të dhënave duke përdorur dy dimensione. Kontribut të rëndësishëm në zhvillimin e vizualizimit të të dhënave u dha edhe në shekujt në vijim. Periudha e shekullit XIX e deri në fillimet e hershme të shekullit XX mund të identifikohet edhe si "Periudha e artë e vizualizimit". Si personazhe më të dalluar të kësaj periudhe janë Uilliam Pleifer, i cili njihet si krijuesi i llojeve më të përhapura të grafikoneve (shiritore, linjore, tërësore, sipërfaqësore) të cilat edhe sot e kësaj dite përdoren, Xhon Snou me krijimin e hartës së kolerës, Çarls Minard krijoi hartën ku ilustron ekspedita ushtarake të Napoleon Bonapartës ndaj Rusisë në vitin 1812, Luixhi Peroco me krijimin e stereogramit tre dimensional, etj.

Gjatë pesë dekadave të fundit, vizualizimi i të dhënave si objekt studimi ka pasur interes të madh, sigurisht edhe duke iu falënderuar zhvillimeve të mëdha harduerike dhe softuerike të cilat ndihmojnë në renderimin e imazheve komplekse. Ditëve të sotme, vizualizimi i të dhënave mund të konsiderohet si degë e statistikës deskriptive moderne, dhe përfshin studimin dhe zhvillimin e teknikave të cilat konvertojnë të dhënat në imazhe. Përpos imazheve, mundësia e interaksionit me vizualizimin, si dhe përshtatja dinamike e tij, definojnë sfidën e ardhshme të vizualizimeve.

1. 1 Lënda e hulumtimit

Sasia e të dhënave që publikohen në internet, qofshin këto përmes medimeve të ndryshme online, gazetave online apo blogjeve nga dita në ditë rriten në mënyrë intenzive. Të gjitha ato informacione që na ofrohen nga to, kanë impakt të madhë për njeriun, lidhur me novitetet që e rrethojnë, veprimtaritë e ndryshme, etj. Duke qenë të

servuar me gjithë këtë sasi të madhe të informacioneve, sigurisht që do ndodh të mendojmë se sa reale janë të gjitha këto informacione?!

Viteve të fundit kemi qenë dëshmitar se në shumicën e medimeve të R. Maqedonisë, kemi mungesë të madhe të objektivitetit. Shpesh ndodh që në përmbajtjet e tyre të vërejmë “pseudo realitetin”, zhargonin special, fshehjen e ndodhive dhe shumë e shumë veprime tjera jo të rregullta.

Janë bërë shumë monitorime mediatike në R. Maqedonisë që dëshmojnë për këto. Ajo që do të jetë si sfidë e kësaj teze do të jetë vizualizimi i këtyre anomalive mediatike. “Një shikim vlen sa një mijë fjalë” është një nga proverbat më esenciale që përshkruan rëndësinë e vizualizimit të të dhënave. Të shikosh dhe kuptosh një fotografi është rezultat i sensit natyral të njeriut, ndërsa për të kuptuar dhe mësuar të dhëna të ndryshme qofshin ato tekstuale apo numerike nevojiten aftësi të caktuara që fitohen që nga moshat e hershme.

Në vendin tonë përdorimi i vizualizimit interaktiv është shumë i rrallë edhe pse gjithmonë do ekzisojë nevoja për tu përdorur gjatë reprezentimit të të dhënave, posaçërisht në periudhë ku shumica e njerëzve kanë qasje në internet.

1.2 Qëllimet e hulumtimit

Te shumica e medimeve online në R. Maqedonisë vërehet anshmëri tek përmbajtjet e materialeve të tyre. Përveç kësaj, si qëllime tjetra plotësues që do të duhen të shqyrtohen janë analiza e përmbajtjeve të medimeve të informacionit, devijimet e informacioneve të marrura, etj.

Qëllimet e hulumtimit janë:

- ✓ Analiza e medimeve online të shkruara
- ✓ Hulumtimi i metodave për vizualizim interaktiv të tekstit të analizuar
- ✓ Propozimi i metodave dhe vizualizimi i informacionit nga teksti i mbledhur

1.3 Hipotezat

Për këtë hulumtim, ne do të parashtrojmë disa hipoteza të cilat gjatë përpunimit të kësaj teme do të vërtetohen në plotësimin ose mosplotësimin e tyre.

Këto hipoteza janë:

- ✓ **Hipoteza I:** Analiza automatike e medimeve në Republikën e Maqedonisë mund të realizohet dhe të nxjerrë në pah mangësitë e medimeve

Për të mbështetur hipotezën e parë kemi këto dy hipoteza shtesë:

- Të gjithë publikimet e medimeve online në Maqedoni mund të mbledhen dhe të analizohen statistikisht
- Vizualizimi i të dhënave mund të nxjerrë në pah njëanshmërinë e medimeve

1.4 Rëndësia e punimit

Qëllimi kryesor i vizualizimit të të dhënave është rritja e njohjes së fenomeneve të ndryshme të përshkraura nëpërmjet të dhënave. Rritja enorme e të dhënave si dhe e teknologjive përmes së cilave ne depërtojmë dhe përpunojmë ato, bëjnë vizualizimin e të dhënave “medium komunikimi të së ardhmes”. Zanafilla e vizualizimit të të dhënave daton prej periudhave shumë më të hershme, duke u përdorur në forma dhe metoda modeste të asaj kohe. Ditëve të sotme me zhvillimin e hovshëm të pajisjeve dhe teknologjive, gjithmonë e më shumë shqyrtohen mundësi se si atë përveç objektivit primar që ka t’i shtohen cilësi të ndryshme për nga mënyra e paraqitjes, të sigurohet qasshmëri nga pajisje dhe teknologji, dinamizëm (animacion) gjatë reprezentimit të të dhënave, interaktivitet, etj.

Rëndësia e këtij punimi mbështetet në këto pika:

- Analizën automatike të medimeve online në Republikën e Maqedonisë
- Vizualizimi i rezultateve të cilat do të tregojnë anshmërinë e medimeve

1. 5 Struktura e temës

Tema është e organizuar në shtatë kapituj, prej ku në *Kapitullin 1* kemi hyrjen, ku kemi tentuar të bëjmë një shikim të shkurtër historisë së vizualizimeve, kronologjikisht kemi përmendur protagonistët më të rëndësishëm të këtij drejtimi, si dhe sfidat e reja të vizualizimit.

Në *Kapitullin 2* do të sqarohet procesi i vizualizimit të të dhënave, ku rëndësi të posaçme do t'i kushtohet rolit të perceptimit human. Gjithashtu, në këtë kapitull do të përmendim interaktivitetin në vizualizime, si dhe do të jepen rekomandime përmes së cilëve do të mundësohet krijim i vizualizimit efektiv.

Sfidat kryesore të matjeve dhe evaluimeve të vizualizimeve efektive do të jepen në *Kapitullin 3*. Do të jepet modeli për dizajnim dhe validim të vizualizimit, si dhe do të përshkruhet standardi DIN EN ISO 9241-12, i cili shërben për prezentimin korrekt të informacionit.

Në *Kapitullin 4* do të përshkruajmë në mënyrë sipërfaqësore vizualizimet e orientuara në internet (VOI), fillimisht duke përmendur Java Applets, Flash, SVG, HTML 5 Canvas, JavaScript, si dhe cilësi plotësuese vizualizimi si interaksioni dhe animacioni.

Në *Kapitullin 5* kemi dhënë përshkrim për librarin më të sofistikuar për krijimin e vizualizimeve interaktive dhe dinamike të orientuara në internet – D3.js, përparsitë e përdorimit të saj si dhe kemi përmendur për atë se si janë bërë vizualizimet e orientuara në internet disa vite më parë.

Në *Kapitullin 6* do të trajtohet vizualizimi i të dhënave tekstuale, ku gjithashtu do të jepen lloje të ndryshme të grafikonëve që do të gjenerohen duke u bazuar në analizat e medimeve të Maqedonisë. Këto analiza do të bëhen me ndihmën e dy aplikacioneve si dhe duke përdorur uebaplikacione tjera.

Kapitulli 7 do të paraqesë pjesën ku do të vërtetohen hipotezat e parashtruara, si dhe do të ofrohen ide për zhvillime të më tutjeshme.

2. VIZUALIZIMI I TË DHËNAVE

Qëllimi kryesor i vizualizimit të të dhënave është rritja e njohjes së fenomeneve të ndryshme të përshkruara nëpërmjet të dhënave. Kjo rritje enorme e të dhënave si dhe e teknologjive përmes së cilave ne depërtojmë dhe përpunojmë ato, bëjnë vizualizimin e të dhënave “medium komunikimi të së ardhmes”. Zanafilla e vizualizimit të të dhënave daton prej periudhave shumë më të hershme, duke u përdorur në forma dhe metoda modeste të asaj kohe. Ditëve të sotme me zhvillimin e hovshëm të pajisjeve dhe teknologjive, gjithmonë e më shumë shqyrtohen mundësi se si atë përveç objektivit primar që ka, t’i shtohen edhe cilësi tjera, si qasshmëri nga pajisje dhe teknologji të ndryshme, dinamizëm gjatë paraqitjeve të të dhënave, interaktivitet, etj.

Në këtë kapitull në mënyrë eksplicite do të bëhet sqarimi i procesit të vizualizimit të të dhënave. Do të përshkruhen faktorët e procesit të vizualizimit, roli i perceptimit human, përparësitë e inkuadrimin të interaktivitetit në vizualizim, si dhe do të ofrohen rekomandime nëpërmjet të së cilëve do të mundësohet krijim i vizualizimit efektiv.

2.1 Faktorët e procesit të vizualizimit

Vizualizimi mund të llogaritet si mekanizëm për transformimin e të dhënave dhe objektivave tona në format më vizual dhe më intuitiv (Ward, Grinstein, & Keim, 2010). Bazuar në këtë ide, mund të nxirren tre faktorë kryesorë të procesit të vizualizimit: *e dhëna, objektivi dhe shfrytëzuesi*.

E dhëna mund të jetë e tipeve të ndryshme, më detajisht e përshkruar në pjesën 2.3. Bazuar në tipin e të dhënës, teknikë e caktuar vizualizimi duhet të zgjedhet për paraqitje apo reprezentim më të mirë të të dhënave. Ky proces quhet mapim. Në këtë proces të dhënat dhe atributet e tyre zëvendësohen me objekte grafike siç janë pika, vija, shirite, pjesë, etj., bashkangjitur këtu cilësi të tyre si madhësinë, pozicionin, orientimin dhe ngjyrën. Për të patur një mapim më efektiv, është shumë e rëndësishme të dihet se

si njerëzit vizualisht e perceptojnë informacionin. Në pjesën 2.2 do të jepet sqarim më i detajuar në sferën e perceptimit human.

Objektivi apo qëllimi mund të llogaritet si faktori kyç gjatë procesit të vizualizimit, definimi i sakt i së cilit mundëson zgjedhjen e duhur të teknikës vizualizuese, përmes së cilës do të paraqitet apo reprezentohet informacioni. Përdorimi i teknikës së duhur të vizualizimit u ndihmon njerëzve që në mënyrë shumë më të lehtë dhe më të qartë të kuptojnë zgjidhjen e problemit. Objektivi si faktor i procesit të vizualizimit, paraqet urë lidhëse me të dhënë si faktor edhe me shfrytëzuesin.

Shfrytëzuesi poashtu paraqet faktor të rëndësishëm të procesit të vizualizimit. Pjesa 2.2 shënon për perceptimit human, konkretisht përshkruan mënyrën se si vizualizimin e percepton truri i njeriut, si dhe rregullat të cilat duhet të merren parasysh gjatë krijimit të vizualizimeve.

Është e pamundur të krijohet një teknikë universale vizualizimi e cila do të ishte efektive për më shumë raste. Dallimet midis shfrytëzuesve, si, aftësitë perceptuale të tyre, kultura, preferencat, etj., si dhe tipet e të dhënave janë disa nga arsyet që e pamundësojnë atë.

Një tjetër komponentë plotësuese e procesit të vizualizimit është interaksioni. Vizualizimi modern tani më konsiderohet si proces dinamik, e jo statik siç ishte më përpara kur e gjenim të shtypur në letër. Kjo teknikë u mundëson shfrytëzuesve të bashkëveprojnë me vizualizimin në forma dhe mënyra të ndryshme. Përdorimi i interaktivitetit në vizualizimet tona nuk është garancë suksesi, por jep shans për të krijuar ato më efektiv (Ward, Grinstein, & Keim, 2010). Ky fakt është i cekur në pjesën 2.5, ku diskutohet për rolin e interaktivitetit për krijimin e vizualizimeve më të mira.

Në mënyrë skematike procesi i vizualizimit mund të reprezentohet me ndihmën e pipeline-it të vizualizimit, i dhënë në figurën më poshtë.

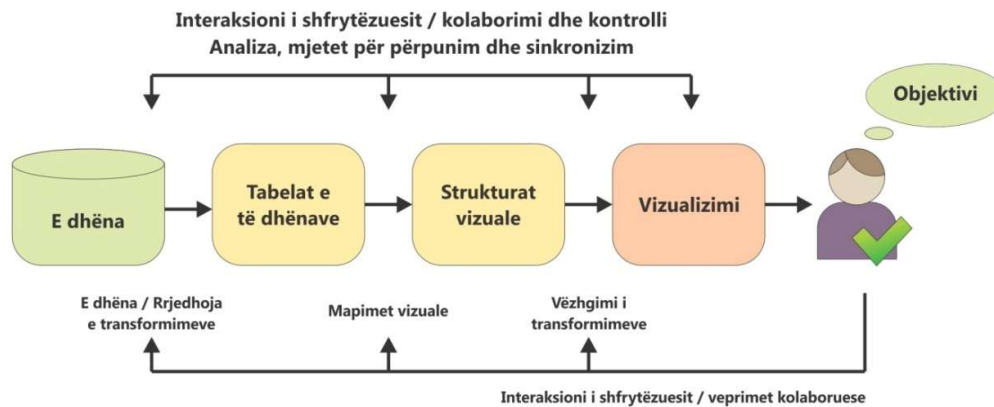


Fig. 1: Shembull i pipeline-it të vizualizimit

Nga kjo që shihet në Fig. 1, në mënyrë përmbledhëse mund të thuhet se pipelin-i i vizualizimit paraqet proces ku merret e dhëna hyrëse e papërpunuar dhe kthehet vizualizimi si informacion dalës. Gjatë këtij procesi, bëhet po ashtu edhe përgatitja dhe filtrimi i të dhënave dhe më pas bëhet mapimi ose zgjedhja e teknikës më të përshtatshme për vizualizim.

2. 2 Shfrytëzuesi si faktor i procesit të vizualizimit

Viteve të fundit, gjithmonë e më shumë rritet përdorimi i vizualizimeve si mjete ndihmëse për zgjerimin e njohurive, meqenëse shfrytëzuesit vazhdimisht mbështeten në to për zgjidhje të problemeve më komplekse. Edhe pse ngjyra dhe teoritë perceptuese mbeten si atributet primare të saj të cilat kërkojnë vëmendje të posaçme për krijim të një dizajni të mirë, zhvillimi i hovshëm i pajisjeve dhe teknologjive të ndryshme mundëson përdorimin e më shumë detajeve tjera të cilat mundësojnë njohje më të përshtatshme të problemeve më komplekse.

Të gjithë ne mendojmë ndryshe. Janë disa aspekte, si, eksperiencia, aftësia njohëse, preferenca, kultura, si dhe shumë tjera, të cilat ndikojnë gjatë mënyrës se si ne e kuptojmë një problem të caktuar dhe si ne i përgjigjemi atij.

2. 2. 1 Rregulla bazuar në perceptim

Aftësitë dhe limitimet e sistemit vizual njerëzor janë faktorë shumë të rëndësishëm të të gjitha vizualizimet. Në rast se qëllimi i vizualizimit është që me saktësi të përcjellë informacionin me fotografi, atëherë duke u bazuar në këtë konstatim vlen të konsiderohet rëndësia e aftësisë perceptuale (Ward, Grinstein, & Keim, 2010).

Stephen Few në veprën e tij "Now You See It." përshkruan disa fakte të rëndësishme rreth mënyrës se si njerëzit pranojnë dhe përpunojnë informacionet vizuale. Bazuar në ato fakte, ai përcakton tre rregulla të bazuara në perceptim, të cilat duhet të konsiderohen për krijimin e një vizualizimi efektiv (Few, Now You See It., 2009).

Fakti i parë tregon se njerëzit nuk mund të vërejnë gjithçka që shohin. Perceptimi vizual është selektiv, dhe këtë e tregon edhe rregulla e parë perceptuese e dhënë nga S. Few. e cila thotë:

1. "Vizualizimi duhet të paraqesë të dhënat në atë formë që ajo që është më e rëndësishme të afishohet, krahas asaj me më pak."

Fakti i dytë tregon se sytë e njerëzve preferojnë më shumë modelet (paternat) e njohura. Njerëzit shohin atë që dinë dhe atë që mund të imagjinojnë. Kështu që, rregulla e dytë perceptuese sipas S. Few thotë:

2. "Vizualizimi është i mirë atëherë kur tregon informacionin përmes modeleve të cilat janë të njohura dhe të parashikuara."

Fakti i tretë i tregon për rolin e memorjes në perceptimit njerëzor (human). Në trurin e njeriut ekzistojnë tre tipe të memorjes: **memorja sensoriale** që memoron atë që sheh syri dhe dërgon atë në **memorjen afatshkurte** (e njohur edhe si memorja punuese) për përpunim e informacioneve të kuptueshme ose të refuzoj ato të pakuptueshmet. Informacionet e kuptueshme dhe të nevojshme për përdorim të mëtutjeshëm nga memorja afatshkurte dërgohen në **memorjen afatgjate**.

Memorja punuese mund të llogaritet si më e rëndësishmja në këtë proces, nga arsyeja se aty bëhet perceptimi i vërtetë. Kjo memorje për nga kapaciteti është shumë e limituar, dhe mund të memorojë nga 3 - 7 pjesë të të dhënave të reprezentuar për kohë. Kështu që, rregulla e tretë perceptuese sipas S. Few thotë:

3. "Vizualizimi duhet t'i shërbejë memorjes punuese".

