

Skriftlig inlämningsuppgift 2

Av Victor Wickström, ywidm03@student.vxu.se.

Fråga 1

En sökmotor såsom Google består tekniskt sett av ett antal komponenter som omnämns i kurslitteraturen. Redogör för dessa huvudkomponenter, dvs. dess funktion och konstruktion samt hur dessa olika delar kan vara uppbyggda och fungerar. Vilken tror du har störst inverkan och är viktigast för dig som utvecklare av optimerade webbplatser? Vilken har störst betydelse för den som söker information?

En sökmotor består av tre komponenter.

- Ett program som hämtar sidor från Internet. Detta program kallas för spindel. Spindeln hämtar information från webbsidor till sökmotorn och sammanställer listor med varje ord från webbsidorna och med sidans webbadress. Spindeln sammanställer sedan ihop informationen ett index. Spindeln följer också länkarna på sidan för att hitta nya sidor att indexera. Den återbesöker också sidorna för att uppdatera indexet. Spindel kan också kallas för robots, crawler och bot.
- En indexerare. Indexeraren går igenom sidan och lägger in alla ord i en katalog där alla ord pekar på det dokumentet som ordet kommer ifrån.
- Sökmodul som söker i indexet eller katalogen.

Det går även att dela upp detta i flera delar. Till exempel i Spindel som laddar ner sidorna, Den vandrade spindeln som kryper runt och följer alla länkar, indexer som delar upp all information som den får av spindlarna i olika sidor, databas som delar upp hemsidorna i olika element och sökmotorsresultat som visar upp resultatet av en sökning.

Jag tror att spindeln är den viktigaste komponenten i en sökmotor för utvecklare. Om inte spindeln hittar din sida så kommer den inte att komma med i indexeringen, om du inte lär dig hur spindeln fångar in informationen kan du inte topp optimera den.

Men det är klart det är bra att ha förståelse för alla komponenter. Kedjan är även beroende av den svagaste länken.

Jag anser att både spindeln och sökmodulen är viktigast för den som söker information. Utan spindeln får den inte in informationen som ska hittas. I en dålig sökmodul får man irrelevant information. Men jag måste nästan säga att det är informationen som är hämtad och algoritmen för att ta fram så relevant algoritm som möjligt är det viktigaste för den som söker information.

Fråga 2

Sökmotorer använder sig av olika sätt att indexera sökord beroende på böjningar mm. Det finns en terminologi som kan användas för att resonera kring sökord och böjningar

av dessa. Redogör för denna terminologi och använd dina egna sökord från tidigare uppgifter som exempel.

Detta kallas för stamning eller engelskans stemming. Stamning bygger på att man söker på sökordets stam. Detta genom att se till att alla ord med samma stam formas genom att den tar bort böjningen från stammen. Många stamningsalgoritmer bygger på en lista med vanligt förekommande ändelser. Algoritmen letar reda på om någon ändelse finns i ordet. Om det finns i ordet ta ändelsen bort om den uppfyller vissa krav. Exempel på krav kan vara att det som är kvar består av ett minst antal bokstäver och att ordet som är kvar har minst en vokal. (Sparck Jones och Willet, 1997).

Ett exempel på detta kan vara att om man söker på ordet blomma får man upp alla sidor med ordet blommor i sig.

Det går även att söka i olika kataloger i vissa sökmotorer då sökmotorn har indexerat olika ord i olika kataloger.

Fråga 3

Metataggar används i olika sammanhang i webbutveckling. Ge exempel på några av de vanligaste. Vilka har du använt dig av i dina uppgifter. Diskutera också baserat på kurslitteraturen och ev. andra källor vilken roll metataggarna har i sökmotoroptimering. Dina egna åsikter och funderingar ska också komma fram här.

De vanligaste meta taggarna är description och keywords.

I description taggen ska det vara en beskrivning om vad sidan handlar om och i keywords ska de nyckelorden, sökorden som du vill att folk ska hitta din sida på vara med i.

Vissa sökmotorer tar det som står i description-taggen och placerar det under titlen vid en sök träff. Många sökmotorer indexerar description-taggen. Det är alltså beskrivningen av vad sidan handlar om och den bör inte vara mer än 250 bokstäver lång och du borde bara maximalt ha med två kopior av samma ord. Annars kan detta leta till spamming och leda till att du bryter mot sökmotorernas regler och kan bli bestraffad utifrån det.

Keywords-taggen handlar om sökorden till din sida denna används av sökmotorernas indexering. Denna tagg var mer viktig förr i tiden men mindre viktig i dagens sökmotorer.

