

Agregator vesti zasnovan na React-u i .Net Framework-u

Pavičić Miloš - SF39/2016

Mentor: Nikolić Siniša

Fakultet Tehničkih Nauka
Univerzitet u Novom Sadu

Sadržaj

- ▶ Uvod
- ▶ Stanje u oblasti
- ▶ Korišćene tehnologije i tehnike
 - ▶ Klijentska aplikacija
 - ▶ Serverska aplikacija
- ▶ Specifikacija
- ▶ Implementacija
- ▶ Demonstracija implementacije
- ▶ Zaključak



Uvod

▶ Podaci u 21. veku

- ▶ “Ove godine sve se vrti oko podataka.” – BBC
- ▶ Digitalni trag (*Digital Footprint*)
- ▶ Načini prikupljanja podataka
 - ▶ Literatura
 - ▶ Ankete
 - ▶ Intervjui
 - ▶ Zapažanja
 - ▶ Dokumenti i zapisi



Uvod

- ▶ Prikupljanje podataka iz literature (veb artikli)
 - ▶ *Data scrapping/web scrapping*
- ▶ Osnovni motiv
 - ▶ Prikupljanje i pružanje podataka korisnicima koji im inače nisu lako dostupni.
 - ▶ Idealno rešenje za korisnika bi bilo da sve tražene informacije s različitih izvora budu dostupne na jednom mestu.
 - ▶ Softver će sve prikazane informacije prikupljati sa različitih izvora što će na kraju pomoći korisniku da ostvari kompletnu sliku o pomenutim događajima/dešavanjima, kao i sagleda situacije iz različitih uglova.



Uvod

▶ Tema rada

- ▶ Tema rada je kreiranje univerzalnog portala za informisanje korisnika kao i pružanje prilike za izjašnjavanje mišljenja na istom.

▶ Cilj rada

- ▶ Izgradnja veb sajta u kom je proces kreiranja sadržaja potpuno automatizovan, bi rešila problem ljudskog uticaja na selekciju bitnosti podataka i pojavljivanje grešaka.



Uvod

- ▶ Veb sajt ili nativna aplikacija
 - ▶ Iako sve veći i veći broj korisnika pristupa internetu putem nativnih aplikacija sa mobilnih telefona, jedan od razloga zašto je odlučeno da se rešenje implementira kao veb sajt umesto mobilne aplikacije je kako bih se omogućilo da korisnici identično mogu da pristupe kompletnom sadržaju kako sa mobilnih tako i desktop/laptop uređaja.



Stanje u oblasti

▶ Softver

- ▶ Softver koji obavlja funkcionalnost prikupljanja internet sadržaja poput online novina, podkasta, blogova i video blogova se naziva agregator.

▶ Agregatori vesti

- ▶ Agregatori pružaju mogućnost pristupa sadržaju sa više različitih sajtova na jednom mestu. Veb stranica agregatora prikazuje trenutno najnoviji sadržaj sa agregiranih sajtova i pri svakoj izmeni sadržaja originalnih stranica vrši se njeno ažuriranje.



