

# Nyttan av search trails

En studie om novisers nytta av en expertgenererad search trail  
för en uppgift av hög komplexitet

Adam Gillström & Stella Millwood

Institutionen för data-  
och systemvetenskap

Examensarbete 15 hp  
Data- och systemvetenskap  
Vårterminen 2021  
Handledare: Preben Hansen  
English title: The Benefit of Search Trails





# Sammanfattning

Denna studie ämnar undersöka vilken effekt en experts (hög domänkunskap) search trail har på novisers (låg domänkunskap) nytta för en uppgift av hög komplexitet. En search trail är en visuell representation av tidigare användares sökinmatningar. En uppgift är av hög komplexitet när den behandlar abstrakta koncept. Tidigare forskning har behandlat uppgiftskomplexitet och expert-novis-relationen vid sökning, men nyttan är understuderad. I denna studie betraktas nyttta vara återanvändning av expertbegrepp. För att angripa frågeställningen ombeds en expert lösa en sökuppgift av hög komplexitet och generera en search trail. Därefter identifieras hen expertbegrepp. Noviser ombeds sedan lösa samma sökuppgift och ta del av expertens search trail. Studien använder sig av forskningsstrategin mixed methods och tar således både ett kvantitativt och kvalitativt perspektiv. Datainsamlingsmetoderna är enkäter och loggfiler, och dessa analyserades med search log analysis, diagram, och en tematisk analys. Search log analysis och diagrammen visar att samtliga noviser återanvände expertbegrepp, men det vanligaste begreppet avsåg inte den komplexa delen av uppgiften. Den tematiska analysen visar att noviserna upplevde att expertens search trail var hjälpsam, men de hade svårigheter för källkritik, och behovet av expertens search trail minskade under sökprocessen. Resultaten visar även att ett antal noviser undviker begrepp de inte förstår, men upplever att de begrepp de använder är hjälpsamma. Slutsatsen är att noviser drar nytta av en experts search trail, men de är osäkra på sina lösningar, och de är inte nödvändigtvis i behov av den under hela sökprocessen. Det är dessutom oklart huruvida expertens search trail är av nytta för uppgifter av hög komplexitet specifikt. Studien uppmuntrar till framtida forskning kring dessa ämnen.

## Nyckelord

search trail, domänkunskap, uppgiftskomplexitet, informationssökning

# Abstract

This study aims to examine what effect an expert's (high domain knowledge) search trail has on novices' (low domain knowledge) benefit for a task of high complexity. A search trail is a visual representation of previous users' queries. A task of high complexity involves abstract concepts. Previous studies have examined task complexity and the expert-novice relationship during search. However, the benefit has not yet been studied extensively. For this study the concept of benefit is measured as reusage of expert terms. In order to answer the research question, an expert is asked to try to solve a search task of high complexity and generate a search trail. Thereafter, the expert identifies terms used in the search trail that are considered to require domain expertise. Novices are then asked to try to solve the same task with the search trail available to them. The study employs a mixed methods approach and uses both quantitative and qualitative data. Methods for data collection include questionnaires and log files. These are analysed using search log analysis, charts and a thematic analysis. The search log analysis and charts shows that all novices reused expert terms, but the most commonly reused term did not refer to the complex part of the task. The thematic analysis demonstrates that the expert's search trail was helpful to the novices, but they experienced difficulties regarding source criticism, and the need for the search trail diminished during the search process. The results also demonstrate that some novices avoided expert terms, but felt the ones they used were helpful. The conclusion is drawn that novices do experience benefit from using an expert's search trail, but they are unsure of their solutions to the search task and they may not need it during the entire search process. It is furthermore unclear whether the search trail was beneficial for a task of high complexity specifically. Further research is needed to explore these issues.

## Keywords

search trail, domain knowledge, task complexity, information seeking

# Synopsis

Bakgrund	I takt med att sökmotorer blivit ett givet verktyg för informationssökning online, så har intresset för deras optimering ökat. Tidigare forskning om search trails har antagit olika perspektiv. Forskning om uppgiftskomplexitet har konstaterat att det är dimensioner (abstrakta koncept) som gör uppgifter komplexa. Dimensioner behandlar abstrakta koncept som inte nödvändigtvis har entydiga svar. De som forskat om expert-novis-relationen har konstaterat att noviser har en positiv inställning till att ta del av experters sökinmatningar. Flera forskare har intresserat sig för denna relation då experter besitter djup kunskap som skulle kunna förmedlas till noviser. Nyttan i sökning har tidigare definierats som återanvändning av begrepp.
Problem	Tidigare forskning har fokuserat på bland annat uppgiftskomplexitet och expert-novis-relationen vid sökning. Det saknas dock omfattande forskning om vilken nytta en search trail kan förse.
Forskningsfråga	Forskningsfrågan är: ”Vad har en search trail genererad av en expert (hög domänkunskap) för effekt på novisers (låg domänkunskap) nytta för uppgifter på hög komplexitetsnivå?”. Denna frågeställning är av intresse att studera eftersom den ämnar fylla en kunskaplucka i tidigare forskning som specificeras i problemet.
Metod	Studien använder mixed methods som forskningsstrategi och behandlar både kvantitativ och kvalitativ data. Datainsamlingsmetoderna är enkäter och loggfiler. Enkäterna analyseras med diagram och en tematisk analys, och loggfilerna analyseras med search log analysis. Deltagare svarar på enkäterna via webben, och loggfilerna genereras utifrån datan i ett open source sökmotorsgränssnitt. Samtliga analyser genomförs via kalkylark. Studien är förankrad i

	vetenskapliga källor och inkluderar reflektioner kring validitet och reliabilitet.
Resultat	Resultaten visar att samtliga noviser återanvände expertbegrepp, men det vanligaste begreppet avsåg inte den komplexa delen av sökuppgiften. Noviserna upplevde att expertens search trail var hjälpsam, men behovet av den minskade under sökprocessen, och noviserna upplevde svårigheter med källkritik. Svaret på frågeställningen är att noviser drar nytta av en experts search trail, men det går inte att konstatera att den är hjälpsam för uppgifter av hög komplexitet specifikt. Nyttan kan även ifrågasättas då noviser inte är säkra på korrektheten kring deras lösningar.
Diskussion	Diskussionen bekräftar delvis tidigare forskning, men även nya koncept introduceras. De viktigaste diskussionspunkterna behandlar huruvida experters search trails är hjälpsamma för uppgifter av hög komplexitet specifikt, hur novisers skepticism kring källor och korrekthet ska hanteras, samt varför behovet av experters search trail minskar under sökprocessen. Studiens begränsningar involverar bortfallet av deltagare och filtrering respektive inkludering av expertbegrepp i loggfilerna. Etiska konsekvenser kan inte identifieras, men studien kan ha samhälleliga konsekvenserna då informationssökning via sökmotorer förekommer dagligen på exempelvis arbetsplatser och inom utbildning. Studien är av nyttा till forskare som studerat olika aspekter av search trails, men inte undersökt dess nyttा.

# Tack

Vi vill rikta ett stort tack till vår handledare Preben Hansen. Preben har varit ett betydelsefullt stöd i examensarbetet, och givit utmärkt handledning. Skaparna av verktyget SearchX har även varit till stor hjälp under arbetet då de via e-postkontakt med oss genomfört förbättringar till systemet. Vi vill även tacka deltagarna som ställt upp i studien och bidragit till insikterna i arbetet. Slutligen riktar vi tack till studenter och lärare som granskat uppsatsen och försett värdefull återkoppling.

# Innehåll

1	Inledning .....	1
1.1	Bakgrund .....	1
1.2	Problemområde .....	1
1.3	Frågeställning .....	2
2	Utökad bakgrund.....	3
2.1	Information seeking .....	3
2.2	Search trails .....	4
2.3	Uppgiftskomplexitet och sökprocessen.....	5
2.4	Relationen mellan experter och noviser.....	6
2.5	Nyttan av search trails .....	7
3	Metod .....	8
3.1	Metodval .....	8
3.1.1	Forskningsstrategi .....	8
3.1.1.1	Mixed methods.....	8
3.1.2	Alternativ forskningsstrategi .....	8
3.1.2.1	Experiment .....	8
3.1.3	Datainsamlingsmetoder.....	9
3.1.3.1	Dokument.....	9
3.1.3.2	Enkät .....	9
3.1.4	Alternativa datainsamlingsmetoder.....	10
3.1.4.1	Intervju .....	10
3.1.4.2	Observation .....	10
3.1.5	Dataanalysmetoder.....	11
3.1.5.1	Search Log Analysis .....	11
3.1.5.2	Diagram.....	11
3.1.5.3	Tematisk analys.....	11
3.1.6	Alternativa dataanalysmetoder.....	12
3.1.6.1	Innehållsanalys.....	12
3.1.6.2	Statistiska test.....	12
3.2	Metodtillämpning .....	13
3.2.1	Urval.....	13
3.2.2	Utförande.....	13
3.2.3	Dataanalys .....	14
3.2.3.1	Search Log Analysis .....	14
3.2.3.2	Diagram.....	14
3.2.3.3	Tematisk analys.....	14
3.3	Forskningsetiska aspekter .....	15
4	Resultat och analys.....	16

4.1	Enkäten som kartlägger noviserna .....	16
4.2	Loggfilerna.....	20
4.3	Enkäten om upplevd nytta.....	22
4.3.1	Kvantitativa resultat för nytta.....	22
4.3.2	Teman från fritextsvaren.....	26
4.3.2.1	Samband mellan låg förståelse för sökuppgiften och låg domänkunskap .....	26
4.3.2.2	Expertens search trail var hjälpsam, men noviser upplevde svårigheter med källkritik .....	26
4.3.2.3	Noviser upplevde att search trailen vägledde dem i början av sökprocessen, och efter ökad domänkunskap minskade behovet av att använda den.....	27
4.3.2.4	Noviser undvek begrepp de inte förstod, men upplevde att begreppen de använde var hjälpsamma .....	28
5	Diskussion .....	30
5.1	Dimensioner .....	30
5.2	Minskat behov av search trailen under sökprocessen .....	30
5.3	Svårigheter med källkritik.....	31
5.4	Noviser undvek begrepp de inte förstod .....	32
5.5	Begränsningar .....	32
5.6	Etiska och samhälleliga konsekvenser .....	32
5.7	Validitet.....	33
5.8	Reliabilitet .....	34
6	Slutsats .....	35
	Referenser .....	36
	Bilaga A – Medgivandeavtal för experten (SearchX) .....	38
	Bilaga B – Medgivandeavtal för noviser (SearchX) .....	40
	Bilaga C – Medgivandeavtal för experten (E-postkontakt) .....	42
	Bilaga D – Medgivandeavtal för den första enkäten (noviser) .....	43
	Bilaga E – Medgivandeavtal för den andra enkäten (noviser) .....	44
	Bilaga F – Kategorier och teman.....	45
	Bilaga G – Expertens search trail .....	46
	Bilaga H – Gränssnittet i SearchX .....	47
	Bilaga I – Den första enkäten.....	48
	Bilaga J – Den andra enkäten.....	49

# Figurer

Figur 1: Information-seeking-faktorer.....	5
Figur 2: Kuhlthaus modell för informationssökningsprocessen.....	6
Figur 3: Svaren grupp A angav på frågan ”Vilka av följande källor brukar du använda när du söker information för en skoluppgift?”.....	17
Figur 4: Svaren grupp B angav på frågan ”Vilka av följande källor brukar du använda när du söker information för en skoluppgift?”.....	17
Figur 5: Svaren grupp A angav på frågan ”Vad gör att du litar på informationen?”.....	18
Figur 6: Svaren Grupp B angav på frågan ”Vad gör att du litar på informationen?”.....	18
Figur 7: Svaren grupp A angav på frågan ”Hur länge söker du information inom domänen bild och design, via en internetbaserad sökmotor?”.....	19
Figur 8: Svaren grupp B angav på frågan ”Hur länge söker du information inom domänen bild och design, via en internetbaserad sökmotor?”.....	19
Figur 9: Svaren grupp A angav på frågan ”När du söker information för en skoluppgift: hur många källor i genomsnitt kollar du på innan du är nöjd?”.....	20
Figur 10: Svaren grupp B angav på frågan ”När du söker information för en skoluppgift: hur många källor i genomsnitt kollar du på innan du är nöjd?”.....	20
Figur 11: En procentuell representation av andel sökinmatningar som innehöll respektive inte innehöll expertbegrepp, i grupp A.....	21
Figur 12: En procentuell representation av andel sökinmatningar som innehöll respektive inte innehöll expertbegrepp, i grupp B.....	21
Figur 13: En representation av expertbegreppen som användes i deltagarnas sökinmatningar, och dess frekvens, i grupp A.....	22
Figur 14: En representation av expertbegreppen som användes i deltagarnas sökinmatningar, och dess frekvens, i grupp B.....	22
Figur 15: Diagrammet visar hur svår deltagarna upplevde att sökuppgiften var att lösa. Svarsalternativen var 1-5.....	23
Figur 16: Diagrammet visar hur säkra deltagarna var på att de lyckades lösa sökuppgiften. Svarsalternativen var 1-5.....	23
Figur 17: Diagrammet visar hur hjälpsam deltagarna upplevde att expertens search trail var. Svarsalternativen var 1-5.....	24
Figur 18: Diagrammet visar andelen deltagare som svarade att det fanns respektive inte fanns expertbegrepp som de inte förstod.....	25
Figur 19: Diagrammet visar under vilken del av sökprocessen som deltagarna drog mest respektive minst nytta av expertens search trail.....	25

# Förkortningar

IS:	Information seeking
IR:	Information retrieval
SLA:	Search log analysis
VPS:	Virtual private server
HTTP:	Hypertext transfer protocol
API:	Application programming interface
JSON:	JavaScript Object Notation
NoSQL:	Non-relational database

# 1 Inledning

## 1.1 Bakgrund

Informationsvetenskap (eng. *Information science*) har historiskt sett tagit an två olika perspektiv. Ursprungligen utvecklades området *information retrieval* (häданefter kallad IR), men detta grenades senare ut till att inkludera *information seeking* (häданefter kallad IS). IR innefattar ett algoritmiskt perspektiv där matematiska modeller används för att generera sökresultat (Hiemstra 2009). IS härstammar istället från det samhällsvetenskapliga området och utgår från användarens behov och förutsättningar (Ingwersen & Järvelin 2005; Case 2007). I kontexten av sökmotorer så har flera verktyg introducerats för att hjälpa användare i informationssökningsprocessen, däribland search trails. Verktyget utvecklades ursprungligen från IR, men ett IS-perspektivet har senare applicerats. Search trails är en viktig komponent i moderna sökmotorer och visar användare sökinmatningar från tidigare användare som försökt lösa liknande sökuppgifter (Capra & Arguello 2019). Funktionen är tänkt kunna underlätta navigering från problem till lösning.

En studie av Capra och Arguello (2019) visade att användares behov av stöd ökade när sökuppgifterna var av hög komplexitet, men att de då hade mindre användning av de search trails som var tillgängliga för dem. Denna paradox utforskade sedan Wang, Zhao och Chen (2020) vidare genom att dela in användarna i noviser (med låg domänkunskap) och experter (med hög domänkunskap). Deras studie visade att noviser upplever att de har användning av experters search trails då det hjälper dem utveckla sina sökstrategier (ibid.).

Vidare menar Shah och Marchionini (2009) att användare tenderar att lösa sökuppgifter på liknande sätt, och att de använder sig av liknande sökinmatningar. Bagley och Levkowitz (2011) skriver att återanvändning av framgångsrika sökinmatningar leder till förbättrad förmåga att söka information och lösa sökuppgifter. För att återkoppla till expert-novis-relationen menar White, Dumais & Teevan (2009) att experter tenderar att lösa sökuppgifter i högre grad än noviser. Under sökprocessen ökar även novisers domänspecifika vokabulär. Författarna menar att en framtida studie skulle kunna studera hur experters sökstrategier kan stödja noviser (ibid.).

