



UPPSALA
UNIVERSITET

InternetExplorers – Delrapport 1

Metod och genomförande

Håkan Selg

Nationellt IT-användarcentrum NITA

Juni 2008

**Internetanvändare i svenska universitet och högskolor 2007
– En framsyn av morgondagens Internetanvändning
Ett projekt finansierat av VINNOVA**

Om projektet

Bland avancerade användare av Internet intar studenter en nyckelroll. Att studera deras beteenden på nätet gör det möjligt för oss att tidigt upptäcka nya tendenser och att förstå vad som ligger framför oss på några års sikt. Vidare utgör studenterna en stor grupp i samhället som snart kommer att föra sin nyvunna kunskap och sitt tänkande in i arbetslivet.

Hur använder de Internet? Vilka webbsidor besöker de, vad läser de om, vilken kunskap har de? Delar de med sig av kunskaper och innehåll, medverkar de i diskussionsgrupper?

InternetExplorers ska belysa hur Internet används inom universitets- och högskolevärlden. Studien ska utgöra ett underlag för en kvalificerad analys och framsyn av morgondagens breddanvändare av Internet. Den ska ingå i en vidare diskussion om nya användarmönster som i sin tur tolkas i skenet av dess potentiella innovationskraft. Målet är att skapa och sprida ny kunskap om centrala förutsättningar för utformning av framgångsrika IT-baserade tjänster.

InternetExplorers genomförs av Nationellt IT-användarcentrum NITA vid Uppsala universitet och finansieras av VINNOVA. Studien sker genom en enkät under hösten 2007 riktad till studenter, doktorander och anställda vid 32 svenska universitet och högskolor. Statistiska Centralbyrån svarar för urval och datainsamling.

Resultaten från undersökningen kommer från och med augusti 2008 att löpande publiceras på projektets webbplats www.internetexplorers.se. Dessa delrapporter, tillsammans med resultaten från fokussamtal, analys och slutsatser kommer att sammanställas i en slutrapport senare under hösten 2008.

För ytterligare information, kontakta Håkan Selg, e-post selg@nita.uu.se, tfn 08-30 00 40, eller gå in på projektets webbplats.

Innehållsförteckning

InternetExplorers: Uppföljning av SUNET-studien	4
Genomförande	5
Urval och datainsamling via SCB	5
Population	5
Urvalsram	5
Viktberäkning	6
Bortfall	6
Uppföljning av urvalet	7
Enkätfrågorna	10
Redovisning av statistiska mått	10

InternetExplorers: Uppföljning av SUNET-studien

SUNET¹ lät år 2003 undersöka den svenska högskolevärldens användning av Internet. Studien genomfördes som en urvalsundersökning och var den första i sitt slag. En bred ansats valdes i syfte att skapa en heltäckande bild av användningen.

Ansvarig för genomförandet var Håkan Selg. Som rådgivare till projektet knöts professor Olle Findahl vid World Internet Institute. Statistiska Centralbyrån svarade för datainsamlingen.

Enkätfrågorna gavs en utformning så att jämförelser med utfallet i World Internet Institutes' undersökningar "Svenskarna och Internet" liksom med andra undersökningar inom World Internet Project skulle vara möjliga. Av dessa jämförelser stod det klart att de mest avancerade Internetanvändarna återfinns inom högskolesektorn. Framför allt gällde det studenter och doktorander inom data- och ingenjörsvetenskapliga utbildningsområden.

Dessa erfarenheter har legat till grund för InternetExplorers. Följaktligen har huvuddragen i genomförandet av SUNET-undersökningen direkt kopierats. Men det finns också viktiga skillnader.

SUNET-studien var en pilotstudie, föregångare saknades i stort sett. Då var det angeläget att skaffa en god överblick av hela det akademiska fältet och att inte missa viktiga användningar och användare. Att identifiera de stora mönstren i användningen blev på motsvarande sätt en viktigare uppgift än att göra djupdykningar inom enskilda delområden.