Shtuar rregullave të bazuara në perceptim, S. Few ka krijuar edhe një ilustrim të quajtur "Ndërtimi i bloqeve për vizualizim të informacioneve kuantitative", i treguar në figurën e mëposhtme:



Fig. 2: Ndërtimi i bloqeve për vizualizim të informacioneve kuantitative

Perceptimi vizual përbëhet nga objekte vizuale bazike (pika, vija, shirite) dhe attribute vizuale (gjatësia, ngjyra, forma). Gjithmonë gjatë procesit të vizualizimit duhet të zgjedhen objektet dhe atributet të cilët do të reprezentojnë në mënyrë më të mirë të dhënën.

Arsyetimi kuantitativ paraqet procesin ku njerëzit vërejnë dhe bashkëveprojnë me të dhënat. Këtu bëhet krahasimi dhe shqyrtimi i relacioneve kuantitative (seritë kohore, shpërndarjet, korrelacionet).

Vizualizimi i informacionit paraqet krijimin e vizualizimit, ku relacionet e rëndësishme kuantitative të të dhënave paraqiten apo reprezentohen me modele vizuale (paterna).

Bloku i fundit, **të kuptuarit e vizualizimit**, eviton keqinterpretimet dhe mundëson njohuri të reja të cilat mund të shërbejnë në vendimarrje. Në rast se bëhet një përmbledhje të kësaj që përmendëm më lart, atëherë shihet qartë se **Perceptimi vizual** mund të llogaritet si blloku kryesorë në procesin e vizualizimit. Kështu që, vlen të shpenzohen më shumë kohë gjatë përzgjedhjes së objekteve vizuale dhe attributeve të tyre gjatë vizualizimit.

2. 2. 2 Atributet vizuale

Colin Ware në veprën e tij (Ware, 2004) ka organizuar atributet vizuale sipas gjatësisë, ngjyrës, pozitës në hapësirë dhe drejtimit. Për dallimet midis **kategorive** përdorimi i llojeve të ndryshme të ngjyrave gjen më tepër përdorim. Scatter plot, grafikonet shiritore dhe linjore poashtu identifikohen bazuar në këto attribute.

Gjatë përshkrimit të dallimeve **kuantitative** më shumë shfrytëzohen atributet si gjatësia dhe pozita 2-D në hapësirë. Jo rastësisht, grafikonet shiritore, linjore dhe scatter plots konsiderohen si teknika më të përsosura vizualizimi, ato përdorin gjatësinë dhe pozitën dy dimensionale në hapësirë për të fituar shkallë më të lartë preciziteti për vlerat e të dhënave të paraqitura.

2. 2. 3 Objektet vizuale bazike

Komunikimi i informacionit është qëllimi kryesor i vizualizimit. Ekzistojnë katër lloje të objekteve bazike të cilat në mënyrë më efektive përshkruajnë vlerat kuantitative: pikat, vijat, shiritet dhe kutitja (boxes). Figura e mëposhtme paraqet ato grafikisht:

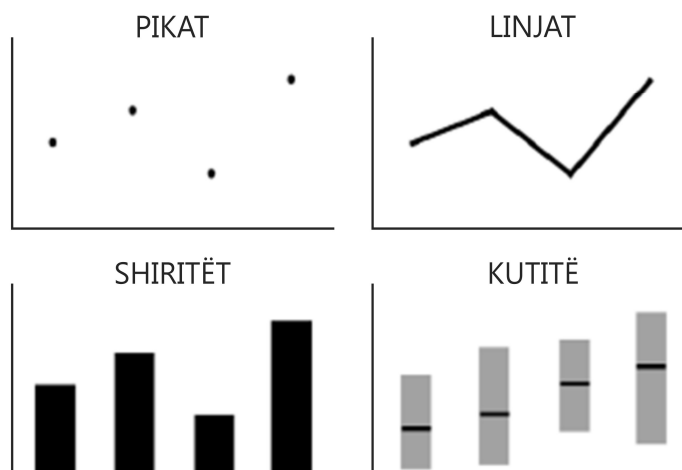


Fig. 3: Katër llojet e objekteve vizuale bazike

Secili lloj ka përparësi dhe mangësi gjatë reprezentimit të të dhënave kuantitative, dhe secili mund të llogaritet si alternativë më e mirë për përcjelljen e ndonjë mesazhi të caktuar kuantitativ (Few, Selecting the Right Graph for Your Message, 2004).

Pikat mund të marrin formën e çdo figure gjeometrike, si, rrethi, vijës, katrorit, trekëndëshit, etj. Përdorimi i pikave ofron dy mundësi primare: 1) ato mund të përdoren për të përshkruar më shumë vlera kuantitative në të njëjtën kohë, si në scatter plot; 2) ato mund të përdoren për të theksuar vlera individuale, me që pikat përdoren në hapësirë 2-D.

Vijat gjithashtu shërbejnë për të përshkruar të dhëna kuantitative (sasiore). Ato përdoren në hapësirë 2-D dhe krijohen nga pika të lidhura që formojnë vijë nga seri vlerash. Këto më së tepërmi përdoren në: 1) lidhjen në seri të pikave të të dhënave individuale si dhe 2) afishimin e trendeve të atyre serive. Vijat janë alternativa më e mirë e ofruar për afishimin e gjendjes së të dhënës në varshmëri nga koha.

Shiritet përshkruajnë të dhënën në mënyrë që ajo sa më shumë të theksojë vlerat individuale. Kjo bëhet nga shkak se shiritet për të përshkruar vlera kuantitative përdorin attribute si: 1) hapësirë 2-D dhe 2) gjatësi për vlerat. Me qenë se shiritet japin aq peshë vizuale për vlerat e reprezentuara, ato nuk janë po aq të mira sa grafikonet vijore kur bëhet afishimi i gjendjes së të dhënave në varshmëri kohore. Por, kur është nevoja për fokusim primar në vlera individuale dhe krahasim të atyre vlerave, ky mund të konsiderohet si më i përshtatshëm.

Kutijat (boxes) janë sikur shiritet, përveç se në to si fundi poashtu edhe fillimi përshkruajnë vlerë kuantitative. Për dallim nga shiritet, kutitë përdoren për të afishuar shpërndarje të një grupi vlerash nga më e vogla deri te më e madhja së bashku me vlerën mesatare.

2. 3 E dhëna si faktor në procesin e vizualizimit

Çdo hulumtim shkencor përfshin një ose më shumë grupe të variablave të cilat hulumtuesit manipulojnë, observojnë dhe masin. Variabla mund të jetë pronë e ndonjë rezultati shkencor i cili mund të ketë vlera të ndryshme. Shpesh ndodh që në hulumtimet shkencore të jetë e nevojshme të bëhen krahasime të ndryshme të rezultateve. Shumica e atyre rezultateve paraqiten apo reprezentohen nga vlerat e variablave të cilat kanë qenë të vendosura gjatë procesit të matjeve. Përpara se të japim sqarim për llojet (tipet) e ndryshme të variablave, është e nevojshme të kuptojmë strukturat e të dhënave dhe pse ato janë aq të rëndësishme.

Për të krijuar një grafikun të mirfilltë, është e nevojshme të definohen karakteristikat kryesore të të dhënës dhe të shqyrtohen limitimet që mund të paraqesin ato. Këto karakteristika janë: 1) struktura e të dhënës; 2) tipi relevant i variablës dhe 3) nivelet e matjeve. (Wallgren, 1996).

2. 3. 1 Struktura e të dhënave

Ekzistojnë dy lloje të strukturave të të dhënave: kros seksionale dhe të serive kohore. Të dallohen këto dy lloje të strukturave të të dhënave nuk është e vështirë, nga arsyeja se në rastin e parë kemi matje të realizuara në një kohë të caktuar, ndërsa te tjetra kemi matje në periudha të ndryshme kohore.

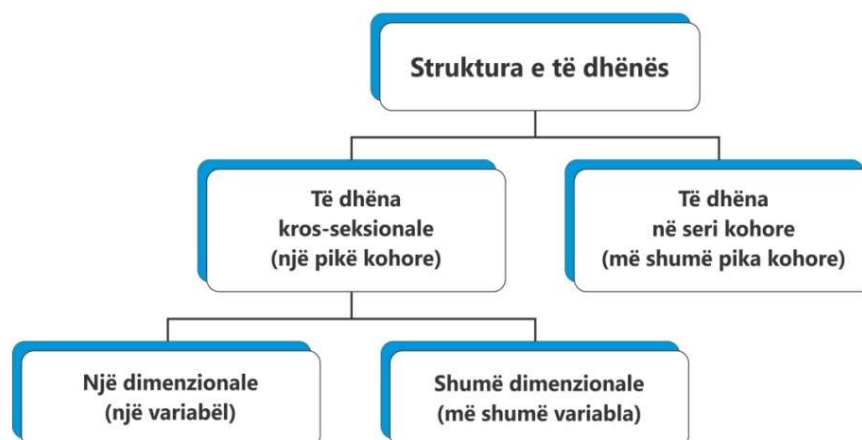


Fig. 4: Hierarkia e strukturës së të dhënave

Të dhënat kros seksionale mund të jenë njëdimensionale (vetëm me një variabël) dhe shumëdimensionale (me më shumë variabla). Figura paraprahe tregon grafikisht strukturën e të dhënës.

2. 3. 2 Tipet e variablave

Ndarja e parë e tipeve të variablave bëhet bazuar në karakteristikat e vlerave të variablave. Vlerat mund të paraqiten duke përdorur numra ose fjalë. Në rastin e parë kemi variabël kuantitative (sasiore) dhe në rastin e dytë kualitative (cilësore). Ndarja e dytë bëhet në mes të variablave kuantitative. Një ndër vetitë më fundamentale të variablave është domeni i vlerave të cilat mund të marrin ato. Bazuar në domenin e vlerave që mund të marrin variablat, ato mund të ndahen në: *diskrete* dhe *të vazhdueshme*. Variablat diskrete shërbehen nga grup i fundëm vlerash të përdorura. Shembull të variablës diskrete kemi në rastin kur ndodh të bëjmë "vlerësim të nxënësve". Vlerat e kësaj variable do të duhej të ishin një nga vlerat: 1, 2, 3, 4, 5. Variablat e vazhdueshme në krahasim me ato diskrete, mund të marrin vlera deri në pakufi. "Gjatësia" mund të llogaritet si shembull të variablave të vazhdueshme: 0.9m, 0.79m, 0.9765m. Në mënyrë përmbledhëse, pjesa e përmendur më lart grafikisht mund të paraqitet përmes kësaj figure:

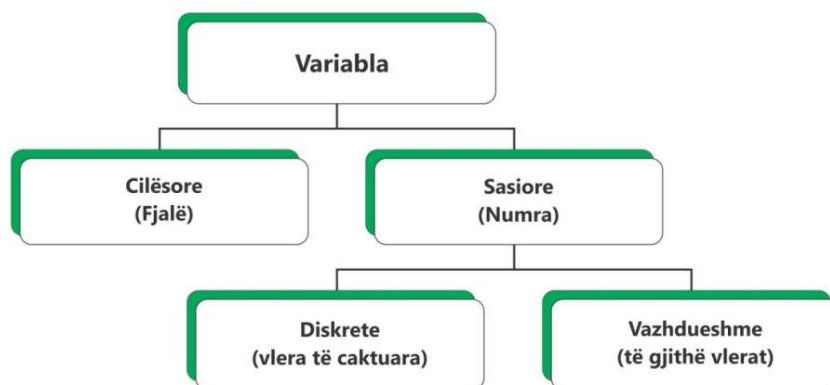


Fig. 5: Hierarkia e tipeve të ndryshme të variablave

2. 4 Objektivi si faktor në procesin e vizualizimit

Objektivi mund të llogaritet si faktori më i rëndësishëm dhe gjithashtu më i vështirë në procesin e vizualizimit. Gjatë kësaj faze gjithçka duhet analizuar dhe definuar mirë, nga arsyeja se bazuar në to zgjedhet teknika e duhur e vizualizimit, e cila teknikë do të duhet në mënyrë sa më të përshtatshme të përcjellë informacionin e përpunuar.

2. 4. 1 Klasifikimi i qëllimeve

Sipas (Zelazny, 2001), për të zgjedhur vizualizim më të përshtatshëm për reprezentim të të dhënave, është e rëndësishme që nga fillimi të definohet mesazhi. Ky mesazh në realitet paraqet edhe objektivi i cili duhet zgjidhur. G. Zelazny thotë që mesazhi i marrur nga vizualizimi i të dhënave gjithmonë do të paraqesë njërën nga pesë llojet elementare të krahasimit: 1) tërësisë, 2) njësive, 3) serive kohore, 4) shpërndarjes frekuede dhe 5) krahasime korelacionale.

Krahasimi i tërësisë (komponentit) ka për qëllim të tregojë madhësinë e çdo pjese si përqindje të totalit. Shembull, ekzistojnë disa komponente, të cilat së bashku formojnë një komponent më të madh të quajtur total. Elementi total gjithmonë reprezenton 100%, kështu që komponentet më të vogla kanë vlera gjithmonë më të vogla se 100% dhe këto në realitet paraqesin pjesë të tërësisë.

Krahasimi i njësive ka të bëjë më shumë me rangimin e njësive (elementeve), gjegjësisht krahasimin kuantitativ të tyre. Njësitë e krahasuara janë të pavarura njëra me tjetrën, në kuptimin se ato nuk përshkruajnë tërësi, siç ishte rasti në krahasimin e tërësisë. Çdo element (njësi) reprezenton vlerë të veten, madhësia e së cilës krahasohet me tjerat.

Krahasimi i serive kohore tregon ndryshimet e vlerave në raport me kohën, qofshin këto trende javore, mujore apo vjetore rritëse, rënëse, apo konstante. Përderisa krahasimet e tërësisë, apo krahasimet e njësive tregojnë raportin e vlerave të të dhënave në një kohë të caktuar, krahasimi i serive kohore tregon ndryshimet e atyre vlerave në varshmëri nga koha.

Krahasimi i shpërndarjes frekvente ka për qëllim të tregojë numrin e elementeve që i takojnë një serie progresive të vargut numerik. Ekzistojnë dy lloje zbatimesh të këtij krahasimi. I pari është përgjithësimi i ngjarjeve të mundshme të një observimi, ku shpërndarja frekvente përdoret për të parashikuar rrezikun, probabilitetin apo shansin, ndërsa zbatimi i i dytë është për të përgjithësuar sasinë e madhe të të dhënave për demonstrimin e relacioneve midis tyre.

Krahasimi i korrelacioneve përshkruan relacionet midis dy variablave apo grupeve të variablave. Konkretisht, tregon nëse njëri grup shkon lart, tjetri grup a do të shkojë lart ose poshtë në mënyrë të duhur.

2. 4. 2 Vizualizimet e rekomanduara

Gene Zelazny veçon pesë lloje të grafikoneve të cilët do të ishin më të përshtatshme për përdorim: pie, bar, column, line dhe dot chart. Në veprën (Zelazny, 2001) shënon që **krahasimi i tërësisë** apo komponentit më mirë mund të demonstrohet përmes Pie grafikoneve. Gjithashtu këtu thuhet se ky lloj grafikoni është më pak praktik se sa të tjerët nga arsyeja se për përshkrimin e vlerave kuantitative, Pie grafikonet përdorin hapësirë 2-D, e cila me krahasimin e një rangui më të madh vlerash bëhet e vështirë për perceptim. Rekomandohet që kur të përdorim Pie grafikonet, gjithmonë të definojmë mirë mesazhin që do përcjellim përmes tyre, dhe poashtu duhet të kemi kujdes në numrin e elementeve që do marrim në shqyrtim.

Krahasimi i njësive më së miri mund të demonstrohet nëpërmjet **Bar** grafikonet (grafikonet shiritore), ndërsa **krahasimet e serive kohore** demonstrohen përmes grafikoneve koloniale (grafikoneve shiritore vertikale) ose grafikoneve linjore (line chart). Sipas G. Zelazny-t, grafikonet të cilët përdoren më së shumti janë: 1) grafikoni linjor (line chart); 2) grafikoni shiritor (bar chart) dhe 3) grafikoni kolonial (column chart), të cilin edhe e quan si "më të mirin, më të vjetrin dhe më të besueshmin". Logjika nga e cila niset G. Zelazny qëndron në atë se grafikonet shiritore, koloniale dhe linjore për reprezentimin e vlerave kuantitative përdorin atributet më efektive, siç janë pozita 2-D (e përshkruar përmes koordinatave X, Y) dhe gjatësia (reprezenton vlerat sasiore), attribute këto të cilat

perceptohen në mënyrë më të qartë, më të lehtë dhe me shkallë më të lartë preciziteti nga njerëzit.

Shpërndarjet frekvente për demonstrimin e relacioneve në mënyrë më të përshtatshme bëhen përmes histogrameve apo shumëkëndëshit frekuent, i cili është lloj i grafikonit linjor (line chart). **Krahasimi i korelacioneve** mundet më mirë të tregohet nëpërmjet grafikonit me pika (Scatter Plot) ose grafikonit shiritor të çiftëzuar (paired bar chart).

Histogramet dhe Scatter Plot ngjashëm si grafikonet shiritore dhe linjore përdorin gjatësinë dhe pozitën 2-D për përshkrimin e vlerave kuantitative, attribute të cilat mundësojnë vizualizimi të perceptohet në mënyrë sa më të qartë dhe saktë.

Të gjitha vizualizimet që rekomandohen nga Gene Zelazny për secilin nga qëllimet krahasuese, janë dhënë në figurën më poshtë në formë matricore.

