

UNIVERSITATEA "ALEXANDRU IOAN CUZA" DIN IAȘI
FACULTATEA DE INFORMATICĂ



LUCRARE DE LICENȚĂ

Aplicație de Detectare Fake News

propusă de

Cristian Tirel

Sesiunea: *Iulie, 2018*

Coordonator științific

Anca Ignat

UNIVERSITATEA "ALEXANDRU IOAN CUZA" DIN IAȘI
FACULTATEA DE INFORMATICĂ

Aplicație de Detectare Fake News

Cristian Tirel

Sesiunea: *Iulie, 2018*

Coordonator științific
Anca Ignat

Avizat,
Îndrumător Lucrare de Licență
Titlul, Numele și prenumele

Data _____ Semnătura _____

DECLARAȚIE privind originalitatea conținutului lucrării de licență

Subsemnatul(a)

.....

domiciliul în

născut(ă) la data de, identificat prin CNP,
absolvent(a) al(a) Universității „Alexandru Ioan Cuza” din Iași, Facultatea de
..... specializarea, promoția
....., declar pe propria răspundere, cunoscând consecințele falsului în
declarații în sensul art. 326 din Noul Cod Penal și dispozițiile Legii Educației Naționale
nr. 1/2011 art.143 al. 4 și 5 referitoare la plagiat, că lucrarea de licență cu titlul:

_____elaborată sub îndrumarea dl. / d-na
_____, pe care urmează să o susțină în fața
comisiei este originală, îmi aparține și îmi asum conținutul său în întregime.

De asemenea, declar că sunt de acord ca lucrarea mea de licență să fie verificată
prin orice modalitate legală pentru confirmarea originalității, consimțind inclusiv la
introducerea conținutului său într-o bază de date în acest scop.

Am luat la cunoștință despre faptul că este interzisă comercializarea de lucrări
științifice în vederea facilitării falsificării de către cumpărător a calității de autor al unei
lucrări de licență, de diploma sau de disertație și în acest sens, declar pe proprie
răspundere că lucrarea de față nu a fost copiată ci reprezintă rodul cercetării pe care am
întreprins-o.

Data azi,

Semnătură student

DECLARAȚIE DE CONSIMȚĂMÂNT

Prin prezenta declar că sunt de acord ca Lucrarea de licență cu titlul „*Titlul complet al lucrării*”, codul sursă al programelor și celelalte conținuturi (grafice, multimedia, date de test etc.) care însoțesc această lucrare să fie utilizate în cadrul Facultății de Informatică.

De asemenea, sunt de acord ca Facultatea de Informatică de la Universitatea „Alexandru Ioan Cuza” din Iași, să utilizeze, modifice, reproducă și să distribuie în scopuri necomerciale programele-calculator, format executabil și sursă, realizate de mine în cadrul prezentei lucrări de licență.

Iași, *data*

Absolvent Prenume Nume

(semnătura în original)

Cuprins

Introducere.....	6
Motivație.....	6
Context.....	7
Obiectiv	7
Contribuții	7
Capitolul I	
I.1 Tehnologii Folosite.....	8
Capitolul II	
II.1 Descrierea Problemei.....	10
II.2 Atacul terorist 9/11.....	11
II.3 Utilizarea Aplicației.....	15
Capitolul III	
III.1 Abordări anterioare	21
III.2 Direcții viitoare în dezvoltarea aplicației	22
Bibliografie.....	25

Introducere

“Newspapers are not made any longer by news or journalism. They are made by sheer weight of money expressed in free gift schemes. They serve not the interests of the many, but the vested interests of the few.”

- (Oswald Mosley)

Știrile false acaparează puternic atenția publicului, ocupând foarte mult spațiu media, având o presupusă putere de neoprit, iar consecințele acestora reprezintă o amenințare foarte mare la adresa structurii societății și a sistemului democratic, acestea sunt considerate a fi o problemă semnificativă cu care se confruntă societatea globală și modul în care aceasta funcționează.

A fost acordată o atenție considerabilă problemei, devenind astfel o problemă modernă, cu toate acestea, istoria demonstrează că știrile false există de multă vreme și au evoluat recent în ceva mult mai invaziv, datorită dezvoltării informațiilor și a tehnologiilor de comunicare care permit o transmitere rapidă și o comunicare eficientă. Știrile false au fost percepute și definite diferit de-a lungul timpului ca instrumente de influență și persuasiune însă acestea nu sunt considerate rădăcina problemei, ci mai degrabă un simptom al unor probleme mai profunde care afectează sferele politice și de informare.

O mare parte din vină a fost îndreptată spre mass-media pentru răspândirea știrilor false la un nivel fără precedent și pe o scară globală. Există multe divergențe între părțile interesate cu privire la definirea și semnificația știrilor false.

Motivație

Sunt numeroase căi de a crea și împrăști știri false, de exemplu, pentru avantaje politice, afectarea reputației unei companii sau persoane, clickbait pentru creșterea veniturilor din publicitate sau pentru căutarea atenției, de aceea consider că este foarte importantă combaterea acestei probleme și este de datoria fiecăruia dintre noi să contribuim la diminuarea fenomenului aflat, din păcate, într-o creștere rapidă.

Context

Știrile false trebuie să fie înțelese în context ca fiind un instrument de modelare a opiniei publice și percepției prin lovirea unei rezonanțe emoționale. În consecință, există elementul de a influența și de a convinge publicul prin știri subiective care conțin o formă de agendă economică sau politică ascunsă. Rezultă că există o intenție de a iniția și de a mobiliza o audiență în vederea realizării scopurilor sau obiectivele comunicatorului.

O altă utilizare a știrilor false devine din ce în ce mai evidentă în actualul mediu de știri și utilizarea este îndreptată spre reducerea opiniei publice privind anumite aspecte. După cum capacitatea de a comunica și de a influența în mod eficient se bazează pe credibilitate, știrile false sunt un mijloc de a ataca marca și reputația adversarilor și, prin urmare, de a le reduce credibilitatea și disponibilitatea de a contesta discursul hegemonic din cauza riscului de a fi ridiculizați și denigrați public.

Obiectiv

Îmi propun realizarea unei aplicații care să vină în ajutorul persoanelor ce doresc o îmbunătățire în detectarea manipulării opiniei publicului cititor, scopul principal al aplicației este de a minimiza răspândirea dezinformării prin oprirea propagării de știri false în rețea. Este o provocare deosebită pentru a atinge acest obiectiv deoarece necesită detectarea de știri false cât mai rapid posibil.

Contribuții

Primul pas în elaborarea lucrării de licență a reprezentat documentarea și informarea pentru a putea găsi cea mai accesibilă cale de a identifica știrile false, pornind de la un URL dat, în acest fel am putut crea o aplicație User Friendly, accesibilă oricui. Pentru realizarea „Aplicației de Detectare Fake News” am găsit o cale originală de a îmbina o mică implementare de parser Google Crawler, pentru a identifica cu ajutorul path-ului din link dacă știrea respectivă a mai fost postată și de alte site-uri asemănătoare, astfel mărimu-mi șansele de a găsi răspunsul la întrebarea: Este această știre falsă ?