- Antal ord bör vara begränsat till mellan 10 och 12 ord.
- Du kan separera ord med hjälp av komma eller mellanslag, du behöver inte använda båda utan bara ett alternativ.
- Se till att de flesta orden i taggen även finns med i webbsidans innehåll.
- Duplicera inte orden en massa.
- Skriv sökorden med små bokstäver.
- Använd inte samma ord på alla sidor.

(Kent, 2004)

Andra taggar som finns är.

- Revist-after – som beskriver hur ofta som sökmotorn ska indexera om sidan.

- Robots – Talar om sökmotorn får eller inte får indexera den.
- Googlebot – Ger information till googles bot hur den ska använda informationen.
- Language – Säger vilket land som sidan är skriven för.
- Copyright – copyright
- Rating – Här kan man skriva om det är barnsäkert.
- Classification – Beskriver om sedan är lämpad för bredband osv.
- Author – Författare.
- Title – titel beskrivning av sidan.
- Hover – nyckelordstext.
- News – nyckelord.

Mina metataggar ser ut på följande sätt.

```
<meta name="robots" content="index, follow" />
<meta name="googlebot" content="index, follow" />
<meta name="description" content="Sökmotoroptimering och sökmotorvänlig
webbdesign. Victor Wickström - Flash design, datavetenskap och medieteknologi." />
<meta name="keywords" content="Flash webbdesign datavetenskap medieteknologi
Victor Wickström Wickstrom" />
<meta name="distribution" content="Global" />
```

I början av kursen verkade som att meta taggarna hade en stor betydelse när jag kollade på sökresultaten bland mina utvalda sökord. De sidorna som rankades högt hade taggarna robots, googlebot, description, keywords och distribution.

Enligt Peter Kent (2004) är det viktigt att descriptiontaggen kommer så högt upp som möjligt i HTML.sidan och att man inte byter plats på content-attributet och name attributet. (Kent, 2004).

Av egna erfarenheter under kursens gång verkar det som att description och keywords taggen har mindre roll nu för tiden, det är i så fall i de sökmotorer där det är descriptiontaggen som syns vid en sökträff. Något annat som jag har märkt är att googlebot kan vara en viktig tagg med attributen content="index, follow" för att få en högre placering. Det kan vara en ren slump att jag fick bättre placering med hjälp av den taggen.

I forumet Webmasters Network har det gått diskussioner om den attributen och revisit-after taggen är meningslös. Detta för att sökmotorerna vill ändå indexera och sidorna och tar med så mycket som möjligt och så ofta som möjligt men då baserat på hur ofta de kan. (<http://www.webmasternetwork.se/>, 2005)

Det verkar också som att meta-taggaras tid har passerat och sökmotorerna bryr sig mer om andra faktorer.

Fråga 4

Css och välkodad (X)HTML kan ha en positiv inverkan på resultaten i sökmotorer. Vad

menas med dessa två begrepp samt, ge exempel på sätt att använda sig av dessa tekniker för att uppnå bättre resultat. Fundera på vilket sätt du använt dig av dessa tekniker på din sida. Hur bör man behandla pdf dokument och javascript kod?

Med hjälp av CSS, (X)HTML och/eller XML kan man dela upp sidan i två delar. En del som har hand om layouten av sidan och en som har hand om struktureringen av texten. Detta gör att du kan strukturera upp informationen på sidan och göra den så ren som möjligt med (X)HTML. Med CSS kan du formatera och säga hur de olika taggarna ska se ut. Med XML kan du skapa dina egna taggar som du namnger och på ett relativt enkelt sätt strukturera upp informationen.

När du sökoptimerar en sida vill du gärna få bort så mycket kod som möjligt och göra så sökmotorboten kan lägga ner sin kraft på att ta hand om informationen.

Jag använder mig i så stor utsträckning som möjligt av strikt CSS och XML. Med detta menas att jag måste följa alla regler till punkt och pricka, om man jämför med vanlig HTML så är HTML mer förlåtande och tillåter än att glömma avsluta taggar och skriva så kallad fulkod. Detta gör jag för att kunna få en enhetlig design över alla sidor, att dela upp koden i information och designlayout samt att kunna få ren och fin strukturering av informationen på sidan.

Det som man bör göra är att ha CSS koden i en separat fil för att få bort så mycket som möjligt av det som inte är information på webbsidan.

Ett exempel:

I (X)HTML-filen finns följande:

```
<h1> Rubrik </h1>
```

I CSS-filen beskrivs det hur h1 taggen ska presenteras.

```
h1 {  
font-family: Verdana, Arial, Helvetica, sans-serif;  
font-size: 16px;  
font-weight: bold;  
color: #3D3D3D  
}
```

Det som hander är att texten och presentationen är uppdelad och CSSen beskriver att taggen h1 ska vara i Verdana familjen, vara 16 punkter hög, fet stil och färgen 3d3d3d.

