

# Sustavno prikupljanje sadržaja s Weba

# Kakav je Internet medij

- Možemo reći da je Internet kao medij po mnogim osobinama poseban, jer povezuje različite oblike usluga koje omogućuju spremanje i razmjenu podataka kao i komunikaciju među korisnicima.

# Tri grupe problema korištenja Interneta

- maliciozni (zlonamjerni) programi usmjereni prema računalu
- neželjeni sadržaji usmjereni prema korisniku
- neprovjereni ili neistiniti sadržaji

# Maliciozni (zlonamjerni) programi

- **Računalni virusi** koji se obično prikaže za neku izvršnu datoteku, a šire se prenošenjem te datoteke na drugo računalo putem nekog medija za distribuciju (spremnički štapić, CD/DVD) ili elektroničkom poštom (e-mail). Posljedice djelovanja virusa variraju od bezopasnog (koji uglavnom dosađuju korisniku) pa sve do teških posljedica kada dolazi do oštećenja datoteke, programa ili cijelog sustava.
- **Crv (Worm)** je program (ili skupina programa) koji je sposoban samostalno se kopirati i raširiti funkcionalne kopije na druga računala preko mreže.
- **Trojanski konji** imaju svojstvo da na prvi pogled izgledaju poput nekog običnog, korisnog programa, ali ipak nanosi štetu računalu na koje se instalira.
- **Spyware (špijunski alat)** je program koji se preko Interneta instalira na računalo i nepoznatoj osobi šalje podatke o korisnikovoj aktivnosti sa ciljem da ukrade broj kreditne kartice ili druge podatke koje može zloupotrijebiti.
- **Adware (oglasni alat)** koji se bez znanja korisnika instalira na računalo i uznemiruje ga nepoželjnim oglasima.

# Neželjeni sadržaji usmjereni prema korisniku

- **spam** je neželjena poruka koja stiže isključivo elektroničkom poštom i u većem opsegu može zatrpiti poštanski sandučić (*Inbox*) elektroničke pošte. Obično su to reklame raznih proizvoda i bezvrijedna pošta sa linkovima na nesigurne stranice.
- **lažna obavijest** (*hoax*) koja nastoji nagovoriti korisnika na neku radnju koja može biti ilegalna ili može uzrokovati oštećenje računala ako korisnik naivno povjeruje.
- **phishing** je postupak krađe tajnih podataka (zlouporaba identiteta) na način da se korisnika navede na otkrivanje povjerljivih podataka vjerujući da se radi o legalnoj instituciji.
- **dialer** je program koji prekida telefonsku vezu sa davateljem internetske usluge i poziv preusmjerava a neku udaljenu zemlju, a posljedica je enormno velik telefonski račun.

# Neprovjereni ili neistiniti sadržaji

- Internet sadrži veliku količinu informacija, ali nisu sve informacije jednako vrijedne jer svatko može napraviti web stranicu.
- Kako između svih tih stranica prepoznati one koje sadrže relevantne informacije, da im možemo vjerovati?

# Kako se zaštititi?

- **Koristiti vatrozid (*firewall*).** Vatrozid (*firewall*) je program koji nadzire (filtrira) prijenos podataka između računala i mreže. Njegova je uloga ograničiti tu komunikaciju tako da zloćudni programi ne mogu bez vašeg znanja pristupiti računalu.
- **Koristiti antivirusni program.** Uloga antivirusnog programa je da na računalu spriječi aktiviranje poznatih zlonamjernih aplikacija, poznatijih pod nazivima virusi, crvi i trojanski konji. Antivirusni program prepoznaje zlonamjerne aplikacije koje su mu poznate uspoređivanjem njihovog koda s bazom takozvanih antivirusnih definicija. Zbog toga je vrlo važno redovito ažurirati (osvježavati) definicije vašeg antivirusnog softvera korištenjem automatskog ažuriranja dostupnog u većini antivirusnih alata (*Automatic Update*).
- **Koristiti redovite softverske nadogradnje.** Ni jedan program, kao ni operacijski sustav nije savršen program i upravo ti propusti su najveći sigurnosni problemi.