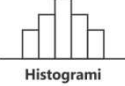
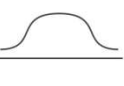
	Grafikoni Pie	Grafikoni Shiritor	Grafikoni Kolonial	Grafikoni Linjor	Scatter Plot
Tërësia Pjesa vs. Tërësia		-	-	-	-
Njësia Si rangohen gjërat	-		-	-	-
Seritë kohore Ndryshimet në raport kohor	-	-			-
Shpërndarjet Frekuenca, rang i të dhënave	-	-	 Histogrami		-
Korrelacionet Raportet e të dhënave	-	 Grafik. shir. i çiftëzuar	-	-	

Fig. 6: Matrica e grafikoneve rekomanduese

2. 4. 3 Grafikone më të avansuara

Përveç pesë teknikave vizualizuese që u përmendën në pjesën 2. 4. 2, ekzistojnë edhe teknika tjera vizualizimi, të cilat po ashtu duhet shqyrtuar. Ato identifikohen me këto emra: grafikone me shirite të grupuara (Grouped Bar Chart), Stacked Bar Chart, Box Plot, Histogram, Parallel Axes, Heat Maps, Scatter Plot, Network Diagrams, Sankey Diagrams, etj.

Grouped Bar Charts paraqet teknikë vizualizimi tek e cila elementet e çdo serie me të dhëna reprezentohen nga shirite të ndara. Të gjitha këto seri pozicionohen krah për krah për të krijuar grupe. Shiritet që identifikojnë elementet e çdo serie gjenden gjithmonë në të njëjtën pozicion të çdo grup, ndërsa dallohen njëri nga tjetri përmes ngjyrës, hijes, etj. Shembull të grafikonit me shirite të grupuara (Grouped Bar Chart) tregohet në figurën më poshtë:

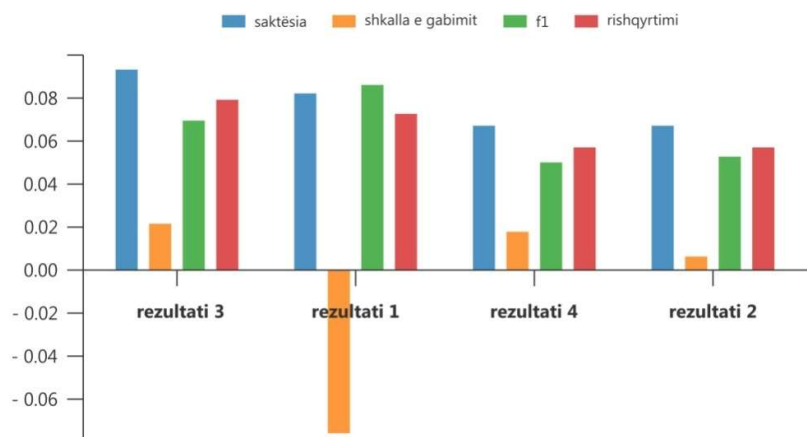


Fig. 7: Shembull grafikoni me shirite të grupuara

Stacked Bar Chart bën grumbullimin e elementeve të një serie me të dhëna në një shirit, për dallim nga grafikoni me shirite të grupuar. Stacked Bar Charts përdoren për të treguar se në ç'masë janë elementet e serive të kategorisë, gjegjësisht, efektin relativ që ka çdo element në seri. Shembull i Stacked Bar Chart jepet në figurën në vijim:

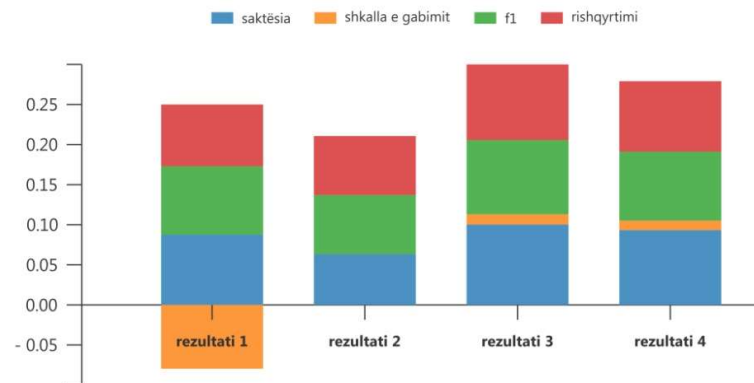


Fig. 8: Shembull grafikoni Stacked Bar

Gjithashtu, me ndihmën e Stacked Bar Chart mund të afishohet edhe përqindja e reprezentuar nga vlerat e të gjitha elementeve të serive me të dhëna.

Box Plot është teknikë shumë efektive vizualizimi në rastet kur duhet krahasuar shpërndarjet në mes të disa grupeve. Qëllimi i saj është të reprezentojë shpërndarjet kuantitative të vlerave të të dhënave me pesë standarde statistikore: vlerën më të vogël, kuartilin e poshtëm, mesataren, kuartilin e sipërm dhe vlerën më të madhe.



Fig. 9: Shembull grafikoni Box Plot

Histogrami paraqet teknikë vizualizimi ku bëhet reprezentimi grafik i të dhënave numerike të shpërndara. Edhe pse kjo teknikë vizualizimi ka ngjashmëri me grafikoneve shiritore (bar chart), dallimet midis njëra-tjetrës për nga spektri i përdorimit të tyre janë

shumë të mëdha. Histogramet përdoren për të treguar shpërndarje të variablave përdorësisht grafikonet shiritore përdoren për të krahasuar variablat.

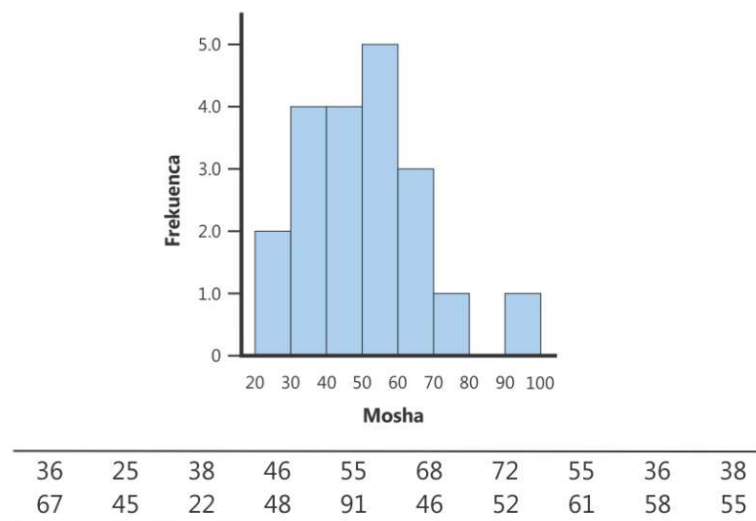


Fig. 10: Shembull grafikoni histogram

Boshtet paralele llogariten si një ndër teknikat më të fuqishme vizualizuese. Bazuar në veprën (Few, Now You See It., 2009), boshtet paralele paraqesin teknikë shumë efektive vizualizuese për reprezentimin dhe analizimin e të dhënave shumë-dimensionale. Këtu çiftet e variablave në vend që të paraqiten në dy dimensione, siç bëheshin në Scatter plot, paraqiten në boshte paralele dhe më pastaj pikat përkatëse do të lidhen me vija. Secila vijë paraqet një rresht të bazës së të dhënës, ndërsa kryqëzimi i vijave ndërmjet dimensioneve shpesh tregon korrelacion invers.

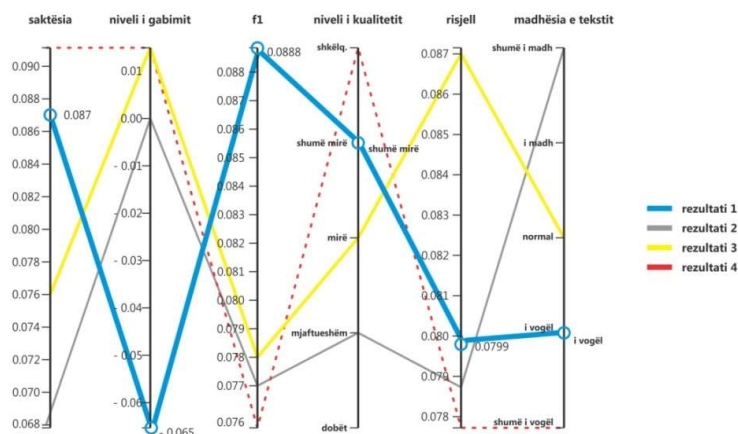


Fig. 11: Shembull i grafikoni me boshte paralele

Heatmaps paraqesin teknikë të vizualizimit të të dhënave, reprezentimit i vlerave të së cilës bëhet nëpërmjet të ngjyrave. Ekzistojnë shumë mënyra të paraqitjes së këtij lloji vizualizimi, por të gjitha ato mënyra kanë një gjë të përbashkët - ato përdorin ngjyrën për të komunikuar lidhjet ndërmjet vlerave të të dhënave, të cilat do të ishin shumë më të vështira për t'u kuptuar në rast se ato prezentoheshin numerikisht në tabelë.

Zakonisht ky lloj vizualizimi përdor një ngjyrë të vetme gjatë paraqitjes së të dhënave, gjegjësisht, shfrytëzon niansat apo variacionet e një ngjyre për të paraqitur vlerat e atyre variablave.

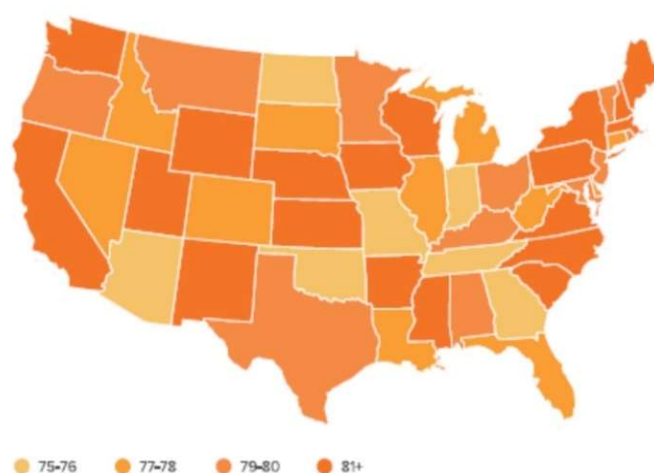


Fig. 12: Shembull grafikoni heatmap

Grafikonet Scatter Plot mund të konsiderohen si një ndër teknikat më të përdorshme vizualizuese, posaçërisht në rastet kur kërkohet që të gjendet korrelacion midis vlerave të observuara kuantitative (variablave). Në rast se nuk ekziston korrelacioni ndërmjet variablave, pikat do të paraqiten në mënyrë arbitrare në boshtin koordinativ, ndërsa në rast se ekziston korrelacion i madh, koncentrimi i pikave gjithmonë e më shumë do të shkojë drejt një vije të drejtë. Kjo vijë e drejtë në realitet definon llojin e korrelacionit.

Në qoftë se variabla vertikale e paraqitur në boshtin Y rritet ashtu si edhe ajo horizontale në boshti x, korrelacioni është pozitiv, ndërsa në rast se këto variabla zvogëlohen apo vazhdimisht shkojnë në proporcion të zhdrejtë, atëherë kemi korrelacion negativ, gjegjësisht nuk kemi korrelacion.

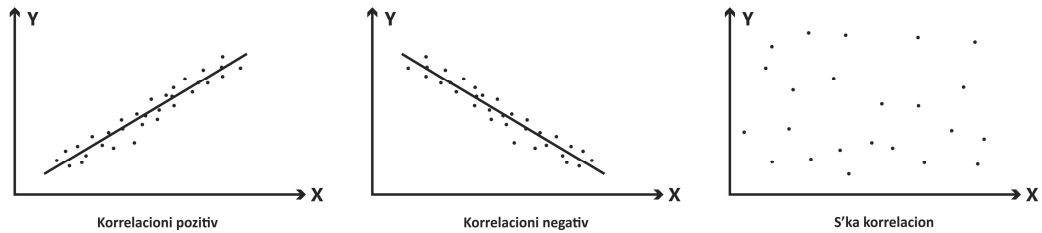


Fig. 13: Shembull grafikoni Scatter Plot

Network Diagramet paraqesin mënyrë të vizualizimit të të dhënave tekstuale, ku fjalët paraqiten si nyje, ndërsa afërsia e tyre paraqet lidhjen midis atyre fjalëve. Trajtimi i tekstit si rrjet hap më tepër mundësi interpretimi dhe lejon atë të manifestohet në polisingularitetin e tij. Kjo nënkupton se kuptimet e shumta që mund të lindin nga kjo teknikë janë nga më të ndryshmet, sidoqoft, çdo formulim i krijuar nga kjo lidhje e tekstit mund të ketë kuptim të caktuar për kohë të caktuar.

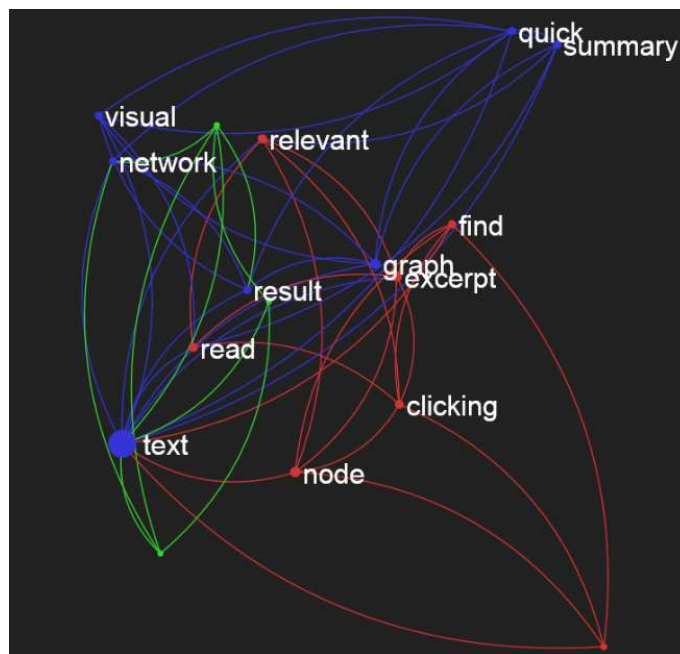


Fig. 14: Shembull të diagramit network

Diagramet Sankey paraqesin lloj të diagrameve rrjedhëse, ku "rrjedha" reprezentohet nëpërmjet shigjetave, trashësia apo gjerësia e të cilave varet nga sasia e vlerave. Këto më tepër përdoren për vizualizimin e energjisë, materialit apo transferimin e kostove dhe gjithashtu gjejnë shumë përdorim në demonstrimin e proporcionalitetit të rrjedhës, ku pjesë të ndryshme të diagramit reprezentojnë kuantitete të ndryshme në sistem.

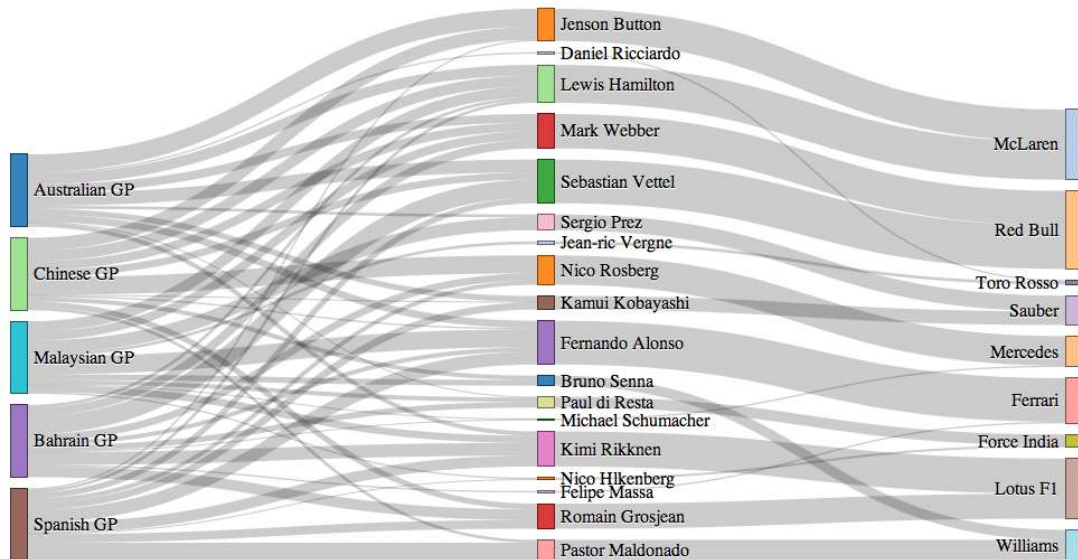
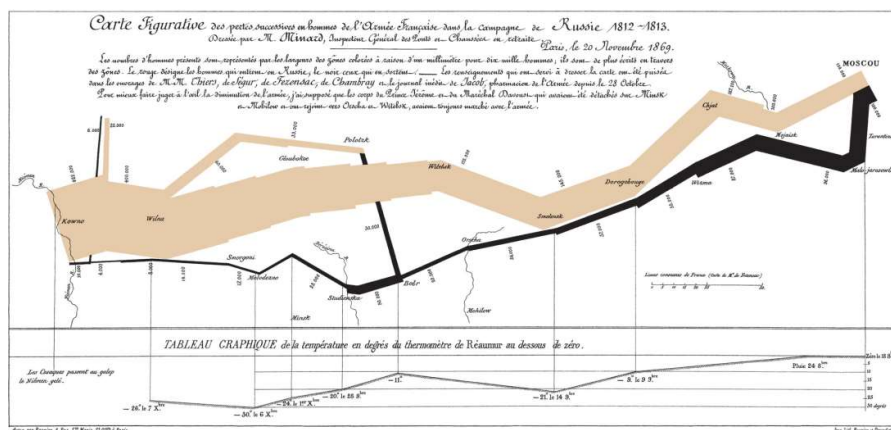


Fig. 15: Shembull të diagramit Sankey

Çarls Minard ishte ndër të parët që në mënyrë më të mirëfilltë bëri përdorimin e këtij lloji të grafikonit, duke përshkruar ekspeditën ushtarake të Napoleon Bonapartës ndaj Rusisë në vitin 1812.