Structura aplicației este concepută să funcționeze treptat, dacă unul din pași nu este îndeplinit, pasul următor nu va apărea, returnând o eroare specifică ce îndrumă userul pentru corectarea pasului anterior.

O listă larg răspândită, provenită din mai multe surse este prezentă pentru a îmbunătăți capacitatea de identificare, la aceasta se poate oricând adăuga noul site ce a fost verificat, indiferent dacă știrea este veritabilă sau nu, existând o listă secundară pentru a acoperi orice situație în cauză.

Unul din pașii identificării este verificarea ultimei date la care site-ul cu pricina a fost modificat, site-urile ce împrăștie știri false sunt active pentru o scurtă perioadă de timp, după intrând într-o stare latentă.

Un alt pas constă în validarea protocolului de securitate, marea majoritate a acestor site-uri nu sunt securizate, publicul lor țintă nefiind de specialitate.

În concluzie, această aplicație va returna un scor calculat după fiecare pas și va determina dacă site-ul conține o știre falsă informând astfel utilizatorul, sporindu-i protecția împotriva manipulărilor de orice fel și evitarea dezinformării.

Capitolul I

I.1

Tehnologii folosite

Pentru dezvoltarea aplicației de detectare “Fake News” au fost folosite următoarele tehnologii:

Eclipse

Window Builder

Limbajul de Programare Java

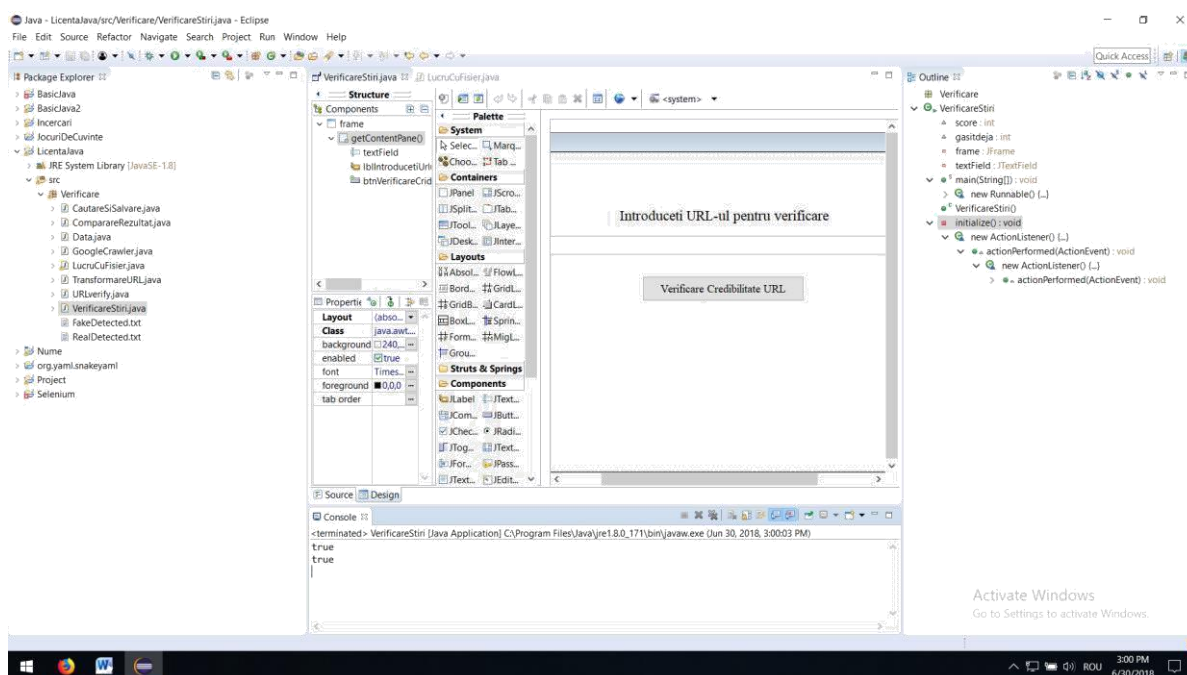
Eclipse

Eclipse este un mediu integrat de dezvoltare (IDE) utilizat în programare și este cel mai utilizat Java IDE. Acesta conține un spațiu de lucru de bază și un sistem plug-in extensibil pentru personalizarea mediului. Eclipse este scrisă mai ales în Java, iar utilizarea sa principală este pentru dezvoltarea aplicațiilor Java, dar poate fi folosită și pentru dezvoltarea de aplicații în alte limbi de programare prin plug-in-uri, inclusiv Ada, ABAP, C, C++, C#, COBOL, D, Fortran, Haskell, JavaScript, Julia, Lasso, Lua, NATURAL, Perl, PHP, Prolog, Python, R, Ruby, Rust, Scola, Clojure, Groovy.

WindowBuilder

WindowBuilder este compus din SWT Designer și Swing Designer elemente ce ușurează procesul de creare a aplicațiilor Java GUI fără a petrece mult timp scriind cod. Utilizează design-ul vizual WYSIWYG și instrumente de aspect pentru a crea de la simple forme la ferestre complexe; codul Java va fi generat pentru dumneavoastră. Adăugați cu ușurință controale utilizând drag-and-drop, adăugați manipuloare de evenimente modificați diferitele proprietăți ale controalelor utilizând un editor de proprietăți.

WindowBuilder este construit ca un plug-in pentru Eclipse și diversele IDE bazate pe Eclipse (RAD, RSA, MyEclipse, JBuilder, etc.). Plug-in-ul construiește un arbore abstract de sintaxă (AST) pentru a naviga în codul sursă și utilizează GEF pentru a afișa și gestiona prezentarea vizuală. Codul generat nu necesită biblioteci personalizate suplimentare care să fie compilate și executate: întregul cod generat poate fi utilizat fără a avea instalat WindowBuilder Pro.



Capitolul II

II.1

Descrierea Problemei

Termenul de știri false reprezintă o problemă la nivel global ce erodează în continuare încrederea deja slabă în mass-media. Foarte multă atenție a fost axată recent pe această problemă, nivelul și natura activității, presupusele riscuri și amenințări care vin împreună cu acestea. Vina pentru tendințele actuale de știri falsificate a fost îndreptată la diferiți țapi ispășitori.

Unii s-au grăbit să dea vina pe tehnologie și prioritățile de bază ale internetului și ale giganților social media precum Google, Facebook și Twitter pentru criză. Alții îndreaptă atenția asupra eșecurilor proprii ale mass-mediei prin imposibilitatea de a se conecta în mod corespunzător cu frustrarea și furia oamenilor și a comunităților.