CSS står för Cascading Style Sheets och är helt enkel till för att designers ska kunna få full kontroll för hur alla element ska se ut. (Kent, 2004)

JavaScript bör man ha i en extern fil som anropas från webbsidan, detta gör också att man kan få ner orelevant information för sökboten. Jag gör det genom att ha med denna taggen `<script src="libs.js" type="text/javascript"></script>` där libs.js är där jag har lagt min JavaScript kod.

För att sökoptimera PDFer kan du antingen konvertera dem till HTML och formatera det med XHTML och CSS eller så kan du lägga in Title, description, keywords taggar osv. De flesta stora spindlarna så som google klarar av att läsa PDFer idag. Jag gjorde ett test att ha en pdf på min sida och sedan ha en sida som jag har konverterat till HTML och CSS. Den sidan som jag har konverterat har en högre ranking i googles sökmotor. Så det kan vara en idé att ha två versioner av pdfen, en i pdf och en i XHTML och CSS.

Fråga 5

Att få med sina sidor i en sökmotor är i de flesta fall önskvärt. Hur bör man gå tillväga för att få sin sida med sin sida i sökmotorer och directories? Behöver man agera över huvud taget?

Hur kan man se till att de inte kommer med och indexerar?

Enligt Peter Kent (2004) kan man få sin webbsida indexerad på två olika sätt, antingen gör man så att folk länkar till ens sida (om de är indexerade själva) eller så registrerar man den själv. Det kan gå mycket fortare att få sidan indexera om man gör så att spindeln hittar sidan genom länkar till en. Att själv registrera den kan ta upp till en vecka om man någonsin blir registrerad. (Kent, 2004)

Här är en lista där man kan registrera sig till olika sökmotorer och directories.

AltaVista.com: <http://addurl.altavista.com/addurl/new>

Google: <http://www.google.com/addurl.html>

Inktomi (via Lycos): http://insite.lycos.com/searchservices/lite_step1.asp

Inktomi (via MSN): <http://submitit.bcentral.com/msnsubmit.htm>

FAST/AlltheWeb: http://www.alltheweb.com/add_url.php

SUNETs Webb katalog <http://katalogen.sunet.se/register-swe.html>

Yahoo directory <http://docs.yahoo.com/info/suggest/>

I filen robots.txt kan du skriva vilka sidor som inte sökmotorerna ska indexera. Detta kan vara bra om man har information som inte vill att andra ska få tillgång till eller om man inte vill att spindeln ska belasta servern på de sidorna.

På webbsidan <http://www.searchtools.com/robots/robots-txt.html> står det hur man kan utforma filen för att sökmotorerna ska kunna läsa och tolka filen.

Det går även att ha metataggar som styr om den aktuella sidan ska indexerar eller inte.

- `<META name="ROBOTS" content="NOINDEX">` - Indexera inte sidan, men följ länkarerna på sida
- `<META name="ROBOTS" content="NOFOLLOW">` - Indexera sidan, men följ inte länkarna på sidan
- `<META name="ROBOTS" content="NOINDEX,NOFOLLOW">` - Indexera inte sidan och följ inte länkar på sidan.

(Search Tools Consulting, 2003)

Fråga 6

Man kan tyvärr råka ut för att ens sidor eller sajter straffas (utesluts) av en eller flera sökmotorer. Ge exempel på faktorer som inte uppskattas av sökmotorer och fundera på

situationer när detta kan leda till problem på dina sajter, dvs. har du ärliga avsikter som du misstänker kan straffas trots goda avsikter? Om du inte hittar egna exempel från dina sajter, tänk dig in i andra situationer då detta kan leda till problem.

Det finns lite olika fall som man kan bli utesluten från sökmotorerna, men de bygger på att man bryter någon av deras regler.

Några av dessa är:

- spamming – Upprepning av ord eller fraser för många gånger.
- cloaking – Är en teknik som visa olika sidor för olika användare så att spindlarna får en sida som optimerad för att få hög ranking och en annan sida till resten.
- dörrsidor – sidor som är gjorda för att skicka tillbaka användarna till den riktiga sidan.
- hallsidor – Är som en sitemap men innehåller dörrsidor.
- spegeldomäner – En teknik som gör att man använder samma sida till flera domäner.

