



UNIVERZITET U NIŠU  
ELEKTRONSKI FAKULTET  
Katedra za računarstvo



# DOCUMENT FINDER

- Kriptografija -

Studenti:

Andrija Petrović

Aleksandar Kostić

Andrej Rakić

Aleksandar Randelović

Jovan Pešić

Katarina Randelović

Đorđe Čikić

Petar Đorđević

Uroš Milivojević

Svetlana Mančić

Mentor:

Prof. dr. Vladan Vučković

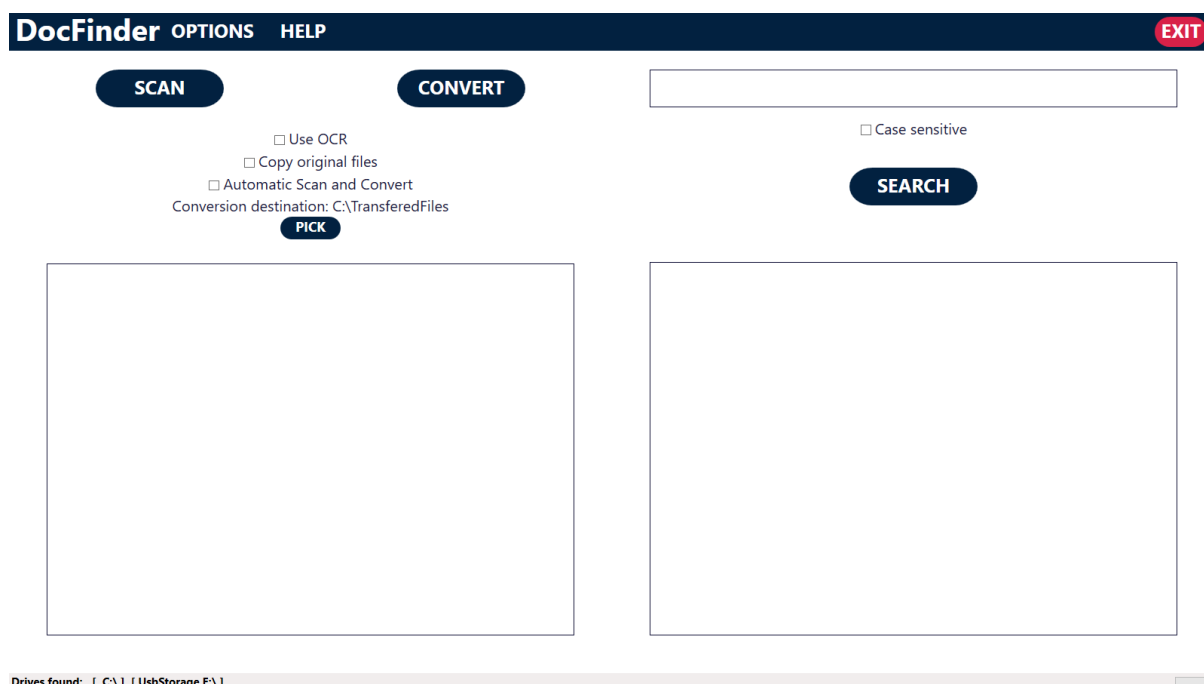
# Uvod

//////Na pocetku bismo mogli nesto da dodamo o već postojećim resenjima? Sta obicno poseduju, da pomenemo neka od njih?

U ovom radu biće reči o aplikaciji koja je implementirana za potrebe predmeta Kriptografija na master akademskim studijama Elektronskog fakulteta u Nišu, modula Računarstvo i informatika. Prvo će biti dat opis aplikacije, njene karakteristike i namena. Nakon toga će detaljno biti opisano korisničko uputstvo i mogućnosti korišćenja.

## Opis aplikacije

Aplikacija **DocumentFinder** omogućava pronalaženje fajlova koji poseduju sledeće extensione: **.txt .doc .docx .pgn** i **.pdf**, a zatim sve fajlove koji nisu tekstualnog tipa, konvertuje u **.txt** format. Fajlovi se objedinjuju u jedan direktorijum, nakon čega je moguće vršiti pretragu teksta unutar svih fajlova, po proizvoljno zadatom terminu.




Slika 1. DocumentFinder početni interfejs

## Funkcionalnost

Interfejs je podeljen na dva funkcionalna dela, sa padajućim menijem pri vrhu. Pri dnu se nalazi traka sa informacijama, na čijoj su levoj strani prikazane sve detektovane particije na sistemu. Novo dodate particije (primer USB), nakon pokretanja aplikacije, bivaju detektovane u fazi skeniranja. Informaciona traka na desnoj strani sadrži informacije o trenutno obrađivanom fajlu, sa indikatorom napretka celokupnog posla. Na levoj strani interfejsa aplikacije mogu se uočiti opcije **SCAN** i **CONVERT**. U fazi skeniranja, sve detektovane particije se pretražuju rekurzivno, pri čemu se pamte izvorne putanje do fajlova, koji poseduju sledeće extensione: **.txt .doc**

**.docx** **.pgn** i **.pdf**. Pritom se tekstualni fajlovi automatski kopiraju u direktorijum **C:\TransferredFiles**. Izvorne putanje pronađenih fajlova se upisuju unutar istog direktorijuma, u log fajl **\_TransferredFilesPaths.txt**. Treba još napomenuti predefinisane putanje: **C:\Windows**, **C:\Recovery**, **C:\Program Files**, **C:\ProgramData**, **C:\\$Recycle.Bin**, koje se zaobilaze tokom skeniranja. To su putanje koje nazivamo sistemskim, zbog čega na njima ne očekujemo da pronađemo fajlove od esencijalne važnosti za krajnjeg korisnika.