Vizualizimet cirkulare paraqesin lloj të grafikoneve të avansuara, ku mund të themi se bëhet gërshetimi i çdo teknike tjetër vizualizuese. Circos është pako softuerike e cila mundëson krijimin e këtyre llojeve të grafikoneve të cilat gjejn zbatim shumë të gjerë posaçërisht në rastet ku kërkohet proporcioni, raporti dhe lidhshmëria e të dhënave shumë shtresore. Circos është fleksibile. Edhe pse e konstruktuar për vizualizimin e të dhënave gjenomike, ajo poashtu mund të krijoj vizualizime për të dhëna të fushave të ndryshme – nga ato gjenomike, mund të bëhet poashtu vizualizimi i migracioneve e deri te artet matematikore.

2. 5 Efekti i interaktivitetit në vizualizim

Është e një rëndësie enorme që shfrytëzuesve t'u mundësohet rregullimi, modifikimi dhe rafinimi i vizualizimeve në mënyrë interaktive, që ato të arrin të realizojnë qëllimet e tyre, si, nxjerrjen e rezultateve, përshkrimin plotësues të përmbajtjes së të dhënave apo zgjedhjen e modeleve përmes së cilave do të paraqiten të dhënat (Ward, Grinstein, & Keim, 2010).

Interaktiviteti paraqet mënyrë komunikimi midis shfrytëzuesit dhe vizualizimit, me që shfrytëzuesi eksploron të dhënat për të fituar njohuri plotësuese.

Në pjesën 2. 2. 1 ku u shtjellua rëndësia e perceptimit human në procesin e vizualizimit, u tregua se analiza e të dhënave përfshinte kërkimin e attributeve interesante dhe studimin e tyre në mënyra të ndryshme, rrugë këto të cilat gjithmonë stimulojnë pyetje të pa parashtruara më parë. Për këto arsye ekziston nevoja për kërkime plotësuese të këtyre attributeve, modeleve apo detajeve tjera në mënyrë efektive. Kjo përkrahje e cila i mundëson të gjitha këto informacione ndihmëse nga vizualizimi, mundësohet përmes përdorimit të interaktivitetit në vizualizim.

Pa interaktivitet, vizualizimi bëhet imazh statik, i cili limiton procesin analitik nga ana e shfrytëzuesit. Sipas Stephen Few, efektiviteti i vizualizimit varet nga dy gjëra: 1) aftësisë së vizualizimeve që në mënyrë të qartë dhe të pastër të reprezentojnë informacionin; 2) aftësisë së njerëzve për të bashkëvepruar me vizualizimin për të nxjerrë dhe kuptuar informacione.

Në veprën (Few, Now You See It., 2009), prezentohen lloje të ndryshme të interaksionit (bashkëveprimit) me të dhëna. Ato të cilat më tepër gjejnë zbatim janë të:

- | | |
|-------------------------|---------------------------------|
| • krahasimit | • agregimit |
| • sortimit | • ri-vizualizimit |
| • shtimit të variablave | • zumimit (zmadhimit të pamjes) |
| • filtrimit | • ri-shkallëzimit |
| • hijezimit | • shënimeve |

Nga llojet e interaksioneve me të dhënat të cilët i përmendëm më lart, vetëm disa nga ato do përshkruhen më detajisht.

Krahasimi si lloj interaksioni me të dhënat nënkupton kërkimin e ngjashmërive dhe dallimeve. Vizualizimi duhet të mbështesë krahasimin për rastet e mëposhtme, si:

- mundësinë e zgjedhjes së grafëve të cilët mbështesin të gjithë spektrin e krahasimeve më shpesh të përdorura (krahasimin e vlerave si dhe krahasimin e modeleve);
- shmangien e nevojës për të lëvizur apo shkuar nga një pamje ekrani në tjetër për të realizuar krahasimin.

Sortimi i mundëson vizualizimit renditjen e vlerave sipas kritereve të caktuar, si, nga më i vogli deri te më i madhi dhe anasjelltas, si dhe formave të tjera.

Filtrimi paraqet reduktimin e vlerave të afishuara në një nëngrup tjetër, sipas ndonjë kriteri të caktuar. Gjithashtu, edhe në këndvështrimin e bazës së të dhënave kjo nënkupton heqjen e disa rreshtave, bazuar në ndonjë kusht të vendosur. Filtrimet në një vizualizim mund të bëhen me ndihmën e radio butonave, çek boksave dhe rrëshqitësve.

Hijezimet ndikojnë në mënyrën e paraqitjes së të dhënave, konkretisht, ato bëjnë që disa nga të dhënat të theksohen më shumë se të tjerat, pa mos shkaktuar ndryshime apo ç'renditje të të dhënave tjera.

Rivizualizimi mundëson ndryshimin e reprezentimit vizual: kalimin nga një tip vizualizimi në tjetër. Të pasurit e aftësive për të bërë këtë në mënyrë të shpejtë dhe lehtë është esenciale. Vizualizimi i cili përmban këtë lloj interaksioni duhet që:

- të mundësoj mënyrë të shpejtë dhe të lehtë të ndryshimit nga një tip vizualizimi në tjetër;
- të mundësoj listë të llojeve të vizualizimeve në dispozicion, të cilat përdoren për ato lloje të të dhënave;
- të pengojë apo vështirësoj përzgjedhjen e vizualizimit i cili do të afishoj të dhënat në mënyrë të papërshtatshme.

2. 6 Vizualizimi efektiv

"Të dhënat tona janë të vlefshme aq sa është edhe aftësia jonë për t'i kuptuar dhe interpretuar ato. Kjo është edhe arsyeja kryesore se pse duhet bërë zgjedhja vizualizimit të duhur". Nëse të dhënat tona keqinterpretohen ose prezentohen në mënyrë joefektive, njohuria apo kuptimi i tyre do të humbet, gjë e cila do të rezultojë negativisht në përcjelljen e mesazhit përmes atij vizualizimi si dhe reputacionit tonë.

2. 6. 1 Këndvështrime në sferën e vizualizimit efektiv

Sipas veprës (Chen, Härdle, & Unwin, 2008), vizualizimi i mirë është ai i cili është efektiv. Mirëpo, çka në të vërtetë nënkupton termi "efektiv" në rastin e vizualizimit të të dhënave? Nënkupton faktin që vizualizimi të jetë mirë i dizajnuar dhe si pasojë e kësaj të çojë deri në shfrytëzim dhe interpretim më të mirë dhe më të lehtë për shfrytëzuesin, apo përfshirjen e të gjitha informacioneve në të dhënat dhe rezultatet të cilat do të duhet të afishohen nga vizualizimi?

Në (Zhu, 2007) autori shënon se edhe pse termi "vizualizimi efektiv" ka qenë i përdorur në shumë publikime, për këtë term asnjëherë nuk është dhënë një definicion i mirëfilltë. Për më tepër, ekzistojnë këndvështrime të ndryshme rreth asaj se çfarë në realitet nënkupton "vizualizimi efektiv". Në po të njëjtin publikim, Ying Zhu thotë se ekzistojnë 3 këndvështrime në sferën e vizualizimit efektiv:

- 1) të dhënës;
- 2) objektivit;
- 3) shfrytëzuesit.

Për nga këndvështrimi i **të dhënës**, efektiviteti i vizualizimit varet nga analogjia ndërmjet vizualizimit dhe të dhënës e cila do të afishohet. Kjo do të thotë se struktura e të dhënës dhe struktura perceptuale e vizualizimit duhet të përputhen.

Sipas këndvështrimit **objektiv (synues)**, efektiviteti i vizualizimit varet nga qëllimi (mesazhi) të cilin do të duhet të përcjellë. Kjo në realitet nënkupton atë se sa vizualizimi

do të mund të ndihmojë në përcjelljen e mesazheve të caktuara, procesin e vendimmarrjeve, si dhe në procesin e leximit dhe analizimit të vizualizimit për marrjen e njohurive të reja.

Këndvështrimi i **shfrytëzuesit** nënvizon se efektiviteti i vizualizimit varet nga kapaciteti i memorjes punuese (afatshkurte) të shfrytëzuesit, njohja e domenit të caktuar dhe eksperiencës me teknikat e vizualizimit.

Në rast se shikojmë edhe një herë pipeline-in e procesit të vizualizimit në pjesën 2. 1, do të shihet qartë se çdonjëri nga 3 këndvështrimet e efektivitetit të vizualizimit fokusohet në një nga komponentët (faktorët) e procesit të vizualizimit.

2. 6. 2 Principet e vizualizimit efektiv

Në këtë punim si definicion për "vizualizim efektiv" kemi zgjedhur atë të Ying Zhu, për arsye se ky tenton të bëjë sintezë të gjithë përfaqësuesve në proces - të dhënës, objektivit (qëllimit, mesazhit) dhe shfrytëzuesit, me idenë që të japi definicion më gjithëpërfshirës. Reprezentimi grafik i këtij definicioni është dhënë në figurën më poshtë:

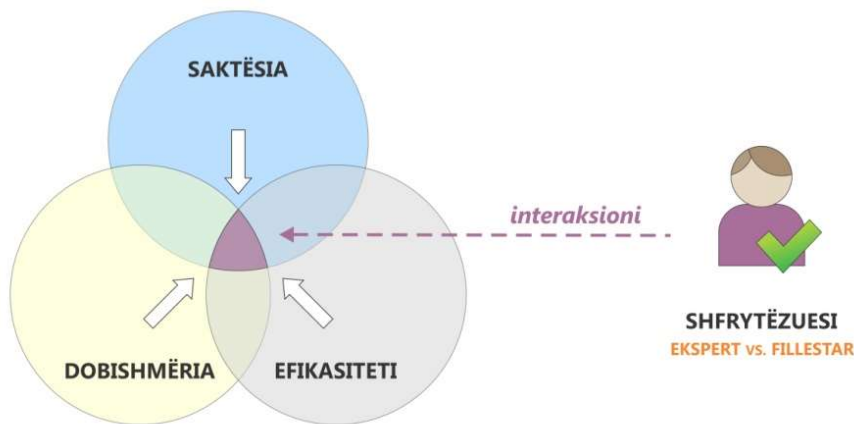


Fig. 18: Definicioni i vizualizimit efektiv

Në (Zhu, 2007), autori tregon se efektiviteti i vizualizimit të të dhënave përcaktohet nëpërmjet tre principeve kryesore: 1) saktësisë, 2) dobishmërisë dhe 3) efikasitetit.

Principi i saktësisë: "Për të qenë vizualizimi efektiv, atributet e elementeve vizuale duhet të përshtaten me atributet e të dhënave, si dhe struktura e vizualizimit duhet të përshtatet me strukturën e grupit të të dhënave" (Zhu, 2007). Principi i saktësisë definon lidhjen midis vizualizimit dhe të dhënës.

Principi i dobishmërisë: "Vizualizimi efektiv duhet t'u ndihmojë shfrytëzuesve të realizojnë qëllimet e tyre" (Zhu, 2007). Principi i dobishmërisë definon lidhjen midis vizualizimit dhe mesazhit të cilin duhet të përcjellë. Vizualizimi mund të krijohet edhe për më tepër qëllime, por rekomandohet që gjithmonë paraprakisht të definohet mirë dhe pastër qëllimi apo mesazhi që duhet përcjellur, për arsye të matjes së dobishmërisë dhe efikasitetit të vizualizimit.

Principi i efikasitetit: "Reduktimi i sasisë së madhe informative për një qëllim specifik në raport me reprezentimet jo vizuale" (Zhu, 2007). Kjo do të thotë që vizualizimet duhet të jenë të lehtë për t'u kuptuar nga shfrytëzuesit dhe të rrisin efikasitetin e qëllimeve që përshkruajnë. Ky princip definon lidhjen midis vizualizimit dhe shfrytëzuesit.

Është e një rëndësie enorme të përmendim që saktësia, dobishmëria dhe efikasiteti i vizualizimit është i ndikuar nga domeni i njohjes së shfrytëzuesit, eksperiencia e shfrytëzuesit me vizualizimet dhe aftësia vizualo-hapsinore. Një domen njohës i shfrytëzuesit dhe eksperiencia në fushën e vizualizimit, vetëm mund të rrisin nivelin e efikasitetit në vizualizime.

3. VALIDIMI DHE EVALUIMI I VIZUALIZIMIT

Në këtë kapitull do të diskutohet për sfidat kryesore të matjeve dhe evaluimeve të vizualizimeve efektive. Gjithashtu, do të jepet modeli për dizajnim dhe validitim të vizualizimit, si dhe do të përshkruajmë standardin DIN EN ISO 9241-12, i cili shërben për prezantimin korrekt të informacionit.

3. 1 Sfidat për vlerësimin e efikasitetit të vizualizimit

Vizualizimi i informacioneve është detyrë sfiduese (Plaisant, 2004). Një nga qëllimet kryesore të tij është zgjerimi i njohurive. Ky qëllim mund të arrihet nëpërmjet shumë mënyrave, si p.sh., vendosjen e interaktivitetit në vizualizimi, përmes "drill down" efektit për marrjen e informacioneve më të detajuara, zumimit apo filtrimit, etj. Njohja gjithashtu mund të zgjerohet duke përdorur afishime plotësuese, të cilat do të plotësojnë njëra tjetrën dhe do të bëjnë të mundur për shfrytëzuesin marrjen e informacioneve më të rëndësishme për kohë më të shpejtë. Interesant fakti që, në të njëjtën mënyrë dhe formë ashtu siç zgjerohen njohjet, po ashtu mund edhe të reduktohen! Në rast se i referohemi shembullit të njëjtë, është tejet e vështirë të definohet se cilat trende (lloje) të interaktivitetit do të kishin qenë të mjaftueshme për të realizuar një vizualizim efektiv. Po ashtu, nuk është gjithmonë e qartë nëse afishimet plotësuese apo aspekte tjera vizuale i ndihmojnë shfrytëzuesit dhe nuk e bëjnë objektivin akoma më të komplikuar.

Tjetër aspekt i rëndësishëm i cili bën procesin e evaluimit të vizualizimit një punë shumë sfiduese është diversiteti i shfrytëzuesve. Ky diversitet nënkupton dedikimin, aftësitë perceptuese si dhe preferencat e njerëzve. Kur bëhet fjalë për atë se kujt i dedikohet vizualizimi, këtu merren parasysh njerëzit e rëndomtë, si dhe fillestarët apo ekspertët e fushës së vizualizimit të informacionit. Aftësitë perceptuese dhe preferencat përfshijnë më shumë aspekte, ndër të cilat më me rëndësi të përmenden janë niveli edukativo-arsimor i njerëzve, mosidentifikimit i ngjyrave dhe strukturave vizuale, si dhe pajisjet në të cilat do të duhet të afishohet vizualizimi.

Në rast se u referohemi principeve të vizualizimit efektiv: saktësisë, dobishmërisë dhe efikasitetit (pjesa 2. 6. 2), të cilët lidhen shumë me komponentet e procesit të vizualizimit (të dhënë, objektivin dhe shfrytëzuesin), nuk është edhe aq e qartë të kuptohet ajo se si të merren parasysh të gjitha ato principe përgjatë procesit të evaluimit të vizualizimit.

Sidoqoftë, ditëve të sotme ekzistojnë qasje të ndryshme për validimin/evaluimin e vizualizimit, të cilat tentojnë të bëjnë të mundur të gjithë këtë punë kaq sfiduese. Në vijim do të paraqesim disa nga ato qasje.

3. 2 Modeli i ndërthurur për dizajnimin dhe validimin e vizualizimit

Tamara Munzner në veprën (Munzner, 2009) prezentoj modelin e ndërthurur për dizajnimin dhe validimin e vizualizimit. Ky model përveç që tregon se si të validojmë dhe evaluojmë vizualizimin, poashtu ndihmon të kuptojmë se kur të bëhet.

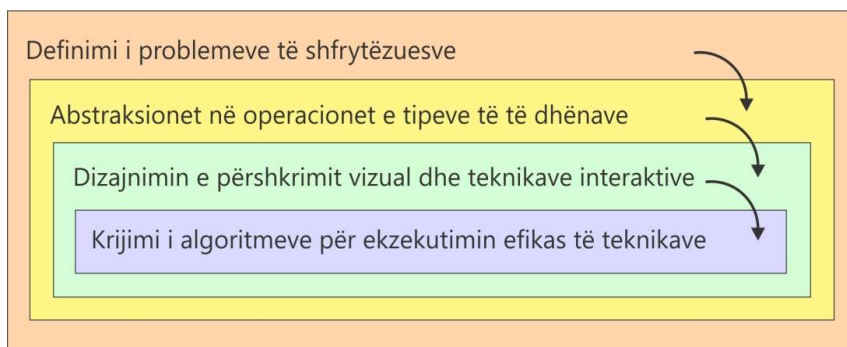


Fig. 19: Modeli i ndërthurur për dizajnim dhe validim të vizualizimit

Ashtu siç shihet nga figura më lart, modeli është i ndërtuar nga katër shtresa. Çdo shtresë ka qëllim të caktuar:

- 1) definimin e problemeve;
- 2) abstraksionet në operacionet e tipeve të të dhënave;
- 3) dizajnimin e përshkrimit vizual dhe teknikave interaktive;
- 4) krijimin e algoritmeve për ekzekutimin efikas të teknikave.

Në **shtresën e parë** është e rëndësishme të definohet qartë objektivat apo qëllimet, të cilat do të duhet të arrihen nga shfrytëzuesit.

Shtresa e dytë paraqet një skenë abstraksioni. Do të jetë e nevojshme të bëhet një përshkrim abstrakt dhe i përgjithshëm i qëllimit që duhet të arrihet apo problemit që do të duhet të zgjidhet dhe tipeve të të dhënave që do të përdoren.

Në **shtresën e tretë** është e rëndësishme të gjendet përshkrimi i përshtatshëm vizual bazuar në tipet e të dhënave dhe qëllimet të cilat duhet të arrihen. Poashtu, në këtë shtresë është e nevojshme të dizajnohen interaksionet, siç janë: sortimi, selektimi apo filtrimi.

Shtresa e katërt është përgjegjëse për krijimin e algoritmit, i cili do të përmbajë përshkrimin vizual dhe interaksionet.