## 1.2 Problemområde

Under de senaste decennierna har stora framsteg skett inom IR, och de kommersiella webbsökmotorerna har blivit allt mer effektiva när det gäller att returnera relevanta sökresultat utefter användares sökfrågor. Sökmotorerna kan exempelvis automatiskt ge svar på enkla faktabaserade frågor (kallas för *look up questions*) direkt, utan användning av andra externa verktyg. Informationssökning i samband med mer komplexa uppgifter är ändå fortfarande ett underutforskat område, och när komplexiteten av sökfrågorna ökar så förser sökmotorerna sämre resultat för användare. Detta anses bero på att komplexa sökuppgifter inte kan formuleras med en enda sökinmatning eller besvaras med kortfattat information (Huurdeman & Kamps 2020). Tidigare studier har bland annat undersökt vilken effekt search trails har på lösning av uppgifter med olika komplexitet (Capra & Arguello 2019), och huruvida experter kan hjälpa noviser förbättra deras sökförståelse

(Wang, Zhao & Chen 2020), men det saknas fortfarande omfattande forskning om vilken nytta som kan utvinnas i dessa kontexter.

## **1.3 Frågeställning**

Studiens frågeställning är: ”Vad har en search trail genererad av en expert (hög domänkunskap) för effekt på novisers (låg domänkunskap) nytta för uppgifter på hög komplexitetsnivå?”. Denna frågeställning bygger på problemet då syftet är att undersöka kunskapsluckan i tidigare forskning. Tidigare forskning om uppgiftskomplexitet, expert-novis-relationen, och nytta i sökning inkluderas i frågeställningen, men den kombinerar koncepten för att vidareutveckla hur de påverkar varandra. Konceptet nytta avgränsas till att behandla återanvändning av expertbegrepp, och expert-novis-relationen definieras utifrån domänkunskap. Andra perspektiv på nytta studeras inte, och deltagarnas sökförmåga påverkar inte deras status som expert eller novis i studien. Vidare analyseras enbart noviserna sökinmatningar i loggfiler. Exempelvis klick, besökta webbplatser, och bokmärken analyseras inte.

# 2 Utökad bakgrund

## 2.1 Information seeking

Forskning inom informationsvetenskap innefattar två olika områden, IS och IR. IR har sitt ursprung i data- och systemvetenskap och behandlar användandet av algoritmer och matematiska modeller för att generera sökresultat, baserat på användares sökinmatningar (Ingwersen & Järvelin 2005). Detta uppnås genom att tillämpa matematiska modeller av information på sökinmatningen, såsom mängdteori, den booleska modellen och vektorrumsmodellen, i syfte att finna de sökresultat som i högsta grad liknar sökinmatningen. De ingående orden i sökfrågan och sökresultatet betraktas som matematiska objekt och likheten mellan dem beräknas som ett nummer. Sökresultaten sorteras sedan utefter deras likhetsvärde (Hiemstra 2009). Kritik som riktats mot IR är att det saknas ett användarcentrerat perspektiv. Mer specifikt, menar Ingwersen och Järvelin, tar IR inte hänsyn till hur användare söker information och använder sökverktyg. IR-system utvecklas med artificiella sökuppgifter, vilket inte är verklighetsenligt för användningskontexten. En användares individuella situation, miljö, eller motiv kan påverka sökningen, och det perspektivet ingår inte i IR (Ingwersen & Järvelin 2005).

IS har sitt ursprung i det samhällsvetenskapliga området (Ingwersen & Järvelin 2005). Användningskontexten är då utgångspunkten för att anpassa informationssökningsprocessen efter användarens förutsättningar och behov. Behovet grundar sig i användarens informationsbrist. Användarens förmåga att upptäcka mönster och åtgärda av brister i redan upptäckta mönster är exempel på ämnen som IS innefattar (Case 2007). Marchionini (1996) skriver att IS visserligen drivs av utveckling utifrån mänskliga behov, men han menar även att behov är ett medel för människor att interagera med deras omgivning. Det är då är väsentligt att organisera och filtrera information så att användare kan utesluta det som är inkorrekt eller av låg kvalité (ibid.).

IS kan vidare delas in i sex olika faktorer som relaterar till varandra, se figur 1 (Marchionini 1996, s. 33). Informationssökaren (eng. *information seeker*) är den aktören som har ett problem eller behov och från det aktiverar mentala bilder eller minnen. Dessa mentala bilder eller minnen utvecklas sedan till definierade problem som lägger grund för en uppgift (eng. *task*). Söksystemet (eng. *search system*) är källan till informationen. Domäner innefattas av olika informationsområden, som till exempel medicin eller lingvistik. Miljön (eng. *setting*) är den situationella och fysiska kontexten. Sökresultaten (eng. *Search outcomes*) är den återkopplingen från systemet, till exempel bilder eller länkar (ibid.).

Inom IS finns det även en diskurs om två olika kommunikationstyper: *asynkron* och *synkron*. I synkron kommunikation interagerar två processer med varandra i realtid, och i asynkron kommunikation är dessa processer tidsmässigt oberoende. Kommunikation som sker asynkront innebär att användare söker information vid olika tider. De kan fortfarande influera varandra, men kommunikationen sker tidsmässigt obundet (Charron-Bost, Mattern & Tel 1996). En search trail är för en användare en interaktiv visuell representation av hur en tidigare användare försökt lösa samma typ av sökuppgift. Informationen som en search trail visualiseras kan inkludera andra användares tidigare sökningar, och sidor som de besökt eller bokmärkt, vilket kan vara till nytta för framtida användares förmåga att söka information (Capra & Arguello 2019). Detta är ett exempel på asynkron kommunikation då sökinmatningarna genererats av användare under olika tidpunkter. När search trails

innefattar kommunikation mellan sökande användare och tidigare sökande användare, så klassas verktyget inom IS som asynkron kommunikation (Charron-Bost, Mattern & Tel 1996).

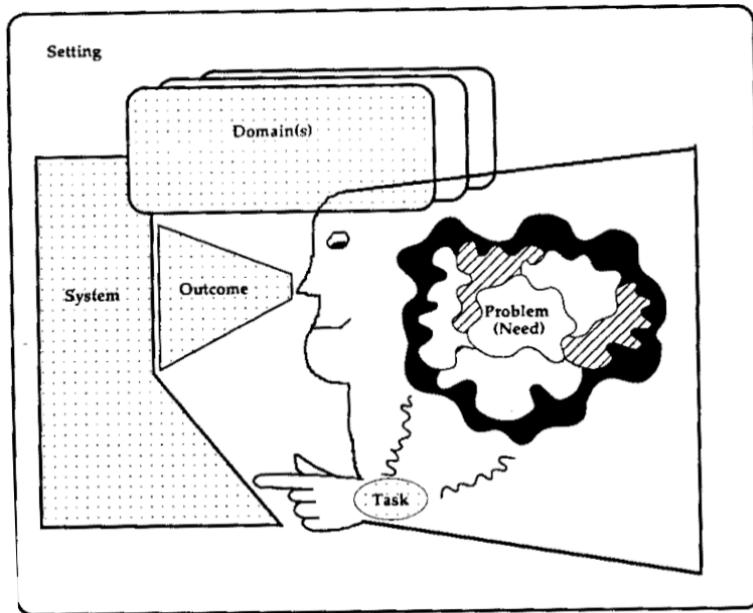
## 2.2 Search trails

Search trails kan positioneras inom både IR och IS beroende på perspektiv på informationssökning. Tidigt i forskningen om informationssökning var perspektivet på search trails IR-fokuserat, och algoritmer utvecklades som behandlade sökfrågor där relevant information från en större informationssamling hämtades. En av de första algoritmerna som användes för detta syfte utvecklades under 1950-talet, känd som den booleska modellen, där en sökfråga betraktades som en logisk sammansättning nyckelord, som användes för att filtrera bort irrelevanta sökresultat. Den booleska modellen lämpade sig väl till att filtrera sökresultat baserat på en sökfråga men saknade möjligheten till att rangordna sökresultaten utefter dess likhet till sökfrågan (Sanderson & Croft 2012). Under 1960-talet utvecklades vektorrumsmodellen som användes för att tilldela likheten mellan en sökinmatning och potentiella sökresultat ett numeriskt värde. Detta numeriska värde användes sedan för att sortera sökresultaten efter relevans (*ibid.*). Samtidigt som vektorrumsmodellen utvecklades vidare för att bli mer precis, började nya perspektiv på search trails växa fram inom IS.

Under mitten av 1980-talet grenades forskningen kring search trails ut till att även innefatta ett IS-perspektiv, och forskningsfokuset skiftade från ett algoritmiskt perspektiv till ett användarcentrerat perspektiv. Under perioden publicerades en artikel av Dervin och Nilan som presenterade ett nytt angrepssätt, som placerade individuella användare och deras behov i fokus. Artikeln blev mycket inflytelserik, och efter dess publicering introducerades fler användarcentrerade tillvägagångssätt och modeller. De fokuserade på informationssökarens kognitiva och mentala processer. Dervin's Sense-Making, Ellis beteendemodell för informationssökande strategier och Kuhlthaus modell för informationssökningsprocessen är exempel på några av de mest centrala utvecklingarna i forskningsområdet (Ingwersen & Järvelin 2005). Kuhlthaus modell för informationssökningsprocessen presenteras i ett senare avsnitt.

För search trails innebar skiften till ett användarcentrerat perspektiv inom IS att användares interaktioner blev det centrala fokuset, snarare än den data de genererade (det algoritmiska perspektivet). Search trails utforskades då som ett potentiellt medel för asynkron kommunikation mellan användare. Tre centrala modeller utvecklades inom IS för studering av search trails: *berrypicking*, *orienteering* och *information foraging*. Modellen information foraging är relevant för bakgrunden till denna studie, då den beskriver *hur* användare använder andra informationssökares tidigare sökinmatningar för att hitta information (Wang, Zhao & Chen 2020). Information foraging bygger på att människan använder sig av särskilda inbyggda mekanismer som stödjer informationssökande. En av dessa mekanismer är till exempel förmågan att söka föda, och människan söker information för att hitta föda. Dessa mekanismer tar utgångspunkt i *optimal foraging theory*, som ursprungligen behandlade hur djur söker efter föda. I information foraging kallas informationssökningsprocessen för *foraging*, alltså eftersökning, och den behandlar hur människan följer en *information scent* för att lösa sin sökuppgift. En information scent avser sambandet mellan informationen som användare tar del av och hur de löser sina sökuppgifter (Pirolli & Card 1999). Denna information används för att beskriva hur en informationssökare tar hjälp av tidigare användares search trails, för att hitta information och lösa sökuppgiften.

## 2.3 Uppgiftskomplexitet och sökprocessen



Figur 1: Information-seeking-faktorer (Marchionini 1996, s. 33).

Tidigare beskrivs användningskontextens sex olika faktorer inom IS. En av dessa faktorer är sökuppgiften som, för denna studie, är särskilt relevant. Informationssökaren har ett behov eller problem, som för att kunna lösas behöver omformuleras till en uppgift (Marchionini 1996). Tidigare studier har undersökt olika aspekter av uppgifters egenskaper, däribland komplexitet. Wildemuth, Freund och Toms (2014) menar att det länge saknats en tydlig definition av komplexitet för uppgifter, och föreslår i sin studie att komplexitet påverkas av tre faktorer (författarna kallar dessa för dimensioner, men eftersom detta begrepp återanvänds med olika innebördar i andra källor så döps de här till faktorer): flera underuppgifter eller steg, flera aspekter, och obestämbarhet (eng. *indeterminability*) (ibid.). Byström och Järvelin (1995) har tidigare utforskat vilken inverkan uppgifters bestämbarhet (eng. *determinability*) kan ha för informationssökningsprocessen. De definierar komplexa uppgifter som osäkerhet kring till exempel vilka processer som är involverade och vilken typ av svar som eftersöks. Enligt författarna avgör bestämbarheten en uppgifts komplexitet. De skriver att enkla uppgifter är de som sker enligt rutin där alla processer, inmatningar och slutresultat är kända. Uppgifterna är då, enligt Byström och Järvelin, bestämbara. Vad gäller komplexa uppgifter är det, enligt författarna, tvärtom (ibid.).

Capra, Arguello, O'Brien, Li och Choi (2018) genomförde en studie där de utgick från att väldefinierade uppgifter skulle uppfattas som mer bestämbara, och mindre komplexa. Resultatet visade dock att typen av innehåll i uppgifterna påverkade bestämbarheten, och därmed komplexiteten (ibid.). Uppgifter som inkluderade *items* uppfattades som mer tydliga (mindre komplexa) än uppgifter med *dimensions*. Items är eller relaterar vanligen till konkreta subjektiv (till exempel antal eller typer), medan dimensioner behandlar abstrakta koncept (till exempel pris eller hållbarhet) (Capra & Arguello 2019). Capra och Arguello (ibid.) genomförde därefter en studie som vidareutforskade detta. Studiens delfråga handlade om hur uppgifters komplexitet kan influera användarens förmåga att utvinna kunskap från search trails. När uppgifterna involverade items så minskade komplexiteten, och när uppgifterna involverade dimensioner så ökade komplexiteten. Paradoxalt så ökade användarnas behov av stöd för uppgifter av hög komplexitet samtidigt som de upplevde att de då hade mindre användning av de tillgängliga search trails (ibid.).

Stages in ISP	Feelings Common to Each Stage	Thoughts Common to Each Stage	Actions Common to Each Stage	Appropriate Task According to Kuhlthau Model
1. Initiation	Uncertainty	General/Vague	Seeking Background Information	Recognize
2. Selection	Optimism			Identify
3. Exploration	Confusion/Frustration/Doubt		Seeking Relevant Information	Investigate
4. Formulation	Clarity	Narrowed/Clearer		Formulate
5. Collection	Sense of Direction/Confidence	Increased Interest	Seeking Relevant or Focused Information	Gather
6. Presentation	Relief/Satisfaction or Disappointment	Clearer or Focused		Complete

Figur 2: Kuhlthaus modell för informationssökningsprocessen (Kuhlthau 1991, s. 367).

Enligt Kuhlthaus modell går användare igenom sex faser när de söker information: uppgiftsinitiering, ämnesselektion, utforskning, fokusformulering, informationsinsamling och sökavslut. Figur 2 visar Kuhlthaus modell för informationssökning (Kuhlthau 1991, s. 367). I den första fasen blir användaren medveten om den information hen saknar för att lösa problemet eller utföra uppgiften. I den andra fasen bestämmer sig användaren för vilken domän som ska analyseras. Den tredje fasen innefattar användarens utforskning av information, och är den fasen som kan medföra osäkerhet och tvivel bland användare. I den fjärde fasen börjar användaren förstå informationen, och i den femte fasen börjar användaren samla in den information som är viktig för att lösa problemet, och i den sjätte fasen avslutar användaren sin sökning och löser problemet. Under dessa faser går användaren igenom olika känslor. Både typerna av känslor kan ändras (till exempel kan användaren till en början uppleva osäkerhet som sedan kan övergå till säkerhet), men även den grad som användaren upplever dem kan ändras (till exempel kan känslan av osäkerhet öka från en fas till en annan) (ibid.).

## 2.4 Relationen mellan experter och noviser

Kuhlthaus modell utgör ett intressant ramverk för det Isenberg tidigare hade identifierat i sin forskning. Isenbergs studerade skillnaderna mellan experter och noviser i kontexten av affärslösningar. Han kunde konstatera att sättet som en uppgift uppfattas påverkas av användarens kunskap, som i sin tur påverkar användarens behov av ytterligare information. Hans studie visade att experter hittar fokusområden fortare än noviser och, för att relatera till Kuhlthaus modell, upplever mindre ångest i processen. Enligt Isenberg har en expert lärt sig arbeta med information som är begränsad och svårtillgänglig, men ändå värdefull, och kan i den situationen utvinna all möjlig information. Samtidigt visade hans studie att de med mer erfarenhet inom affärslösningar hade mindre användning av den tillgängliga informationen än de med mindre erfarenhet (Ingwersen & Järvelin 2005).