Stanje u oblasti

- ▶ **Vrste agregatora**

- ▶ Manualni Agregatori
- ▶ HTML Agregatori
- ▶ RSS Agregatori

- ▶ **Efekat na sajtove sa kojih se prikuplja sadržaj**

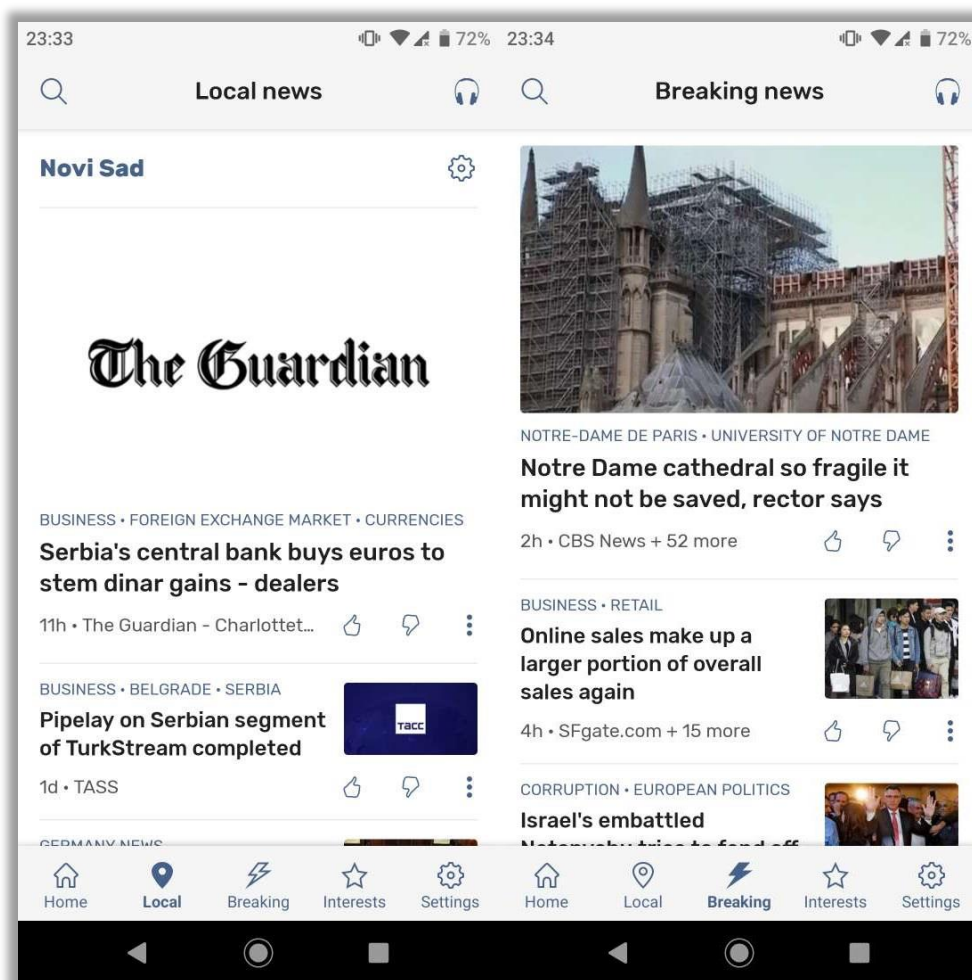
- ▶ Uvek je postojala polemika da li su agregatori dobro ili loše rešenje što se tiče prikupljanja sadržaja i efekta koji imaju na izvorni sajt.