Pra, siç shihet edhe në figurën e ardhshme, të gjitha nivelet e modelit janë të ndërthurura. Si pasojë e kësaj, dalja (output) nga një nivel më lart, paraqet hyrje (input) të nivelit më poshtë, gjë kjo e cila kërkon shumë kujdes gjatë dizajnit. Në rast se bëhet gabim në nivel më lart, ato gabime përcillen në të gjitha nivelet vijuese. Në rast se në fazën e abstraksionit nuk është bërë zgjedhje e duhur, atëherë as edhe përshkrim vizual perfekt si dhe dizajni i algoritmit nuk do të krijojnë sistem vizualizimi që do të zgjidhë problemin e caktuar (Munzner, 2009).

Modeli i ndërthurur ofron instruksione për përcaktimin e qasjeve të duhura validuese gjatë identifikimit të kërcënimeve eventuale të nivelet. Këto kërcënime ndahen në katër kategori:

- 1) problem i gabuar: shfrytëzuesi nuk ka problem të tillë;
- 2) abstraksion i gabuar: gabim në tipet e të dhënave apo qëllimet
- 3) përshkrimi vizual / interaksion i gabuar: mënyra e vizualizimit të të dhënave joefektive
- 4) algoritëm i gabuar: kodi është shumë i ngadaltë

Figura në vijim tregon kërcënimet dhe qasjet e mundshme valide në mënyrë përmblledhëse në çdo nivel.

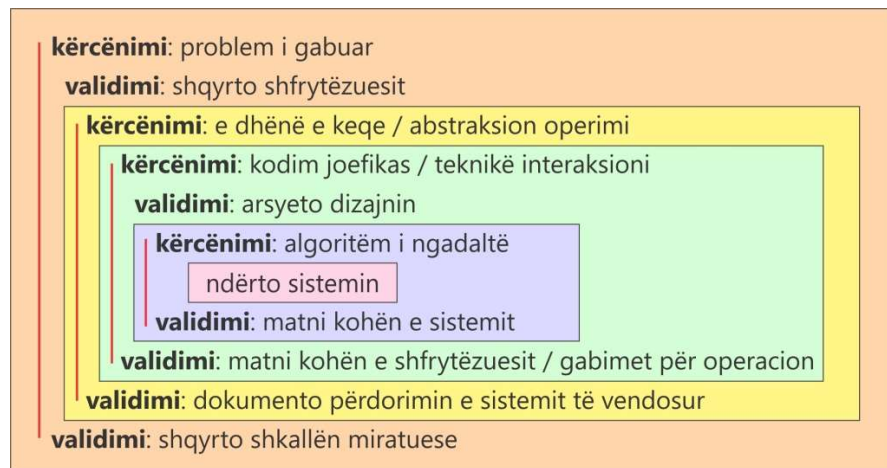


Fig. 20: Kërcënimet dhe validimet në modelin e ndërthurur

Nga figura shihet qartë se modeli i ndërthurur jep idenë e një procesi i cili në mënyrë iterative vjen duke u përsosur, ku validimi bëhet rregullisht gjatë procesit të dizajnit të sistemit. Përveç këtij modeli, gjithashtu, është e mundur që të bëhet evaluim i vizualizimit duke përdorur standard. Pjesa e ardhshme jep një pamje të përgjithshme të DIN EN ISO 9241-12, i cili ofron rregulla të rëndësishme për prezentimin korrekt të informacionit.

3. 3 Standardi DIN EN ISO 9241-12 "Information Presentation"

DIN EN ISO 9241 është standard shumë pjesësh nga International Organization for Standardization (ISO) e cila përmban ergonominë e bashkëveprimit njeri-kompjuter. Ky standard përbëhet nga 17 pjesë, ku një nga to është përgjegjës për paraqitjen e informacionit (Europäisches Komitee für Normung, 1996-07). Standardi DIN EN ISO 9241-12 definon kërkesat e mëposhtme për prezentimin vizual të informacionit:

- 1) **Qartësinë.** Informacioni duhet të interpretohet qartë dhe shpejt.
- 2) **Aftësia për të dalluar informacionin.** Informacioni i afishuar duhet qartë të dallohet.

3) **Kompaktësinë.** Të tregohet vetëm informacioni i cili është i nevojshëm për qëllimin e caktuar.

4) **Konzistent.** Informacioni i cili ka kuptim të njëjtë, duhet të prezantohet në mënyrë të njëjtë.

5) **Vëmendja.** Vëmendja e shfrytëzuesit duhet të shkoj gjithmonë në informacionin e rëndësishëm.

6) **Lexueshmëria.** Informacioni duhet lehtë të lexohet.

7) **Aftësia për të kuptuar informacionin.** Domethënia e informacionit duhet që lehtë të kuptohet dhe interpretohet.

Gjithashtu, është e rëndësishme të përmendet se të gjithë standardet e serisë ISO 9241 bazohen në tre principe kryesore: efektivitetin, efikasitetin dhe satisfaksionin e shfrytëzuesit.

Efektiviteti përshkruan aftësinë e shfrytëzuesve për të realizuar një qëllim të caktuar duke përdorur sistemin, si dhe kualitetin e rezultateve.

Efikasiteti paraqet nivelin e resurseve të konsumuara për kryerje të qëllimit. Ato resurse mund të jenë ose kohore apo njohëse të bëra nga shfrytëzuesit.

Satisfaksioni i shfrytëzuesit është faktor shumë i rëndësishëm i cili tregon reaksionet subjektive të shfrytëzuesve përderisa përdorin sistemin.

4. VIZUALIZIMET E ORIENTUARA NË INTERNET

Meqenëse ditëve të sotme informacionet dixhitale rriten në përmasa të mëdha, nevoja për vizualizimin e tyre bëhet jashtëzakonisht e rëndësishme, sepse ato mundësojnë interpretim më të lehtë dhe më të përshtatshëm të të dhënave dhe analizave më komplekse. Me ndihmën e vizualizimeve, faktet e fshehura për të dhënat mund të paraqiten me ndihmën e modeleve (patterns) apo trendeve tjera. Vendime strategjike mund të arrihen me njohuritë e fituara nga eksplorimet e të dhënave në mënyrë interaktive.

Zhvillimet e fundit në teknologjinë e web browser-ve mundësojnë krijimin e vizualizimeve interaktive të orientuara në web, me idenë e zgjerimit të audiencës. Në këtë temë do të shqyrtohet përdorimi i JavaScript-it në web aplikacione. Do përshkruhen teoritë dhe strukturat (framework) kryesore të vizualizimit dhe do të realizojmë praktikisht ato.

4. 1 Pse vizualizim të orientuar në internet?

Zhvillimi dhe avansimi i shfletuesve të internetit viteve të fundit bën që vizualizimet e orientuara në internet të jenë më të përzgjedhurat kohëve të fundit. Interaksioni dhe përshtatshmëria në pajisje të ndryshme janë ndër cilësitë më të rëndësishme që mundësohen nga internet shfletuesit. Përparësitë që ofrohen nga këto janë të shumta, si p.sh., mundësia që në mënyrë të rregullt dhe pa instalime të ndryshme programore të bëhet shikimi i vizualizimit nëpër pajisje mobile, tabletë apo kompjuterë, mundësia e ndërhyrjes në vizualizim dhe nxjerrja e informacioneve plotësuese e të dhënave apo rezultateve, mundësitë për filtrime, etj. Rritja e kërkesave për interaksionet në kohë reale bën që internet shfletuesit e kohëve të sotme të bëjnë edhe përpunimin e të dhënave, gjë të cilën vite më parë e realizonin serverat.

Meqë JavaScript është gjuha e internet shfletuesve, kjo poashtu shfrytëzohet edhe për vizualizimin e të dhënave. Client-side JavaScript, apo JavaScript interpretuesi i vendosur në internet shfletuesin, bashkon skriptimin në JavaScript së bashku me Document Object Model (DOM). Client-side JavaScript u mundëson shfrytëzuesit të kenë nën kontroll përmbajtjen dhe dukjen e dokumentit, interaksionin me HTML format dhe shfrytëzuesit tjerë. Kjo qasje ka edhe avantazhe edhe disavantazhe. Meqenëse kodi në JavaScript ekzekutohet në kompjuterin e shfrytëzuesit, përpunimi realizohet shumë shpejtë, sepse në server nuk përpunohen të dhënat. JavaScript gjithashtu mundëson funksionalitet të zgjeruar për web faqet, me idenë e përshtatjes së vizualizimeve dhe sjelljes së web faqeve me to. Edhe pse janë të shumta avantazhet e përdorimit të JavaScript, po ashtu është e nevojshme të tregohen edhe disa aspekte negative. Strukturat (frameworks) të cilat përdoren për zgjerimin e funksionaliteti, mund të rrisi çështjet e sigurisë, sepse ato ekzekutohen në client-side.

Për realizimin e vizualizimeve interaktive të orientuara në internet ekzistojnë mjete të ndryshme. Para se të avansoheshin internet shfletuesit, apletet e vogla ishin mjete më të përdorura për vizualizim. Këto kërkonin instalimin e programeve plotësuese si Java Applets apo Adobe Flash. Edhe pse ekzistojnë mjete të cilat ende kërkojnë instalime të tilla, mjete të vizualizimit të orientuar në internet gjithmonë e më shumë tentojnë të bëhen si pjesë përbërëse e internet shfletuesit.

4. 2 Mjetet e vizualizimit të orientuar në internet

Në këtë pjesën vijuese do të bëhet përshkrimi i mjeteve për vizualizim të orientuar në internet (VOI), më saktësisht, do të përshkruhen mundësit që ofrojnë ato, përparësitë dhe mangësitë e tyre, këndvështrimin e sigurisë, si dhe zhvillimet e fundit që lidhen me këto teknologji.

4. 2. 1 Java Applet

Përdorimi i animacionit në web faqe, gjithmonë ka qenë punë mjaft sfiduese dhe gjithashtu e vështirë për zhvilluesit. Në fillimet e hershme, efektet dinamike të web faqeve jepeshin nëpërmjet formatit të animuar fotografik ".gif". Me që këto nuk ishte i mundur interaktiviteti, vite më vonë, për të sofistikuar dhe për të tentuar të sigurohet kontroll më i madh i këtyre përmbajtjeve, filluan të përdoren të ashtuquajturat java applets.

Edhe pse fatkeqësisht java kërkon aftësi programuese, që në realitet paraqet problem shtesë për të krijuar applet-et tona, ekzistojnë edhe applet-e që ofrohen falas të cilat mundësojnë përshtatje. Kështu që, në vend që të mësojnë programimin në java, njerëzit mund lehtësisht të përdorin applet-et e gatshme. Applet-et shpesh kanë qenë edhe të kritikuar për faktin që disa nga ato janë të programuara që eventualisht të shfrytëzojnë të gjithë resurset e kompjuterit, dhe kjo rezulton me "ngrirje" të punës së shfletuesit.

4. 2. 2 Flash

Flash (më parë e quajtur edhe Macromedia Flash) është platformë multimediale e zhvilluar nga Adobe, e cila përdoret për krijimin e përmbajtjeve të ndryshme, si, animacioneve, grafikës vektoriale, video materialeve, lojërave, aplikacioneve kompjuterike dhe mobile, etj. Përmbajtjet e krijuara në Flash, mund t'i vendosim në HTML faqet dhe po ashtu në programe të posaçme që mundësojnë afishimin e përmbajtjeve të këtilla. Për dallim nga applet-et, krijimi i animacioneve në Flash është shumë më i lehtë sepse mundet të bëhet edhe pa kod. Për krijimin e animacioneve më të avansuara, ku vjen në shprehje përdorimi i interaktivitetit, është e domosdoshme të përdoret gjuha programore e Flash-it e quajtur ActionScript.

Me qenë se Flash-i është thellë i ngulitur në web shfletuesit tonë, aplikacionet dhe faqet e internetit të cilat i vizitojmë, çështja e sigurisë është një nga pikat që i shtyn njerëzit të largohen nga përdorimi i tij dhe të mendojnë se si të zëvendësojnë atë. Bazuar në blogun e kompanisë Heimdal (Heimdal), rreth 110 milion faqe interneti që i bie 11% e

të gjithë website-ve në botë përdorin flash, shtuar këtu edhe 1 miliard shfrytëzuesve të Adobe Air e cila është e krijuar në po të njëjtën platformë, bën këto që vazhdimisht të jenë të çenuar nga sulme të ndryshme.

Lidhur me kërcënimet që mund t'i hasin shfrytëzuesit gjatë përdorimit të kësaj platforme, ky blog jep statistikën e cila tregon se 34% e kërcënimeve bëjnë lejimin e ekzekutimit të kodit nga burime të dëmshme, 21% nga kërcënimet krijojnë overflow e cila paraqet anomalinë e sjelljes jo normale të programit, gabimin e rezultateve, rënies së programit, etj., 18% bëjnë sulme të mohimit të shërbimeve - DoS (Denial of Service) attacks, 16% bëjnë korruptimin e memorjes, 5% mundësojnë bajpasin për qasje deri te kompjuteri apo resurset, 3% mundësojnë marrjen e informacioneve dhe 3% të sulmeve e përbëjnë XSS (Cross Site Scripting) 2% dhe CSRF (Cross Site Request Forgery) 1%.

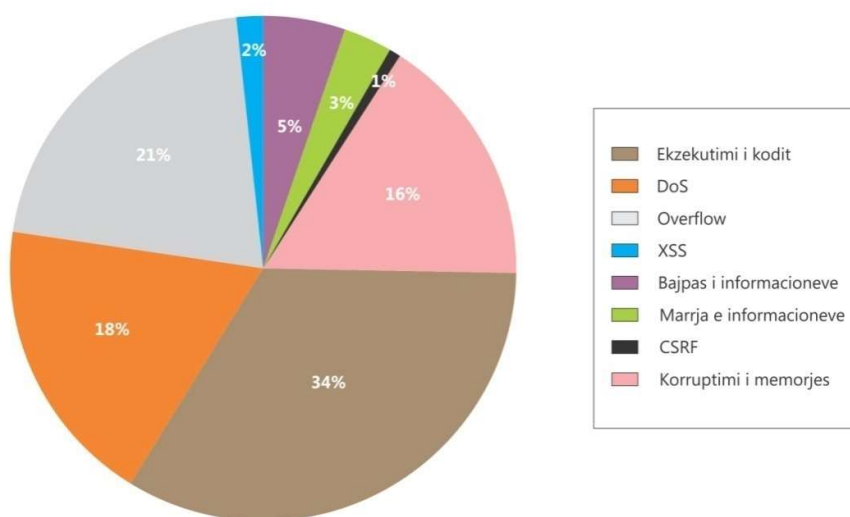


Fig. 21: Statistikë e kategorisë së kërcënimeve në Flash

4. 2. 3 Scalable Vector Graphics (SVG)

SVG është format i grafikës vektoriale i bazuar në XML, i cili përkrah interaksionin dhe animacionin. Konsorciumi World Wide Web (W3C) në mënyrë eksplicite rekomandon përdorimin e SVG-së në shfletuesit e internetit (W3C, 2011).

SVG përdoret për të definuar grafikë vektoriale 2D në internet duke u bazuar në formatin XML. Meqenëse është format vektorial, madhësia e imazhit mund të

ndryshohet pa mos humbur në kualitet, gjë kjo e cila indirekt nënkupton edhe aftësinë për zumim. Mundësia e këtij formati që të integrohet me W3C standardet tjera, siç është Document Object Model (DOM), i bën të mundur që nëpërmjet JavaScript-it të vendosen kode të ndryshme te objektet, siç mund të kemi rast kur shfrytëzuesi të klikojë apo të kalojë në objekt.

Çdo figurë e vizatuar konsiderohet si SVG objekt. Kur shfrytëzuesi ndryshon atributet e një objekti, web shfletuesi e renderon figurën automatikisht. Kështu që, përdorimi i SVG gjen zbatim shumë të gjerë në aplikacionet ku kemi hapësira të mëdha renderuse si Google Maps. Edhe pse ky format sjell përfitime të mëdha si aftësinë për shkallëzim, duhet gjithmonë me kujdes të punojmë dhe të shohim reagimin dhe funksionimin e tij. Arsyeja për këtë qëndron në atë se jo të gjithë internet shfletuesit e shfaqin një lloj, dhe posaçërisht duhet nënvizuar edhe fakti tjetër që në rast se kemi të krijuar objekte më komplekse, atëherë web shfletuesi renderimin e bën më ngadalë.

4. 2. 4 HTML 5 Canvas

Elementi Canvas është pjesë e HTML 5 dhe mundëson renderimin dinamik të figurave 2D. Ky element me atributet e tij për gjerësi dhe gjatësi (width, height) krijon katëror përmbajtës (container) për figurat, dhe nëpërmjet JS bëhet vizatim i grafikëve.

Canvas API paraqet nivel të ulët të modelit të orientuar në piksel. Për dallim nga SVG objektet, një herë pasi përfundojmë me vizatimin e figurës, sistemi "harron" faktet rreth mënyrës së krijimit, gjegjësisht nuk jepet mundësia që të bëhet editim të objekteve të vizatuara. Parim të ngjashëm pune poashtu na mundëson edhe programi i njohur për vizatime *paint*. Dallim tjetër krahasuar me SVG është në mënyrën e vendosjes së event handler-ve për elementet e canvas, p.sh., koordinatat duhet patjetër të përputhen në mënyrë manuale për t'u përcaktuar eventet me klikim të miut.

4. 2. 5 JavaScript

JavaScript (JS) konsiderohet si gjuha programore e internetit. Përdorimin kryesor JavaScript e gjen në anën e klientit, ku shfrytëzuesit e kontrolljnë dhe modifikojnë përmbajtjen e dokumentit në web shfletues.

Programimi në JavaScript ofron mundësinë e manipulimit të përmbajtjes së dokumentit nëpërmjet Document object dhe Element objects që përmban. Dizajnimi apo prezantimi i përmbajtjeve mund të modifikohet duke skriptuar stile në CSS. Shtuar kësaj, përmes përmbajtësve të eventve (event handlers) mund të ndryshojmë sjelljen (behaviour) e dokumentit. Kombinimi i materialit skriptues, prezantimit dhe sjelljes së dokumentit quhet Dynamic HTML ose shkurt DHTML (Flanagan, 2011).