Frågorna fick därför en allmän utformning och täckte ett stort antal olika användningsområden såväl studie- som arbetssituationer som i den privata sfären. Vidare var ett viktigt syfte var då att upptäcka eventuella systematiska skillnader i Internetanvändningen med avseende på personalkategorier, utbildningsområden och högskolekategorier.

Med en sådan långtgående uppdelning i undergrupper² riskerar precisionen i de statistiska skattningarna att bli lidande. Varje skattat antal eller värde omges av ett större osäkerhetsintervall än om man haft färre strata. Detta följer emellertid av målsättningen att i första hand uppnå en heltäckande bild av landskapet och att inte missa viktiga mönster i användningen.

I jämförelse är InternetExplorers mycket mer fokuserad på 1) interaktiva användarmönster bland 2) avancerade Internetanvändare. Detta får dels betydelse för inriktning på frågorna. Dessa är nu avgränsade till interaktiva tillämpningar som bloggar, nätgemenskaper (eng. *communities*), öppen källkod och öppna databaser etc., och med en hög detaljeringsgrad. Vissa allmänna frågor om Internetanvändning har dock behållits för att skapa jämförbarhet med SUNET-studien.

Även urvalet har påverkats. Av SUNET-studien framgick att användarmönstren bland universitets- och högskoleanställda inte nämnvärt avvek från svenskarna i allmänhet. Därför har urvalet den här gången koncentrerats till studenter och doktorander, och med extra höga urvalsstorlekar bland de kategorier som i SUNET-studien identifierades som mest avancerade. Syftet är att inom mycket kvalificerade och nydanande tillämpningar, där ännu så länge procentuellt endast små andelar av populationen är aktiva, ändå kunna hitta så pass många respondenter att statistisk bearbetning ska vara möjlig.

¹ Swedish University Computer Network, med organisatorisk koppling till Vetenskapsrådet.

² Stratum, pl. strata

Fortfarande innehåller InternetExplorers data från ett mindre urval bland universitetsanställda, och då enbart forskande och undervisande personal. Syftet är dels att skapa en jämförelsegrupp till 1) studenter/doktorander, 2) svenska Internetanvändare generellt och 3) motsvarande personalkategorier i SUNET-studien.

För att sammanfatta: Ändamålet med SUNET-studien var att skapa bred överblick, medan InternetExplorers syftar till fördjupning på avancerade användarmönster.

Genomförande

Urval och datainsamling via SCB

Studien har genomförts i form av en urvalsundersökning. Under sommaren 2007 lades ett uppdrag på Statistiska centralbyrån (SCB) att ansvara för insamlingen av uppgifter. Undersökningen genomfördes som en postenkät inklusive ett följebrev (missiv) utformat av NITA, med tre påminnelser under perioden november 2007 – januari 2008.

Population

De objekt som man vill kunna dra slutsatser om fanns i tre, delvis överlappande, delpopulationer:

1. Personal vid universitet och högskolor
2. Registrerade på forskarutbildning
3. Registrerade på grundläggande högskoleutbildning

Endast studerande och personal vid de 32 högskolor som omfattades av 2003 års studie ingick i undersökningen.

Urvalsram

Urvalsramarna till de tre delpopulationerna togs från registren ”Personal vid universitet och högskolor” och ”Universitets- och högskoleregistret”. Aktualiteten i registerstatistiken skilde sig mellan delpopulationerna.

1. Personal: Ramuppgifterna avsåg 2006 års uppgifter. I urvalsramen ingick tjänstekategorierna: professorer, forskarassistenter, lektorer, adjunkter, ”annan undervisande och forskande personal” och forskarstuderande. Antalet i urvalsramen var 23 319.
2. Forskarstuderande: Ramuppgifterna avsåg uppgifter för höstterminen 2006. I urvalet ingick doktorandstuderande med en aktivitet om minst 10 procent. Antalet i urvalsramen var 15 393.
3. Grundläggande högskoleutbildning: Ramuppgifterna avsåg preliminärdata för höstterminen 2007. Endast heltidsstuderande, definierat som programstuderande eller registrering på minst 15 kurspoäng (nya poängsystemet) t.o.m. oktober 2007, ingick i urvalsramen. Antalet i urvalsramen var 253 442.