(SEOdesign, 2005)

Andra metoder kan vara:

- Osynliga texter – Läger text i samma färg som bakgrunden eller gömmer den på något annat sätt.
- Missledande titel/beskrivning/taggar.
- Missledande länkar
- Överregistrering – Detta kan ske vid överanmälning till en massa gratistjänster som anmäler sidor till sökmotorer. Går in i kategorin spamming.
-

Min sida låg uppe på en domän först och sedan flyttade jag den till en annan domän. Problemet var att jag inte tog bort den första domänen och det gjorde så att jag blev utesluten från SUNETs webbkatalog och Google. När jag ändrade om den ”gamla” sidan så att den fick en redirectiontag så blev jag indexerad i google igen. Det tog inte många dagar innan de upptäckte detta och av SUNET fick jag ett mail av att jag var utesluten från deras kataloger på grund av att den var för lik den andra sidan. Jag fick inget meddelande av google men de skötte allt automatiskt och jag behövde inte göra något för att få den indexerad igen.

Fråga 7

Ge din bild av området sökmotoroptimering. Vilka steg ser du i processen? Vilka delar är viktigast? När bör detta tankesätt komma in i utvecklingsprocessen? Basera detta på källor från kurslitteraturen, kursen och ev. andra källor av hög kvalitet.

Analys

Detta är det första steget. Före du sätter igång med allting bör du fundera på vad din webbplats ska innehålla, vilka sökord du vill att folk ska hitta din sida med och hur man vill synas.

Du vill förmodligen ha så mycket trafik som möjligt men ändå ha relevanta besökare som blir nöjda när de besöker din sida. Du behöver tänka på vilken information som du kan erbjuda till dina besökare. Om du till exempel ska göra en hemsida för ett företag som säljer vägsalt kanske du inte ska sökoptimera sidan för sökordet sex som får tusentals träffar varje dag. Du kanske väldigt många besökare men de är inte intresserade av ditt företag och vägsalt.

Du kan gå in på <http://www.wordtracker.com/> som tar fram hur många sökträffar olika sökord ger. (SEOdesign, 2005)

Bygg din hemsida

När du har kommit på vilka sökord och vad din webbplats ska handla om ska bygga den och om du har information ska du ut med den.

När du bygger webbplatsen måste du bygga den så att spindlarna kan läsa den och tolka informationen på sidan. Om du har byggt upp din sida i Flash kan det vara en idé att tänka om för det kan vara svårt att få en topp placering i sökmotorer med en sida som är skapt endast med Flash. Detta är för att de flesta spindlarna inte kan läsa innehållet i Flash filerna. Det du kan göra är att skapa en identisk sida i HTML som går att sökoptimera.

Du bör inte bygga upp din sida i frames heller för ofta kan inte spindlarna gå vidare efter första sidan. (SEOdesign, 2005)

På min webbplats hade jag en frame som startsida först men detta gjorde att det bara var den första sidan som blev indexerad. Det som man kan göra är att bygga upp resten i no frames taggen och skriva in all information, länkar och en länk till en sitemap som gör det möjligt för spindeln att följa.

Använd inte JavaScript länkar och placera JavaScripten i en extern fil.

Bygg up det med strikt CSS och XHTML detta för att spindeln ska kunna läsa och strukturera informationen på ett bra sätt.

Använd titel och meta taggarna på ett vettigt sätt. Glöm inte alt taggen i imagetaggen, om du inte har med den kan inte sökmotorn veta vad det är för bild. Gå igenom resten av taggarna och använd H1, H2 och bold så att sökorden finns med i dem och du kan ändra utseende med CSS om du inte vill att de ska se ut som de gör. Ha med länktexten i länkarna och döp dina sidor med relevanta namn.

Detta är bara ett axplock med saker som går att göra för att få en sökoptimerad sida. Det finns massor av finesser. Men för att hålla det kort tog jag med dessa för att jag tycker att det är grejer och tips som har till grunden att tänka på när man bygger en sökoptimerad sida.

Skapa en roots.txt

Tala om för spindlarna vilka sidor som den inte får besöka.

Sökordsdensiteten

Se till att sökorden förekommer ett visst antal gånger och det med hänsyn till hur många mycket text det är. Men se till att det inte blir så kallad spamming. Alltså att något ord förekommer för mycket. (SEOdesign, 2005)

Namn

Se till att du har ett bra domännamn som helst innehåller något av dina sökord som du vill optimera. Se också till att namnen på sidorna har ett beskrivande namn och helst något av de sökorden som du optimerar för.

Registrering

När du är klar med innehållet av sidan och optimerat sidan är det dags för att registrera sidan i sökmotorer och directories.