Përdorimi i JavaScript shton aftësi eksploruese, duke mundësuar krijimin e animacioneve, lëshimin e audio dhe video materialeve dhe validimin e vlerave hyrëse.

4. 3 JavaScript dhe internet shfletuesit si platformë vizuale

JavaScript është gjuhë programuese dinamike e cila më tepër përdoret për editimin e përmbajtjes së dokumentit të afishuar në internet shfletues. JavaScript mund të ndryshojë elementet e DOM-it (Document Object Model) të HTML-së, atributet dhe stilet.

Një nga pjesët më të mira të JavaScript-it konsiderohet JavaScript Object Notation (JSON). Me që JSON objekti është në format tekstual, ai mundet lehtë të lexohet. Përveç kësaj mundësie që ofron JSON, vlen të nënvizohet fakti që ky nuk është i varur në ngjuhë programore (Crockford, 2008), gjë e cila mundëson të dhënat të lexohen dhe përdoren nga gjuhë të ndryshme programore.

Meqenëse është shumë e vështirë të përmbash kërkesat të ndryshme të shfletuesve të internetit, ekzistojnë shumë librari të ndryshme të JavaScript për zhvillim të tyre. Megjithëse strukturat (frameworks) e JavaScript-it kanë karakteristika të përbashkëta, si manipulimi në DOM, çdo strukturë (framework) ka funksionalitet të ndryshëm. P.sh., jQuery, një nga libraritë më të përdorshme në website-et (SimilarTech,

2015), përdor CSS selektorët për të manipuluar me HTML elementet, përderisa D3 është librari e JavaScript-it e cila manipulon dokumentet që bazohen në të dhëna duke ofruar komponente të fuqishme vizualizimi.

4. 3. 1 Modeli referues dhe JavaScript

Modeli referues është i rëndësishëm për kuptimin dhe zbatimin e cilësive të JavaScript-it për vizualizimin e të dhënave. Figura në vijim paraqet versionin e modifikuar të modelit referues për vizualizim me JavaScript.

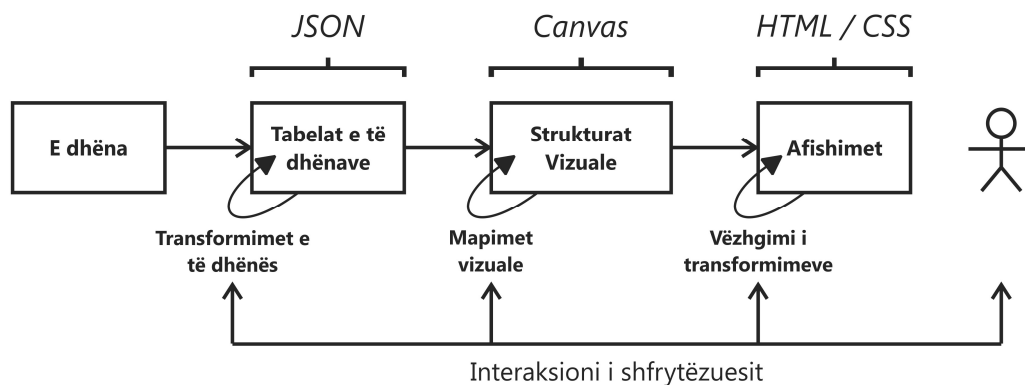


Fig. 22: Modeli referues i adaptuar për JS vizualizim të orientuar në internet

JSON korrespondon me tabelat e të dhënave, me që jep atributet e metadatave. Strukturat vizuale, si grafikoni pjesor (pie chart), paraqesin vizualizime të cilat interpretohen në mënyrë direkte nga njeriu. Elementet canvas të HTML 5 përdoren për të vizatuar elemente në web faqe. View Transformation në mënyrë interaktive modifikon strukturat vizuale për të krijuar vizualizime interaktive (p.sh., zumim dhe fish eye view). Views identifikon shikimin final i cili krijohet përmes HTML elementeve që formatohen përmes CSS.

Modeli referues mund të realizohet shumë lehtë nëpërmjet JavaScript. Kjo gjuhë përfshin struktura të dedikuara, si, JSON, SVG, Canvas dhe HTML, të cilat përshtaten në tabelat e të dhënave, strukturat vizuale dhe shikimin e modelit.

4. 3. 2 Interaksioni dhe animacioni

Robert Spence në veprën (Spence, 2007), për rëndësinë përdorimit të interaksionit në vizualizime thotë: "Interaksioni ndërmjet njeriut dhe kompjuterit është **në zemër të vizualizimit modern**, edhe atë për arsyen e vetme thelbësore: përfitimet e mëdha të cilat mund të rriten nga të qenit **në gjendje për të ndryshuar pikëpamjen e dikujt** për një korpus me të dhëna". Ky korpus zakonisht është aq i madh sa që përmes një shikim gjithëpërfshirës vështirë se do të kishte mundur të arrijmë deri në njohje. Ato të cilët dëshirojnë të fitojnë njohuri, patjetër duhet të eksplorojnë në mënyrë interaktive nëngrupet e atij korpusi për të gjetur rrugën drejt një shikimi i cili do të përmbante "a ha!" faktorin.

JavaScript ofron përkrahje të shkëlqyeshme për interaksion në internet shfletuesit. Client-side JS strukturat (frameworks) mund të përdoren për shtimin e sjelljeve dinamike, siç janë interaksioni dhe animacioni. Shembull, SVG grafikonet shiritore me lëvizje dhe interaksion nga shfrytëzuesi mund të gjenerohet me struktura të ndryshme të JavaScript (p.sh., D3.js).

4. 3. 3 Kompatibiliteti dhe aksesibiliteti

Vizualizimet e orientuara në internet shikohen nga shfletues të kompanive të ndryshme, siç është Google Chrome, Mozilla Firefox, Opera, etj. Është punë mjaft sfiduese krijimi i JS vizualizimit i cili do të shfaqej në mënyrë identike te të gjithë web browser-at. (Flanagan, 2011) rekomandon përdorimin e librarive plotësuese për zgjidhjet e problemeve që lidhen me kompatibilitetin.

Aksesibiliteti (qasshmëria) mund të konsiderohet si një ndër faktorët më të rëndësishëm të VOI. (PewResearchCenter, 2015) tregon se në tetor të vitit 2014, mbi 60% e popullatës së moshuar amerikane është e pajisur me telefona të mençur (smartphones) dhe mbi 40% me tabletë. Kjo statistikë në mënyrë implicite sinjalizon se përveç kompjuterëve, vizitimi i faqeve të internetit bëhet poashtu edhe përmes pajisje të tjera, që d.t.th. se është e një rëndësie enorme që afishimi i faqeve të internetit si dhe aksionet të cilat do të ofrohen tek ato web faqe duhet të jenë të pavarura në pajisje.

4.3.4 Client-side Frameworks

Client-side frameworks ndërtojnë nivel më të lart API (Application Programming Interface) krahasuar me API-t që ofrohen për web shfletuesit. Kodi në JavaScript duhet të jetë i shkruar për API të definuar nga struktura e zgjedhur (Flanagan, 2011). Një strukturë e shkruar në mënyrë të rregullt ndikon po ashtu në çështjet që kanë të bëjnë me kompatibilitetin dhe aksesibilitetin, të cilat u sqaruan në pjesën më lart. Një nga përparësitë më të mëdha që ofrojnë strukturat është zgjerimi i funksionalitetit me më pak kod.

Për vizualizimin e të dhënave ekzistojnë më shumë struktura (framework). Ndër më të njohurat janë:

- **Processing.js:** Kjo strukturë paraqet JS port të Processing programming language të dizajnuar për vizualizime në web. Ajo konverton kodin nga Processing në JavaScript dhe e ekzekuton. Ky modul bën të mundur përdorimin e elementit canvas të HTML 5 për renderim të vizualizimit 2D dhe 3D. Processing API përfshin metoda të ndryshme si canvas dhe manipulim me ngjyra, forma gjeometrike, vizatime dhe funksione matematikore.
- **Raphaël.js:** Kjo librari mundëson API për manipulim me SVG. Kjo më së shumti përdoret për vizatim të grafikës vektoriale në internet. Përveç këtij qëllimi, objektiv tjetër i këtij API është zgjidhja e problemeve që lidhen me kompatibilitetin e versioneve të internet shfletuesve që përdorin VML (Vector Markup Language). Web faqe si Washington Post dhe Times Online përdorin këtë strukturë nga arsyeja e përparësive kompatible.
- **Protovis.js:** Bostock dhe Heer (Bostock & Heer, Protovis: A Graphical Toolkit for Visualization, 2009) krijuan librarinë Protovis JS për gjenerimin e SVG grafikëve. Në Protovis, vizualizimet dizajnohen në formë hierarkie shënimesh (hierarchy of marks) dhe poashtu këtu mundësohen cilësi si trashëgimia, shkallëzimi dhe organizimi. Limitimi i mundësisë së animacionit dhe lëvizjeve bënë autorët e Protovis-it të zhvillojnë pasuesin e Protovis, D3.js, për të siguruar animacion dhe interaksion më të avansuar.

- **D3.js:** Data-Driven Document (D3) është librari e JS, e cila si qëllimi kryesor ka krijimin e vizualizimeve interaktive dhe responsive duke lejuar animacionin. D3 JavaScript bën të mundur përdorimin e SVG, HTML 5 dhe CSS standardeve për krijim të vizualizimeve më të avansuara.

Ekzistojnë librari të ndryshme të JavaScript për të bërë më të lehtë zhvillimin e JavaScript aplikacionit. Çdo strukturë e JavaScript-it ka funksionalitet të definuar si manipulimin me DOM, zhvillimin e GUI, zhvillimin grafik dhe vizualizues dhe gjithashtu web aplikacionin.

5. PLATFORMA D3.js

D3.js (ose vetëm D3 për *Data-Driven Documents*) është librari e JavaScript për krijimin e vizualizimeve interaktive dhe dinamike të orientuara në internet. Krijuesi i D3.js Mike Bostock, përpara se të krijoj D3.js, ka kontribuar edhe në zhvillimin e librarive tjera për vizualizim të të dhënave, si Protovis dhe Polymaps, të cilat më vonë do i shërbenin si motiv për të bërë një hap më të madh, si krijimin e një librarie akoma më të sofistikuar e cila do të fokusohet në standarde dhe web shfletues modern. Dizajni i fuqishëm i kësaj librarie mundëson që kjo librari të mos përdoret vetëm për krijimin e grafikoneve, por edhe për shumë lloje tjera vizualizimi, si krijimin e hartave, diagrameve interaktive, dashbordeve, si dhe gjithçka tjetër që shohim në web. Kjo në realitet, jep kuptimin më të mirë për atë se çfarë në të vërtetë paraqet vizualizimi.

Kohë më parë, krijimi i aplikacioneve apo përmbajtjeve më të avansuara në internet, nuk ishte i mundur në rast se ato nuk ishin të realizuara nëpërmjet Flash-it apo Java applet-it. D3.js mundëson të njëjtat performansa, por të integruara në web standarde dhe DOM në thelb të HTML.

Edhe pse aftësia për të krijuar lloje të ndryshme të grafikëve është një nga pikat më të fuqishme të D3.js, niveli i lartë i interaktiviteti mund të llogaritet si aftësi më e rëndësishme që u mundësohet web zhvilluesve modern. Nëpërmjet D3.js, çdo element i çdo grafikoni, mund të bëhet interaktiv në të njëjtën mënyrë.

5.1 Qëllimet kryesore

Bazuar në (Bostock, Ogievetsky, & Heer, D3: Data-Driven Documents, 2011), vizualizimet interaktive duhet të përmbajnë web teknologji të ndryshme, si HTML, CSS, JavaScript dhe SVG. Ato thonë që veglat e vizualizimit të cilat nuk kanë aftësi për manipulim direkt të DOM-it ekzistues, nuk mund të përdorin web teknologji të ndryshme (Bostock, Ogievetsky, & Heer, D3: Data-Driven Documents, 2011).

Me qëllim të tejkalimit të mangësive në krahasim me veglat më të hershëm për vizualizim, D3.js zhvilluesit kategorizuan qëllimet si vijon:

- **Kompatibilitetin:** D3.js mundëson komponente të ripërdorshme për avansimin e aksesibilitetit dhe efikasitetit. Poashtu, librari plotësuese të JS mund të përdoren për të qenë kompatibil për versione më të vjetra internet shfletuesish.
- **Debugimin:** D3.js thjeshtëson debugimin për arsye të JS konzolës në veglat e zhvilluesit. Zhvilluesit mund të bëjnë debugimin në mënyrë interaktive te shfletuesit e internetit nëpërmjet JS.

5. 2 Krijimi i dizajnit

Funksioni atomik i D3.js mund të konsiderohet operacioni *selection*. Elementet mund të ndikohen nga dokumenti duke përdorur *selection* metodën. Pasi të selektohen elementet relevante, operatorët mund të zbatohen për t'u modifikuar përmbajtja (p.sh. atributet, stilet, HTML). Event handlers mund të përdoren për krijimin e interaksioneve, ku poashtu me ndryshimin e attributeve mund të vij në shprehje efekti i animacionit.

5. 2. 1 Selektorët

D3 bën të mundur përdorimin e CSS 3 për selektim të elementeve. Ekzistojnë dy funksione selektimi: **select** dhe **selectAll**. Elementet mund të filtrohen dhe selektohen sipas tageve ("tag"), klasave (".class"), identifikatorëve ("#id"), attributeve ("[name=value]"), si dhe tipareve tjera (Bostock, Ogievetsky, & Heer, D3: Data-Driven Documents, 2011). Në vazhdim do të bëjmë një krahasim të mënyrave se si bëhet një modifikim i thjeshtë (p.sh., ndryshimi i ngjyrës së tekstit të paragrafëve) në W3C DOM API dhe D3.

```
var paragraphs = document.getElementsByTagName("p");
for (var i = 0; i < paragraphs.length; i++) {
    var paragraph = paragraphs.item(i);
    paragraph.style.setProperty("color", "white", null); }
```

Siç mund të shihet modifikimi i dokumentit duke përdorur W3C DOM API është pak sa i lodhshëm për atë se emrat e metodave janë me shumë fjalë dhe qasja e domosdoshme kërkon iteracion manual dhe vëzhgim të gjendjes temporale. Ndërsa për të bërë ndryshimin në D3, këtu kemi:

```
d3.selectAll("p").style("color", "white");
```

Poashtu, këtu mundësohet edhe manipulimi i nyjeve individualisht:

```
d3.select("body").style("background-color", "black");
```

Formati selektor, i definuar nga W3C Selectors API, është kompatibil me web shfletuesit modern. Pasi të selektohen elementet, operatorët mund të aplikohen për manipulim me atributet, stilet, vetitë, HTML dhe përmbajtjet tekstuale të elementeve të selektuara.

5. 2. 2 Operacionet me të dhënat

Në D3.js, të dhënat janë të lidhura në DOM elementet. D3 lejon ripërdorimin duke u mundësuar shfrytëzuesve punë me grupe të dhënash të ndryshme, pa mos u kërkuar veprime plotësuese. D3 pranon tipe të ndryshme të të dhënave si vargjet, stringjet dhe objektet. Shtuar këtyre, kjo mundet të trajtoj edhe grupe të dhënash shumë të mëdha në formatin JSON apo CSV skedarët (fajllat).

5. 2. 3 Interaksioni dhe Animacioni

D3.js më tepër fokusohet në reprezentimin dinamik se sa në atë statik të të dhënave. Kështu që, kjo kornizë përmban përkrahjen për tranzicion (lëvizje). Elementeve të selektuara mund t'iu vendosen event listener-a (dëgjues / reagues ngjarjesh) që të ofrojnë ngjarje apo mundësi të caktuara të cilat do të përkrahen dhe lejohen nga web shfletuesi (psh., click, mouseover, submit).

Tranzicionet dërgojnë ngjarjet (events) te objekti i adresuar dhe si rrjedhojë elementeve u mundësohet lëvizja. D3 automatikisht menaxhon me orarin e tranzicioneve, duke u siguruar elementeve ekskluzivitetin, efikasitetin dhe rrjedhën e caktuar kohore (Bostock, Ogievetsky, & Heer, D3: Data-Driven Documents, 2011).

6. VIZUALIZIMI I TË DHËNAVE TEKSTUALE.

RAST STUDIMI: ANALIZA E MEDIUMEVE NË R. MAQEDONISË

Në dekadën e fundit, interesimi për vizualizimin e të dhënave tekstuale ka qenë gjithmonë e më shumë në rritje. Arsytet për këtë janë të shumta, por megjithatë, disponueshmëria e sasive të mëdha të të dhënave tekstuale heterogjene (të shkaktuara nga popullariteti i rrjeteve sociale si dhe web faqeve të ndryshme) dhe mundësia e zbatimit të algoritmeve të ndryshme për procesim teksti, mund të identifikohen si dy arsye më të forta.

Këto dy arsye, poashtu ishin edhe inspiruesit kryesor të kësaj teme, ku si objektiv kryesor kemi analizën e medimeve online në Maqedoni dhe vizualizimin e rezultateve përmes së cilave do të fitohen njohuri të reja. Për të analizuar këto lajme, ne kemi përdorur shërbimin <http://timemachine.truthmeter.mk>¹.

Kjo platformë na mundëson që nëpërmjet shërbimeve të web shërbimeve (REST shërbime) të mund të filtrojmë të gjitha lajmet e publikuara online në Maqedoni. Në të janë të mbledhura lajme nga 20 medime në gjuhën shqipe dhe 33 medime në gjuhën maqedonase. Sistemi lejon që lajmet të filtrohen sipas parametrave më poshtë:

gjuhës, burimit, datës (prej - deri), numri i lajmeve të kthyer dhe fjalës.