Wang, Zhao och Chen (2020) genomförde en studie som vidareutvecklar Capra och Arguellos resonemang avseende den påverkan som uppgifters komplexitet har på användares benägenhet att ta del av search trails. De applicerar detta på expert-novis-relationen, med noviser som uppgav att de

upplevde svårigheter med att hitta rätt sökresultat. Studien visar att noviser har en allmänt positiv inställning till att använda search trails som har utformats av experter, och att de är hjälpsamma för noviser att utveckla sina sökstrategier (*ibid*). Capra och Arguellos (2019) studie visade att användare söker hjälp när de har en komplex uppgift, men att de då samtidigt inte har tillräckligt med användning av de search trails som är tillgängliga. Denna paradox i deras resultat leder till frågan: När är det lämpligt att visa search trails? (*ibid*). Wang, Zhao och Chen (2020) skriver att noviser har användning av search trails när de är utformade av experter. Detta skulle kunna introducera ett svar på Capra och Arguellos fråga.

## 2.5 Nyttan av search trails

En studie utförd av Shah och Marchionini (2009) visade att användare som försöker lösa liknande uppgifter i flera fall använder sig av liknande sökinmatningar. När användare försöker lösa samma sökuppgift, och har samma informationsbehov, så tenderar de att formulera sina informationsbegär på liknande sätt. Författarna skriver att detta kan användas som en drivkraft för att utveckla forskning inom informationssökning (*ibid*.). Vidare skriver Bagley och Levkowitz (2011) att återanvändning av framgångsrika sökinmatningar (det vill säga sökinmatningar som tidigare lett till lösta uppgifter) leder till bätttrad förmåga att söka information och lösa sökuppgifter.

I en studie av Jenkins, Corritore och Wiedenbeck (2003) diskuterar författarna skillnaden i hur användare söker information beroende på deras domän- och webbkunskap (författarna syftar på sökkunskap via internet i denna kontext). De skriver att domänexperter var mer tydliga i vilken typ av information de sökte än domännoviser, och de hade en bättre förståelse för relationen mellan ämnesspecifika begrepp. De skriver även att domänexperter kunde i högre grad än domännoviser förklara varför de ansåg att källor var trovärdiga, och var mer bekväma med att använda källor som för dem var främmande (*ibid*.). Vidare visar en studie av White, Dumais och Teevan (2009) att domänexperterna löste uppgifter i högre grad än domännoviserna, och intressant att notera är att novisernas användning av ett domänspecifikt vokabulär, under studiens gång, ökade. Författarna skriver att studiens resultat kan användas i framtiden för att undersöka hur experters sökstrategier kan stödja noviser i att lära sig mer om vokabulär och resurser (*ibid*.).

White, Dumais och Teevan (2009) menar att experter löser sökuppgifter i hög grad än noviser. Deras sökinmatningar kan således betraktas som mer framgångsrika än novisers. Enligt Bagley och Levkowitz (2011) leder återanvändning av framgångsrika sökinmatningar till bätttrad förmåga att söka information och lösa sökuppgifter. För att återkoppla till Capra och Arguello, så skulle då återanvändning av tidigare framgångsrika sökinmatningar, genererade av experter, kunna vara en utgångspunkt för att vidare utforska problemen kring nyttan av search trails för uppgifter av hög komplexitet, med noviser som informationssökare.

# 3 Metod

## 3.1 Metodval

### 3.1.1 Forskningsstrategi

#### 3.1.1.1 Mixed methods

Frågeställningen behandlar konceptet nyta. För denna studie används både en kvalitativ och kvantitativ ansats för att mäta nyta och således besvara frågeställningen. Den kvantitativa datan kan förse information om vad noviser gör, medan den kvalitativa datan förser kontexten bakom deras agerande - varför de gör så. Frågeställningen implicerar en utforskande studie snarare än en förklarande, och då är kombinationen av kvalitativ och kvantitativ data intressant. Flera av de tidigare studier som belyst bakgrunden till frågeställningen har använt kvantitativa ansatser, och studien ämnar vidare utforska ämnet genom att använda kvalitativ data som komplement. *Mixed methods* är en forskningsstrategi som lämpar sig när en studie är problemdriven och ämnar se på problemet ur olika perspektiv, vanligen med en både kvalitativ och kvantitativ ansats (Denscombe 2014). Detta är en problemdriven studie som grundar sig i en paradox kring nyttan av search trails. Genom att, i enlighet med mixed methods-strategin, angripa problemet från olika perspektiv i form av olika metoder för datainsamling och analys kan studien förse en mer heltäckande bild av nyttan av search trails (*ibid.*). Tidigare studier har inte använt kvalitativa ansatser, och genom att tillämpa en både kvalitativ och kvantitativ ansats på denna studie bör resultatet visa att tidigare forskning bekräftas, men sedan vidareutvecklas med det kvalitativa perspektivet.

Studien genomförs sekventiellt vad gäller datainsamlingsmetoderna. I denna studie undersöks novisers nyta, och novisernas utförande i testet baseras på resultaten från en experts deltagande. Noviserna får först svara på en kvantitativ enkät, sedan försöker de lösa en sökuppgift som dokumenteras kvantitativt i form av loggfiler, och till sist får de svara på en andra enkät som är både kvantitativ och kvalitativ. I en sekventiell studie så samlas kvalitativ och kvantitativ data in i en viss ordning. Ordningen avgör hur datan från olika källor kompletterar varandra. En hierarki måste även definieras kring vilken data som prioriteras (Denscombe 2014). Motivering bakom ordningen i denna studie lyder att den första enkäten förser en kartläggning av användargruppen genom att inkludera frågor om ålder, modersmål och domänkunskap. Den andra datainsamlingen dokumenterar novisers sökbeteenden och fungerar som en referens för den tredje datainsamlingen. Den andra enkäten (tredje datainsamlingen) bygger på den andra genom att inkludera frågor om den upplevda nyttan (kvalitativt) och användarnas upplevda ageranden (kvantitativt). Den första datainsamlingen har lägre prioritet än de andra två eftersom den inte direkt besvarar frågeställningen. Syftet med de andra två är att mer direkt svara på frågeställningen. Den tredje datainsamlingsmetoden prioriteras samtidigt högre än den andra, då den förser en utförlig beskrivning av nyttan. Eftersom det är novisers nyta som frågeställningen behandlar och inte expertens, så prioriteras expertens data lägre än novisernas.

### 3.1.2 Alternativ forskningsstrategi

#### 3.1.2.1 Experiment

Experiment övervägdes som alternativ forskningsstrategi. I ett experiment undersöks den kausala relationen mellan beroende och oberoende variabler. För att isolera variablerna som undersöks så

kontrolleras miljön. Det innebär att faktorer som potentiellt kan påverka den beroende variabeln antingen hålls konstanta, elimineras, eller inte förändras alls. När faktorerna hålls oförändrade så introduceras istället den beroende variabeln i miljön (Denscombe 2014). Frågeställningen för studien behandlar vad en experts search trail har för effekt på novisers nytta (för uppgifter på hög komplexitetsnivå). Effekten hade kunnat vara fokuset för ett experiment. Då hade effekten kunnat mätas i form av den kausala relationen mellan en experts search trail (oberoende variabel) och novisers nytta (oberoende variabel).

Experiment valdes bort som forskningsstrategi eftersom studiens frågeställning är utforskande snarare än förklarande. Därmed lämpar sig inte ett teoretiskt ramverk för att formulera en hypotes. Experiment är en förklarande forskningsstrategi som utgår från teorier för att förklara en relation mellan variabler (Denscombe 2014). Denna studie syftar istället till att besvara frågeställningen genom att belysa olika perspektiv på ämnet, utan förbindelse till tidigare forskning med tillhörande teorier.

### **3.1.3 Datainsamlingsmetoder**

#### **3.1.3.1 Dokument**

Dokument i form av loggfiler används som datainsamlingsmetod. I kontexten av search trails används loggfiler för att samla data om till exempel sökinmatningar, klick, och besökta webbplatser. Loggfiler kan även identifiera och särskilja användare, vilket innebär att de kan visa vilken användare som utför vad. En av fördelarna med loggfiler är datainsamlingen är automatiserad. En forskare kan välja vilken typ av data som loggfilerna dokumenterar, men det är inte forskaren som manuellt dokumenterar datan. För denna del av studien eftersöks konkret rådata om användarnas agerande, som sedan fungerar som komplement till den andra enkäten som förser kontext till rådatan. Detta återkopplar till syftet med att kombinera kvalitativ och kvantitativ data som beskrivs under avsnittet om val av forskningsstrategi. Påverkan av den mänskliga faktorn i datainsamlingsmetoden dokument minimeras (Jansen 2008), till skillnad från i exempelvis observationer.

Ytterligare en fördel med loggfiler är att de lämpar sig för syftet att mäta återanvändningsfrekvens (Jansen 2008) - något som är av betydelse för studiens frågeställning. Vidare möjliggör metoden för insamling av data på distans. Forskare och deltagare behöver inte vara närvarande på samma geografiska plats (ibid.). Eftersom deltagarna i studien inte hade möjlighet att närvvara på samma geografiska plats så var detta fördelaktigt.

#### **3.1.3.2 Enkät**

Innan loggfilerna dokumenterar data (då användare söker och försöker lösa sökuppgiften) så tilldelas noviserna en första enkät. Denna enkät är av kvantitativ karaktär och behandlar enkla frågor om till exempel ålder, språkliga kompetenser och källkritiska förmågor. Syftet med denna enkät är att kartlägga novisgruppen. Frågor om till exempel ålder ställs i syfte att säkerställa att noviserna är myndiga. Frågor om källanvändning och källkritik ställs i syfte att kartlägga gruppens domänkunskap som referens i analysen.

Loggfiler kan samla in entydiga data om användares interaktioner med system, exempelvis klick, och sökinmatningar. Loggfiler förser däremot inte data om *varför* användare utfört dessa handlingar. Därför tilldelas noviserna även en andra enkät som komplement till loggfilerna. De svarar på enkäten efter att loggfilerna dokumenterat data och den består av frågor som berör deltagares upplevda nytta av en experts search trail. Den andra enkäten är både kvantitativ och kvalitativ. De kvantitativa frågorna

efterföljs av kvalitativa frågor. Den upplevda nyttan kan då mätas med tydliga parametrar (de kvantitativa frågorna) till följd av efterfrågad kontext till svaren (kvalitativa frågor, alltså fritext).

Frågeställningen för denna studie behandlar novisers nytta av en experts search trail. Det är inte en specifik grupp i samhället eller på ett företag som undersöks, utan noviser i generell bemärkelse. Daten som eftersöks bör kunna generera slutsatser som kan standardiseras för noviser, och enkäter lämpar sig i detta syfte. Vidare är enkäter lämpliga alternativ när ämnet som studeras inte är kontroversiellt (Denscombe 2014). Eftersom nyttan av search trails inte betraktas vara ett kontroversiellt ämne så bedöms datainsamlingsmetoder (såsom intervjuer) av personlig karaktär inte vara nödvändiga i studien. Förutsättningen för enkäter är dessutom en öppen sociala miljön där respondenter kan vara ärliga i sina svar (ibid.). Eftersom noviserna meddelas att testet är orelaterat till deras utbildning, och att deras lärare inte kommer att få se deras individuella svar så har denna förutsättning för enkäter som datainsamlingsmetod tillämpats. Ytterligare en fördel med enkät som datainsamlingsmetod är att den andra enkäten besvaras direkt efter sökprocessen (och loggfilsdata samlats in). Denscombe (ibid.) skriver att enkäter lämpar sig när respondenter har goda förmågor att minnas händelserna. Detta är av betydelse för denna studie eftersom förväntningen är att det finns en kort tidsram då respondenterna minns deras sökinmatningar. En nackdel med enkäter är att det kan vara utmanande att säkerställa en hög svarsfrekvens (Johannesson & Perjons 2014). För att undvika detta inkluderar instruktionerna förklaringar om när noviserna ska vidare till nästa steg i testet. De förses även i den inledande fasen av testet med en introduktion om vilka delar testet består av.

### **3.1.4 Alternativa datainsamlingsmetoder**

#### **3.1.4.1 Intervju**

Intervjuer fördelaktiga att använda när komplex information, såsom respondenternas känslor och åsikter efterfrågas. Datainsamlingsmetoden betraktas vara särskilt effektiv för detta ändamål, jämfört med andra datainsamlingsmetoder (Johannesson & Perjons 2014). Eftersom denna studie ämnar undersöka nyttan av en experts search trail utifrån novisers åsikter så hade intervju som datainsamlingsmetod varit ett alternativ.

Studiens frågeställning behandlar dock noviser i generell bemärkelse. Det är alltså inte en specifik grupp i samhället, eller från ett företag som frågeställningen berör. Intervjuer lämpar sig när respondenterna besitter djup och unik kunskap om ett område. De är även användbara när frågorna som ställs är av känslig karaktär (Denscombe 2014). Eftersom frågeställningen dels inte behandlar ett kontroversiellt ämne, och dels inte berör en specifik grupp med unik kunskap, så är intervjuer inte lämpliga för denna studies datainsamling. Om denna datainsamlingsmetod hade valts så hade strukturerade intervjuer varit mest lämpliga. Enligt Johannesson och Perjons (2014) är strukturerade intervjuer och enkäter likartade eftersom de involverar ett antal fördefinierade frågor. Eftersom enkäter genererar data som är mer användbar för att besvara frågeställningen (avseende generalisering), så valdes det alternativet.

#### **3.1.4.2 Observation**

Ytterligare en datainsamlingsmetod som övervägdes var observation. När en observation genomförs, så observeras studieobjekten i deras naturliga miljö, och deras handlingar iakttas och antecknas (Denscombe 2014). Fördelen med denna metod är att den möjliggör för forskare att finna sig i kontexten som studeras. I denna studie hade då beteenden kunnat dokumenteras. Till exempel hade observationer kunnat behandla gester och ansiktsuttryck. Analysen hade kunnat förse ett perspektiv på

nytta med fokus på tolkningar av dessa uttryck. Nackdelen med observationer är att de inte vanligen brister i kontexten till det som studeras, och beteenden förenklas (*ibid.*). Tidigare forskning som legat till grund för frågeställningen visar tydligt att det är kontexten till nytta av search trails som är understuderad. Det är även av betydelse för studien att inte förenkla datan som dokumenteras.

Eftersom det är ett understuderat ämne så är det väsentligt att datainsamlingen inte präglas av mänskliga tolkningar då det saknas teoretiska ramverk i tidigare forskning avseende nytta av search trails.

### **3.1.5 Dataanalysmetoder**

#### **3.1.5.1 Search Log Analysis**

Loggfilerna analyseras med *search log analysis* (häданefter kallad SLA), som är en metod som undersöker utvald data i informationssökningsloggfiler. Denna studie utforskar novisers nytta av en expertgenererad search trail, och nytta behandlar den grad som noviser återanvänder innehållet i en experts search trail. SLA lämpar sig för att mäta återanvändningsfrekvensen i denna kontext, och har därmed valts som analysmetod. Källan för denna dataanalys är loggfilerna, i samband med genomförandet i SearchX. SLA erbjuder en konkret metod för att studera empirisk data som är insamlad från användare i en verklighetenlig kontext (Jansen 2008).

#### **3.1.5.2 Diagram**

Svaren på de kvantitativa frågorna i den första och andra enkäten analyseras med hjälp av deskriptiv statistik i form av diagram. För varje kvantitativ fråga presenteras svaren i form av diagram, där x-axeln visar svaret (till exempel 1-5 på en fråga med skaltyp intervall) och y-axeln visar antal som angett svaret. Syftet med diagrammen är att komplettera analysen av svaren på de kvalitativa frågorna. Svarsanalysen på kvantitativa frågor är lägre prioriterad än svarsanalysen på de kvalitativa frågorna. Anledningen till detta är att de kvantitativa frågorna fungerar som komplement till de kvalitativa. Frågeställningen för denna studie behandlar nytta noviser har av en experts search trail. För att noviserna ska kunna adekvat uttrycka sina upplevelser är de kvalitativa frågorna väsentliga. De kvantitativa frågorna ställs dels i syfte att förstå fritextsvaren och undvika tolkningar av låg kvalitet, och dels för att komplettera analysen av återanvändningsfrekvenserna identifierade i loggfilerna.