Știrile false, ca instrument de manipulare a comunicării nu sunt ceva nou, un exemplu fiind faptul că Președintele Thomas Jefferson s-a plâns în 1807 că cineva care nu citește un ziar este mai bine informat despre evenimentele petrecute decât cineva care citește ziarele.

Aproximativ un deceniu în urmă, unii academicieni au încercat să caracterizeze astfel de știri și formate de comunicare ca satiră și talk show-uri, mai recent, definiția acestora s-a mutat la una care include elementele de înșelăciune și dezinformare.

Ce este nou și în schimbare este modul în care sunt concepute știrile false, de aceea este nevoie ca metodele de detectare a acestor instrumente să fie îmbunătățite constant.

II.2

Atacul terorist 9/11/2001

Înainte de a începe țin să precizez că nu susțin nicio parte a acestei dispute ci doar încerc să prezint problema Fake News, cu cele mai certe date legate de acest subiect astfel încât decizia bilanțului celor întâmplate să rămână la alegerea fiecăruia.

Imediat după macabra zi din data de 9/11/2001 au luat naștere din ce în ce mai multe teorii ale conspirației și împrăstieri de știri false, ținta principală fiind guvernul Statelor Unite ale Americii, deoarece multe dintre întrebări nu au primit nici până în ziua de astăzi un răspuns, 1 din 7 americani este convins că acest atac terorist nu a fost real ci un “inside job”. Acest subiect de știri este atât de subiectiv încât întrebarea: Noi ce sau pe cine am putea crede? nu are un răspuns cert, fiind foarte ușor să spui un lucru și să-l faci să fie credibil în ochii anumitor oameni. Însă până și cei ce cred povestea originală și anume că 9/11 a fost cu adevărat un atac terorist, rămân reci la declarația Guvernului American ce susține că 19 teroriști devotați Al-Qaeda au fost responsabili.

În ciuda presiunilor puse de către presă atât prin știri reale cât și prin știri false, FBI-ul nu a dorit niciodată să ofere informații despre pașaportul în sine, sau imagini de pe camerele CCTV din aeroport pentru a dovedi că teroriștii au urcat cu adevărat la bordul avioanelor, de aici unii oameni trag concluzia că într-adevăr teroriștii au fost forțați să comită aceste fapte, sau li s-a înscenat totul.

Cea mai importantă întrebare din acea zi este dacă într-adevăr teroriștii islamiști au pus stăpânire pe cele două avioane, un prim argument al celor ce considerau că Guvernul a fost în spatele atacului este dovada găsirii unui pașaport al unui presupus terorist, în stare perfectă la locul prăbușirii celor două turnuri.

Întrebarea acestora este cum a reușit un pașaport să fie neafectat, având în vedere că două clădiri de 110 etaje din fier nu au rezistat, aici știrile false (sau reale) susțin că pașaportul a fost „plantat” pentru a-i putea învinui și a manipula populația în vederea declarării războiului din Irak, teorie susținută de tatăl presupusului terorist, ce a comunicat într-un interviu că fiul său nu ar fi avut legături cu organizații de terorism și că pașaportul acestuia a fost furat în timp ce acesta era considerat dispărut la momentul respectiv.

Scepticii susțin cu dovezi ilustrate în pozele de mai jos, că impactul avionului care a lovit Pentagonul a produs pagube minore în comparație cu cele două turnuri, în răspuns Guvernul susține că acesta a lovit doar o parte proaspăt renovată ce a fost fortificată, însă în detrimentul lor vine o întrebare repetitivă în această problemă: De ce dovezile video sunt atât de puține și de o așa proastă calitate ?



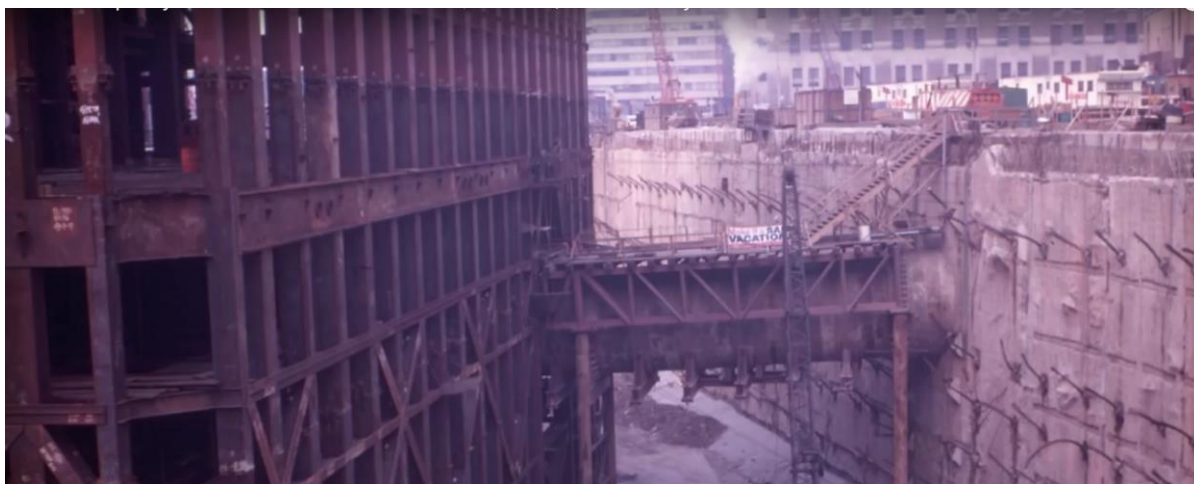


Scepticii susțin în continuare știrile false (sau reale) cu ajutorul a câteva videoclipuri ale prăbușirii ambelor turnuri unde se arată valuri de ejecții de masă orizontale care se răsucesc pe fețele clădirilor, ținând aproape pasul cu materialul care se află în afara clădirii, cu mult sub zona de distrugere în sine. Ejecțiile par să vină de la mai multe etaje în același timp, ceea ce nu este în concordanță cu ideea că ejecțiile au constat din resturile suflate la etaj. În plus față de valurile masive de ejecții, există numeroase fotografii și videoclipuri care prezintă ejecții individuale, concentrate și de mare viteză ale materialelor pe mai multe etaje, sub punctul de colaps. Acestea sunt ușor explicate ca ejecții explozive, ele nu sunt explicate în mod convingător ca jeturi de aer comprimat care scapă.

Lipsa unei cauze suficiente pentru prăbușire a fost documentată, disputată, raționalizată și obfuscată. Combustibilul cu jet ar fi ars în primele zece minute. Cea mai mare parte a combustibilului a fost arsă într-o minge de foc în afara clădirii, mai ales în cazul Turnului de Sud, unde avionul a pierdut în cea mai mare parte coloanele de bază. Incendiile din clădiri, dincolo de primele câteva minute, erau, în esență, incendii de birou și nu foarte mari, aprinse de combustibilul cu reacție, ca un lichid mai ușor pe cărbune. Jetul de combustibil este kerosen, iar temperaturile de la incendiile de kerosen sau de la birou sunt insuficiente pentru a topi, sau chiar a slăbi catastrofal, coloanele masive de oțel care dau spre miezul clădirii.