Për më tepër rreth API të sistemit mund të shihet në:

<http://timemachine.truthmeter.mk/api>

Me qëllim që të mundemi të nxjerrim të dhëna me vlerë, ne zhvilluam një aplikacion, i cili do të shërbej për analizën e të dhënave nga shërbimi i përmendur:

¹ timemachine.truthmeter.mk është sistem që bën indeksimin e medimeve në Maqedoni. Për më tepër: <http://timemachine.truthmeter.mk/about-project>

Fig. 23: Aplikacioni 1

Pasi që sistemi të dhënat i kthen në JSON, është e nevojshme të bëhet serializim i të dhënave në C#. Për këtë qëllim, zhvilluam dy klasa:

```
class NewsSource
{
    public int id { get; set; }
    public string name { get; set; }
    public string rss_feed { get; set; }
    public string homepage { get; set; }
    public string source_language { get; set; }
    public string created_at { get; set; }
    public string created_by { get; set; }
    public string updated_at { get; set; }
    public string updated_by { get; set; }

    public string delted_at { get; set; }
    public string delted_by { get; set; }
}

class NewsArticle
{
    public int id { get; set; }
    public int source_id { get; set; }
    public string url { get; set; }
    public string title { get; set; }
    public string short_content { get; set; }
```

```

        public string content { get; set; }
        public string media { get; set; }
        public string screenshot_url { get; set; }
        public string author { get; set; }
        public string category { get; set; }
        public string rss_time { get; set; }
        public string retrieved_time { get; set; }
        public bool is_processed { get; set; }
        public object parent_id { get; set; }
        public int scheduler_id { get; set; }
        public string lang { get; set; }
        public List<object> similar { get; set; }
    }
}

```

Më pas, leximi i të dhënave bëhet në këtë mënyrë:

```

HttpRequest request =
(HttpWebRequest)WebRequest.Create("http://timemachine.truthmeter.mk/api/
v1/getSources?X-
Authorization=a0dcec4a54d55029cb0ad5856d7f9aca6becfa76&lang=" +
cmbLanguages.Text);
WebResponse response = request.GetResponse();
using (Stream responseStream = response.GetResponseStream())
{
StreamReader reader = new StreamReader(responseStream, Encoding.UTF8);
string rez = reader.ReadToEnd();
List<NewsSource> results =
JsonConvert.DeserializeObject<List<NewsSource>>(rez);
// Manipulohet me te dhenat e listes results
}

```

Përpos këtij aplikacioni, ne poashtu shfrytëzuam edhe web aplikacionin <http://wordle.net>, te i cili manualisht futëm përmbajtjet sipas ndonjë kriteri të caktuar dhe nga këtu morrëm teknikë të duhur vizualizuese.

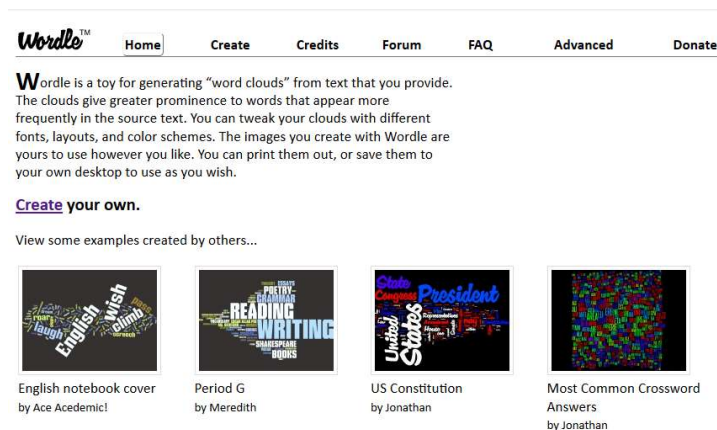


Fig. 24: Web aplikacioni wordle.net

Sfidë tjetër poashtu e rëndësishme në këtë temë ishte përdorimi i vizualizimeve të njohura si “circos”. Për realizimin e tyre ne zhvilluam një aplikacion i cili nga fajlla tekstual krijon model dataset-i i cili më pas duhet të bëhet upload në web aplikacionin <http://mkweb.bcgsc.ca/tableviewer/visualize/> dhe prej këtu nxirret vizualizimi.

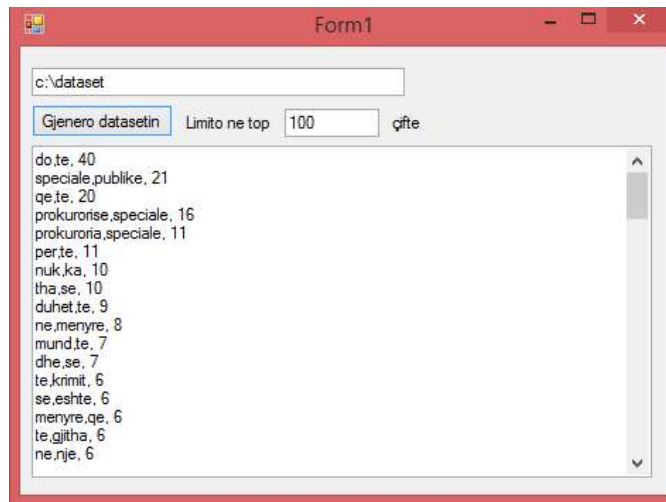


Fig. 25: Aplikacioni 2

Pjesa vijuese jep pjesën më të rëndësishme të kodit:

```
//fillimisht i ndajm rreshtat
char[] line_delim = { '.', '!', '?' };
string[] reshtet = all_content.Split(line_delim);

//Krijojme nje strukture dictionary qe mban çiftet e
fjaleve ne formatin (f1, f2) dhe frekuencen

foreach (string resht in reshtet)
{
    char[] word_delim = { ',', ' ', ';', ':', '"', '\',
        "'"};

    //shkaku limitimit te ASCII
    string[] fjalet = resht
        .Replace('ë', 'e')
        .Replace('ç', 'c')
        .ToLower().Split(word_delim);

    //elimino fjalet qe perdoren shpesh (stopwords)
    fjalet = fjalet.ToList().Except(stop_words).ToArray();

    fjalet = fjalet.Where(x => x.Length >= 2).ToArray();
    fjalet = fjalet.Select(x => x.Trim()).ToArray();

    foreach (string f in fjalet)
```

```

    {
        if (labels.Contains(f) == false)
        {
            labels.Add(f);
        }
    }
    //Tani na duhen te gjithe kombinimet e fjaleve
    List<string> kombinimet = new List<string>();

    for (int i = 0; i < fjalet.Length - 1; i++)
    {
        try
        {
            string komb = fjalet[i] + "," + fjalet[i + 1];
            kombinimet.Add(komb);
        }
        catch (Exception)
        {
        }

    }

}

foreach (string s in kombinimet)
{
    if (s.Contains(","))
    {
        if (res.ContainsKey(s))
        {
            res[s]++;
        }
        else
        {
            res.Add(s, 1);
        }
    }
}

int limit = Convert.ToInt32(txtLimit.Text);
foreach (KeyValuePair<string, int> kv in
res.OrderByDescending(x=>x.Value).Take(limit))
{
    txtResult.Text += kv.Key + ", " + kv.Value +
    Environment.NewLine;
}

```




Fig. 26: Web aplikacioni mkweb.bcgsc.ca/tableviewer/visualize/

6. 1 Rezultatet e fituara

Me përdorimin e aplikacionit të përmendur, si dhe duke iu referuar edhe rezultateve tjera të fituara nga aplikacione shtuese, në vijim do bëjmë vizualizimin e tyre duke përdorur teknika të ndryshme vizualizimi. Secila teknikë e përdorur do të mvaret nga mesazhi të cilin do të duhet ta përcjellë. Në vazhdim do të shënojmë rezultatet e fituara dhe vizualizimet përkatëse:

Rezultati 1: *Frekuenca e përdorimit të emrave dhe mbiemrave të politikanëve shqiptar dhe maqedonas te mediumet online në gjuhën shqipe, nga 20 mars 2016 - 20 mars 2017, si dhe nga data 27.04.2017 – 12.05.2017:*

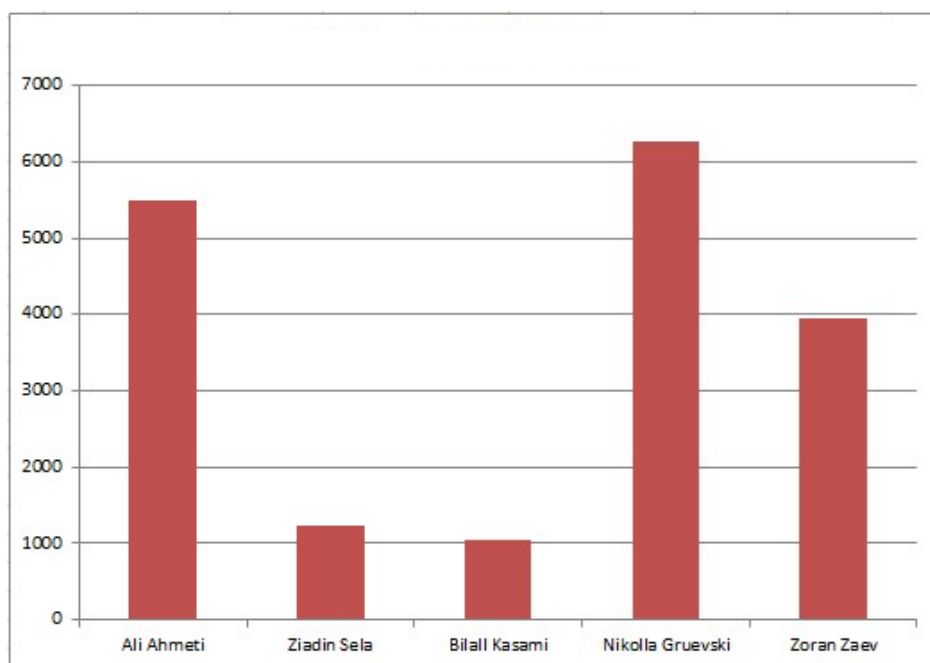


Fig. 27: Frekuenca e përdorimit të emrave dhe mbiemrave të politikanëve të R. M. - 1

Për realizimin e këtij grafikoni përdorëm aplikacionin i cili në mënyrë automatike bën analizën e medimeve online në Maqedoni. Në rastin e mësipërm analiza është bërë duke u bazuar në termet që kemi shënuar, ku konkretisht gjejmë frekuencën e përdorimit të tyre, e cila bazohet në gjuhën e zgjedhur të medimeve dhe poashtu zgjedhjen e datave se prej kur e deri kur të bëhet ajo analizë. Është interesante të shihet në

vizualizimin e mësipërm se lideri politik Ali Ahmeti ndan përfshirje pothuajse të njëjtë me dy liderët maqedonas, ndërsa dy liderët tjerë shqiptarë Bilal Kasami dhe Ziadin Sela përfaqësohen me rreth 20% të totalit të përfaqësimit të liderit Ali Ahmeti.

Vizualizimi ndryshon mjaft nëse ndryshohet intervali kohor! Për ilustrim, bëme një analizë të përfshirjes së liderëve politik pas ngjarjeve në parlamentin Maqedonas të datës 27 prill. Frekuenca e përfshirjes së liderëve politik në Maqedoni pas kësaj dite e deri në 12 maj, dallon bindshëm. Grafikoni i mëposhtëm qartë ilustron këtë.

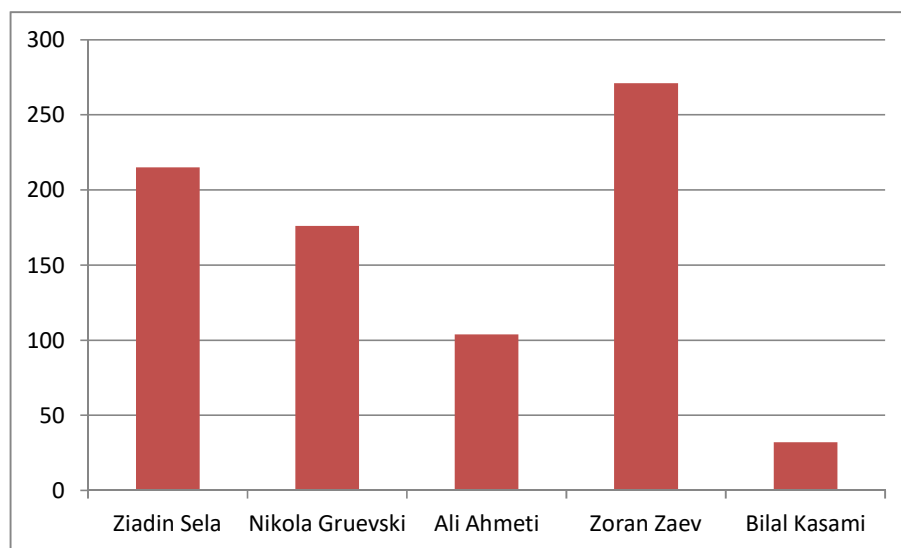


Fig. 28: Frekuenca e përdorimit të emrave dhe mbiemrave të politikanëve të R. M. - 2

Rezultati 2: *Fjalët më të përdorura nga data 10.11.2016 – 30.11.2016 nga mediumi online në qjuhën shqipe “TV-Shenja”:*

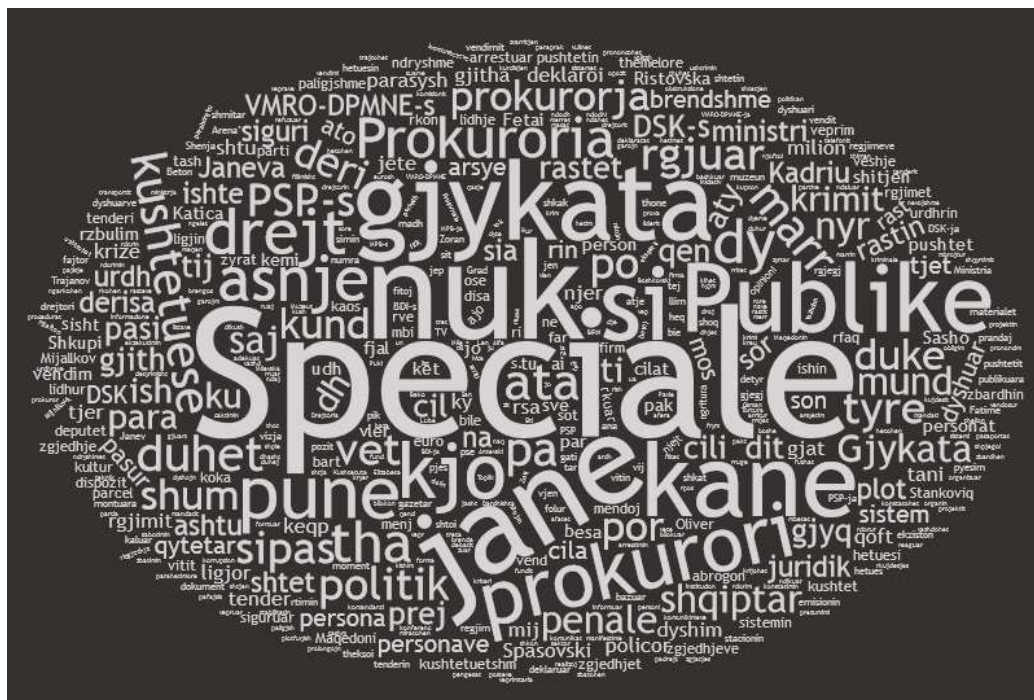


Fig. 29: Fjalët më të përdorura në mediumin “TV-Shenja”

Vizualizimi i njohur si “word cloud” u realizua nëpërmjet një ueb aplikacioni i quajtur wordle. Kjo analizë u bë në mënyrë manuale, duke marrë të gjitha lajmet e publikuara nga mediumi “TV-Shenja”, prej në datën 10.11.2016 – 30.11.2016. Analiza që u bë nxjerr në pah fjalët më të përdorura të të gjithë lajmeve të publikuara në ato data. Ky model vizualizimi përdor madhësinë e shkronjave për të determinuar se cilat kanë qenë fjalët më shpesh të përdorura. Duket qartë nga vizualizimi i mësipërm se cilët kanë qënë temat kryesore për të cilat mediumi i zgjedhur ka raportuar. Sipas vizualizimit, mediumi ka raportuar shumë për gjykatat, gjykimet dhe prokurorinë speciale publike. Kjo është edhe një rezultat i pritur nëse analizohen të dhënat nga konotacioni kohor dhe situata politike në vend.

Rezultati 3: Frekuenca e përdorimit të fjalëve në platformat e partive politike VMRO-DPMNE dhe LSDM në zgjedhjet parlamentare të vitit 2016:

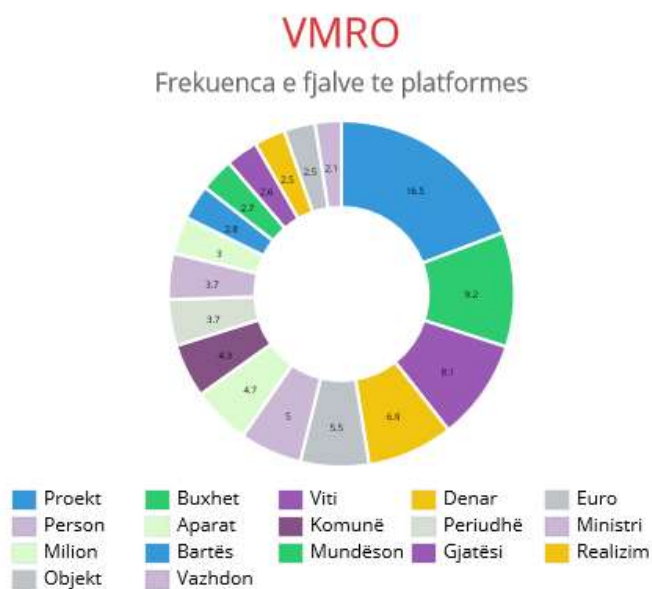


Fig. 30: Frekuenca e përdorimit të fjalëve në platformën e VMRO-DPMNE

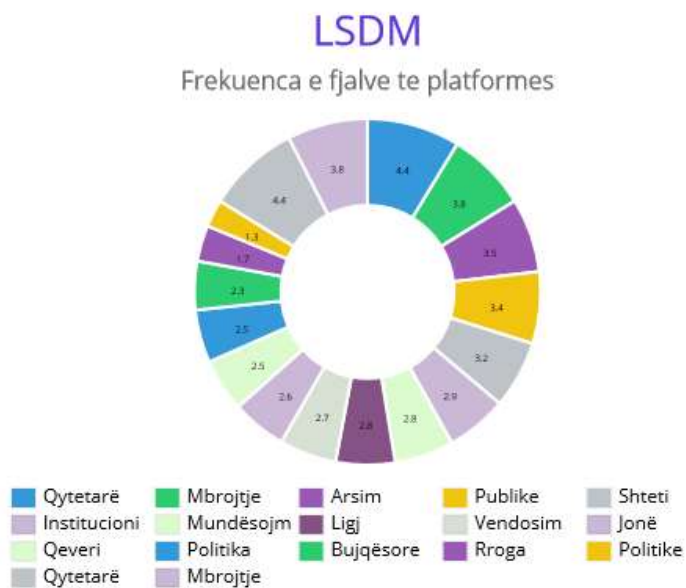


Fig. 31: Frekuenca e përdorimit të fjalëve në platformën e LSDM

Një studim interesant që mendonim se do jetë me vlerë është edhe krahasimi i platformave të dy partive politike Maqedonase². Në këtë grafikun paraqitëm fjalët më shpesh të përdorura në platformat e zgjedhjeve parlamentare të vitit 2016 të dy partive maqedonase VMRO-DPMNE dhe LSDM. Këto të dhëna janë publike dhe të lejuara për përdorim, dhe ideja jonë ishte që të bënim krahasimin e tyre nëpërmjet një grafikoni modern interaktiv, i cili u realizua nëpërmjet një ueb aplikacioni.

