

Digitala guldgrävare



**INFORMATION
OCH UPPGIFTER
OM STORA
DATAMÄNGDER
PÅ NÄTET**

Digitala guldgrävare



Innehållsförteckning

Introduktion	3
Vad, vem och hur?	
Tre frågor om big data	4
Vad är big data?	4
Vem samlar in data?	5
Hur samlas data in?	5
Från informationsmassa till mervärde	
– hur big data utnyttjas?	8
Värdefull data	8
Uträtta ärenden enkelt och snabbt	9
Riktad marknadsföring	9
Ute på okänt vatten	
– är intimiteten hotad?	11
Big data och datasekretess	11
Låt bli mitt privatliv!	12
Kontrollera din egen data	13
Aktivera dig på nätet	13
Tips om applikationer	14

KANSALLINEN AUDIOVISUAALINEN INSTITUUTTI
NATIONELLA AUDIOVISUELLA INSTITUTET
NATIONAL AUDIOVISUAL INSTITUTE



Materialet är producerat av Nationella audiovisuella institutet, FiCom RF, Konkurrens- och konsumentverket, Lastenlinkit.fi, Finlands Lionsförbund RF, Dataombudsmannens byrå, Verke och Kommunikationsverket.

Materialet har kommenterats av utbildningschef Heini Ala-Vannesluoma från centralen för vidareutbildning vid Tammerfors tekniska universitet och överinspektör Emil Asp från Kommunikationsministeriet.

Illustrering: Emmi Jormalainen



Introduktion

Hur kan reklamerna på Facebook veta vart jag vill åka på semester? Hur kan enskilda sökord hjälpa med att förutspå en influensaepidemi? Hur kan den information jag delar på nätet gynna miljontals andra?

Information som internetanvändare ger om sig själva och som samlas om dem utgör en enorm mängd data, så kallad *big data*. Att kunna analysera denna information som finns utspridd på nätet medför stora möjligheter, men väcker också många frågor. Under nutidens guldrusch gräver man efter guld bland big data. Guiden Digitala guldgrävare innehåller information och uppgifter om big data. Fokus ligger främst på insamlandet av information om internetanvändare, och på informationens användningssyften.

Med hjälp av materialet kan man reflektera över den egna rollen som informationsgivare och -nyttjare. Målet är att stöda ett aktivt och ansvarsfullt medborgarskap, samt att erbjuda verktyg för informationshantering och utvecklandet av livskompetensen i dagens digitala värld.

Anvisningar till den som använder materialet

Materialet lämpar sig som mediefostringsmaterial för ungdomar i högsta-
diåldern och äldre, t.ex. i skolor, bibliotek, inom ungdomsarbetet, eller för mediefostran under fritiden.

Guiden består av fyra olika delar, som introducerar fenomenet big data sett ur internetanvändarens perspektiv.

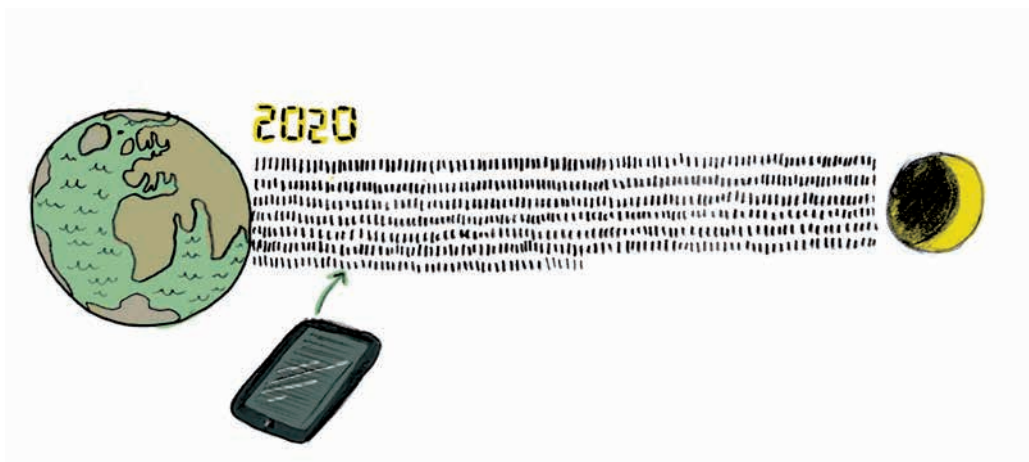
- Vad, vem och hur? Tre frågor om big data
- Från informationsmassa till mervärde - hur big data utnyttjas?
- Ute på okänt vatten – är intimiteten hotad?
- Kontrollera din egen data