Det totala antalet i urvalsramen var 284 524, dvs. färre än summan av de tre grupperna. Orsaken är att det fanns personer som ingick i mer än en av de tre grupperna. Vanligaste kombinationen var doktorandstuderande som även fanns i personalkategorin forskarstuderande.

För de två studerandegrupperna bildades åtta strata uppdelade på ämnesinriktning som grupperats enligt SUN2000. Ett nionde stratum bildades för personal som inte var registrerade i forskarutbildning.

Stickprovets storlek i dessa strata bestämdes utifrån ett precisionskrav: felmarginalen för ett 95-procentigt konfidensintervall för en skattad andel på 50 procent inom strata skulle vara högst ± 6 procent. Det gav en urvalsstorlek på approximativt 400 personer i varje stratum. I två strata – Data- och systemvetenskap samt Tekniska utbildningar – sattes ett högre precisionskrav vilket gav urvalsstorlekar på approximativt 600 personer. Inom just dessa ämnesområden återfanns i 2003 års undersökning en stor del av de Internetanvändare som uppvisar de mest avancerade användarmönstren.

Tabell 1 Population och urval fördelat på stratum

	Population	Urval
TOTALT	284 524	4 000
Humaniora och teologi	22 503	400
Samhällsvetenskap och juridik	80 897	400
Matematik och naturvetenskap	12 390	400
Datavetenskap och systemvetenskap	6 494	600
Teknisk utbildning	42 008	600
Läroarbetsutbildning	38 462	400
Vård- och medicinutbildning	43 668	400
Övriga utbildningar	21 844	400
Personal	16 258	400

Viktberäkning

Vikter har tagits fram så att resultatet kan redovisas för hela populationen och inte bara för de svarande. Vid viktberäkningen har hänsyn dels tagits till urvalsdesignen, dvs. urvalets storlek för en viss grupp – t ex studenter i humaniora – i relation till populationen. Hänsyn har också tagits till skillnader i svarsbenägenheten mellan olika grupper (t ex mellan kvinnor och män).

Bortfall

Populationen utgjordes av studerande och personal vid universitet och högskolor. Totalt besvarade 2 047 personer frågeblanketten, vilket utgör 50 procent av urvalet (vägd svarsandel). 134 svarande var dock övertäckning, d.v.s. att de inte längre studerade eller arbetade inom högskolan. Tidigare erfarenheter, bl.a. från Studentspegeln, pekade mot en svarsfrekvens på 55 procent. Utan en separat bortfallsundersökning är det inte möjligt att förklara det nu uppnådda lägre utfallet. Vi kan dock konstatera att utfallet i 2003 års undersökning, till den delen det gällde studenter, också låg strax under 50 procent. Det kan alltså inte uteslutas att utfallet är kopplat till ämnet för undersökningen.

För de enskilda frågorna kan sedan ytterligare bortfall tillkomma till följd av att vissa svarsalternativ inte fyllts i, eller fyllts i på ett felaktigt sätt. Genomgående är dock detta *partiella bortfallet* i anslutning till de enskilda frågorna mycket lågt.

Frågan att ställa sig är hur bortfallet kan tänkas påverka undersökningens resultat? Vi vet t.ex. att svarsbenägenheten är åldersrelaterad, och att bortfallet generellt är högre bland yngre personer än bland äldre. Eftersom undersökningsmaterialet klart visar på att Internetanvändning i många avseenden är generationsfråga, ligger det nära till hands att misstänka att bortfallet resulterar i en underskattning av nya mönster i Internetanvändningen.