Stanje u oblasti

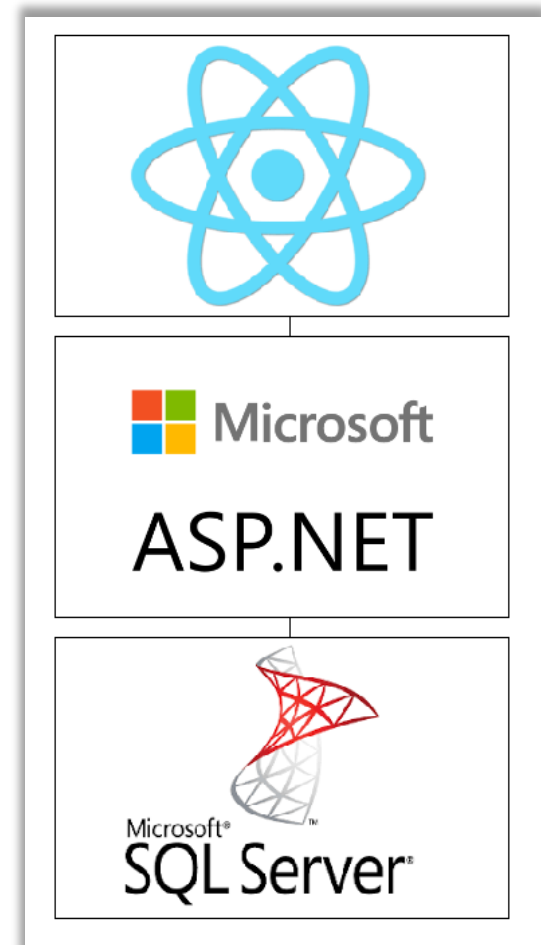
► Postojeći agregatori

- DnevneNovine.rs
- Naslovi.net
- Svevesti.com
- Global-Report.com
- News360



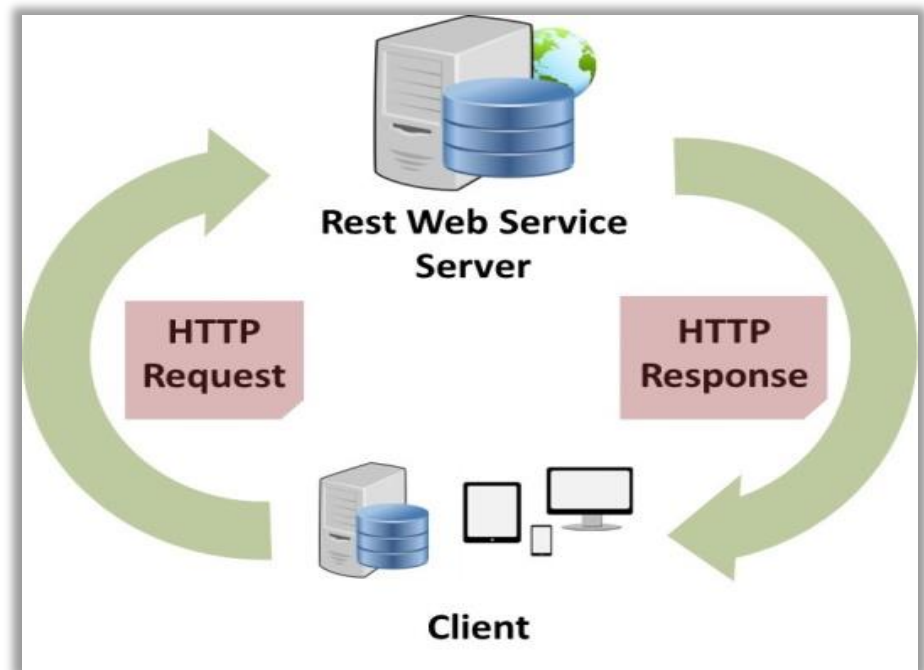
Korišćene tehnologije i tehnike

- ▶ Struktura ovog sistema je sačinjena iz tri celine; klijentske strane, serverske strane i baze podataka. Klijentska i serverska strana su povezani pomoću REST-a.
- ▶ Klijentska strana
 - ▶ React
 - ▶ Axios



Korišćene tehnologije i tehnike

- ▶ Serverska strana
 - ▶ ASP .Net Framework
 - ▶ Entity Framework
 - ▶ *HtmlAgilityPack*
 - ▶ *Selenium*
- ▶ Baza podataka
 - ▶ Microsoft SQL Server



Specifikacija

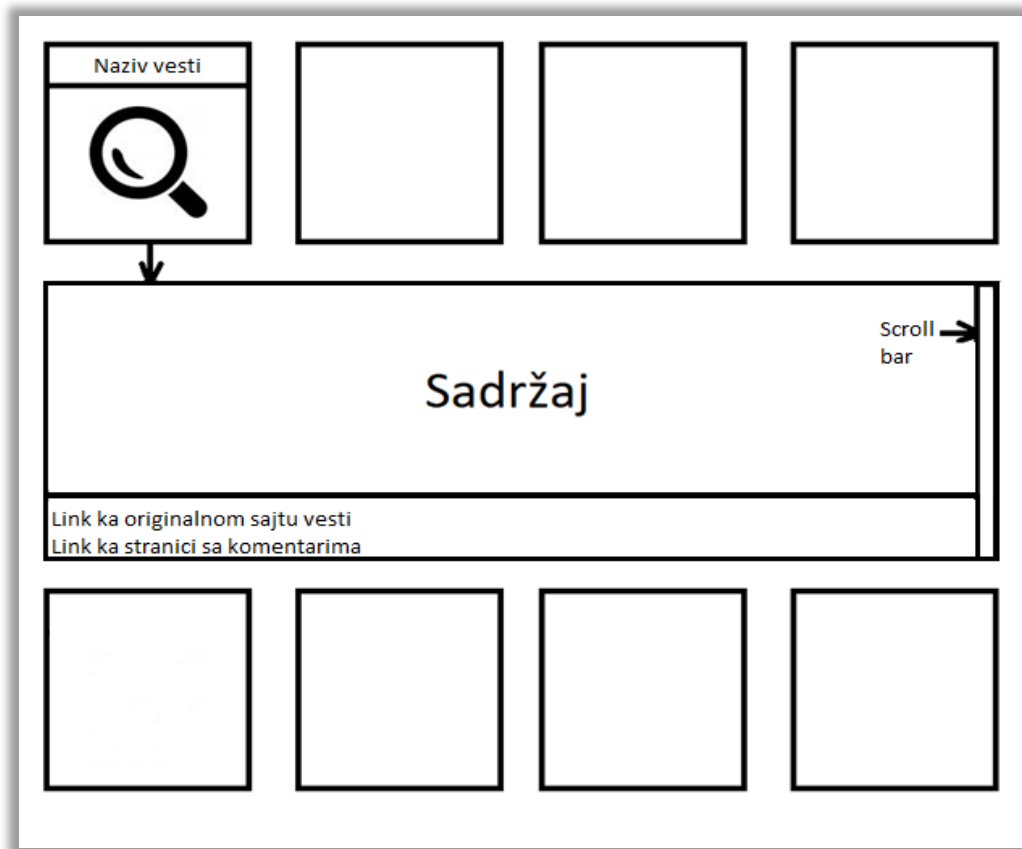
- ▶ Kreiran je sajt koji sadrži vesti koje dolaze sa različitih portala dnevnih novina u Srbiji (Kurir, B92, Blic, Telegraf, Srbija Danas).
- ▶ Podaci koji se preuzimaju sa originalnih sajtova:
 - ▶ Naslov vesti
 - ▶ Link ka vesti
 - ▶ Slika vesti
 - ▶ Tag