I slutet av varje del finns övningsidéer för eleverna/ungdomarna. De kan utnyttjas som sådana, eller anpassas så att de lämpar sig för den egna gruppen. För att utföra en del av övningarna behövs datorer och internetuppkoppling.

Vad, vem och hur? Tre frågor om big data

VAD ÄR BIG DATA?

Big data innebär att man med hjälp av internet och modern teknologi samlar in, lagrar och analyserar stora mängder oorganiserad information i olika format och från olika källor. Typiskt för big data är att informationen skapas och sprids med en väldig fart. Vårt digitala universum – all den data vi skapar och delar – växer kontinuerligt med 40 procent per år. Forskningsinstitutet IDC har uppskattat att det för lagring av all ackumulerad data år 2020 skulle behövas en så stor mängd tunna surfplattor att om man radar dem på varandra skulle högen bli mer än 6,5 gånger så hög som avståndet till månen.



Big data ackumuleras t.ex. från lokaliseringsinformation, informationsökningar, innehåll i sociala medier, samt insamlad information från olika mätinstrument och sensorer. Olika former av information från många olika källor, t.ex. bilder, lokaliseringsinformation och diskussioner på sociala medier, kan på många olika sätt kombineras för analys och förädling. Big data har egentligen inget egenvärde, men informationen kan förädlas och sedan utnyttjas på många olika sätt. Därför kan big data liknas vid artefakten Sampo i finsk mytologi. Kalevalas Sampo var en mirakelmaskin som skapade rikedomar ur tomma intet.

DEM SAMLAR IN DATA?

Den enorma mängden information ackumuleras oftast utan att internetanvändaren ens lägger märke till det. Varje tangenttryckning, sidladdning och nyhetsartikel kan få vårt digitala fotavtryck att växa. Idag hänger också allt fler vardagliga saker ihop med nätet. Dels ackumuleras en hel del information, men information samlas också in aktivt.

Exempel på sådana som samlar information:

- aktörer inom sociala medier (t.ex. användarprofiler)
- administratörer av sökmotorer och nättjänster (sökhistorik, loggfiler)
- företag (transaktioner, kundregister, kundernas internetbeteende)
- offentlig förvaltning (bl.a. hälsovården, myndigheternas register, trafikplanering och -övervakning)
- leverantörer av karttjänster (bl.a. geografiska data, lokalisering av människor och trafikmedel)

HUR SAMLAS DATA IN?

Insamlingen av information om internet- och applikationsanvändare sker genom många olika processer.

Information samlas t.ex. då man skapar eller använder profiler på sociala medier, eller då man i samband med en produktbeställning på nätet anger information.

En sökmotor kommer ihåg dina tidigare sökningar genom att lagra din sökhistorik, och en webb-TV kan rekommendera intressanta filmer eller TV-serier på basen av vad du tidigare tittat på. Reseplanerarnas och karttjänsternas funktionalitet baserar sig både på geografiska data man fått från den enskilda användaren och de val större grupper personer fattat i trafiken.

En apparat som är ansluten till nätet kan identifieras med hjälp av IP-adressen. Med hjälp av kakor får en applikation eller webbsida och dess administratör information om vad en person gör på nätet. Tekniska data kan också fås direkt från t.ex. webbläsaren: ju mer webbläsarens användare modifierat webbläsaren genom att exempelvis installera expansioner, desto lättare är det att identifiera en person som rör sig på nätet.

INTERNETTERMINOLOGI

Webbläsare (i vardagligt tal ofta browser) är ett datorprogram som användaren kan läsa och skicka texter, se bilder, samt granska annat WWW-innehåll. Exempel på webbläsare är bl.a. Firefox, Chrome, Internet Explorer och Safari.