#### **3.1.5.3 Tematisk analys**

Svaren till de kvalitativa frågorna analyseras med en tematisk analys. Denscombe döper analysmetoden till kodning, men den mer konkreta metoden som Clarke och Braun (2006) förklrar är en utveckling av detta. Clarke och Braun skriver att en tematisk analys inleds med att forskare bekantar sig med datan (*ibid.*). Denscombe (2014) beskriver detta som processen av att kommentera datan, och därmed bekanta sig med den. Därefter kodas datan. Kodning av enkätdatan utförs genom att systematiskt tilldela datasegment olika begrepp eller fraser. Syftet är att identifiera mönster och intressanta uppfattningar i datan (Clarke & Braun 2006). Som en del av förberedelserna inför datanalysen är beslutet om vilken typ av information som kodas nödvändigt för att kunna tillämpa metoden systematiskt. Typen av information kan vara till exempel känslor, åsikter, eller användning av specifika ord (Denscombe 2014). Detta är särskilt viktigt för denna studie då frågeställningen behandlar novisers personliga uppfattningar. Loggfilerna förser enbart kvantitativ information om användares ageranden, utan kontexten bakom dem. För denna studie togs beslutet att främst åsikter och upplevelser kodas. Datans kan även kodas latent eller semantiskt (Clarke & Braun 2006), och i denna studie valdes semantisk kodning som tillvägagångssätt. Eftersom frågeställningen behandlar ett relativt underutforskat ämne, nämligen nytta av search trails, så var syftet att undvika teorier i form

av latent kodning tidigt, och låta motsägelser bli tydliga. Eftersom det saknas tidigare forskning i området togs beslutet att inte teoretisera tidigt.

När koderna har genererats så kategoriseras de. Clarke och Braun (2006) skriver att det är i denna fas som forskare söker teman, men Denscombe (2014) beskriver detta som kategorisering av data.

Kategoriseringen (tematiseringen) möjliggör för identifiering av samband mellan koder. Det är i detta skede som Denscombes beskrivning av analysprocessen skiljer sig från Clarke och Brauns beskrivning. Det som Denscombe beskriver som steg i analysprocessen ingår i de nästkommande faserna av Clarke och Brauns tematiska analys. När kategorierna (temana) har genererats så revideras de och reduceras i antal. Syftet är att att säkerställa att de inte överlappar eller är inkonsekventa. Därefter definieras de och döps (Clarke & Braun 2006). I denna fas ingår Denscombes (2014) steg att kontinuerligt jämföra den utvecklande analysen med kategorierna och koderna. Det som Clarke och Braun sedan kallar för definierung och döpning av teman, kallar Denscombe för identifiering av nyckelkoncept.

### **3.1.6 Alternativa dataanalysmetoder**

#### **3.1.6.1 Innehållsanalys**

Svaren på de kvalitativa frågorna i den andra enkäten hade kunnat analyseras med hjälp av innehållsanalys. Den huvudsakliga fördelen med innehållsanalys är att metoden kan kvantifiera innehållet i en text med en metodik som är tydlig och upprepbar, i form av språkliga enheter (Denscombe 2014). Denna analysmetod lämpar sig därmed väl för att analysera all form av text och skulle då kunna appliceras på den kvalitativa delen av studiens andra enkät. En specifik möjlighet med innehållsanalys är att den låter analytiker upptäcka betydelser i texter som inte var skribentens avsikt (*ibid.*). Detta är inte relevant för frågeställningen i studien då forskningsområdet är relativt understuderat. Analysmetoden skulle kunna vara mer lämpad i framtiden när konventioner har etablerats kring nyttan av search trails, och dessa utmanas. Vidare är en kritik mot innehållsanalys att kontexten inte tas till hänsyn (*ibid.*). Frågorna i den andra enkätens kvalitativa del ger betydelsefulla kontexter till svaren, och därfor är det särskilt viktigt med en analysmetod som kan inkludera dem, såsom tematisk analys.

#### **3.1.6.2 Statistiska test**

För svaren på de kvantitativa frågorna i enkäterna övervägdes statistiska test för dataanalysen. Mer specifikt övervägdes Spearmans korrelationstest och regressionsanalys. Spearmans korrelationstest hade kunnat undersöka styrkan på relationen mellan variabeln upplevd nytta (som behandlas i de kvantitativa frågorna i den andra enkäten) och variabeln återanvända expertbegrepp (som behandlas i dokumenten, alltså loggfilerna). Vidare hade regressionsanalysen kunnat komplettera korrelationstestet genom att undersöka eventuell kausalitet mellan variablerna. För att resultaten på statistiska test ska kunna påvisa stark signifikans så krävs det ett litet konfidensintervall som, med säkerhet, visar att studiens stickprov kan förutse den övergripande populationens medelvärde. Konfidensintervallet är alltså ett mått på resultatets osäkerhet. Ett litet konfidensintervall förutsätter att studien använder ett tillräckligt stort stickprov, eftersom risken för ett stort konfidensintervall ökar ju mindre stickprovet är (Hjerm, Lindgren & Nilsson 2014). Studien hade inte kunnat använda statistiska test som analysmetod, då den använder ett litet urval som inte hade kunnat generera resultat med stark signifikans.

## **3.2 Metodtillämpning**

### **3.2.1 Urval**

Denna studie använder ett ändamålsenligt urval, där testanvändarna är medvetet valda baserat på deras betydelse för frågeställningen som studien ämnar besvara. Ett ändamålsenligt urval är ett ickesannoliketsurval som bäst lämpar sig när en viss kunskap och relevans för studien efterfrågas (Denscombe 2014). Eftersom frågeställningen behandlar domänkunskap är kunskap ett centralt kriterium för urvalsgruppen, och ändamålsenligt urval möjliggör för specificering av detta.

Som en representation av expert-novis-relationen har elever och lärare valts ut på gymnasial nivå för studien. Läraren undervisar eleverna på en gymnasieskola norr om Stockholm i Sverige. Eleverna studerar samhällsvetenskapliga programmet. Genom e-post-kontakt och förenkäter kontrolleras ålder, språklig nivå, och domänkunskap. Samtliga deltagare är minst 18 år gamla. Därmed är de myndiga och kan skriva under medgivandeavtal utan vårdnadshavare. De är även samtliga flytande i svenska i tal och skrift, vilket är nödvändigt då testet utförs på svenska. Gymnasieläraren har mer än en universitetsexamen och betraktas därmed ha hög kunskap inom hens domän. Domänen är bild och design. Eleverna är noviserna i studien och saknar universitetsexamen. Domänkunskaperna kontrolleras via e-postkontakt med experten, och via en förenkät för noviserna. Det är en expert och 24 noviser som deltar i studien. Varken specificering av kön, etnicitet eller geografiska plats begärs av deltagarna då detta är irrelevant för frågeställningen. E-postkontakten och förenkäten fokuserar på deltagarnas domänkunskap då kunskap om detta är en förutsättning för att kunna besvara frågeställningen.

### **3.2.2 Utförande**

Sökgränssnittet SearchX implementeras med hjälp av en Docker-avbild som instansieras på en Ubuntu VPS, tillhandahållen av företaget Digital Ocean. Digital Ocean förser även VPS:en med en publik IP-adress, och en extern brandvägg som konfigureras genom företagets webbplats. När Docker-avbilden instansieras så skapas ett antal mjukvaru-”containers”, vilket är isolerade miljöer som kan exekvera en eller flera applikationer, och varje container inrymmer en nödvändig komponent för SearchX. Bland dessa komponenter finns gränssnittets front- och back-end som underbyggs av Node.js, och en datalagring i MongoDB som är en NoSQL databas. För att tillåta extern åtkomst till gränssnittet integreras SearchX med HTTP-servern NGINX, som sedan exkluderas från den externa brandväggen. Slutligen så utökas SearchX huvudkod med ett API-anrop till Microsoft Azure API *Bing Web Search API*. Detta möjliggör för reell webbsökning inom gränssnittet, med hjälp av Bing. En bild på gränssnittet i SearchX är bifogad som bilaga H.

Inledningsvis i testet ombeds experten besvara frågor om hens domänkunskap via e-postkontakt. Frågorna behandlar till exempel expertens utbildning. Baserat på vilken domän expertens besitter hög kunskap inom, så utformas sökuppgiften. Instruktioner skickas till experten via e-post, där hen blir ombedd att genomföra en sökning i SearchX, och därefter identifiera expertbegrepp i sökinmatningarna. Efter detta genomförs ett pilottest med två noviser. Pilottestet genomförs likadant som det huvudsakliga testet, men det sker dagen innan. Eventuella problem i pilottestet tas till hänsyn och det huvudsakliga testet ändras utifrån dem.

När experten har genomfört sitt test matas hens search trail in i varje groupID för noviserna. Noviserna får tillgång till sina instruktioner via Google Drive. Dessa instruktioner innehåller vardera en länk till ett gemensamt Microsoft Teams-samtal och en länk till SearchX med ett unikt gruppID till varje novis.

Syftet med att deltagarna ingår i ett Microsoft Teams-samtal är att kunna förse hjälp vid eventuella behov under testet. Samtalet dokumenteras inte. Noviserna svarar på den första enkäten som skapas via Survey&Report. Den första enkäten i testet är bifogad som bilaga I. I enkäten skriver noviserna in deras gruppID. Därefter söker de i SearchX och försöker lösa sökuppgiften. Expertens search trail är då tillgänglig. Sedan svarar noviserna på den andra enkäten som även den skapas genom Survey&Report. Den andra enkäten i testet är bifogad som bilaga J. Även i den andra enkäten skriver noviserna in deras gruppID. Experten och noviserna får läsa igenom och skriva under medgivandeavtal till samtliga delar av testen (se bilagor A, B, C, D och E). När datan är insamlad och ska analyseras används expertens lista av identifierade expertbegrepp som referens.

### **3.2.3 Dataanalys**

#### **3.2.3.1 Search Log Analysis**

Loggfilerna analyseras med SLA och mäter återanvändningsfrekvensen i hela datasetet genom att betrakta varje teckensträng (inom en sökinmatning) som är separerad med ett blanksteg som ett begrepp (Jansen 2008). Loggfilerna som genereras av SearchX lagras i en MongoDB-databas, där varje sökinmatning representeras i form av ett JSON-dokument som innehåller fälten *query*, *userId*, *groupID* och *created*. Fältet *query* innehåller sökinmatningens innehåll och fältet *created* avser klockslaget den matades in medan *userID* och *groupID* innehåller en autogenererad kod som självständigt kan identifiera vilken användare som skrev sökinmatningen respektive vilken session som den genomfördes inom.

SLAn utförs i Google Sheets, där MongoDB-databasens innehåll betraktas som analysens dataset, och varje expertbegrepp jämförs med innehållet i samtliga JSON-dokuments query-fält. För detta används sökfunktionen i Google Sheets, där varje expertbegrepp klistras in som sökord. Detta möjliggör för de expertbegrepp som förekommer i datasetet att särskiljas med blanksteg och varje expertbegrepps förekomst inkluderas i den totala återanvändningsfrekvensen. För att se till att experten som genererar search traillen inte påverkar studiens resultat, så används fältet *userId* för att exkludera expertens sökinmatningar från samtliga *groupID*. Efter att expertens sökinmatningar filtreras bort från datasetet används fältet *groupID* för att identifiera en specifik novisdeltagares resultat, i MongoDB och studiens två enkäter. Därmed kan svar och sökningar kopplas till enskilda noviser. Efter att expertens sökinmatningar filtreras bort så består alltså varje session enbart av novisernas data. Slutligen genomförs en manuell kontroll av hela datasetet avseende stavfel, plural av expertbegrepp, och skillnader i tempus, i syfte att även inkludera återanvända expertbegrepp som innehåller språkliga fel.

#### **3.2.3.2 Diagram**

Svaren på de kvantitativa frågorna i den första enkäten och den andra enkäten behandlas med hjälp av verktyget Google Sheets och presenteras i form av diagram. Fördelen med Google Sheets är att det är molnbaserat, vilket möjliggör för kollaborativt arbete i realtid. I Google Sheets motsvarar varje rad en enskild deltagare svar på enkäterna, och *groupID* inkluderas. Kolumnerna visar vilken fråga svaren avser. Genom att markera specifika kolumner kan Google Sheets automatiskt generera diagram som visar samtliga svar på en specifik fråga. Dessa diagram presenteras sedan i resultatavsnittet.

#### **3.2.3.3 Tematisk analys**

Den tematiska analysen utförs med hjälp av verktyget Google Sheets. I denna kontext är fördelen med verktyget att det under kodningen förser en överblick av datan. Raderna motsvarar varje deltagares enskilda svar på fritextfrågorna, där kolumnen anger frågan. I kolumnen bredvid läggs koderna in till

svaret på raden. Därefter placeras koderna in i olika kategorier, även kallade teman. Kolumnerna anger kategorierna (temana), och i raderna under läggs koderna in. Denscombe (2014) skriver att kategoriseringen utförs i syfte att identifiera samband mellan koder. Google Sheets grafiska gränssnitt med rader och kolumner möjliggör för en överblick av hur ett datasett blev en kod, och hur koden blev en kategori (ett tema). Dessa koder och kategorier reduceras sedan i antal och revideras. För att underlättar denna process tilldelas kategorier (teman) olika färger i Google Sheets, och det reviderade temat som utvecklats från kategorin (temat) får samma färg. Detta möjliggör för att visuellt se hur analysen har utvecklats. Även de definierade och döpta temana färgkodas. Detta tillvägagångssätt underlättar för att kontinuerligt granska koder, kategorier och teman när de slutgiltiga temana identifieras.

### **3.3 Forskningsetiska aspekter**

En fördelaktig aspekt av hur SearchX genererar loggfilerna är att deltagarna kan inte identifieras med IP-address utan får ett autogenererat grupp-ID. Detta är att föredra eftersom användarnas data tillfälligt lagras på en Ubuntu-server under tiden som studien genomförs. Vidare tilldelas samtliga deltagare, inklusive experten, medgivandeavtal för de olika delarna av genomförandet. Experten tilldelas ett medgivandeavtal för e-postkontakten och ett separat avtal för genomförandet i SearchX. I den första respektive andra enkäten ingår medgivandeavtal till noviserna, och utöver det får noviserna även enskilda medgivandeavtal för genomförandet i SearchX.

Instruktionerna till experten skickades direkt via e-post. Instruktionerna till noviserna däremot lades upp på Google Drive, i en mapp som var tillgänglig för alla med länken. Syftet med detta var att undvika att skicka instruktionerna till elevernas (novisernas) e-postadresser, eftersom detta hade kunnat identifiera dem. Läraren (experten) skickade Google Drive-länken till eleverna där de kunde hitta sina instruktioner. Under genomförandet behandlades alltså varken elevernas namn, e-postadresser, eller IP-adresser.