Chiar dacă flăcările și temperatura aerului s-au ridicat la maxim, masa mare de oțel ar îndepărta căldura și nu ar ridica suficient temperatura oțelului. Pentru ca temperatura oțelului să se apropie de temperatura de topire a oțelului, incendiile ar trebui să fie de lungă durată, dar aceste incendii erau foarte scurte, de ordinul unei ore.

Există fotografii și înregistrări video ale unei femei înclinată pe o grilă și fluturând în gaura unde unul dintre avioane se prăbușise așteptând să fie salvat. Aceasta ar părea a fi o mărturie directă că incendiile de pe podele unde impactul și combustibilul cu reacție au avut efectul cel mai mare au scăzut, iar temperaturile aerului și oțelului au fost destul de moderate pentru ca oamenii să se plimbe și să atingă oțelul: insuficient pentru a provoca defectarea coloanelor din oțel.



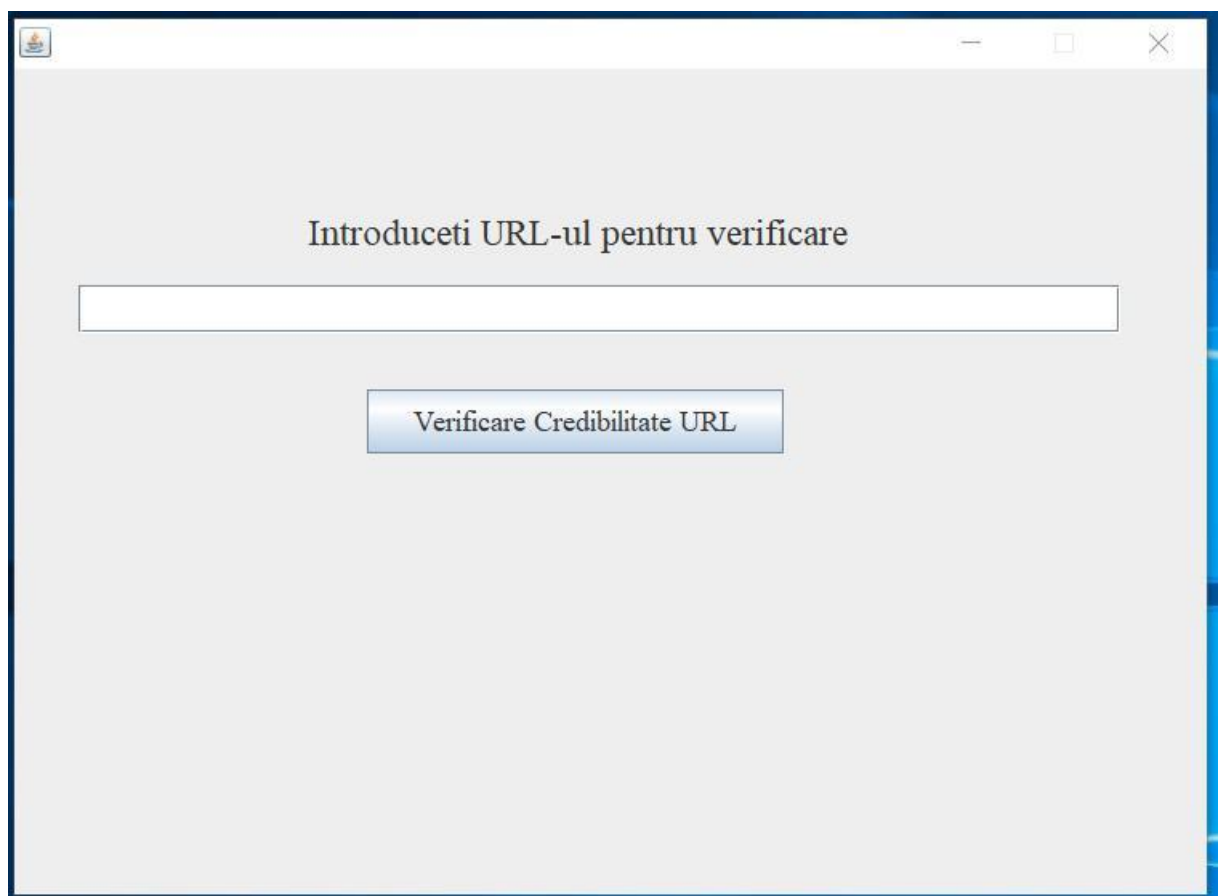
Din păcate nu vom ști niciodată răspunsul la această dilemă ce se dorește a fi îngropată în istorie, dar un lucru este cert, fie că scepticii au dreptate, fie că Guvernul Statelor unite spune adevărul pur în declarațiile sale, atât știri false cât și știri reale au fost împânzite populației, iar îmbunătățirea metodelor de descoperire a înșelăciunii ar putea pune capăt viitoarelor întrebări fără răspuns și oferind o rază de lumină în căutarea adevărului.

II.3

Prezentarea Aplicației

Java este un limbaj de programare cu scop general care este concurent, bazat pe clasă, orientat obiect și conceput special pentru a avea cât mai puține dependențe de implementare posibil. Este destinat să permită dezvoltatorilor de aplicații să "scrie o singură dată, să ruleze oriunde", ceea ce înseamnă că codul Java compilat poate rula pe toate platformele care suportă Java fără a fi nevoie de recompilare.

Aplicația de Detectare Fake News a fost concepută în limbajul Java și constă în introducerea unui URL pentru a putea fi verificat dacă site-ul respectiv este considerat a fi o sursă de răspândire a dezinformării sau dacă știrea selectată este una falsă.



Primul pas în acest demers constă în verificarea existenței URL-ului prin apăsarea butonului “Verificare Credibilitate”, și de asemenea dacă acesta este securizat, majoritatea site-urilor ce răspândesc știri false nu sunt securizate, publicul lor țintă nefiind unul de specialitate. Acest pas a fost realizat cu ajutorul `HttpURLConnection`, o conexiune URL cu suport pentru caracteristici specifice HTTP.

Fiecare instanță `HttpURLConnection` este utilizată pentru a face o singură solicitare, dar conexiunea de rețea la serverul HTTP poate fi partajată în mod transparent de alte instanțe.

```
public class URLverify {  
  
    public static boolean checkURLexistance(String targetUrl) {  
        HttpURLConnection URLconn;  
        try {  
            URLconn = (HttpURLConnection) new URL(targetUrl).openConnection();  
            URLconn.setRequestMethod("HEAD");  
  
            return (URLconn.getResponseCode() != HttpURLConnection.HTTP_NOT_FOUND);  
        } catch (Exception e) {  
            System.out.println("Error: " + e.getMessage());  
            return false;  
        }  
    }  
}
```

Indiferent dacă forma URL-ului dumneavoastră nu este completă, exemplu: “`https://www.timesnewroman.ro/`”, sau este scrisă sub forma “`timesnewroman.ro`” aplicația este capabilă să vă găsească adresa dorită.