En applikation (eller app) kan vara ett datorprogram, ett program som installeras på en smarttelefon, ett spel, eller en expansion till en webbläsare. Applikationer är ofta verktyg ämnade för ett begränsat syfte, t.ex. att överföra filer, dela bilder, dölja sin egen IP-adress, eller kommunicera med familj och vänner.

Administratör är en person eller en grupp som övervakar en webbtjänst och ser till att allt fungerar. De flesta webbtjänster med registreringsmöjlighet har administratörer, eller åtminstone någon som äger webbadressen. Man kan vara i kontakt med administratören i frågor som gäller användningen av webbtjänsten. Ifall administratörens kontaktuppgifter inte uppges på en finländsk .fi-webbsida kan man på adressen domain.fi få reda på ägarens kontaktuppgifter.



VAD DÅ FÖR "KAKOR"?

En webbsida kan använda sig av kakor (cookies). En kaka är en fil som skickas till internetanvändarens dator, och som sedan lagras där.

Med hjälp av kakorna kan en internetanvändares uppgifter tillfälligt eller långvarigt sparas (så länge webbläsaren är igång, eller ännu längre). T.ex. kan lösenord eller innehåll i en webbutiks varukorg sparas så att de finns kvar efter att webbsidan eller webbläsaren stängts.

Kakor kan ofta underlätta internetanvändningen, då sidorna lagrat användarens uppgifter sedan tidigare. Å andra sidan kan man med hjälp av kakor iakta

internetanvändare. Kakor kan t.ex. avslöja via vilken webbadress man kom till sidan och hur länge man stannade kvar på sidan. Sidans ägare eller administratör kan också ge informationen vidare till en tredje part, t.ex. i reklamsyfte eller till företag som sysslar med att analysera sådana data.

I webbläsarens inställningar kan man inaktivera kakorna eller begränsa dem. Vissa tjänster, exempelvis nätbanker, behöver dock kakor för att fungera rätt. Kontrollera möjligheten att inaktivera kakor i din webbläsares inställningar och anvisningar.



Idéer på övningar 1.

Börja med att självständigt göra upp en lista över vilka alla saker du använder nätet till. Granska din lista och fundera på hur olika apparater, webbsidor och applikationer du använder samlar information om dig. Jämför svaren i gruppen och fundera på hurudan information som kan ha samlats in.

Tips: I samband med uppgiften kan deltagarna bekanta sig med en australiensisk bloggares lista "What does Facebook know about you". Bloggaren räknar upp saker som Facebook kan räkna ut om sina användare. Det är år 2030. Du är alltså 15 år äldre än nu. Fundera på vilka alla tjänster och föremål som kan samla in information om dig i framtiden. Välj en livssituation eller -händelse då du tror att sådan information som samlats in om dig kunde gynna dig och medföra mervärde eller bekvämlighet. Rita en bild av eller berätta en berättelse om dig själv i just den situationen.



Från informationsmassa till mervärde – hur big data utnyttjas?



VÄRDEFULL DATA

Att utnyttja enorma informationsmassor kräver mycket utvecklad analysering och förmågan att koppla ihop enstaka data med övrig tillgänglig information. Insamlad data har också direkt monetärt värde då företag som specialiserat sig på att analysera big data säljer information om internetanvändare till andra företag och aktörer.

Förädlad data kan vid sidan av företag och marknadsförare också gynna den enstaka internetanvändaren eller t.ex. någon som konstruerar en webbtjänst. Analysering av stora datamängder kan också främja vetenskaplig forskning och hjälpa den offentliga förvaltningen eller företag att t.ex. förbättra smidigheten och säkerheten i trafiken.



Sökmotorerna på internet kan ge intressant och nödvändig information om nationella eller världsomfattande fenomen. Genom att föra statistik över olika sökningar som görs och jämföra deras frekvens kan man få information om t.ex. epidemier, konsumtionsvanor eller fluktuationer i ekonomin. Sökningarna reflekterar främst aktuella händelser och fenomen, men kan ibland förutspå händelser i den närmaste framtiden.

En sökmotor kan förutspå en influensaepidemi

Då man blivit sjuk, fått hög feber och ont i halsen kanske man börjar göra symptomrelaterade sökningar på nätet. Googles influensasida samlar in denna information och kan på så sätt förutspå sjukdomsspridningar innan de blivit egentliga epidemier.