# 4 Resultat och analys

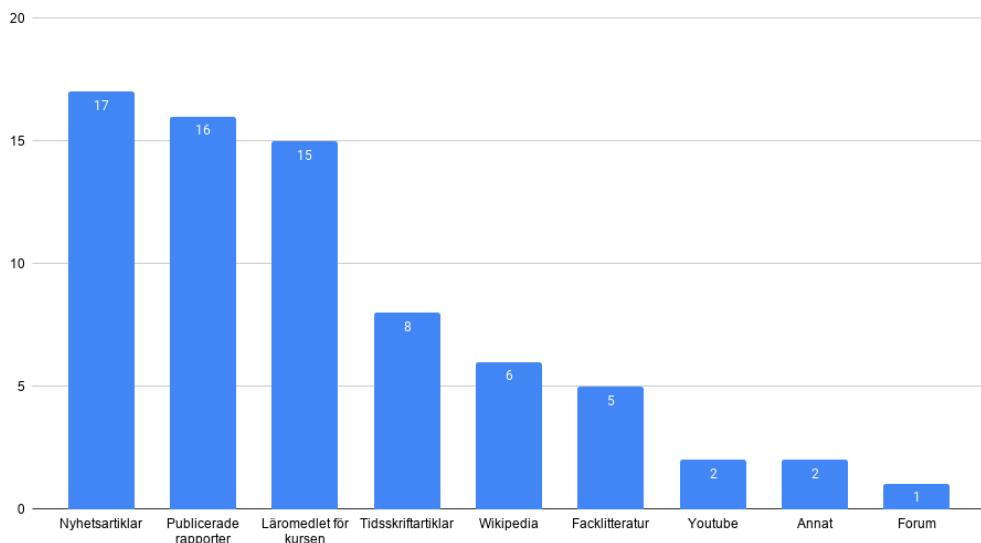
Eftersom experten besitter hög kunskap inom domänen bild och design så blev sökuppgiften följande för både experten och noviserna: ”Vad finns det för olika färgsystem och hur skiljer de sig i användningslämplighet?”. Expertens search trail är bifogad som bilaga G. Pilottestet visade att det fanns en del tekniska hinder för noviserna, som behövde åtgärdas innan det huvudsakliga testet. Testet visade att en deltagare inte kunde logga in i Teams-mötet eftersom hen använde sin mobil istället för en stationär eller bärbar dator. Det framkom även att det fanns ett begränsat antal webbläsare som SearchX var kompatibel med, nämligen Firefox och Chrome. Deltagaren från pilottestet använde webbläsaren Safari. Pilottestet ledde till en redigering av instruktionerna så att de inkluderade information om vilken typ av webbläsare noviserna kunde använda, och att testet behövde genomföras via en dator.

Under det huvudsakliga testet var det ett antal deltagare som inte svarade på den andra enkäten. Eftersom svaren på den andra enkäten är betydelsefulla för kontexten till loggfilerna så har datan från gruppen som deltagit i samtliga delar av testet särskilts från gruppen bestående av data från samtliga deltagare. När resultaten från den första enkäten och loggfilerna presenteras så visas två diagram, ett för respektive grupp. Gruppen med samtliga deltagare kallas häданefter grupp A. Gruppen med de deltagare som även svarade på den andra enkäten kallas hädanefter grupp B. I grupp A deltog 24 personer i utförandet, varav 20 personers data kunde användas i studien. 4 personer exkluderades eftersom de antingen inte var myndiga eller skrev in felaktiga gruppID. Av de 20 personerna i grupp A var det 14 personer som svarade på den andra enkäten. Dessa 14 personer utgör grupp B. Sammanfattningsvis svarade grupp A på den första enkäten och sökte i SearchX där datan sparades i loggfiler. Grupp B svarade på den första enkäten, sökte i SearchX där datan sparades i loggfiler, och svarade även på den andra enkäten. Grupp A består av grupp B. Resultatavsnittet belyser skillnaderna mellan de två grupperna för den första enkäten och loggfilerna. Resultaten från den andra enkäten avser enbart grupp B.

## 4.1 Enkäten som kartlägger noviserna

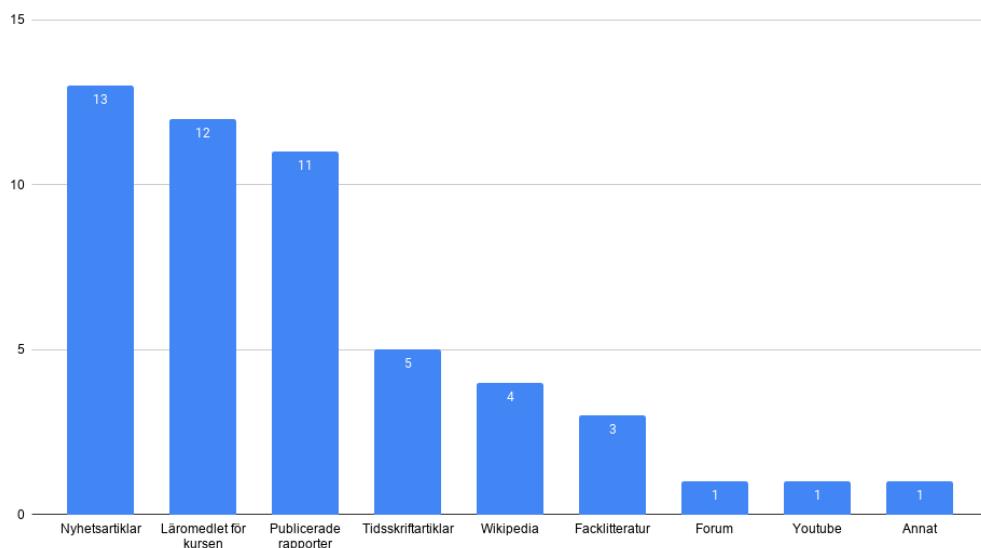
Samtliga deltagare var myndiga, flytande i svenska i tal och skrift, och hade inte studerat eftergymnasiala studier. De vanligaste källorna deltagarna använde sig av var nyhetsartiklar, publicerade rapporter och läromedlet för kursen. I kontrast var det ett fåtal deltagare som angav att de använde Youtube och forum. Två svar från grupp A dokumenterades på ”Annat” där de angav ”pålitlig källa” och ”litar inte på den helt, utan tar faktan och är objektiv i det jag skriver”. Den deltagare som angav »pålitlig källa” under ”Annat” ingick även i grupp B. Figur 3 och 4 visar svaren på frågan ”Vilka av följande källor brukar du använda när du söker information för en skoluppgift?”.

Vilka av följande källor brukar du använda när du söker information för en skoluppgift?



Figur 3: Svaren grupp A angav på frågan "Vilka av följande källor brukar du använda när du söker information för en skoluppgift?".

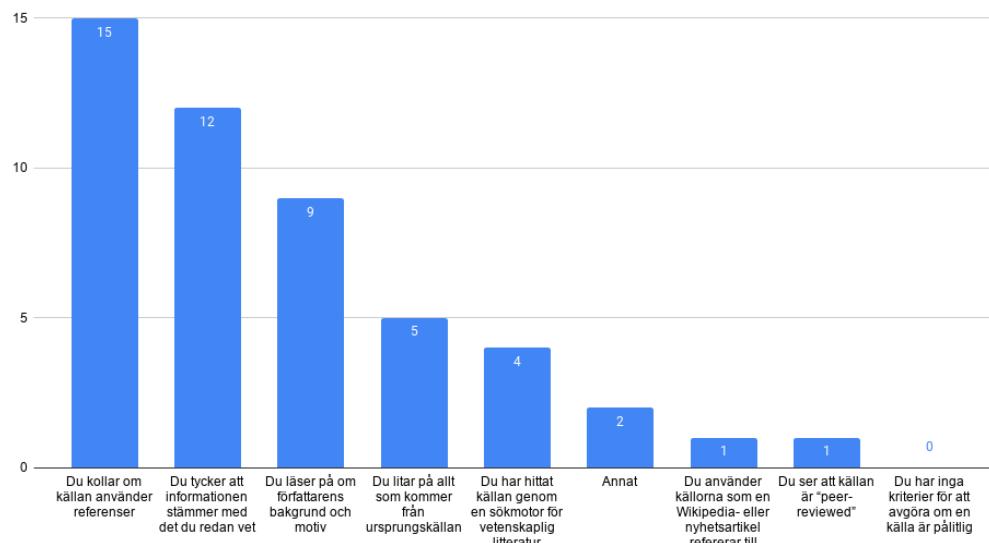
Vilka av följande källor brukar du använda när du söker information för en skoluppgift?



Figur 4: Svaren grupp B angav på frågan "Vilka av följande källor brukar du använda när du söker information för en skoluppgift?".

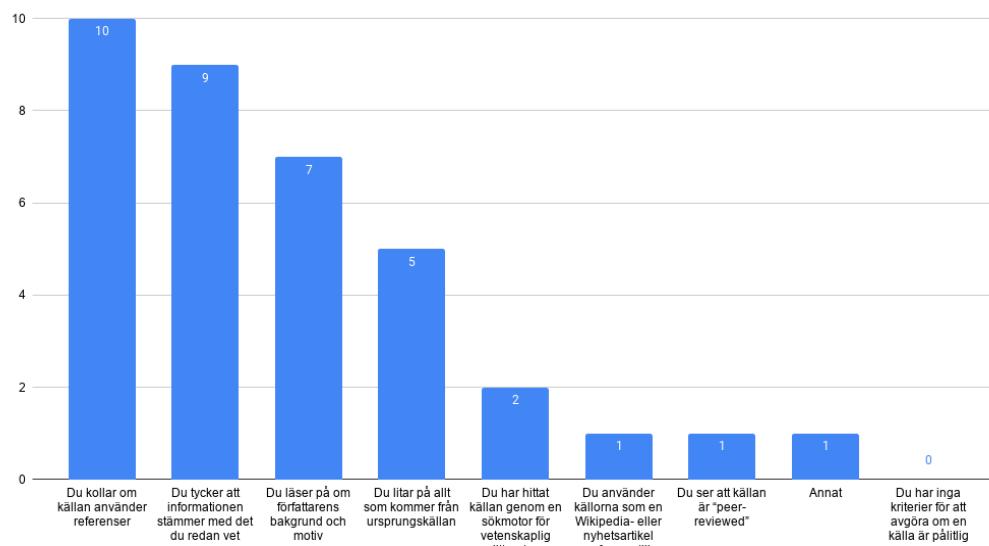
För att avgöra om källor är pålitliga svarade majoriteten av deltagarna att de ser att källan i fråga använder referenser, informationen stämmer med det de redan vet, och att de läser på om författarens bakgrund och motiv. Figur 5 och 6 visar resultaten avseende deltagarnas förmåga till källkritiskt tänkande, med svaren på frågan "Vad gör att du litar på informationen?". Resultaten från grupp A och grupp B är snarlika, där de två vanligaste alternativen kan betraktas som metoder av låg kvalité för att bedöma källors pålitlighet. Det tredje vanligaste svaret är dock en god metod för att avgöra källors pålitlighet.

Vad gör att du litar på informationen?



Figur 5: Svaren grupp A angav på frågan "Vad gör att du litar på informationen?".

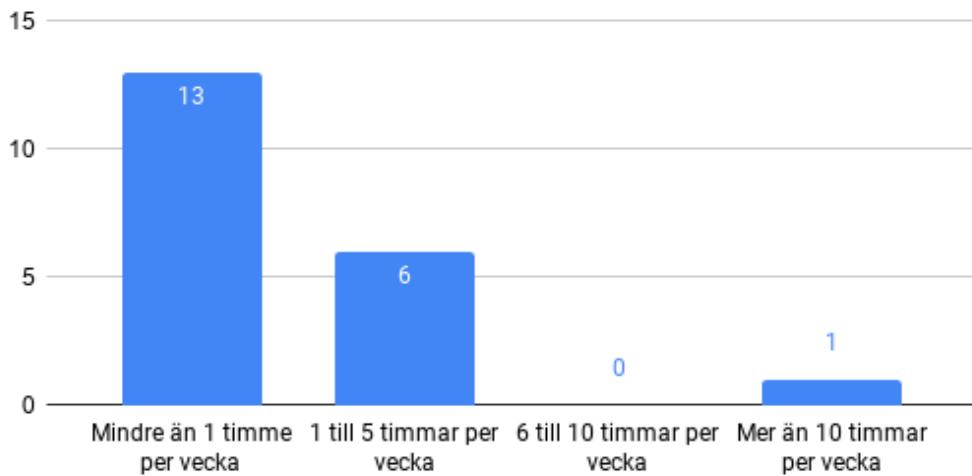
Vad gör att du litar på informationen?



Figur 6: Svaren Grupp B angav på frågan "Vad gör att du litar på informationen?".

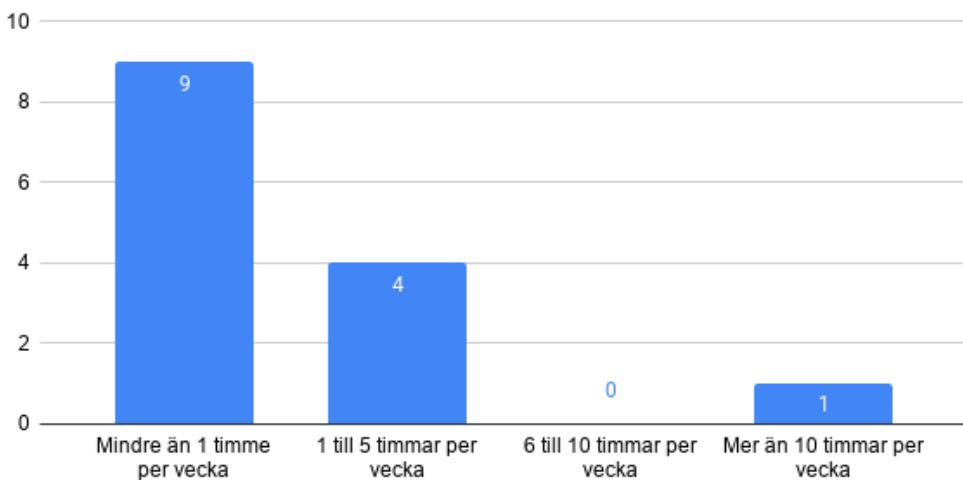
Deltagarna sökte sällan på information inom domänen, där majoriteten svarade att de söker på informationen mindre än 1 timme per vecka. Det stämmer för både grupp A och grupp B. Resultaten från dessa diagram bekräftar att noviserna sällan söker inom domänen och de tyder på novisernas låga domänkunskap. Resultaten redovisas i figur 7 och 8, där frågan är "Hur länge söker du information inom domänen bild och design, via en internetbaserad sökmotor?".

## Hur länge söker du information inom domänen bild och design, via en internetbaserad sökmotor?



Figur 7: Svaren grupp A angav på frågan "Hur länge söker du information inom domänen bild och design, via en internetbaserad sökmotor?".

## Hur länge söker du information inom domänen bild och design?

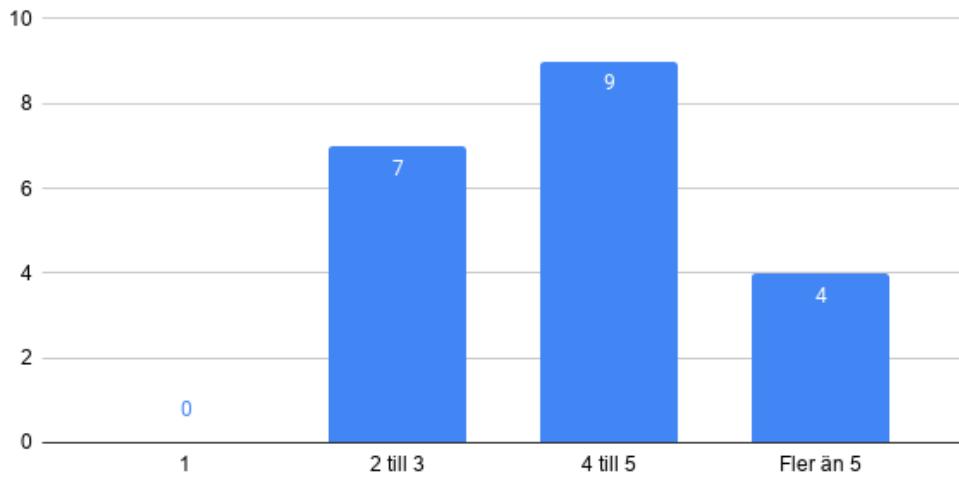


Figur 8: Svaren grupp B angav på frågan "Hur länge söker du information inom domänen bild och design, via en internetbaserad sökmotor?".

Vidare visade resultaten att deltagarna vanligen sökte i genomsnitt på ett flertal källor innan de var nöjda med sökresultaten. I grupp A var det sju personer som svarade två till tre källor, nio personer som svarade fyra till fem källor, fyra personer som svarade fler än fem källor, och ingen som svarade att de i genomsnitt sökte på en källa. I grupp B var det sex personer som svarade två till tre källor, fem personer som svarade fyra till fem källor, tre personer som svarade fler än fem källor, och ingen som svarade att de i genomsnitt sökte på en källa. Båda grupper inkluderade deltagare som svarade att de använde fler än fem källor, vilket kan betraktas som ett högt antal för domännoviser. Frågan specificerar dock inte kvalitén på dessa källor, utan syftar enbart på källor i generell bemärkelse.