Nga grafikonet e sipërme dallohet qartë orientimi politik i këtyre dy partive. Përderisa partia LSDM përqëndrohet në konceptin qytetar, arsimimin, rroga etj., partia në pushtet VMRO përqëndrohet në projekte, buxhet, realizim projektesh etj.

² Arsyeja e zgjedhjes së partive politike maqedonase ishte për shkak se partitë politike shqiptare në Maqedoni nuk kanë platforma të publikuara ose në rast se kanë, ato janë mjaft të varfëra.

Rezultati 4: Trendi i përdorimit të fjalëve nga 1 janar 2016 deri në 29 mars 2017:

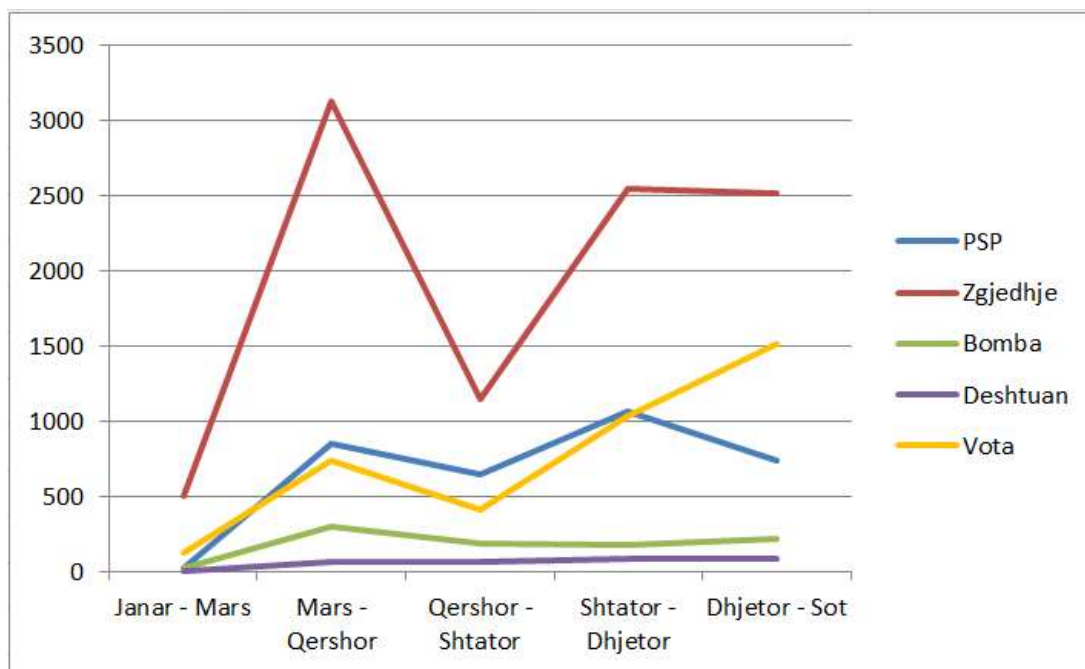


Fig. 32: Trendi i përdorimit të fjalëve

Vizualizimi i njohur si “*line chart*”, u përdor për treguar se me sa frekuencë në të gjithë mediumet e Republikës së Maqedonisë janë përdorur termet “PSP”, “Zgjedhje”, “Bomba”, “Dështuan” dhe “Vota” nga janari i vitit 2016 e deri në ditën e sotme. Analiza e këtyre fjalëve është bërë me ndihmën e aplikacionit për analizën automatike të medimeve online në R. Maqedonisë. Frekuencat e të gjitha këtyre fjalëve i merrnim duke u bazuar në data të caktuar (për çdo tre muaj), me qëllimin që të zbulohen trende dhe njohuri të reja.

Nga grafikoni i mësipërm ajo që lehtë mund të vërehet është se prej të gjithë fjalëve të marra në shqyrtim, termi “Zgjedhje” është bindshëm më i përdorur në krahasim me tjerët, posaçërisht në kuartilin e dytë dhe dhe të katërt të vitit, ku në njërin u paralajmëruan zgjedhjet parlamentare dhe nuk u mbajtën dhe në periudhën tjetër kur u mbajtën. Për të plotësuar akoma këtë, kemi edhe termin “Vota” e cila gjithmonë e më shumë është përdorur në kuartilin e fundit të vitit dhe kjo në të vërtetë paraqet një identifikator plotësues për zgjedhjet e mbajtura në këtë periudhë.

Rezultati 5: *Krahasimi i medimeve Sitel, Kanal 5, Telma dhe Shenja lidhur me përmbajtjet e lajmeve të tyre nga data 21.03.2017 – 28.03.2017:*

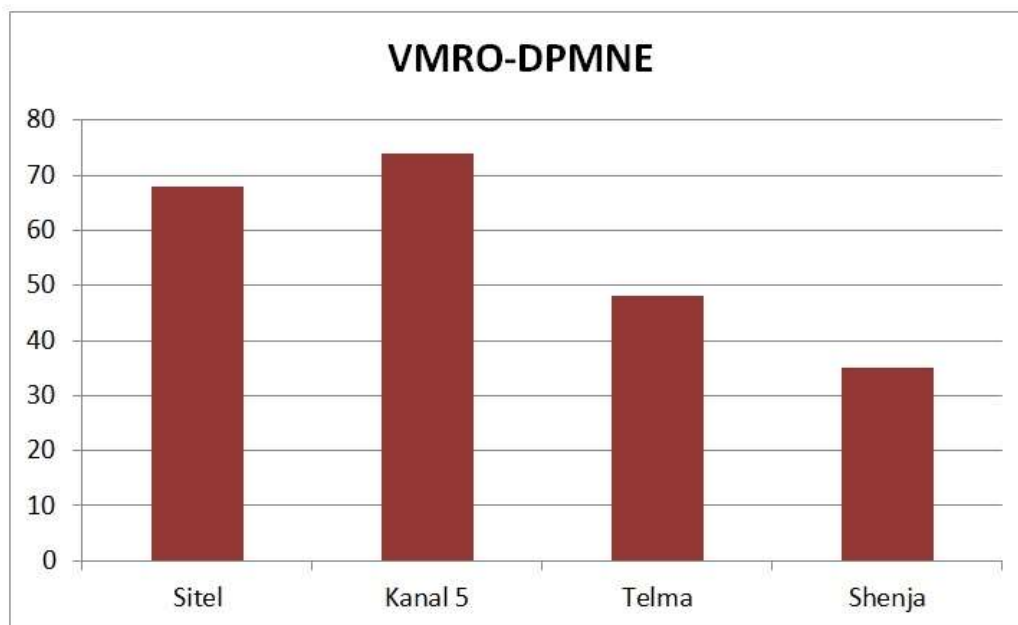


Fig. 33: Frekuenca e përdorimit të termit VMRO-DPMNE në përmbajtjet e katër medimeve

Ky vizualizim tregon se sa herë në përmbajtjet e lajmeve të publikuara nga mediumet online, “Sitel”, “Kanal 5”, “Telma” dhe “TV-Shenja” nga data 21.03.2017 – 28.03.2017, si temë kanë pasur partinë VMRO-DPMNE.

Nga ky grafik mund të shihet se dy mediumet “Sitel” dhe “Kanal 5” dukshëm trajtojnë temat të cilat lidhen me këtë parti. Numri i përmbajtjeve të cilat ato publikuan brenda këtyre datave është sa dyfishi apo ndoshta edhe më tepër krahasuar me dy mediumet tjera.

Rezultati 6: Analiza e përmbajtjeve të shkrimit të analistëve politik A. Bushati dhe M. Nazarko nga periudha 01.05.2017 – 15.05.2017:



Fig. 34: Vizualizimi nëpërmjet teknikës “word cloud”

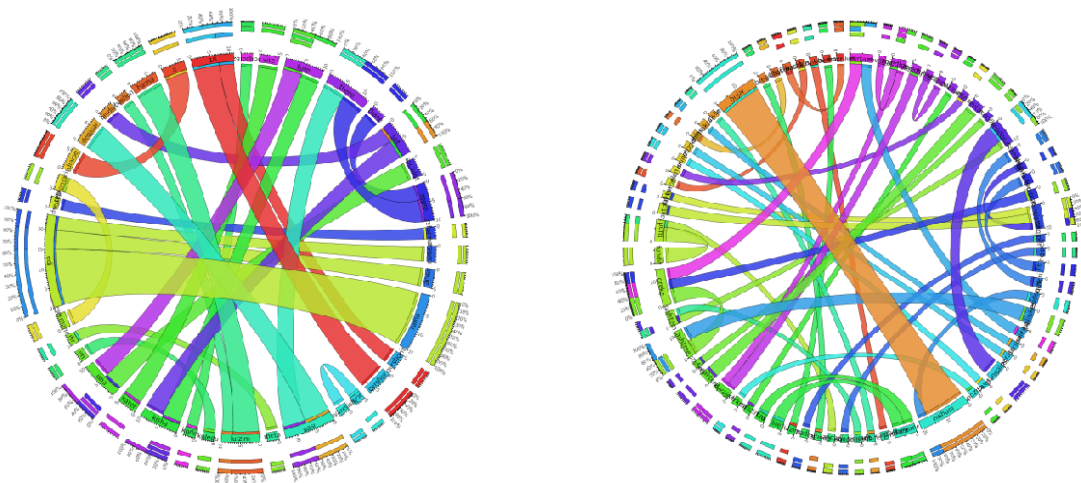


Fig. 35: Vizualizimi nëpërmjet teknikës “circos”

Një nga hipotezat e kësaj teze ishte se ne do të mund të “shohim” anshmërinë që medime të caktuara e mbajnë ndaj një ideje, partie politike ose koncepti. Për këtë arsye zgjodhëm që të analizojmë të gjithë shkrimet e dy analistëve politik (Mentor Nazarko dhe Andi Bushati) të shkruara gjatë muajit Maj të 2017. Përzgjedhja e analistëve ishte e rastësishme, por kishim kujdes që të dy të jenë analistë politikë me qëllim që temat e përfshira të kenë ngjashmëri.

Për këtë qëllim ne përdorëm dy teknika të ndryshme vizualizimi, të cilat në njërin na jepet frekuenca e fjalëve më të përdorura në publikimet e tyre për datat referuese, ndërsa vizualizimi tjetër përveç që mundëson lidhjen e fjalëve të përdorura, ajo poashtu na tregon se cila nga ato fjalë është përdorur më shpesh dhe me cilën.

7. PËRFUNDIMI

Në këtë temë ne arritëm t'i japim një shpjegim të mirfilltë vizualizimit të të dhënave, duke u nisur nga zanafilla e tij e deri në zhvillimet që ka hasur ditëve të sotme. U trajtuan elemente shumë esenciale të tij, si për nga mënyra se si perceptohet nga njeriu, e deri te standardet e rekomanduara për paraqitjen e tij.

Objektiv tjetër të rëndësishëm të cilin arritëm ta realizojmë ishte edhe krijimin e një aplikacioni i cili do të bëjë analizën automatike të medimeve online në R. Maqedonisë si dhe poashtu përdorëm aplikacione tjera shtuese me qëllim që të nxjerrim më tepër modele vizualizimi.

Hipotezat e këtij punimi ishin:

- ✓ **Hipoteza I:** Analiza automatike e medimeve në Republikën e Maqedonisë mund të realizohet dhe të nxjerrë në pah mangësitë e medimeve.

Mbështetur në hipotezën e parë edhe dy hipotezat alternative se:

- Të gjithë publikimet e medimeve online në Maqedoni mund të mblidhen dhe të analizohen statistikisht;
- Vizualizimi i të dhënave mund të nxjerrë në pah njëanshmërinë e medimeve.

Hipotezën e parë e vërtetuar me ndihmën e aplikacionit i cili bënë analizën automatike të medimeve online në R. Maqedonisë dhe poashtu duke përdorur aplikacione shtesë, me qëllim që rezultatet e fituara nga këto aplikacione të mund të vizualizohen në sa më shumë teknika të përshtatshme vizualizimi.

Për të vërtetuar hipotezën e dytë, ne përdorëm sistemin vetëm se të implementuar - "timemachine.truthmeter.mk"

7. 1 Udhëzim për zhvillim të mëtutjeshëm

Viteve të fundit, zhvillimi i hovshëm i teknologjisë, gjithmonë e më shumë bën që mediumet përpos mënyrës së zakonshme të shfaqjes të tyre, të mendojnë se si ato të njejtën ta bëjnë edhe në internet. Përparësitë që na ofrohen nga përdorimi i teknologjisë janë të shumta, por sigurisht që duhet të merret parasysh edhe fakti që e tregon thënia: “Ne jemi të pasur me të dhëna, por të varfër me informata”.

Ideja kryesore e kësaj teme ishte që ne të mund të analizojmë publikimet e mediumeve online në R. Maqedonisë dhe të vizualizohen rezultatet e marra nga ato, me tendencë që të pasurojmë këtë “varfëri të informatave”.

Sistemin të cilin e krijuam për analizën e mediumeve online në R. Maqedonisë, sigurisht që akoma mund ta sofistikojmë dhe ta bëjmë më efikas. Këtë përmirësim që mund të hasi ky sistem e shoh në dy pikëpamje të ndryshme. Nga njëra anë nëse i referohemi mënyrës se si funksionon aktualisht, do ishte mirë që si plotësim aplikacioni të posedonte një hapësirë të posaçme në të, ku aty do nxirrte lloj vizualizimi të zgjedhur nga shfrytëzuesit. Si zhvillim tjetër që mendoj se do të kishte qenë pozitive për një numër të madhë njerëzish, është nëse një i tillë zhvillohet si “shtesë” (add-on, extension) për web shfletuesit, i cili në kohë reale shfrytëzuesve do u jepte informacione lidhur me përmbajtjen e faqes së hapur.

LITERATURA

- Bostock, M., & Heer, J. (2009). Protovis: A Graphical Toolkit for Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 1121-1128.
- Bostock, M., Ogievetsky, V., & Heer, J. (2011). D3: Data-Driven Documents. *IEEE Transaction on Visualization and Computer Graphics*, 2301-2309.
- Chen, C.-h., Härdle, W., & Unwin, A. (2008). *Handbook of Data Visualization*. Berlin: Springer-Verlag.
- Crockford, D. (2008). *JavaScript: The Good Parts*. Beijing: O'Reilly.
- Few, S. (2004). Selecting the Right Graph for Your Message.
- Few, S. (2009). *Now You See It*.
- Flanagan, D. (2011). *JavaScript: The Definitive Guide, 6th Edition*. Sebastopol, CA: O'Reilly.
- Heimdal. (2016, April 24). www.heimdalsecurity.com/blog/adobe-flash-vulnerabilities-security-risks/. Gjetur April 24, 2016, nga www.heimdalsecurity.com: <https://heimdalsecurity.com/blog/adobe-flash-vulnerabilities-security-risks/>
- Liu, S., Zhou, M. X., Pan, S., Qian, W., Cai, W., & Lian, X. (2009). Interactive, topic-based visual text summarization and analysis. *CIKM* (fv. 543-552). New York: ACM.
- Lu, H., & Liu, G. (2011). Text Visualization and Visual Analytics Based on Multi-layer Topic Maps. *J. Information and Computational Science*, 2459-2464.
- Munzner, T. (2009). A Nested Model for Visualization Design and Validation.
- PewResearchCenter. (2015). *Mobile Technology Fact Sheet*. Gjetur December 11, 2015, nga Pew Research Center: <http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/>
- Plaisant, C. (2004). The Challenge of Information Visualization Evaluation.
- SimilarTech. (2015, December 7). *JavaScript Technologies Market Share and Web Usage Statistics*. Gjetur 12 7, 2015, nga SimilarTech: <https://www.similartech.com/categories/javascript>
- Spence, R. (2007). *Information Visualization - Design for Interaction, 2nd Edition*. Essex, UK: Pearson Education Limited.

- W3C. (2011). *Scalable Vector Graphics (SVG) 1.1 (Second Edition)*. W3C.
- Wallgren, A. (1996). *Graphing Statistics and Data. Creating Better Charts*.
- Ward, M., Grinstein, G., & Keim, D. (2010). *Interactive Data Visualization: Foundations, Techniques, and Applications*.
- Ware, C. (2004). *Information Visualization: Perception for Design*. Morgan Kaufmann.
- Zelazny, G. (2001). *Say It With Charts: The Executive Guide To Visual Communication*. London: McGraw-Hill.
- Zhu, Y. (2007). Measuring effective data visualization. (fv. 652-661). Berlin: Springer-Verlag.