Odată ce pagina a fost găsită, urmatorul pas constă în căutarea cu ajutorul unui parser Google Crawler atât a domeniului pentru a afla care domenii sunt asemănătoare cât și a path-ului, pentru a înțelege care site-uri au postat o știre asemănătoare, astfel îmbunătățindu-ne șansele de a găsi în listele de date unul din site-urile căutate de noi și a înțelege dacă acesta a fost catalogat deja ca fiind fals sau adevărat.

Google Crawler este un parser foarte simplu ce reprezintă pagina rezultatelor motorului de căutare și definește poziția în rezultatul căutării.


```

public static String getSearchContent(String path) throws Exception {
    final String agent = "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)";
    URL url = new URL(path);
    final URLConnection connection = url.openConnection();
    connection.setRequestProperty("User-Agent", agent);
    final InputStream stream = connection.getInputStream();
    return getString(stream);
}

public static List<String> parseLinks(final String html) throws Exception {
    List<String> result = new ArrayList<String>();
    String pattern1 = "<h3 class=\"r\"><a href=\"/url?q=\"";
    String pattern2 = "\">";
    Pattern p = Pattern.compile(Pattern.quote(pattern1) + "(.*?)" + Pattern.quote(pattern2));
    Matcher m = p.matcher(html);

    while (m.find()) {
        String domainName = m.group(0).trim();

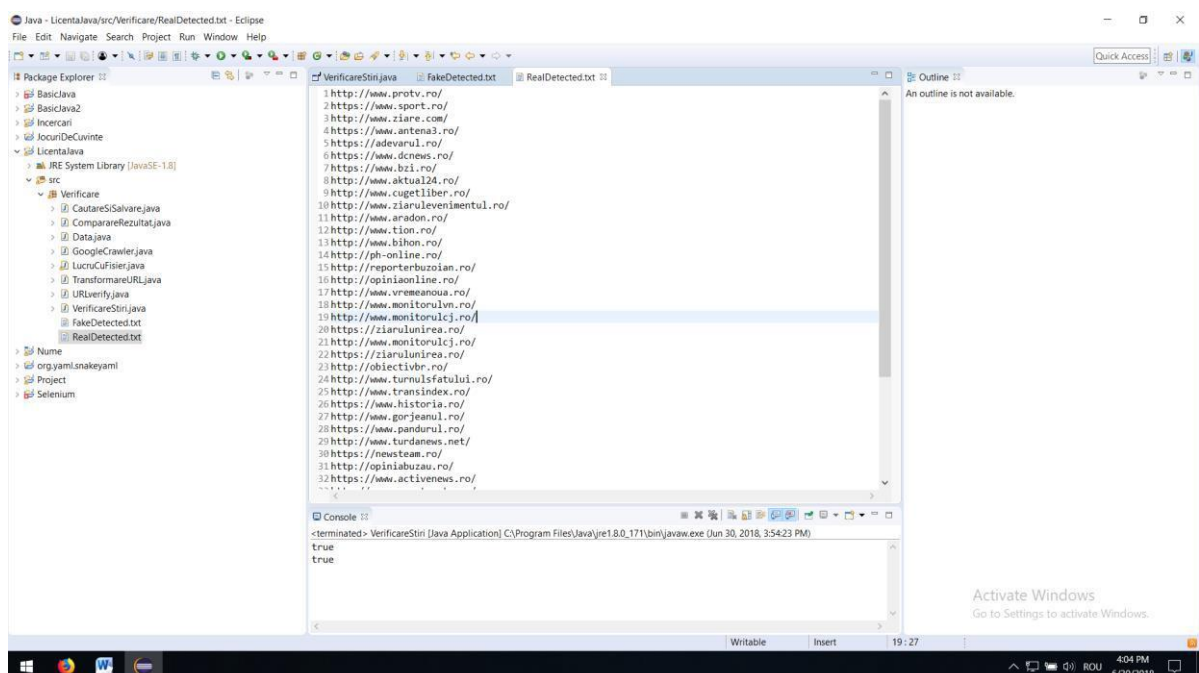
        /** remove the unwanted text */
        domainName = domainName.substring(domainName.indexOf("/url?q=") + 7);
        domainName = domainName.substring(0, domainName.indexOf("&"));

        result.add(domainName);
    }
    return result;
}

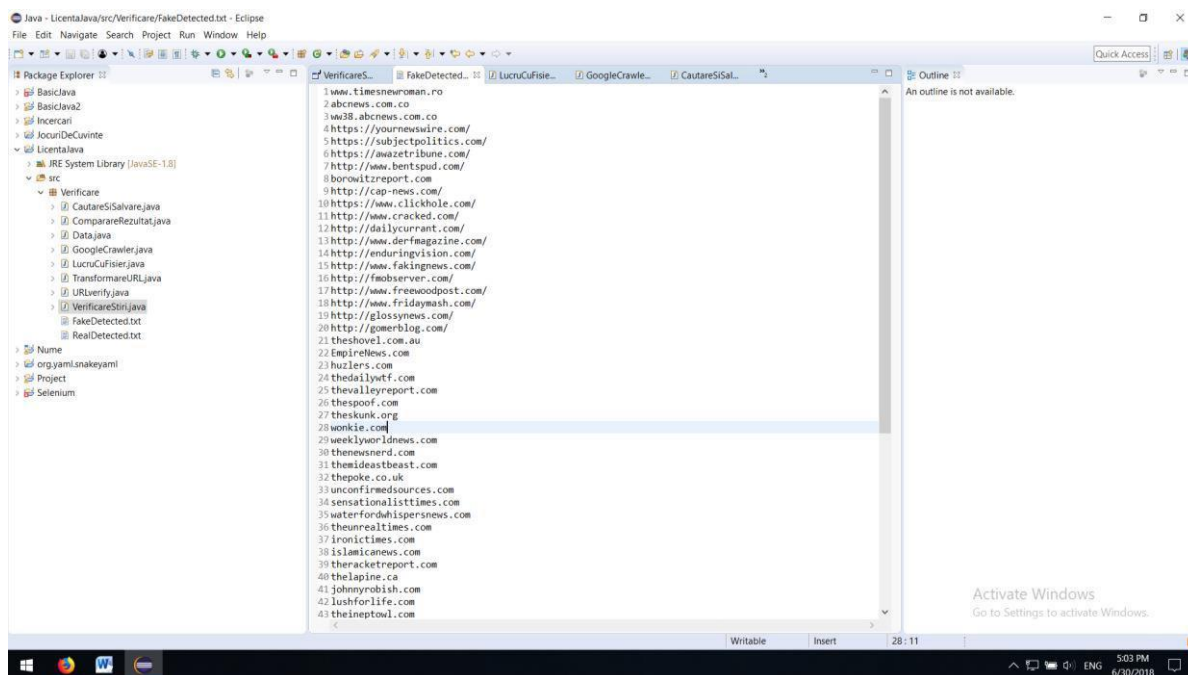
```

Înainte de a efectua căutările, link-ul introdus suferă anumite modificări, variabila „SLDTLD” va memora adresa site-ului, iar variabila „calea” va memora path-ul. Rezultatele găsite pentru fiecare vor fi salvate în două liste, ListaCale și Lista SLDTLD, acestea fiind căutate în următoarele liste de date ce conțin URL-uri identificate deja.