Hittills har hälsomyndigheter oftast fått tillförlitlig information om sjukdomsepidemier först flera dagar i efterhand, då människor sökt sjukhusvård på grund av symptomen. Därför hoppas man att sökmotorernas realtidsdata kunde fungera som verktyg för tidigare identifiering av epidemier, så att sjukdomen kan förebyggas. Enligt undersökningar stämmer sökmotorernas förutspådda epidemier och verkliga influensaepidemier väl överens, men modellen kan inte ännu ses som helt pålitlig.

UTRÄTTA ÄRENDEN ENKELT OCH SNABBT

Den vanliga internetanvändaren märker av förändringen av big data i form av tjänster som utvecklas snabbt och som blir lättare att använda. Genom att följa med vad användare gör och hur de betar sig kan tjänsteleverantörerna få veta vilka delar av tjänsten som behöver åtgärdas och vilka egenskaper som uppskattas mest.

Inköp, bank- och myndighetstjänster, samt många andra ärenden uträttas idag delvis eller helt och hållet på nätet. T.ex. reseplanering i realtid, en webbutiks varukorg som sparas, eller lagrade användaruppgifter är möjliga på grund av att information samlas in, lagras och utnyttjas för utveckling av olika tjänster.

RIKTAD MARKNADSFÖRING

Olika formulär på webbsidor är ett exempel på informationsinsamling för riktad marknadsföring. Samtidigt som sidan hjälper dig med att sortera olika potentiella alternativ och att hitta en produkt du vill ha, får webbsidan en hel del information om dig och om vad du är intresserad av.

Laura längtar efter värmen

Vädret är uselt och jag har inte ännu tagit ut alla semesterdagar. Nu längtar man verkligen efter ett varmt resmål. Jag söker sista minuten-resor på sidan galetbilligaflyg.nu. Borde man åka till Spanien, är det redan varmt i Grekland? Jag fyller i olika datum och söker efter trevliga resmål. Sidan erbjuder också förmånliga hotellnätter och jag kruxar för vilken typ av övernattningsställe jag är intresserad av. Borde man boka nu genast?

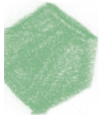
Nästa dag dyker det upp reklamer på min Facebooksida där webbsidan superbiligaresor.com erbjuder charterresor till Grekland, Spanien och Bulgarien. Hur kunde Facebook veta att jag vill åka på semester?



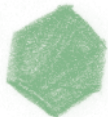
En webbsida kan överlämna eller sälja sådan information som samlats in åt så kallade tredje parter, alltså t.ex. marknadsförare, webbtjänster eller företag som ägnar sig åt analysering och profilering av internetanvändare. Ett besök på en webbsida eller några klick ger inte marknadsförarna någon exakt bild av dina konsumtionsvanor, men ju fler sidor du besöker och ju fler sökningar du gör kan man erbjuda dig desto mer riktad marknadsföring.



Idéer på övningar 2.



Dela in er i grupper om t.ex. fyra personer. Ena paret i gruppen söker med hjälp av en sökmotor efter texter och artiklar med sökorden "big data OCH fördelar", medan det andra paret använder sökorden "big data OCH nackdelar". Läs igenom texter ni hittat i ca en halv timme och skriv upp för- eller nackdelar som nämns. Jämför sedan resultaten med det andra paret och fundera på om för- eller nackdelarna väger högre. Ni kan också ordna en debattövning där ena sidan förhåller sig positivt till fenomenet big data och den andra sidan förhåller sig negativt.



Hur kan Google ge rätt sökningsförslag då du bara skrivit några bokstäver? Bekanta er med sökmotorens verktyg automatisk ifyllning (autocomplete) med hjälp av [Mediekunskapskolans övning](#).



Ute på okänt vatten – är intimiteten hotad?

Sett ur den vanliga internetanvändarens synpunkt finns det en hel del frågetecken med big data, speciellt då det gäller integritetsskydd eller missbruk av information. Brukar du läsa webbsidornas eller applikationernas användarvillkor? Har du tänkt på ifall någon skulle kunna missbruka din information på nätet, eller orsaka skada?