Scanning: 113. C:\TransferredFiles\install-log.6.txt



*Slika 2. Faza skeniranja*

Prilikom konverzije **.doc** **.docx** i **.pdf** fajlova, kreiraju se tekstualni fajlovi unutar već pomenutog direktorijuma, imenovani identično izvornom fajlu, ali sa **.txt** ekstenzijom i tekstualnom **ASCII** sadržinom.

Converting: 12. Alberto Fernandez Villan - Mastering OpenCV 4 with Python.pdf




*Slika 3. Faza konverzije*

Na korisniku je da odabere način izvršavanja pomenutih procesa. Naime, moguće je koristiti optičko prepoznavanje znakova, izborom **Use OCR** opcije, prilikom konverzije **.pdf** dokumenata koji nisu u potpunosti pretraživi. Biranjem opcije **Copy original files**, izvorni **.pdf** **.doc** i **.docx** fajlovi bivaju kopirani u direktorijum, na predefinisanoj putanji **C:\TransferredFiles**, koja je lako izmenljiva, izborom direktorijuma klikom na dugme **PICK**. Objedinjavanje procesa skeniranja i konverzije (*automatizacija*), postiže se izborom **Automatic Scan and Convert** opcije, kao i **Scan and Convert** opcije iz padajućeg menija. Polje za tekst, ispod dugmeta **PICK**, služi za prikaz putanja do izvornih fajlova, pronađenih nakon okončane faze skeniranja.

Desni deo interfejsa aplikacije služi pretragu **.txt** fajlova, dobijenih nakon faze skeniranja i konvertovanja, i prikaz rezultata pretrage. Na vrhu je polje za unos teksta koji se traži, a ispod njega opcija **Case sensitive**, koja ukazuje na to da li se veličina unetih karaktera ignoriše ili ne. Pretraga počinje klikom na dugme **SEARCH** i može se stopirati u bilo kom trenutku klikom na dugme **STOP**, koje se nalazi na vrhu.

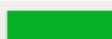
Na samom početku se proverava da li je obavljeno skeniranje i konverzija i da li je unet tekst, jer je to preduslov za pretragu. Ako su navedeni uslovi ispunjeni, pretražuju se svi **.txt** fajlovi, koji se nalaze u direktorijumu **C:\TransferredFiles**. Takođe, proverava se izbor opcije **Case sensitive**. Primer uticaja ove opcije na rezultat pretrage može se uočiti na slikama 4 i 5, gde se vidi da se broj pronađenih fajlova razlikuje.

Finished search for: Dark. Found 4 files.



*Slika 4. Pretraga "Dark"*


Finished search for: dark. Found 8 files.



*Slika 5. Pretraga "dark"*

Provera se vrši za svaki fajl. Ukoliko tekst koji se pretražuje sadrži više reči, za svaku od njih se vrši provera.

Searching: 2107/3352 Snort Intrusion Detection and Prevention Toolkit by Caswell B., Baker A., Beale J. (z-lib.org).txt



*Slika 6. Faza pretrage*

Lista pronađenih fajlova je sortirana najpre po broju pronađenih reči iz unetog teksta, a zatim po nazivu fajla i prikazana u polju ispod **SEARCH**.

operativni sistemi

☐ Case sensitive

SEARCH

0.1 - 0 predmetu 2014.txt Words found: operativni, sistemi  
 1. Uvod i pregled operativnih sistema.txt Words found: operativni, sistemi  
 10. Upravljanje datotekama.txt Words found: operativni, sistemi  
 2. Upravljanje procesima.txt Words found: operativni, sistemi  
 3. Niti.txt Words found: operativni, sistemi  
 4. Konkurentnost - uzajamno isklucivanje i sinhronizacija.txt Words found: operativni, sistemi  
 5. Konkurentnost - uzajamno blokiranje i gladovanje.txt Words found: operativni, sistemi  
 6. Upravljanje memorijom.txt Words found: operativni, sistemi  
 7. Virtuelna memorija.txt Words found: operativni, sistemi  
 8. Jednoprocesorsko planiranje.txt Words found: operativni, sistemi  
 9. Upravljanje UI i planiranje diska.txt Words found: operativni, sistemi  
 AppCache132925201749764963.txt Words found: operativni, sistemi  
 Distribuirani sistemi - skripta 2019.txt Words found: operativni, sistemi  
 GitHub.txt Words found: operativni, sistemi  
 Izveštaj sa stručne prakse.txt Words found: operativni, sistemi  
 Operativni sistemi - pitanja za usmeni.txt Words found: operativni, sistemi  
 Operativni sistemi - SKRIPTA (Pravna) .txt Words found: operativni, sistemi

*Slika 7. Rezultat pretrage*

## Dodatna funkcionalnost