Lista cu site-uri de știri reale :



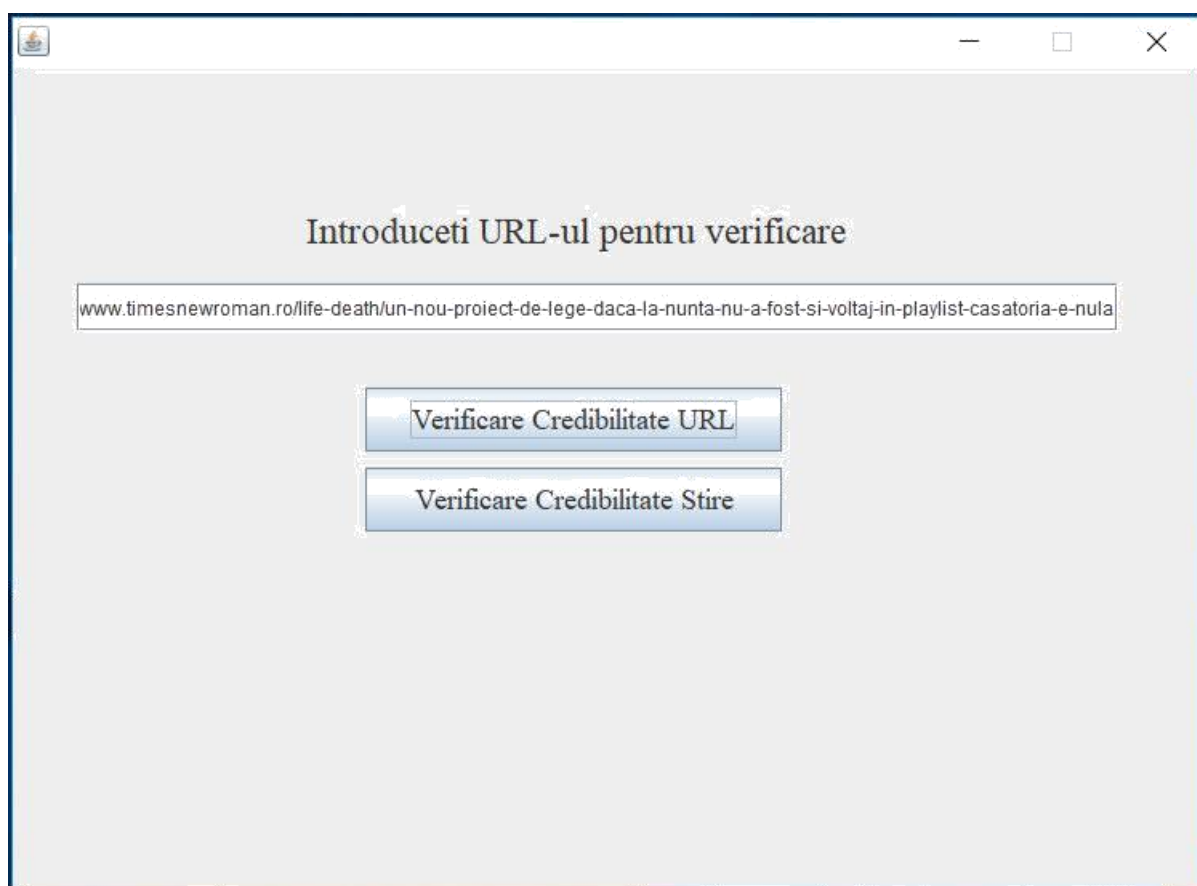
Lista cu site-uri de știri false:



Motivul implementării a două liste, este de a grabi căutarea, în cazul în care site-ul nu este securizat, prima listă în care se va efectua căutarea va fi lista cu conținut raportat ca fiind fals.

Un alt pas constă în cunoașterea activității paginii respective, în marea lor majoritate site-urile ce răspândesc conținut dezinformator au un punct culminant după care ori dispar cu totul, ori devin inactive. Prin identificarea ultimei modificări aduse paginii, putem înțelege gradul de activitate a paginii și ne putem îmbunătăți concluzia finală.

Pentru efectuarea pașilor de mai sus, după verificarea URL-ului, butonul “Verificare Credibilitate Știre” devine activ.



În momentul apăsării acestui buton Aplicația de Detectare Fake News va întoarce un rezultat fie acesta pozitiv sau negativ, bazat pe calcularea unui scor, ce scade sau crește după anumite criterii:

Site-ul nu este securizat : +50

Site-ul nu a fost modificat în ultimele 3 zile: +30

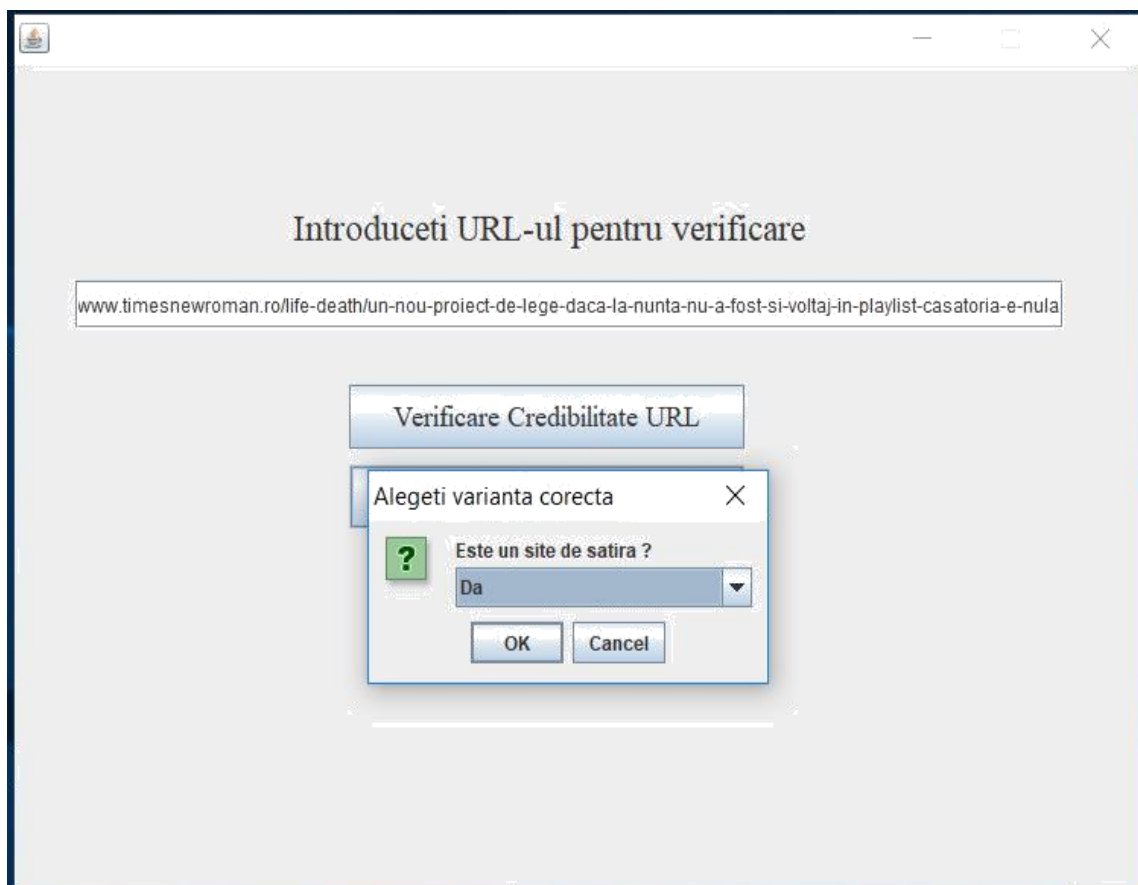
Site-ul este găsit în baza de date a știrilor false: +100

Site-ul este găsit în baza de date a știrilor reale: -100

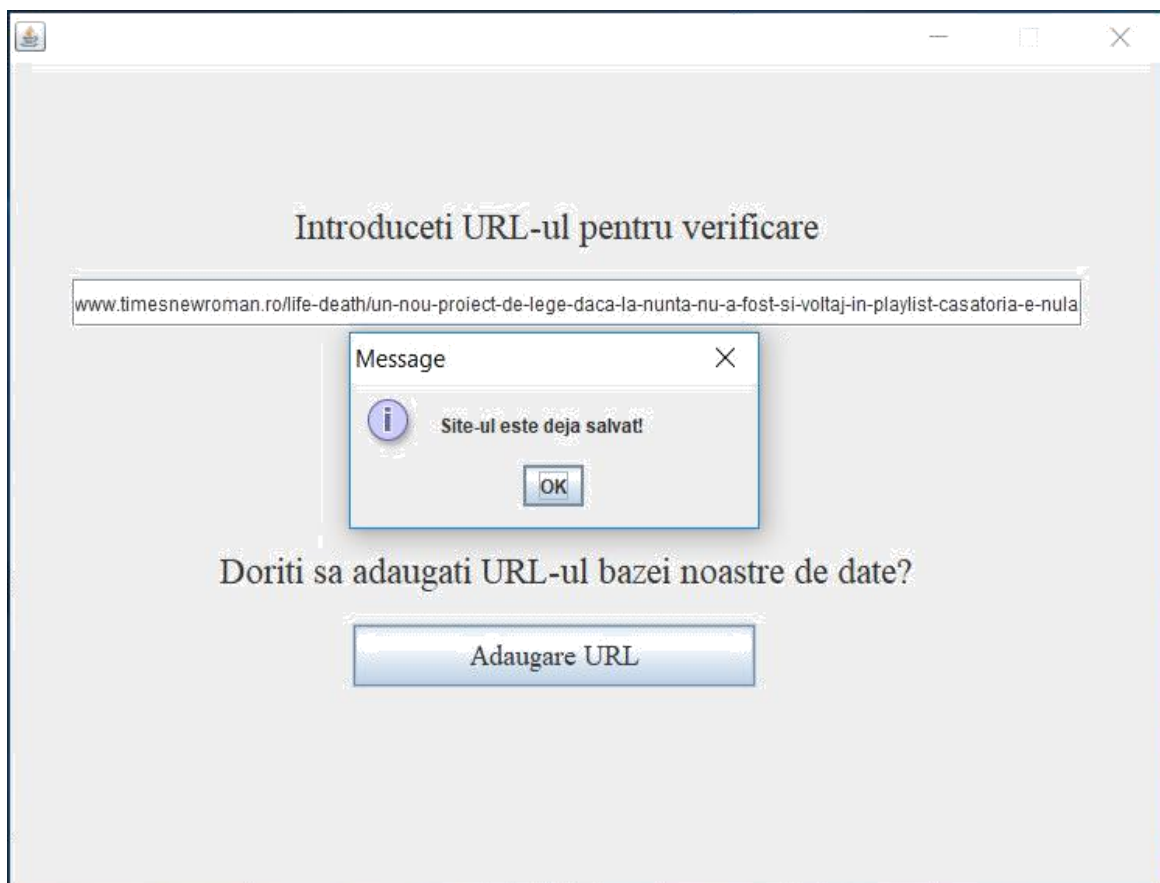
Dacă scorul este mai mare de 0 știrea respectivă este catalogată ca fiind falsă, iar dacă scorul este sub 0, aceasta este considerată reală.

În cazul în care datele găsite cu ajutorul Google Crawler-ului nu sunt suficiente, iar aplicația nu poate da un răspuns automat cert (scorul este 0), este nevoie de intervenția Userului, acesta este întâmpinat de o serie de întrebări de tipul Da sau Nu, ce vor ajuta la identificarea naturii site-ului.

Fiecare întrebare valorează +10 sau -10 puncte în funcție de răspunsul acordat și sunt într-un număr negativ, pentru a asigura un răspuns final.