Samtidigt finns sådan information om olikheter i svarsbenägenhet inbyggd i viktberäkningen, dvs. bortfallets effekter ska i princip ha eliminerats vad gäller åldersfaktorn. Motsvarande gäller för kön och ämnestillhörighet. Men, eftersom svarsfrekvensen generellt är lägre än för t.ex. Studentspegeln, där samma metod för kalibrering av vikterna tillämpas, har vi troligen här att göra med en annan – okänd – faktor. Därmed går det heller inte att ha någon uppfattning om bortfallet slår i den ena eller andra riktningen på enkätresultaten. Det enda som är klart är att vi vid punktskattning av en viss variabel får ett bredare osäkerhetsintervall än vad ursprungligen var målet (± 6 procent).

Uppföljning av urvalet

I frågeformuläret ombads respondenterna att ange inom vilket ämnesområde man studerade eller arbetade under höstterminen 2007 (Fråga 3). Totalt fanns fjorton fasta svarsalternativ och ett öppet. Det har gjort det möjligt att jämföra utfallet med vad som förväntats utifrån de registeruppgifter som legat till grund för det stratifierade urvalet. Svaren har sedan kodats enligt följande:

Tabell 2 Uppföljning av ämnesinriktning. Kodning av svarsalternativen

	Angivet ämnesområde (Fråga 3)	Kod	Benämning
1	Språk och språkvetenskap	101	Humaniora
2	Humaniora utom språk och språkvetenskap	101	
3	Teologi	101	
4	Samhällsvetenskap inkl. beteendevetenskap	102	Samhällsvetenskaper
5	Ekonomi	102	
6	Juridik	102	
7	Matematik	103	Naturvetenskaper
8	Naturvetenskap	103	
9	Datavetenskap	104	Datavetenskaper
10	Civilingenjörsutbildning	105	Teknik
11	Annan teknisk utbildning	105	
12	Läroarbetsutbildning	106	Läroarbetsutbildning
13	Läkare/tandläkare/apotekarutbildning	107	Medicin/Vård
14	Vårdutbildning	107	

Antal öppna svar uppgick till 272. Med stöd av Svensk utbildningsnomenklatur (SUN 2000) har de öppna svaren sedan kodats. Utöver koderna 101-107 ovan finns 108, ”Övrig ämnesinriktning” och 109 ”Personal”. I de relativt få fall det inte varit möjligt att koppla samman ett öppet svar med SUN har den ursprungliga stratumkoden fått gälla, liksom i de två fall där ifyllda uppgifter saknades.

Även inom gruppen ”Personal” har ett tjugotal omkodningar gjorts och som bygger på respondenternas egna uppgifter.

På så sätt har en ny variabel ”Stratumutfall” skapats. Denna har sedan korstabulerats mot den ursprungliga stratumvariabeln:

Tabell 3 Uppföljning av ämnesinriktning. Jämförelse mellan urval och utfall. Antal personer

Ursprungligt stratum	Stratumutfall									Total
	Humaniora	Samhällsvetenskap	Naturvetenskap	Datavetenskap	Teknik	Läro- utbildning	Medicin/Vård	Övriga inriktningar	Personal	
Humaniora och teologi	181	14	3	0	2	10	1	0	2	213
Samhällsvetenskap och juridik	8	150	2	0	2	1	7	0	1	171
Matematik och naturvetenskap	1	14	162	7	5	3	4	3	1	200
Data- och systemvetenskap	6	18	5	190	14	0	0	0	3	236
Teknik	3	8	7	16	239	2	0	2	1	278
Läro- utbildning	4	5	6	0	1	184	0	0	0	200
Medicin/vård	1	10	14	0	2	3	164	5	3	202
Övrig ämnesinriktning	63	28	15	5	9	36	2	25	4	187
Personal	0	0	1	0	1	0	0	0	223	226
Total	267	247	215	218	275	239	178	36	238	1 913

Av tabellen framgår att Övriggruppen kraftigt reducerats genom omkodningen. Framför allt förklaras det av att ett stort antal medie- kommunikationsrelaterade ämnesområden, liksom utbildningar inriktade mot formgivning, design och pedagogik förts över till grupperna Humaniora, Samhällsvetenskap och Läro- utbildning. Kvar i Övriga inriktningar återfinns främst utbildningar inriktade mot tjänster samt vissa SLU-utbildningar inom skogsbruk, trädgård och veterinärmedicin. Även Datavetenskap har tappat en del, främst till Samhällsvetenskap och Teknik.