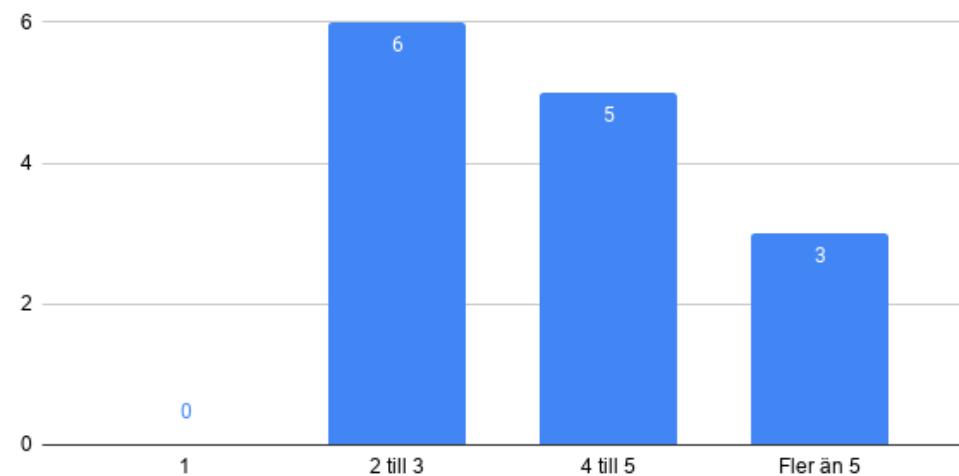
Figur 9 och 10 visar svaren till frågan ”När du söker information för en skoluppgift: hur många källor i genomsnitt kollar du på innan du är nöjd?”.

När du söker information för en skoluppgift: hur många källor i genomsnitt kollar du på innan du är nöjd?



Figur 9: Svaren grupp A angav på frågan ”När du söker information för en skoluppgift: hur många källor i genomsnitt kollar du på innan du är nöjd? ”.

När du söker information för en skoluppgift: hur många källor i genomsnitt kollar du på innan du är nöjd?

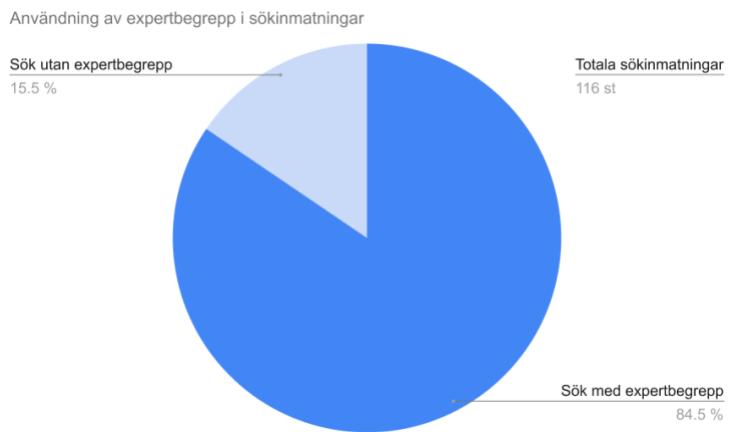


Figur 10: Svaren grupp B angav på frågan ”När du söker information för en skoluppgift: hur många källor i genomsnitt kollar du på innan du är nöjd? ”.

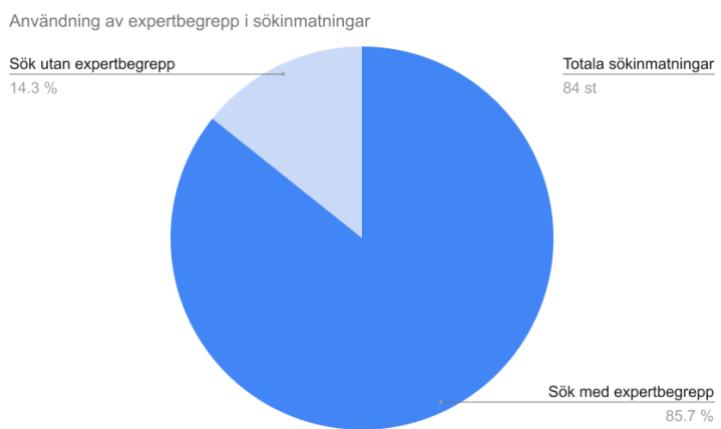
## 4.2 Loggfilerna

Loggfilsdata för grupp A visade att 84,5 procent av sökinmatningarna innehöll expertbegrepp, av de totala 116 sökinmatningar i SearchX. För grupp B var det 85,7 procent av sökinmatningarna innehöll något eller flera av expertens begrepp, av totalt 84 sökinmatningar i SearchX. Båda grupperna visar att majoriteten av sökinmatningarna innehöll expertbegrepp. I kontrast var det 15,5 procent av sökinmatningarna som inte innehöll expertbegrepp i grupp A, och för grupp B var det istället 14,3

procent. I figur 11 och 12 presenteras dessa resultat i form av cirkeldiagram. För båda grupper kan det även konstateras att samtliga deltagare använde expertbegrepp vid något tillfälle i sökinmatningarna.

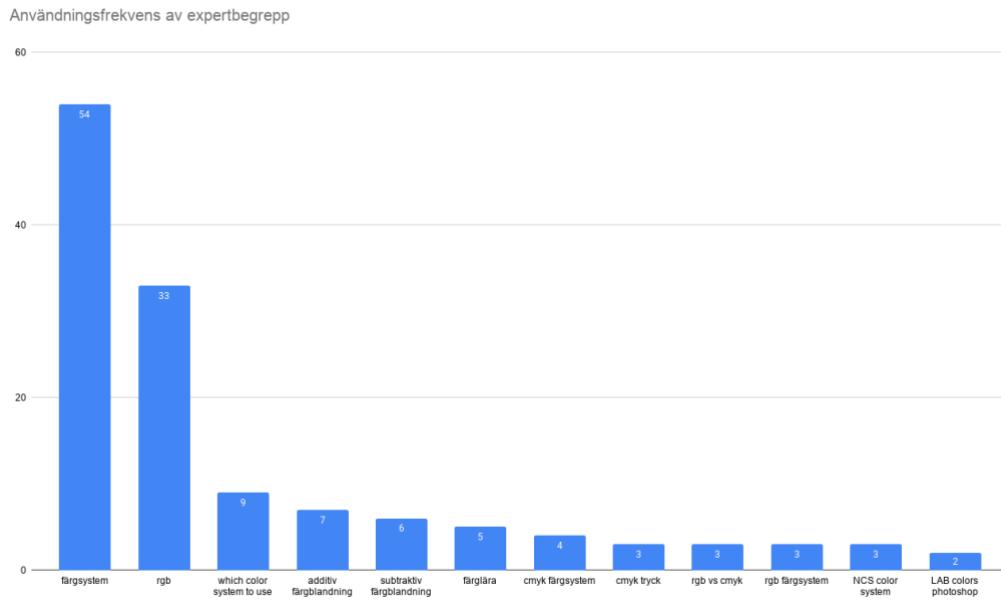


*Figur 11: En procentuell representation av andel sökinmatningar som innehöll respektive inte innehöll expertbegrepp, i grupp A.*

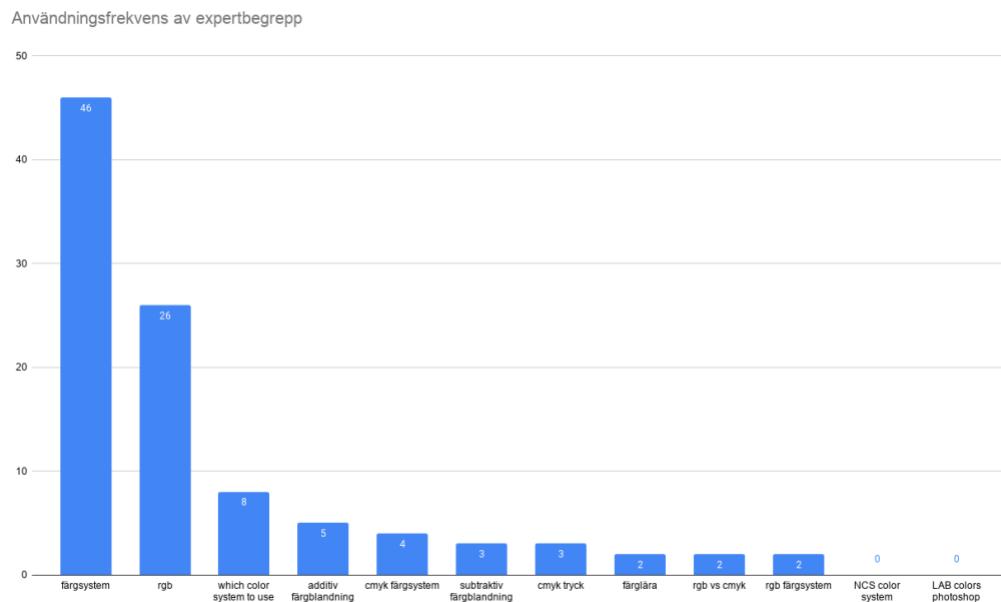


*Figur 12: En procentuell representation av andel sökinmatningar som innehöll respektive inte innehöll expertbegrepp, i grupp B.*

I figur 13 och 14 visas de expertbegrepp som deltagarna använde och dess frekvens. Begreppet ”färgsystem” förekom 54 gånger i sökinmatningarna för grupp A, och för grupp B var det 46 gånger. Detta var det vanligaste expertbegreppet deltagarna använde, och det är även detta begrepp som förekom i den första delen av sökuppgiften. Resultaten visar att begreppet ”LAB colors photoshop” förekom sällan i sökinmatningarna för grupp A, och aldrig för grupp B. Detta begrepp ingår heller inte i sökuppgiften.



Figur 13: En representation av expertbegreppen som användes i deltagarnas sökinmatningar, och dess frekvens, i grupp A.



Figur 14: En representation av expertbegreppen som användes i deltagarnas sökinmatningar, och dess frekvens, i grupp B.

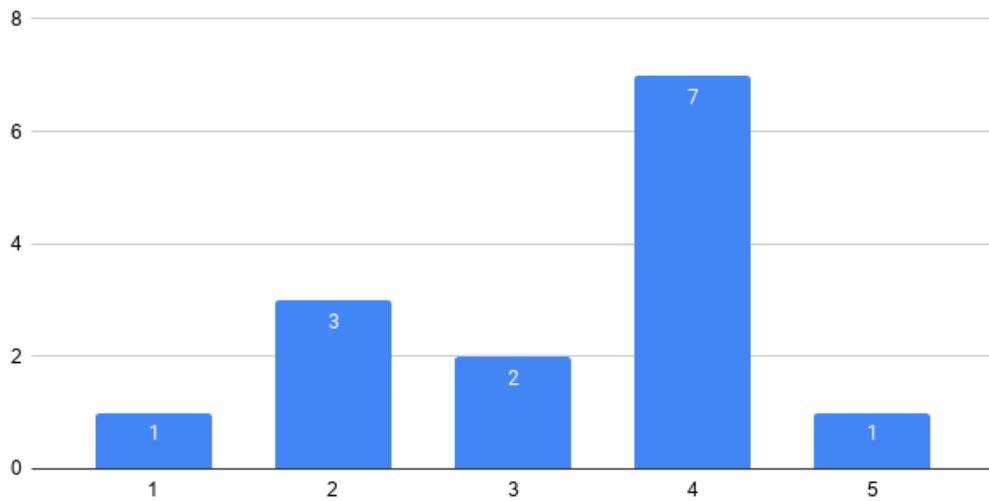
## 4.3 Enkäten om upplevd nytta

### 4.3.1 Kvantitativa resultat för nyttan

Deltagarna upplevde att sökuppgiften var svår, och figur 15 visar att för svarsalternativ mellan 1-5 var 4 det mest förekommande svaret. Det var få deltagare som svarade 1 eller 5 vilket tyder på att sökuppgiften inte var extremt enkel eller svår för dem. Gruppens svar polariseras i frågan om deltagarna hade återanvänt några av expertens sökinmatningar. Hälften svarade ”ja”, och hälften ”nej”.

Detta är något motsägelsefullt i relation till loggfilsdata för grupp B som visade att 100 procent av deltagarna hade återanvänt expertbegrepp.

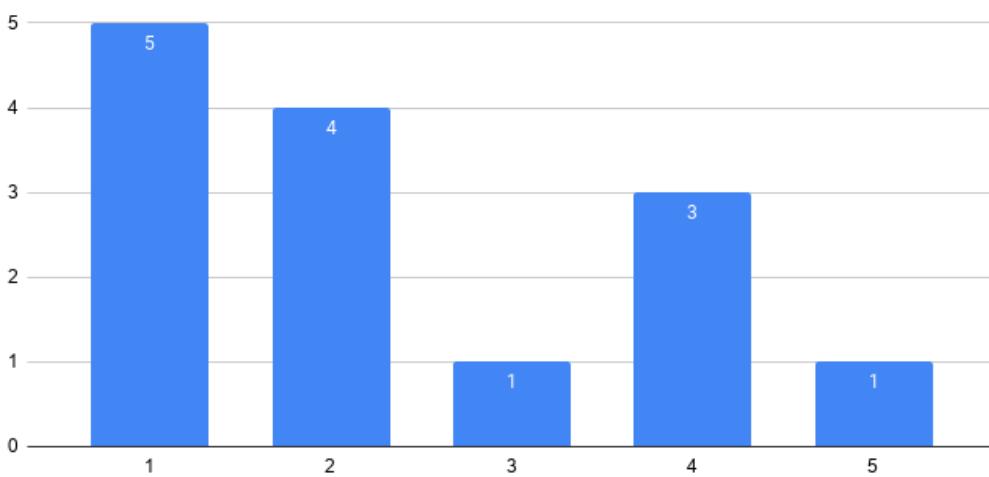
På en skala 1-5, hur svår tyckte du att sökuppgiften var att lösa?



Figur 15: Diagrammet visar hur svår deltagarna upplevde att sökuppgiften var att lösa. Svarsalternativen var 1-5.

En övervägande del av deltagarna var osäkra på om de lyckats lösa sökuppgiften, detta förtydligas i figur 16, med svar på frågan ”På en skala 1-5, hur säker är du på att du lyckades lösa sökuppgiften?”. För svarsalternativen 1-5 var det nio deltagare som svarade 1 eller 2 och fyra deltagare som svarade 3, 4 eller 5. Detta tyder på att det fanns en viss osäkerhet bland deltagarna, med avseende på om de lyckats lösa sökuppgiften.

På en skala 1-5, hur säker är du på att du lyckades lösa sökuppgiften?

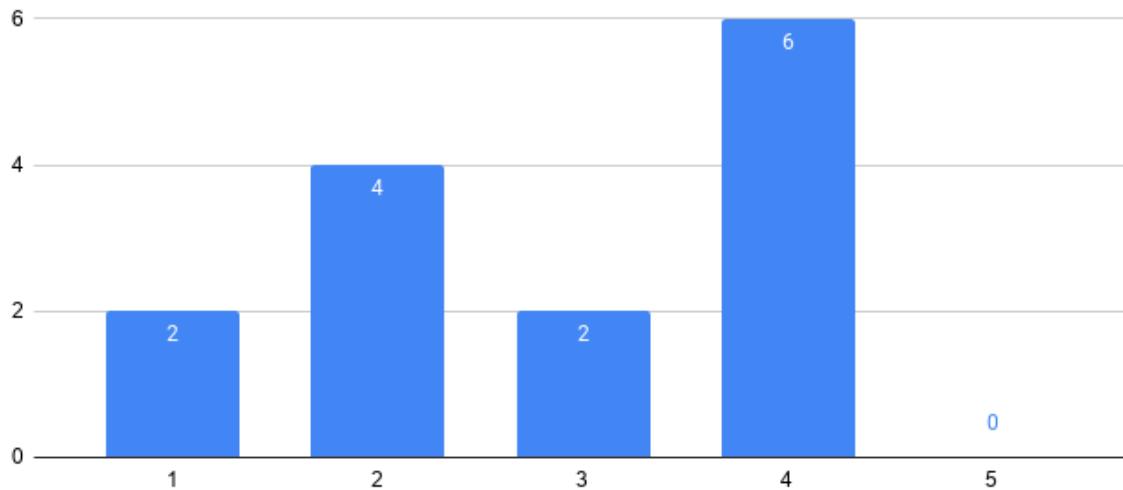


Figur 16: Diagrammet visar hur säkra deltagarna var på att de lyckades lösa sökuppgiften. Svarsalternativen var 1-5.