Specifikacija

► Funkcionalnosti sajta

- Proširivanje vesti
- Sekcije sortirane po popularnosti
- Sekcije sortirane na osnovu korisnikovih pregleda
- Pretraživanje
- Kategorizacija
- Čuvanje vesti



Specifikacija

► Vrste korisnika:

- Gost
- Običan korisnik (ulogovan)
- Administrator

► Prava pristupa

Proces	Gost	Korisnik	Administrator
Čitanje vesti	Da	Da	Da
Pretraga	Da	Da	Da
Kategorije	Da	Da	Da
Komentari	Čitanje	Čitanje, Postavljanje komentara, Izmena i brisanje sopstvenih komentara	Čitanje, Postavljanje komentara, Izmena i brisanje svih komentara
Čuvanje vesti	Ne	Da	Da
Stranica sačuvanih vesti	Ne	Stranica sa sopstvenim sačuvanim vestima	Stranica sa sopstvenim sačuvanim vestima
Stranica za upravljanje korisnicima	Ne	Ne	Da



Implementacija

► Klijentska aplikacija

► Rutiranje

```
1. function App() {  
2.   return (  
3.     <Router>  
4.       <Route path="/" exact component={() => <Layout><Homepage /></Layout>} />  
5.       <Route path="/home" component={() => <Layout><Homepage /></Layout>} />  
6.     </Router>  
7.   </Fragment>  
8. );}
```

► Ispis vesti

```
1. export default const Newest = (props) => {  
2.   const [posts, setPosts] = useState([]);  
3.  
4.   useEffect(() => {  
5.     getNewestData(setPosts); //API Call  
6.   }, []);  
7.  
8.   return (  
9.     <Fragment>  
10.      <h2>Najnovije vesti</h2>  
11.      <div className="row">  
12.        {posts && posts.map((item, i) => <Post key={i} {...item} /> )};  
13.      </div>  
14.    </Fragment >  
15.  )  
16. }
```

Implementacija

► Serverska aplikacija

► Primena selenium-a

```
1.  private static HtmlDocument GetHtmlDocumentFromSel(IEnumerable<SelInstruction> instructions, string url,
    string encoding)
2.      {
3.          ChromeOptions options = new ChromeOptions();
4.          options.AddArgument("headless");
5.          IWebDriver _driver = new ChromeDriver(options);
6.          _driver.Navigate().GoToUrl(url);
7.          IWebElement element = null;
8.          HtmlDocument doc = new HtmlDocument();
9.          foreach (var inst in instructions)
10.             {
11.                 element = _driver.FindElement(By.XPath(inst.XPath));
12.                 if (inst.Action == "INPUT")
13.                 {
14.                     element.SendKeys(inst.Text);
15.                 }
16.                 else
17.                 {
18.                     element.Click();
19.                     Thread.Sleep(1000);
20.                 }
21.             }
22.          doc.LoadHtml(_driver.PageSource);
23.          _driver.Dispose();
24.          return doc;
25.      }
```