Följande tabell anger hur stor andel av det faktiska utfallet som sammanfaller med ursprungligt stratum. Genomgående ligger andelarna i intervallet 82 – 92 procent. Extremfallen utgörs av Övriggruppen, 13 procent och Personal, där endast ett mindre antal omkodningar gjorts, 98 procent.

Tabell 4 Uppföljning av ämnesinriktning. Jämförelse mellan urval och utfall. Andelar i procent

Ursprungligt stratum	Stratumutfall									Total
	Humana	Samhälls-	Natur-	Data-	Teknik	Läro-	Medicin/	Övriga	Personal	
	niora	vetenskap	vetenskap	vetenskap		utbildning	Vård	inriktningar		
Humaniora och teologi	85%	7%	1%	0%	1%	5%	0%	0%	0%	100%
Samhällsvetenskap och juridik	5%	87%	1%	0%	1%	1%	4%	0%	1%	100%
Matematik och naturvetenskap	1%	7%	81%	4%	3%	2%	1%	1%	0%	100%
Data- och systemvetenskap	2%	8%	2%	81%	6%	0%	0%	0%	1%	100%
Teknik	1%	3%	2%	6%	86%	1%	0%	1%	0%	100%
Läroutbildning	2%	2%	3%	0%	1%	92%	0%	0%	0%	100%
Medicin/Vård	0%	5%	7%	0%	1%	2%	81%	2%	2%	100%
Övrig ämnesinriktning	34%	15%	8%	3%	5%	19%	1%	13%	2%	100%
Personal	0%	0%	1%	0%	1%	0%	0%	0%	98%	100%

Den viktigaste praktiska konsekvensen av uppföljningen är att gruppen ”Övriga inriktningar” som en följd av att den blivit så starkt decimerad inte kommer att ingå i den vidare resultatredovisningen.

En fråga att ställa i det sammanhanget är om det över huvud taget är motiverat att låta Övriggruppen ingå i urvalet? Bortsett från den iögonfallande höga felprocenten i urvalet handlar det om 400 enkäter till en kategori där utfallet på enkätfrågorna inte är möjligt att uttolka.

Det som ändå motiverar att behålla Övriga inriktningar är det önskvärda att kunna göra jämförelser med tidigare undersökningar. Försvinner Övriggruppen påverkar detta studentpopulationen totalt, något som får betydelse för respektive andelar för kvinnor och män, andelarna för olika ålderskategorier osv.

En annan fråga är vilken betydelse omkodningen får på möjligheten att göra korrekta punktskattningar av olika variabler. Omkodningen kan innebära att en person flyttas från en urvalsgrupp där han eller hon har hög vikt till en urvalsgrupp med låga vikter. Svaret på frågan är att ingenting tyder på att felen i urvalen är systematiska, dvs. felen är slumpmässiga, bör de felaktiga vikterna jämnas ut varandra.

Enkätfrågorna

Utformningen av enkätfrågorna har skett i samarbete med svenska och utländska forskare, detta som ett led i att försäkra sig om frågornas relevans givet aktuella frågeställningar på den internationella forskningsfronten.

Undersökningsområde	Forskare
Generellt om tillgång och användning av Internet	Olle Findahl, World Internet Institute
Öppen källkod	Cristina Rossi, Politecnico di Milano
Eget företagande	Lars Frederiksen, Linus Dahlander, Imperial College, Londond
Öppna läresurser	Jan Hylén, Metamatrix
Mobiltelefon som gränssnitt	Tove Jaensson, SICS/Mobile Life

Enkätstrukturen utformades i samarbete med Harald Theorin, SCB. En preliminär version av enkäten granskades av en särskild expertgrupp inom SCB och med ett stort antal förslag till förbättringar som resultat. Med några få undantag justerades enkäten enligt rekommendationerna.