BIG DATA OCH DATASEKRETESS

Att ge ut egna personuppgifter på nätet och att tjänsteleverantörerna samlar in personuppgifter har under en kort tid blivit mycket vanligt. Man kanske inte lägger märke till hur omfattande insamlingen av information är (t.ex. med hjälp av kakor). Den big data som ackumuleras på nätet berättar en hel del om internetanvändaren och möjliggör en allt noggrannare profilering. Det finns farhågor om att insamlandet av big data är ett hot mot människors privatliv och datasekretess.

För att trygga människors rätt till privatliv i enlighet med Finlands lag bör den som samlar in eller använder big data minnas att:

- Det bör finnas en laglig grund för all användning av personuppgifter, t.ex. att personen gett sitt samtycke.
- Webbtjänstens användare bör öppet informeras om all insamling, användning och utlämning av personuppgifter. Generellt får personuppgifter inte behandlas i andra avseenden än de som nämns i villkoren.
- En person har generellt rätt att kontrollera sina egna uppgifter i olika register. Denna rätt gäller också sådan ny information som skapats genom att kombinera olika data.

På webbaserade tjänster hittar man oftast en dataskyddsbeskrivning och användarvillkor. Bekanta dig med dem före du ger dina personuppgifter och börjar använda tjänsten. På så sätt vet du vad du förbinder dig vid och vad dina uppgifter används till. Samtidigt får du veta vem du har att göra med och från vilket land tjänsteleverantören kommer. Lagstiftningen som berör internetanvändarnas rättigheter varierar mellan länder och områden.

Mer information om datasekretess och dataskyddsbeskrivningarna hittar du på dataombudsmannens webbsida tietosuoja.fi.

Krångliga användarvillkor?

Användarvillkoren och dataskyddsbeskrivningarna kan ibland kännas som långa och krångliga, men speciellt då det gäller sidor du ofta besöker lönar det sig att bekanta sig med dem noggrant. Med webbsidor och applikationer är det ofta så att man genom att använda tjänsten godkänner tjänstens villkor, som i exemplet nedan.

”Genom att använda Webbsidan förbinder du dig vid att följa gällande användarvillkor. Eftersom Webbsidans ägare kan göra förändringar i användarvillkoren uppmanar vi dig att regelbundet bekanta dig med dessa villkor. Ifall du inte godkänner användarvillkoren ber vi dig att inte använda Webbsidan.”

Webbtjänster gör rätt så ofta förändringar i användarvillkoren, så t.ex. sekretessinställningarna kan det löna sig att uppdatera med jämna mellanrum. Det är också bra att komma ihåg att all privat information man laddar upp på nätet kan bli offentlig någon dag.



LÅT BLI MITT PRIVATLIV!

Att lagra privat information på olika nätbaserade databaser och att dela information offentligt t.ex. på sociala medier ökar risken för olika typer av missbruk. En form av bedrägeri som blivit vanligare de senaste åren är identitetsstöld på nätet. En svindlare kan exempelvis skapa en skenprofil för en annan person på ett socialt medium, eller beställa produkter under en annan persons namn från en webbutik.

Ju mer information man kan hitta om dig på nätet desto lättare är det att utföra en identitetsstöld. Så lite som ett telefonnummer och födelsedatum kan leda till en hel del krångel. Bekanta dig alltså med webbtjänsternas sekretessinställningar och kontrollera hur pass detaljerade uppgifter du ger om dig själv offentligt.



Idéer på övningar 3.

Känner du till användarvillkoren på de sociala medier du använder mest? Bekanta er tillsammans med användarvillkoren. Hittar ni några överraskningar? Tycker du villkoren är tydliga, eller svåra att förstå?

Sök reda på en nättidnings, en spelsidas eller ett socialt mediums registerbeskrivning. Granska och fundera tillsammans på hurudan information webbsidan samlar in om användarna. Hurudan information samlar webbsidan in utan att användaren aktivt förser sidan med information? Vad behövs informationen till? Vilken information vill du ge åt dessa webbsidor? Diskutera: hur noggranna är ni med er egen sekretess?

Kontrollera din egen data

Att data ackumuleras och aktivt samlas in är något man inte själv kan hindra, förutsatt att man inte helt avstår från modern teknologi. Man kan dock påverka vilken information om sig själv och om ens aktiviteter som samlas in.