# Praktični savjeti

- Ne otvarati privitak elektroničke pošte ako ne poznajemo osobu koja ga je poslala.
- Lozinke koje se koriste za različite pristupe ne smiju biti predvidljive jer se lako probiju.
- Kada ostavljate osobne podatke na web-obrascima provjerite da li se radi o sigurnim web stranicama (imaju slovo "s" iza protokola - <https://>). Sjetite se da na internetu, surfajući ili dopisujući se nikad niste sami. Stoga prije nego ostavite osobne podatke dvaput razmislite jer mogu biti upotrijebljeni u zlonamjerne svrhe.
- Izbjegavati sumnjive lokacije na Internetu gdje bi mogli biti izloženi zlonamjernim programima.
- Ne otvarajte *PopUp* prozore jer često predstavljaju zamku.
- Nikad ne instalirajte nepoznat program.
- Informirajte se o računalnoj sigurnosti.



# Preuzimanje sadržaja s weba

- **Tražilice** (*Search Engine*) su sustavi za pretraživanje koji za izdvajanje stranica koriste ključne riječi koje korisnik upiše u okvir traženja.
- **Tematski katalozi** sadrže veliku količinu podataka koji su organizirani po temama. Također, kao i tražilice, omogućuju pretraživanje stranica.
- U novije vrijeme nastaju web sjedišta koja objedinjuju različite informacije iz većeg broja izvora i nazivaju se **portali**. Portali često sadrže tematske kataloge i tražilice radi lakšeg dohvata informacija i ljudi ih rado koriste u prvom redu zbog vijesti.

# Spremanje cijele web stranice

- Ako spremate cijelu web stranicu postupak je sljedeći:
  - Učitajte web stranicu u Internet Explorer.
  - Kliknite na Datoteka (*File*) → Spremi kao (*Save As*) nakon čega će se otvoriti dijaloški okvir Spremi kao.
  - Otvorite mapu spremanja.
  - Po potrebi izmijenite naziv datoteke.
  - Kliknite na gumb Spremi.

# Spremanje dijela teksta s web stranice

- Učitajte web stranicu u Internet Explorer.
- Označite željeni dio teksta (povlačenjem miša).
- Postavite pokazivač miša na označeni dio i kliknite desnim klikom miša.
- Iz skočnog izbornika izaberite naredbu Kopiraj (*Copy*).
- Otvorite program za obradu teksta (*Word*) i umetnite tekst naredbama: Uređivanje (*Edit*) → Zalijepi (*Paste*).

# Spremanje slika s web stranice

- Postavite pokazivač miša na sliku i desnim klikom otvorite skočni izbornik.
- Odaberite naredbu Spremi sliku kao (*Save Pictures As*) koja će otvoriti dijaloški okvir za spremanje.
- Odaberite lokaciju za spremanje i kliknite na gumb Spremi.

# Ispis web stranice

- Ukoliko želite ispisati cijelu web stranicu na pisač:
  - Kliknete na izbornik Datoteka (*File*) → Ispis (*Print*).
  - U dijaloškom okviru Ispis (*Print*) birate hoćete li ispisati sve ili samo određene stranice.
- Za ispis samo označenog teksta na web stranici:
  - Označite tekst koji želite ispisati.
  - Desnim klikom na označeni tekst otvorit će se skočni izbornik.
  - U skočnom izborniku kliknite na opciju Ispis (*Print*).
  - U dijaloškom okviru za ispis označite Odabir (*Selection*).

# Pojmovi

- **Web** (*World Wide Web*) – usluga pregledavanja web-stranica na Internetu.
- **Virus** – računalni program koji je napravljen kako bi bez korisnikovog znanja ušao u njegovo računalo i načinio neku štetu.
- **Antivirusni program** – program za zaštitu računala od virusa. Viruse otkriva pri pokušaju ulaska u računalo i uništava ih.
- **Neželjena poruka** (*Spam*) – bezvrijedna poruka elektroničke pošte (*e-mail*) od nepoznatog pošiljatelja.
- **Web-preglednik** (*Web-browser*) – program za pregledavanje sadržaja na web stranicama.
- **Web-tražilica** (*Search engine*) – usluga za pretraživanje sadržaja na Internetu pomoću ključnih riječi.
- **Ključna riječ** – pojam koji najbolje opisuje sadržaj koji tražimo na Internetu.
- **Tematski katalozi** (*Web directory*) – usluga za pretraživanje sadržaja na Internetu putem web stranica gdje su sadržaji razvrstani po kategorijama i podkategorijama.
- **Download** – preuzimanje sadržaja sa Interneta i spremanje na lokalno računalo.
- **Upload** – slanje podataka sa lokalnog na udaljeno računalo.