Enkäten översattes till engelska av Jennifer Sundberg som också medverkade i 2003 års undersökning. En delgrupp, 257 utländska medborgare, av urvalet fick både en svensk och en engelsk enkät vid första utsändning och vid enkätpåminnelser. 101 av dem svarade på enkäten, varav 46 på den engelska varianten.

Med så pass få respondenter av den engelska versionen kan det naturligtvis ifrågasättas om översättningen är motiverad. För att sprida kunskap om undersökningen bland utländska forskare har den engelska version dock visat sig vara av avgörande betydelse. Som ett resultat har ett antal propåer om forskarsamarbete kommit in.

Resultatet från insamlingen har redovisats i form av avidentifierade mikrodata för bearbetning, samt en teknisk rapport.

Redovisning av statistiska mått

De statistiska storheterna (parametrarna) utgörs t.ex. av totaler och kvoter. De senare består vanligtvis av procentuella andelar. Statistiken kan redovisas för hela målpopulationen, men mer information fås genom att dela upp den på olika redovisningsgrupper.

Erfarenheterna från 2003 års undersökning pekade på fyra faktorer med betydelse för observerade variationer:³

- *Kön*, där män genomgående utnyttjar Internet i större utsträckning än kvinnor
- *Ålder*, där utnyttjandegraden är högst i yngre åldrar för att sedan successivt avta.
- *Ämnesområden*, där Internet används mer bland studenter inom samhälls-, natur- och datavetenskaper samt tekniska utbildningar, än inom humaniora, lärar- och medicin/vårdutbildningar.

³ Selg, Håkan, Internet i den svenska högskolan våren 2003. Användarstudier inom SUNET. Rapport 2003-10-29. <http://basun.sunet.se/anvandare.pdf>

- *Fildelning*, där tunga fildelare genomgående utnyttjar Internet i större utsträckning och för mer kvalificerade ändamål än måttliga fildelare vilka i sin tur uppvisar mer avancerade användarmönster än dem som inte fildelar alls.

Utfallet från 2003 års undersökning visade också att användarmönstren bland den högskolans anställda personal ganska väl överensstämde med vad som kunnat iakttas bland svenskarna i allmänhet.⁴ Därför kommer även en redovisning att göras med avseende på personalkategori, där lärare/forskare ges rollen som kontrollgrupp. I tabellbilagorna redovisas därför utfallet för enkätfrågorna med beräkning av svarsandel på följande kategorier:

Personalkategori:	Studenter/doktorander Lärare/forskare	Ämnesområde:	Humaniora Samhällsvetenskap Naturvetenskap Datavetenskap Teknik Läroutbildning Medicin/Vård
Kön:	Kvinnor Män	Fildelarkategori:	Tunga fildelare Måttliga fildelare Före detta fildelare Ej fildelare
Åldersgrupp:	– 25 år 26 – 30 år 31 – 35 år 36 – 45 år 46 – 55 år > 55 år		

Svarsandelarna beräknas genom korstabulering. Som signifikansmått används chitvå-test (χ^2), eller där så är möjligt, Fishers exakta test⁵. Signifikansgränsen är satt vid $p < 0.050$. Det ska utläsas som att sannolikheten att de uppmätta skillnaderna mellan olika kategorier, t.ex. kvinnor och män, skulle vara felaktig, är mindre än 5 procent. För de fall $p \geq 0.050$ anges endast p-värdet, men inte svarsandelarna. Vidare anges antalet svar i urvalet som beräkningen är baserad på, t.ex. $n=1913$.

⁴ Jämförelse gjordes med data från World Internet Institute, www.wii.se

⁵ 2x2-tabeller