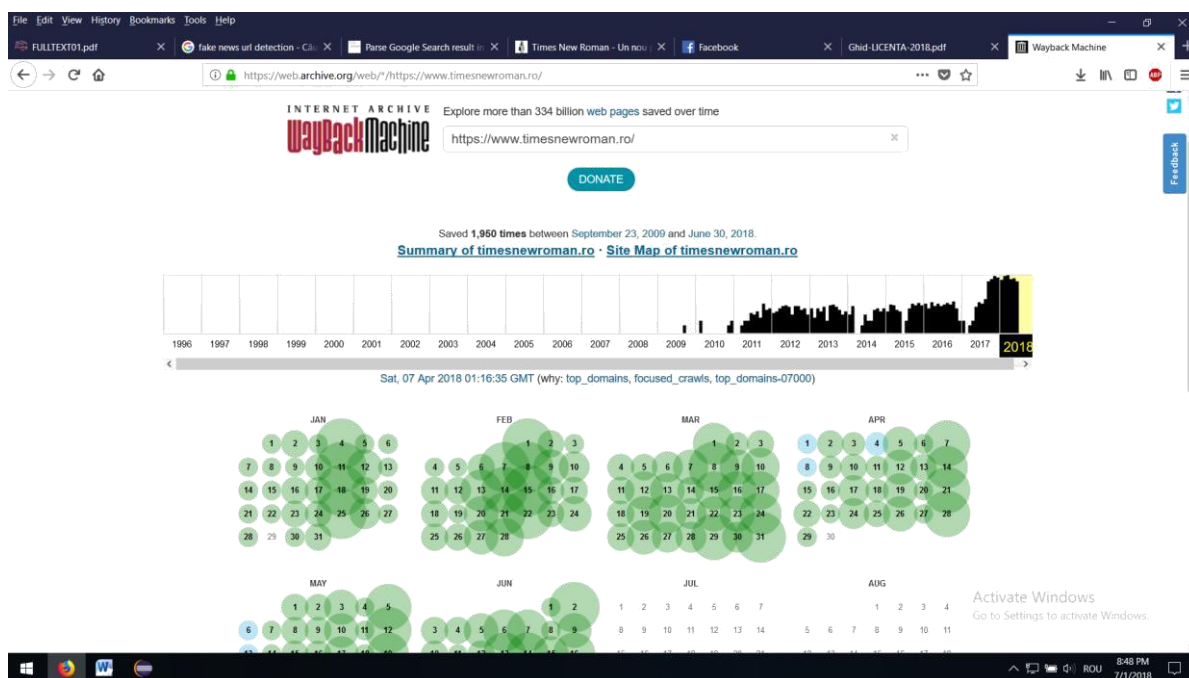
Pentru a continua îmbunătățirea aplicației am implementat un ultim buton, „Adăugare URL”, unde, dacă Userul este de accord poate adăuga noul site bazei noastre de date stocate în cele două liste, în funcție de rezultatul identificării. Acesta devine activ imediat ce pasul anterior este finalizat, odată ce un URL se dorește a fi adăugat, se fac verificările necesare în pașii anteriori pentru a observa dacă acesta nu este deja în baza de date. Pentru a evita dublurile, site-ul poate fi adăugat doar o singură dată, un mesaj de stare este afișat în caz contrar.



Abordări anterioare

În unele cazuri, căutările mele pentru îmbunătățirea Aplicației de Detectare Fake News, nu au adus rezultatele dorite, unul din acestea fiind descoperirea aplicației web, The Wayback Machine, ce este o arhivă digitală a World Wide Web și a altor informații de pe Internet. A fost lansată în 2001 de către Internet Archive, o organizație nonprofit cu sediul în San Francisco.

https://web.archive.org/web/*/https://www.timesnewroman.ro/



Această aplicație nu conferă un răspuns cert în identificarea ultimei date în care site-ul a fost modificat, ci data ultimului snapshot. În cazul unei lipse masive de informații, ar putea fi un punct de cercetare bun, dar în acest caz, nu ajungem la un rezultat favorabil căutării noastre.

Capitolul III

III.1

Direcții viitoare în dezvoltarea aplicației

Factorul uman este evident cel hotărâtor în a lua decizia dacă o știre este demnă de încredere sau nu. Consider că ar trebui îmbunătățit conceptul de "trustworthiness", adică "demn de încredere", tocmai pentru a ajuta utilizatorii să decidă singuri concluziile, nefiind nici științific și nici etic ca un algoritm să ia această decizie în locul lor.

Implementarea unei metode de procesare naturală a limbajului, adică automatizarea eficientă a multor acțiuni umane care țin de investigație amănunțită:

Colectarea de date masive
Indexarea lor
Procesarea automată a conținutului articolelor
Extragerea de cuvinte cheie

O altă direcție importantă ar fi identificarea primului articol de știri care răspândește o anumită informație, fiind mai ușor astfel să o catalogăm. Modul în care putem găsi pacientul zero este prin descoperirea de legături între articole, găsirea de similarități între acestea și crearea unui grafic care conține toate articolele care scriu despre același subiect.

III.2

Concluziile Lucrării

În concluzie, știrile false sunt povești false concepute pentru a crea propagandă sau a influența publicul într-o problemă politică sau sociologică, cu ajutorul mass-mediei sociale, a internetului și a surselor de știri stabilite. Vestea falsă servește în mod negativ diverselor scopuri ale societății, cum ar fi influențarea și convingerea oamenilor prin mijloace materiale false și aducând daune unei persoane, eveniment, idee, concept sau oamenilor în general.

Nu este clar ce intervenții, dacă există, ar putea inversa această tendință spre minciună. Există puține dovezi că oamenii își schimbă opinia, deoarece văd că un site de verificare a faptelor respinge, de exemplu, una dintre credințele lor. Etichetarea de știri false ca atare, într-o rețea socială sau într-un motor de căutare, poate face puțin pentru ai descuraja.

După cum se arată, cele trei cauze principale sunt: impactul social al mass-mediei și al internetului, creșterea surselor de știri nesigure și căderea de știri autoritare, și în cele din urmă anonimatul din spatele creării de știri false și dezinformare.

Pe scurt, mass-media socială pare să amplifice sistematic falsitatea în detrimentul adevărului și nimeni - nici experți, nici politicieni, nici companii de tehnologie – nu pot, cel puțin momentan, să inverseze această tendință. Este un moment periculos pentru orice sistem de guvernare bazat pe o realitate publică comună.

Bibliografie

Fake News: as the Problem or a symptom of a Deeper Problem?

<https://uu.diva-portal.org/smash/get/diva2:1181745/FULLTEXT01.pdf>

Curs practic de Java, Cristian Frăsinaru

https://profs.info.uaic.ro/~acf/java/Cristian_Frasinaru-Curs_practic_de_Java.pdf

Eclipse, The Platform for Open Innovation and Collaboration

<http://www.eclipse.org/>

WindowBuilder

<https://www.eclipse.org/windowbuilder/>

Lista Fake News

<https://github.com/Aloisius/fake-news/tree/master/lists>

Lista Real News

<http://www.trafic.ro/vizitatori/top-siteuri-stiri-massmedia/saptamana-pg3>

Clasa HttpURLConnection Java

<https://docs.oracle.com/javase/7/docs/api/java/net/URLConnection.html>

Google Crawler

<https://www.admfactory.com/how-to-parse-google-search-result-in-java/>

Wayback Machine

<https://archive.org/web/>

Concluzii Fake News

<https://thespidermancyclist.wordpress.com/2017/12/07/conclusion-of-fake-news/>

9/11: Conspiracy Theories & The Unanswered Questions |

Documentary <https://www.youtube.com/watch?v=rzsV-rehvhc>