Figur 17 visar svaren på frågan ”På en skala 1-5, hur hjälpsamma var expertens sökinmatningar för att lösa sökuppgiften?”. Det framgår tydligt majoriteten ansåg att expertens search trail hjälpsam, då sex personer valde alternativ 4. Samtidigt var det sex deltagare som valde alternativen 1 eller 2, vilket medför att resultatet är något tvetydigt. Orsaken till att dessa sex deltagare inte upplevde

sökinmatningarna som hjälpsamma kan möjligen delvis förklaras med hjälp av figur 18 som berör novisernas förståelse för expertbegreppen. Svaren avser frågan ”Var det några begrepp i expertens sökinmatningar som du inte förstod?”. Det var nämligen 35,7 procent av deltagarna som svarade att det fanns expertbegrepp som de inte förstod, och loggfilerna visade att det var flera expertbegrepp som återanvändes sällan.

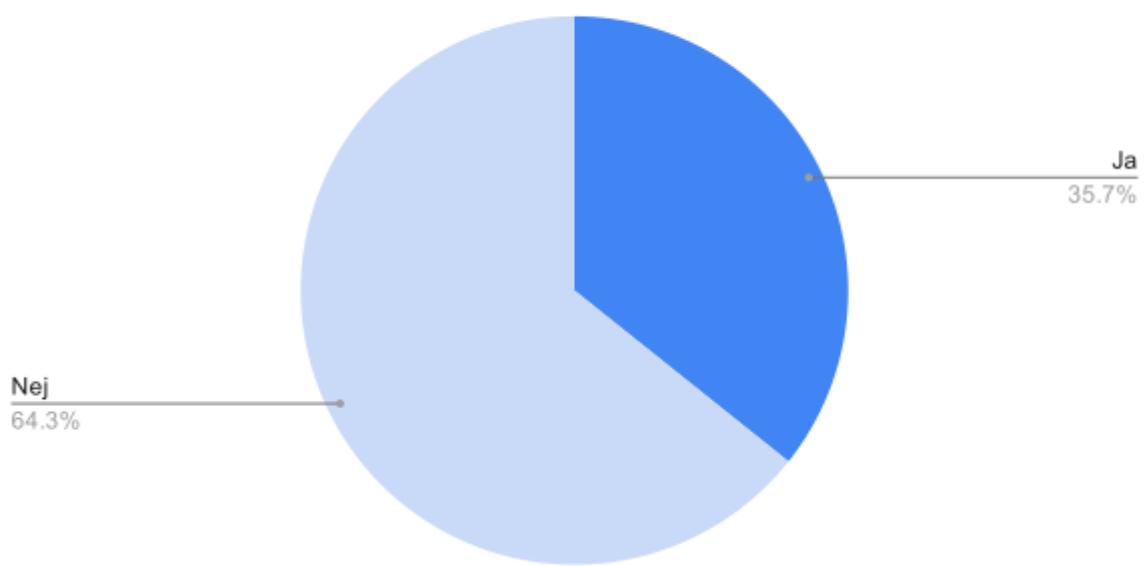
**På en skala 1-5, hur hjälpsamma var expertens sökinmatningar för att lösa sökuppgiften?**



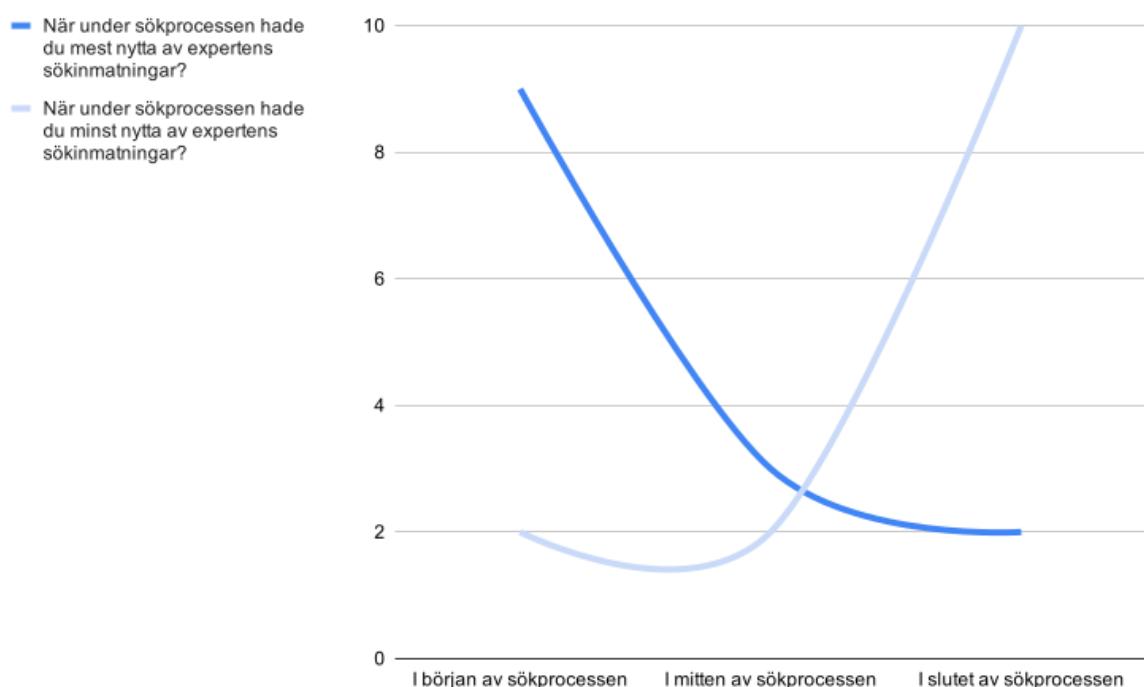
Figur 17: Diagrammet visar hur hjälpsam deltagarna upplevde att expertens search trail var. Svarsalternativen var 1-5.

Figur 18 visar att det var en något övervägande del som svarat att det inte var några expertbegrepp de inte förstod. Detta kan alltså omformuleras till att de som angett ”Nej” som svar förstod samtliga expertbegrepp. Figur 19 visar samtidigt att majoriteten av deltagarna upplevde att expertens search trail var som mest hjälpsam vid början av sökprocessen. Daten som figur 18 och 19 visar är särskilt intressant i relation till den kvalitativa datan som presenteras i ett senare avsnitt. Enligt den kvalitativa datan förändras deltagarnas förståelse för expertbegreppen under sökprocessen.

Var det några begrepp i expertens sökinmatningar som du inte förstod?



Figur 18: Diagrammet visar andelen deltagare som svarade att det fanns respektive inte fanns expertbegrepp som de inte förstod.



Figur 19: Diagrammet visar under vilken del av sökprocessen som deltagarna drog mest respektive minst nytta av expertens search trail.

## **4.3.2 Teman från fritextsvaren**

Den andra enkäten innehöll frågor som låt deltagare svara i sina egna ord, det vill säga i fritext. I detta avsnitt ingår citat från deltagarna, och deltagarna benämns här med deras nummer från utförandet (som deras lärare tilldelade dem). Till exempel har vi ett citat från deltagare 24. Detta är alltså en elev som blivit tilldelad numret 30 av sin lärare, och det är orelaterat till antalet deltagare som svarat på den andra enkäten. Inledningsvis granskades all data genom att kommentera den. Detta ledde till en överblick av datan och tidiga tankar om innehållet. Svaren kodades sedan med fokus på data som är relevant för att besvara frågeställningen. Därefter kategoriseras koderna efter deras relation.

Kategorierna kallas för teman av Clarke och Braun. Temana reviderades sedan för att säkerställa att de inte överlappade, och att datan var tillräcklig för att motivera dem. Slutligen definierades temana mer konkret i syfte att tydliggöra de huvudsakliga koncepten i datan. Bilaga F visar de reviderade temana, samt de definierade temana. Under analysprocessen togs både koderna och den råa datan till hänsyn för att minska risken för kontexten att misstolkas. I detta avsnitt utgör rubrikerna de definierade temana, och analysen bakom dem beskrivs.

### **4.3.2.1 Samband mellan låg förståelse för sökuppgiften och låg domänkunskap**

Kodningen visade att det var flera deltagare upplevde sökuppgiften som svår och hade mer specifikt svårigheter med dimensionen i sökuppgiften. Denna data tilldelades koder såsom *svår sökuppgift* och *dimensioner är svåra* som var vanligen förekom i deltagarnas svar. Vidare identifierades koderna *läg domänkunskap* och *brist på erfarenhet i domänen*. Dessa koder förekom främst vid svar på enkätfrågan om problemet bakom att deltagarna använde sig av expertens search trail, och de förekom vanligen i samband med deras uppfattning om sökuppgiftens komplexitet.

Koderna *läg domänkunskap* och *brist på erfarenhet i domänen* sammansättades därefter till en egen kategori *läg kunskap och erfarenhet i domänen*. Koderna *svår sökuppgift* och *dimensioner är svåra* sammansättades till kategorin *sökuppgiften var utmanande*. Ursprungligen kombinerades dessa kategorier med andra kategorier om källkritik, och utgjorde det reviderade temat *läg domänkunskap med svårigheter för källkritik och förståelse för sökuppgift*. Efter undersökning av den råa datan, däremot, separerades källkritik från detta tema, och det nya definierade temat blev *samband mellan låg förståelse för sökuppgiften och låg domänkunskap*. Datans visade att deltagare upplevde att dimensionen i sökuppgiften var svår att förstå, medan de skrev att de hade låg domänkunskap. Källkritiken var ett separat tema. Följande är två citat från samma deltagare som beskriver sambandet. Det första citat avser fritextsvaret till frågan: ”När under sökprocessen hade du mest nytta av expertens sökinmatningar? (Vänligen motivera ditt svar.)”, och det andra citatet till frågan: ”På en skala 1-5, hur svår tyckte du att sökuppgiften var att lösa? (Vänligen motivera ditt svar.)”.

”För att jag hade själv inte så bra koll på det här ämnet.” - Deltagare 18

”Jag tror jag hittade svar om vad det fanns för färgsystem, men inte om hur det skilde sig i användarslämpligheten.” - Deltagare 18

### **4.3.2.2 Expertens search trail var hjälpsam, men noviser upplevde svårigheter med källkritik**

Analysen visade att en vanligt förekommande kod var *svårt med källkritik*. Deltagarna svarade att det var utmanande att hitta sökresultat som de uppfattade som trovärdiga och pålitliga. Även koderna *ovanliga sökresultat* och *osäkerhet kring korrekthet* förekom. Vad gäller den sistnämnda koden var sökuppgiften kontexten. Deltagarna upplevde alltså att sökresultaten var ovanliga och svåra att bedöma som pålitliga, och de var osäkra på huruvida de löst sökuppgiften. Dessa koder ingick senare i

kategorin *källkritik*. Kategorin ingick ursprungligen i det reviderade temat *bristande domänkunskap med svårigheter för källkritik och förståelse för sökuppgift*. Under denna del av analysen tolkade vi att detta tema innefattade kategorierna *sökuppgiften var utmanande, källkritik, och bristande kunskap och erfarenhet i domänen*. Citatet nedan är ett fritextsvar till frågan: ”På en skala 1-5, hur svår tyckte du att sökuppgiften var att lösa? (Vänligen motivera ditt svar.)”.

”Jag upplever att jag hittade rätt svar, men är osäker på om det faktiskt är det rätta.”

- Deltagare 11

Flertalet koder behandlade även hjälpsamheten av expertens search trail, då koder såsom *relevanta sökinmatningar och positiv till begrepps användningens påverkan* identifierades. Dessa koder inkluderades i kategorin *hjälpsamt att använda search trail*. Detta blev en av kategorierna som ingick i det reviderade temat *search trail var hjälpsam och vägledande*. Under processen av att definiera teman undersöktes den råa datan och dess koder återigen, dels för att öka förståelsen för kontexten, och dels för att se sammanhanget av vad deltagarna svarat på olika frågor. Utifrån detta utvecklades analysen till att deltagare ansåg att expertens search trail var hjälpsam, men att de upplevde svårigheter med källkritiken. Därmed definierades temat till *expertens search trail var hjälpsam, men noviser upplevde svårigheter med källkritik*. Nedan presenteras ett citat som exemplifierar kontexten bakom detta resonemang. Citatet är ett fritextsvar till frågan: ”På en skala 1-5, hur hjälpsamma var expertens sökinmatningar för att lösa sökuppgiften? (Vänligen motivera ditt svar.)”.

”De var hjälpsamma, men det var svårt att få upp trovärdiga källor som man kände igen för några referenser var däremot ovanliga och såg inte trovärdiga ut.” - Deltagare 18

#### **4.3.2.3 Noviser upplevde att search trailen vägledde dem i början av sökprocessen, och efter ökad domänkunskap minskade behovet av att använda den**

I den andra enkäten blev deltagarna tillfrågade när under sökprocessen de hade mest nytta av expertens search trail. I den kvantitativa delen av enkäten skrev deltagarna in när de hade mest respektive minst nytta av expertens search trail, och till följd av detta fick de tillfället att utveckla sina svar i fritext. När deltagarna svarade att de hade mest nytta av expertens search trail i början av sökprocessen, så svarade de även att den var vägledande. Datot från motiveringarna till frågan kodades till bland annat *search trail gav vägledning och search trail underlättar formulering på sökinmatningar*. Dessa koder ingick sedan i kategorierna *hjälpsamt att använda search trail och search trail var vägledande*. Dessa kategorier kombinerades sedan med andra för att utgöra det reviderade temat *search trail var hjälpsam och vägledande*.

Då deltagarna svarade att de hade minst nytta av expertens search trail i slutet av sökprocessen, så förekom koderna *minskat behov av expertbegrepp efter ökad domänkunskap, förstod okända expertbegrepps innebörd under sökning* vanligen i motiveringarna. Dessa koder ingick sedan i kategorierna *domänkunskap ökar under sökprocessen och förståelse för expertbegrepp ökar under sökprocessen*. Analysen utvecklades till att kombinera dessa två under det reviderade temat *domänkunskap ökar under sökprocessen*, eftersom datan visade att ökad förståelse för expertbegrepp under sökprocessen relaterade till ökad domänkunskap.

När temana skulle definieras undersöktes koderna igen. Analysen visade att de deltagare som svarade att expertens search trail var vägledande i början av sökprocessen senare svarade att deras domänkunskap ökat i slutet av sökprocessen, och att behovet av expertens search trail då hade minskat. Därmed blev det definierade temat *noviser upplevde att search trailen vägledde dem i början*

*av sökprocessen, och efter ökad domänkunskap minskade behovet av att använda den.* Följande är två citat av samma deltagare som exemplifierar temat. Det första citatet är ett fritextsvar till frågan ”När under sökprocessen hade du mest nytta av expertens sökinmatningar? (Vänligen motivera ditt svar.)”, och det andra citatet är ett fritextsvar till frågan ”När under sökprocessen hade du minst nytta av expertens sökinmatningar? (Vänligen motivera ditt svar.)”.

”Dem ger ju absolut en idé över vad det är man ska börja söka efter.” - Deltagare 9

”När man har samlat på sig tillräckligt med egen fakta, behövs inte andra sökinmatningar.”  
- Deltagare 9

#### **4.3.2.4 Noviser undvek begrepp de inte förstod, men upplevde att begreppen de använde var hjälpsamma**

Detta tema utvecklades genom fler iterationer än de andra. Detta beror på att vissa koder tolkades som motsägelsefulla, och det var otydligt huruvida datan var relevant för att besvara frågeställningen. Under analysprocessen, ändå, visade utvecklingen att temat är relevant när kontexten tas till hänsyn.