Länkbyten

Se till att andra sidor länkar till en. Detta är ett viktigt steg att få en högre ranking och om du inte får högre ranking så har du i alla fall lockat folk till din sida genom länkarna. Se till att länkarna är rätt utformade. (SEOdesign, 2005)

Dessa stegen anser jag är de viktigaste i processen. Men om man ska välja de tre viktigaste grejerna skulle jag säga att det är att få ut informationen, strukturera information och byta länkar. Detta är för att om du väl ha skrivit sökoptimerad text som andra länkar till så kommer sökmotorerna att hitta till sidan förr eller senare.

Med andra ord är det bra att tänka på sökoptimeringen före utformningen av webbplatsen och under hela utvecklings processen och sedan att underhålla den med länkar och ny mer information. Men om man har byggt en bra grund är det mycket lättare att sökoptimera webbplatsen.

Fråga 8

Vilka fem faktorer anser du vara viktigast för att en webbplats ska rankas väl i en sökmotor ,motivera ditt svar. Här ska du även dra paralleller till och inkludera associationer till din sökmotordagbok där du gjort observationer av några konkurrerande webbplatser under en period.

1. Länkbyten

Nästan alla toppplaceringar under tiden som jag har bedrivit min sökmotorsdagbok har haft hög pageranking. De bygger på att man har sidor som länkar till en. Om man har webbsidor som länkar till ens sida hittas den lättare av spindlar. Jag har märkt att andra sökmotorer och spindlar har hittat mig en väg jag har registrerat mig till.

2. Använd taggar

Nästan alla sidor som har rankats bra bland mina sökord har haft med något av sökorden i titel taggen och meta-taggar. Vissa sökmotorer verkar kolla mer på detta än andra om man utvärderar min dagbok. Jag har även sett att det är bra med många länkar och en bra relevant beskrivning av länkarna. Det är även bra att ha med H1 taggen och markera sökorden. Det sista har jag inte sett så mycket av bland de sidorna som jag har undersökt.

3. Ut med all information

Under de senaste dagarna har jag sett via google analytics att det är många som har sökt att och kommit in på mina webbsidor genom att använda sökord som jag inte hade räknat med från början. Detta är för att jag har lagt ut uppsatser som folk sedan har sökt folk sökt på ord som finns med i uppsatserna och hitta dem. Detta har nästan dubblerat mitt besöksantal. Man vet aldrig vad folk söker på, det var en som har sökt på "information om vad gör man om bilen fastnar i snö?" och det är absolut inte vad min sida handlar om men den personen som har sökt på hade surfat runt på flera sidor av min webbplats också.

4. Ha sökorden högt upp

När jag söker på sidor så brukar jag leta efter den informationen jag ser först, i många fall orkar jag inte skrolla utan istället går jag tillbaka och går in på en ny träff. Det verkar som att några sökmotorer tänker det samma såsom google. Att den värdesätter informationen som kommer högt upp på sidorna och sedan skalar av neråt.

5. Strukturering av informationen

Jag har själv märkt att mina webbsidor rankas högre när jag kör strikt CSS även om jag nästan hade det innan. Detta vet jag inte om det stämmer men det verkar som att mina webbsidor rankas högre än andra mycket tack vare detta. Att många inte delar upp och strukturerar informationen samt att de inte delar upp sidorna så att CSSen tar hand om designen så att det bara finns den allra nödvändigaste koden i HTML filen.

Något som även skulle vilja ha med är namn, att man har en egen domän och har valt att ha med något sökord i domän namnet samt döper sina HTML filer till sökorden. Detta också för att ge ett seriösare intryck till besökarna.

Referenser

DVA030_HT05 Webbdesign för sökoptimering (2006) *LUVIT Learning Centre – Education*

Kräver inlogging

Tillgänglig <<http://bth.luvit.se/luvit/main.asp>> (2006-01-10)

Jones Sparck, Peter Karen, Peter Willet (1997) *Readings in Information Retrieval*, Morgan. San Fransisco, California. Kaufmann Publishers, Inc.

Kent Peter. (2004) *Search.Engine.Optimization.for.Dummies* USA. Wiley Publishing inc

Search Tools Consulting (2003) *Search Indexing Robots and Robots.txt - SearchTools.com*

Tillgänglig: <<http://www.searchtools.com/robots/robots-txt.html>> (2006-01-10)

SEOdesign (2005) *SEOdesign - Sökmotoroptimering och Webbdesign. Marknadsföring på Internet*

Tillgänglig: <<http://www.seodesign.se>> (2006-01-10)

Webmasters Network (2005) *Webmasters Network*

Tillgänglig: <<http://www.webmasternetwork.se/>> (2006-01-10)