Implementacija

► Autorizacija

► Atribut za autorizaciju

```
1. public class AuthorizeUserAttribute : AuthorizeAttribute
2. {
3.     public string AccessLevel { get; set; }
4.     private readonly UserManager _userManager;
5.
6.     public AuthorizeUserAttribute()
7.     {
8.         _userManager = new UserManager();
9.     }
10.
11.     protected override bool AuthorizeCore(HttpContextBase httpContext)
12.     {
13.         var isAuthorized = httpContext.User.Identity.IsAuthenticated;
14.         if (!isAuthorized)
15.         {
16.             return false;
17.         }
18.         var username = httpContext.User.Identity.Name;
19.         var currentUserRole = _userManager.GetByUsername(username).Type;
20.         if (AccessLevel == null || AccessLevel == currentUserRole)
21.         {
22.             return true;
23.         }
24.         return false;
25.     }
26. }
```



Implementacija

► Upravljanje neželjenim pristupom

```
1.  protected override void HandleUnauthorizedRequest(AuthorizationContext filterContext)
2.      {
3.          var isAuthorized = filterContext.HttpContext.User.Identity.IsAuthenticated;
4.          if (!isAuthorized)
5.          {
6.              filterContext.HttpContext.Response.StatusCode = 401;
7.              filterContext.HttpContext.Response.End();
8.          }
9.          else
10.         {
11.             filterContext.HttpContext.Response.StatusCode = 403;
12.             filterContext.HttpContext.Response.End();
13.         }
14.     }
```

► Upotreba atributa

```
1.  [AuthorizeUser(AccessLevel = "ADMIN")]
2.      public JsonResult Index(int page = 1)
3.      {
4.          var users = _userManager.UsersPerPage(page);
5.          var lastPage = _userManager.CountPages();
6.
7.          return Json(new { users, lastPage }, JsonRequestBehavior.AllowGet);
8.      }
```



Demonstracija implementacije

► Stranica za prijavu ili registraciju

[Klikni za povratak na početnu stranicu](#)

[Klikni za povratak na početnu stranicu](#)

Dobrodošli na News

[Loguj se](#)

[Nemaš nalog? Registruj se.](#)

Dobrodošli na News

[Registruj se](#)

[Imaš nalog? Loguj se.](#)


► Početna stranica

News


Dobrodošli na News portal, ovde možete pročitati vesti prikupljene sa svih većih portala u Srbiji.

[Pogledaj najnovije »](#)


Najpregledanije vesti u poslednjih 6 sati




Bogdan promašio trojku za pobjedu u Bostonu, nova slaviše Toronto i Majamija



Fond za inovacije u medijskoj sferi




Metju Peri priznao u koju kolegicu iz serije "Prijatelji" je bio zaljubljen




Nakon Albanije jak zemljotres pogodio i BiH


Personalizovane vesti




Vučić potvrdio: Na Sretenje će biti usvojena posebna DEKLARACIJA O SRBIJI



Prelistavanje dnevnih novina za 24. novembar 2019. godine



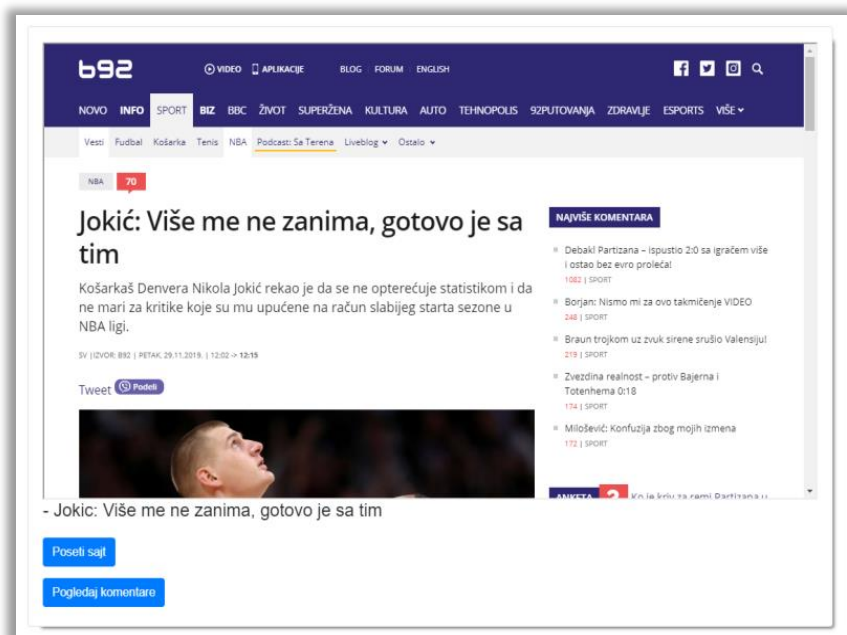
UHAPŠEN ZBOG ZLOČINA NAD SRBIMA: Upravnik logora Osman Osmanović PAO na prelazu Sremska...