Den egna informationen kan man göra någonting produktivt med eller frivilligt överlåta i forskningssyfte.



AKTIVERA DIG PÅ NÄTET

Ett sätt att utnyttja big data är att erbjuda en flexibel prissättning av produkter beroende på efterfrågan, tidpunkt och konsumentbeteenden. En flexibel (eller dynamisk) prissättning används t.ex. vid försäljning av konsert- och flygbiljetter. Då man vill köpa biljetter kan priserna variera en hel del från dag till dag eller t.o.m. enligt klockslag. Man kan använda sig av vakter som meddelar dig då de biljetter du vill ha är på specialerbjudande.

Att överlåta egna användaruppgifter till olika företag kan också ses som frivillig- och gratisarbete. Om företagen får intäkter tack vare dina användaruppgifter, borde du då erhålla en andel av intäkterna? Vissa tjänsteleverantörer kan betala för att få tillgång till dina användaruppgifter, och betala dig då de säljer uppgifterna anonymt till tredje parter. För tillfället rör det sig dock på sin höjd om några tiotals euro per år åt den enskilda användaren.

Vid sidan av att dela information om internetbeteende kan privatpersoner också dela med sig av sin egen dators databehandlingskapacitet, t.ex. för att främja forskning. Vid Stanfords universitet forskar man i olika behandlingsformer för degenerativa minnessjukdomar, och vid Oxfords universitet försöker man lösa problemen med klimatförändringen.

<http://folding.stanford.edu/>

<http://www.climateprediction.net/>

Med hjälp av dessa anvisningar kan du radera kakor som iakttar ditt internetbeteende. Om du inte vill att kakorna lagras på din dator kan radera dem regelbundet eller använda läget privat surfning.

Chrome: https://support.google.com/chrome/answer/95647?hl=sv&ref_topic=3421433

Firefox: <https://support.mozilla.org/sv/kb/rensa-webbhistorik-kakor-sparad-data-firefox-OS>

Explorer: <http://windows.microsoft.com/sv-se/internet-explorer/delete-manage-cookies#ie=ie-11>



TIPS OM APPLIKATIONER

Du kan aktivera din webbläsares **Do Not Track-egenskap**, som skickar begäran om att webbsidan inte skall följa med vad du gör på sidan. För tillfället struntar dock de flesta webbsidorna i begäran och samlar oberoende in information om besökarna.

Webbläsarnas **inkognito- eller privat surfning-läge** innebär inte att du surfar anonymt på nätet. Webbsidorna du besöker kan fortfarande se din IP-adress och din webbläsares egenskaper. Inkognito-läget innebär dock att webbhistoriken inte syns på din dator, och att kakorna blockeras. Detta kan vara praktiskt om man t.ex. letar efter en överraskningspresent med familjens gemensamma dator, men läget skyddar inte dig märkbart på nätet.

Vissa webbläsare (t.ex. Firefox och Chrome) har expansioner man kan använda för att följa med hur egna uppgifter sprids på nätet eller för att begränsa spridningen. T.ex. expansionerna **Lightbeam** och **Collusion** avslöjar kakor från tredje parter som sparas på din dator. **Adblock Plus** hindrar reklamer i webbläsaren och i videor.

På nätet finns också klassificeringssystem (t.ex. finländska Web of Trust) för webbsidor eller -butiker som baserar sig på användarnas omdömen. Denna typ av applikationer påverkar i och för sig inte den information som samlas in om användaren, men kan vara till hjälp då man skall ta ställning till huruvida man vill uppge sina uppgifter. Anvisningar om säker användning av internetjänster finns bl.a. på [Kommunikationsverkets hemsidor](#).



Idéer på övningar 4.

Titta tillsammans på videon "[Valvova silmä](#)" (riktad till vårdnadshavare, på finska). Huruvida tankar väcker videon? Vilka bra och dåliga sidor med den övervakning dagens teknologi möjliggör hittar ni? Sök information på nätet både i inkognitoläge och med vanliga inställningar. Får du likadana sökresultat t.ex. med Google? Rekommenderar YouTube samma videor i bägge webbläsarfönstren? Jämför resultaten i grupp eller parvis.