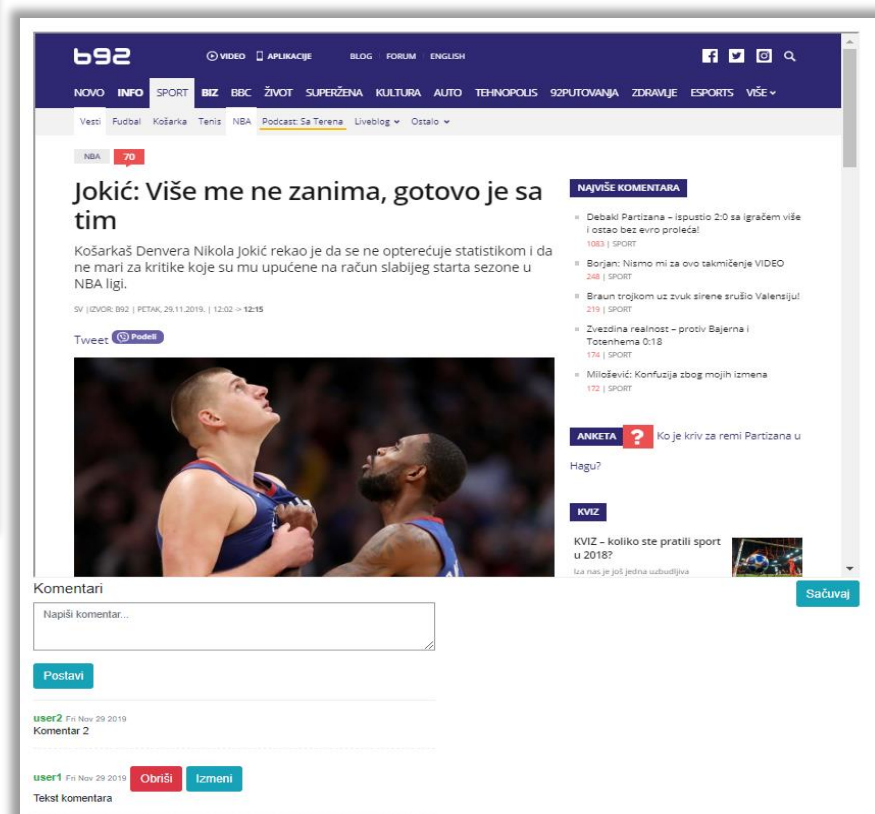
ANASONOVIĆ OTKRILA ISTINU O PROTESTIMA: Dilas davao pare, Borko vikao na mene, a Vojteh...

Demonstracija implementacije

► Proširena vest



► Stranica jedne vesti



Demonstracija implementacije

► Stranica sa dodatnim informacijama

Stranica je napravljena sa namerom da postignemo što efikasnije prikupljanje informacija širom Srbije i sveta.

Ovde možete poslati bilo kakav komentar ili preporuku.

Unesite vaš email

Ne delimo vaš email sa drugim licima.

Napiši komentar, preporuku...

☐ Prijavi me da dobijam novosti o portalu.

Pošalji

► Stranica za upravljanje korisnicima

#	Korisničko ime	Email	Tip korisnika	Opcije
1	1234	1234@test.com	USER	<div>Admin</div> <div>Obriši</div>
2	user1	user1@gmail.com	ADMIN	<div>Korisnik</div> <div>Obriši</div>

1

Zaključak

- ▶ U ovom radu je opisan višeslojni sistem koji demonstrira jedan od načina za prikupljanje podataka sa stranih sajtova i njihovu manipulaciju.
- ▶ Realizovanjem ovog projekta rešio bi se problem snabdevanja korisnika informacijama koje nisu tako lako dostupne i sprečio svaki tip manipulacije podacima pa samim tim i krajnjim korisnikom.



Zaključak

▶ Mogućnost proširivanja

- ▶ Stranica na kojoj bi uz pomoć osnovnog poznavanja HTML-a i XPath-a bilo moguće dodati još izvornih stranica sa kojih bi se prikupljali članci.
- ▶ Stranica na kojoj bi svaki korisnik mogao da zabeleži tagove postova koje ne želi više da pronalazi.
- ▶ Optimizacija sajta za konverziju u PWA aplikaciju.



► Kraj