Från datan identifierades koden *undvek okända expertbegrepp*. Det var två deltagare som förklarade att de valde att inte använda begreppen från expertens sökinmatningar som de inte förstod. Koden behandlar ett unikt ämne för datan, som andra dataset inte inkluderar. Därför utgjorde denna kod en egen kategori, nämligen *undvek okända begrepp*. En deltagare som svarade att de undvek okända begrepp skrev även att de använde de begrepp de vanligen söker med. Detta kodades som *sökte efter vana*, och utgjorde kategorin *sökstrategi baserad på vana*. De två kategorierna *sökstrategi baserad på vana* och *undvek okända begrepp* reviderades till temat *undvek att ta hjälp av search trail*. Detta berodde delvis på att personen som sökte efter vana skrev att de använde expertens search trail i låg grad.

När temat skulle definieras så var det två deltagares svar som var särskilt betydelsefulla, eftersom kontexten i deras svar förklarade sambandet mellan de koder som beskrivits hittills i detta tema. Genom att följa deltagarnas svar på de olika enkätfrågorna kunde vi förstå att deltagarna undvek expertbegrepp som de inte förstod, samtidigt som de upplevde att de begrepp de använde var hjälpsamma. Som svar på frågan ”Var det några begrepp i expertens sökinmatningar som du inte förstod? (Vänligen beskriv hur du hanterade detta.)” skrev både deltagare 24 och 11 att de undvek okända expertbegrepp. Som svar på frågan ”Hur påverkade begreppsanvändningen din förmåga att söka relevant information?” svarade deltagare 24 att de expertbegrepp hen hade använt var hjälpsamma i sökprocessen. Vidare skrev de båda som svar på frågan ”Tycker du att detta har varit ett användbart sätt att söka information? Alltså, att ha en experts sökinmatningar tillgängliga? Vänligen motivera ditt svar.” att de upplevde att expertens sökinmatningar var vägledande. Följande tre citat kommer från samma deltagare. Det första citatet är ett svar på frågan ”Var det några begrepp i expertens sökinmatningar som du inte förstod? (Vänligen beskriv hur du hanterade detta.)”, det andra citatet är ett svar på frågan ”Hur påverkade begreppsanvändningen din förmåga att söka relevant information?”, och det tredje citatet är ett svar på frågan ”Tycker du att detta har varit ett användbart sätt att söka information? Alltså, att ha en experts sökinmatningar tillgängliga? Vänligen motivera ditt svar.”. Det är nödvändigt att påpeka att svaret till frågan ”Hur påverkade begreppsanvändningen din förmåga att söka relevant information?” skrevs efter att deltagaren svarat att ja, hen hade återanvänt något av expertbegreppen.

”Valde att ej använda dessa begrepp” - Deltagare 24

”påverkade min förmåga positivt” - Deltagare 24

”väldigt användbart, ger riktlinjer för mig att ta fram information om ett ämne jag inte kunde något om från början” - Deltagare 24

# 5 Diskussion

I detta avsnitt förs en diskussion om resultaten som presenterats i det tidigare avsnittet. Daten från samtliga delar av testet diskuteras i relation till varandra. Den kvalitativa datan från den andra enkäten används som utgångspunkt i diskussionen, och den kvantitativa datan från loggfilerna och enkäterna används som komplement för att förstå den kvalitativa datan. Temana från resultatet döps om i detta avsnitt. Anledningen till detta är att namnen på de definierade temana är långa och därför inte lämpar sig som rubriker. Syftet med att döpa om dem i detta avsnitt är att öka läsbarheten.

## 5.1 Dimensioner

Under analysen av den kvalitativa datan, identifierades ett samband mellan deltagarnas låga förståelse för sökuppgiften och deras låga domänkunskap. Flera av de deltagare som hade svårigheter med sökuppgiften, specificerade även att det var dimensionen i sökuppgiften som var utmanande. Av denna anledning uppfattades sökuppgiften som komplex och svårbesvarad, vilket överensstämmer med Capra et al. (2018) förklaring om att sökuppgifter uppfattas som mer komplexa när dimensioner ingår. Vidare, var det enligt datan från loggfilerna främst expertbegrepp som inte avsåg dimensionen som återanvändes, trots att experten skrev sökinmatningar som avsåg dimensionen. Wang, Zhao och Chen (2020) utvecklar att flera användare har en positiv inställning till användandet av expertinmatningar för att vidareutveckla sina egna sökstrategier. Denna positiva inställning framgår även i denna studie då flera deltagare anger att de är positiva gentemot search trailen och att den är både vägledande och hjälpsam.

Detta leder till frågan om all återanvändning av expertbegrepp kan tyda på nytta för uppgifter av hög komplexitet? Om noviser inte återanvänder expertbegrepp som avser dimensionen i sökuppgiften, och du upplever att dimensionen är en utmanande del, så kan hjälpsamheten ifrågasättas. Deltagarna förklarade att expertens search trail var hjälpsam, men det går inte att fastställa att detta gäller för uppgifter av hög komplexitet. För att vidare undersöka detta skulle den insamlade datan behöva analyseras med en särskiljning av vilka sökinmatningar som avser dimensionen i en sökuppgift och vilka inte. En sådan studie skulle mer specifikt kunna identifiera nytta av search trails för en uppgift av hög komplexitet specifikt.

## 5.2 Minskat behov av search trailen under sökprocessen

Den insamlade datan visar att det var i början av sökprocessen som noviserna upplevde att de hade mest nytta av expertens search trail. Den kvantitativa datan från den första enkäten visade detta tydligt i figur 19, och som kontext visade den kvalitativa datan att expertens search trail ursprungligen var vägledande, men att efter ökad domänkunskap så minskade behovet av att använda den. För att förstå varför expertens search trail var vägledande är det lämpligt att återkoppla till en tidigare studie. Jenkins, Corritore och Wiedenbeck (2003) skriver att domänexperter är innehållsdrivna och använder deras höga kunskap för att göra tvetydiga sökproblem entydiga, välja lämpliga sökinmatningar och utvärdera relevant information. Till skillnad från domännoviser förstår även domänexperter relationen mellan olika ämnesspecifika sökbegrepp (ibid.). Även Isenberg påpekade att

experter har lärt sig utvinna information som är värdefull, men svårtillgänglig och begränsad. De hittar fokusområden fortare än noviser och upplever mindre ångest i sökprocessen (Ingwersen & Järvelin 2005). Noviserna i denna studie har uttryckt att expertens search trail underlättade deras förmåga att formulera sökinmatningar. Datat tyder på att noviserna använde expertens sökinmatnignar för att de var välformulerade och problemlösningssinriktade. Datat visade även att noviser förlitade sig på expertens höga domänkunskap, och eftersom de ansåg att expertens search trail var hjälpsam, så kan denna tillit ha haft en positiv effekt. Denna förlitan kan ha grundats i novisers vilja att minska ångest och öka effektivitet, i kontexten av Isenbergs påståenden om experters sökförmåga. Det är vidare intressant att användarnas behov av expertens sökinmatningar minskar mot slutet av sökprocessen.

White, Dumais och Teevan (2009) noterade i deras studie att domännovisers vokabulär ökar under sökprocessen. De kunde konstatera i deras resultat att novisers vokabulär ökade mer än experters vokabulär (ibid.). I denna studie visar resultaten från den kvantitativa datan i den andra enkäten att noviser hade mindre nytta av expertens sökinmatningar mot slutet av sökprocessen. Den kvalitativa datan visade att detta berodde på att noviserna då hade ökad domänkunskap. Samtidigt visar loggfilsdatan från grupp B att 85,7 procent av sökinmatningarna innehöll expertbegrepp. Datat visar inte att användningen av expertbegrepp minskar vid slutet av sökprocessen. Noviserna använde alltså inte färre begrepp vid slutet av sökprocessen, men de upplevde att de hade mindre användning av expertens search trail då. Den kvantitativa datan från den andra enkäten visar även att enbart hälften av deltagarna ansåg att de hade återanvänt expertbegrepp, medan loggfilsdatan visar att samtliga deltagare återanvänt expertbegrepp. Detta skulle kunna bero på att novisernas vokabulär hade ökat under sökprocessen, och att begreppen då hade internaliseringats. Om begreppen hade internaliseringats då, så kan detta vara en förklaring till varför noviserna mot slutet av sökprocessen upplevde att de sökte mer självständigt. Detta skulle kunna peka mot att expertens search trail är en metaforisk språngbräda för noviser, och att de efter ökad domänkunskap själva upplever att de kan orientera sig i sökprocessen och förstå vilken information de behöver för att lösa sökuppgiften. Den kvalitativa datan visar att expertens search trail var hjälpsam, men den visar inte att användningen av den direkt ledde till rätt svar på sökuppgiften.

## 5.3 Svårigheter med källkritik

Loggfilsdatan visade att samtliga deltagare återanvände något eller flera av expertens sökbegrepp. Den kvalitativa datan från den andra enkäten visade som komplement till detta att noviserna ansåg att expertens search trail var hjälpsam. Problematiskt var samtidigt novisers förmåga att tänka källkritiskt. De upplevde att expertens search trail ledde dem till källor som inte var trovärdiga. Deltagare 11 förklarade att hen hade funnit svar på sökuppgiften, men att hen var osäker på om svaret stämde. Deltagare 18 skrev även att källorna som expertens search trail ledde till inte var trovärdiga, att de var ovanliga, och att deltagaren inte kände igen dem sedan tidigare.

Jenkins, Corritore och Wiedenbeck (2003) skriver att domänexperter kan i högre grad förklara varför de anser att källor är trovärdiga, och att de är mer bekväma med att använda främmande källor än vad domännoviser är. I kontrast föredrar domännoviser att använda källor de är bekanta med (ibid.). Datat från denna studie överensstämmer med dessa slutsatser från tidigare forskning. Deltagarna kopplar trovärdighet med bekantskapen av källor. Detta leder till frågan om hur det påverkar novisers nytta av en experts search trail. I studien av Wang, Zhao och Chen (2020) förslog noviserna inkludering av experters utvärdering av information i search traillen. På så sätt skulle noviserna kunna se vilka källor experten använde och varför (ibid.). Detta förslag är intressant i relation till datan i denna studie

eftersom det framgår att enbart experters sökinmatningar är otillräckliga för noviser att hitta trovärdig information. Om noviser, som till exempel deltagare 18, inte kan vara säkra på att lösningen till sökuppgiften stämmer, så kan nyttan av expertens search trail ifrågasättas.

## 5.4 Noviser undvek begrepp de inte förstod

Resultaten visade att deltagare undvek begrepp de inte förstod. Samtidigt svarade samma deltagare att de ansåg att deras förståelse för begreppen de använde ökade under sökprocessen, och att de var hjälpsamma. I Isenbergs studie undersöktes skillnaderna mellan experter och novisers sökbeteende i kontexten av affärslösningar. Han konstaterade att noviserna hade ett större behov av informationen som var tillgänglig än experterna (Ingwersen & Järvelin 2005). För att koppla detta till studien i fråga så undvek deltagarna begrepp de inte förstod i expertens search trail (den tillgängliga informationen), trots att de, enligt Isenberg, hade ett större behov av dem. Eftersom de samtidigt får en ökad förståelse för expertbegrepp när de använder dem, så kan vi identifiera utmaningen att uppmuntra noviser att använda begrepp de inte förstår, eftersom de troligen kommer att förstå dem under användning, och få nyttan av dem.

## 5.5 Begränsningar

Resultatet i studien begränsas av bortfallet. Ursprungligen var det 24 deltagare i utförandet, men detta förminkades till 20 deltagare på grund av svaren i den första enkäten. Vidare var det enbart 14 av dessa 20 deltagare som svarade på den andra enkäten. Detta begränsar givetvis resultaten eftersom förväntningen var att den insamlade datan skulle omfatta 24 deltagare. För att hantera detta presenterades datan från både grupp A och B eftersom återanvändningen av expertbegrepp ändå framkom i grupp A. Eftersom nyttja betraktas som återanvändning av expertbegrepp i denna studie var detta resultat relevant för frågeställningen. Diskussionen kring loggfilsdatan understöds även av grupp A, eftersom denna grups sökningar (med sex fler deltagare) i hög grad stämmer överens med sökningarna från grupp B.

I studiens SLA togs även beslut om vilka expertbegrepp som skulle inkluderas respektive exkluderas. Stavfel, plural av expertbegrepp, och skillnad i tempus av dem inkluderas i resultaten. Exempelvis betraktades *additiv färgbladning* som *additiv färgblandning*. För att identifiera dessa begrepp krävdes en manuell genomgång av loggfilerna, vilket på grund av den mänskliga faktorn kan ha lett till felmarginaler. Översättningar av expertbegreppen exkluderas. Ett exempel är när en deltagare skrev *color theory*. Detta skulle kunna vara en översättning av begreppet *färgteori* eller expertbegreppet *färglära*. Eftersom det inte är tydligt huruvida deltagaren syftade på färglära eller färgteori så kan *color theory* inte betraktas som ett expertbegrepp. Vidare förekom det att expertbegrepp var skrivna på svenska, och att deltagare översatte dem till svenska. De expertbegrepp som novisen använde räknades alltså enbart om språket de skrevs på överensstämde med expertens. Detta ledde till att ett antal begrepp som skulle kunna betraktas som expertbegrepp utelämnades.

## 5.6 Etiska och samhälleliga konsekvenser

Inga negativa etiska konsekvenser kan förutses med denna studie. Detta beror på att flertalet åtgärder togs för att anonymisera deltagare, och de blev även ombedda att skriva under medgivandeavtal. Deltagarna är således medvetna om vad deras deltagande i studien innebär, och de kan inte

identifieras. De samhälleliga konsekvenserna av studien är att den kan påverka framtida utveckling av informationssökning i sökmotorer. Studien kan vara till nytta för forskare inom informationssökning, och det är bland annat dessa personers arbeten som bidrar till utvecklingen av informationssökningssystem. Sökmotorer har utvecklats till att bli effektiva och ändamålsenliga medel för att söka information. Huurdeman och Kamps (2020) skriver att det krävs mer än enbart en sökinmatning för att lösa sökuppgifter vid uppgifter av hög komplexitet, och det är i denna kontext som synen på sökmotorers utveckling nu förändras. Vid komplexa sökuppgifter krävs dynamiskt stöd för användare i deras sökprocess (ibid.). Denna studie förankras i tidigare forskning som beskriver förslag om hur dessa stöd kan se ut. Studien utvecklar diskursen om vilken nytta användare har av search trails, och förser en grund för diskussionspunkter som kan användas vid utveckling av sökmotorer. Utmaningarna som identifierats, och frågorna som utvecklats, är användbara steg i forskningen om search trails.

## 5.7 Validitet

Validitet avser huruvida datan som samlats in är av rätt typ för att besvara frågeställningen, och om den har mätts korrekt (Denscombe 2014). Tidigare forskning har visat att det finns en relation mellan bristfällig källanvändning och källkritisk förmåga och låg domänkunskap. Utifrån detta utvecklades frågorna till den första enkäten. Resultaten visade att novisernas källanvändning och källkritiska förmåga inte var bristfällig. Det var alltså rätt typ av frågor som ställdes i det avseende att de behandlade källor, men validiteten kan ha ökat om de hade ställts på ett annat sätt. Vidare utvecklar Denscombe (ibid.) att validiteten ökar när *check questions* ställs. Den andra enkäten innehöll en fråga om när expertens search trail var till mest nytta, och en annan om när den var till minst nytta. Dessa frågor är *check questions* eftersom de hade kunnat identifiera motsägelser. Resultaten visade dock att det inte var några motsägelser i svaren, se figur 19.

Denscombe (2014) menar även att datans validitet kan minska vid manuell datainmatning. För att öka validiteten genomförs felkontroller i syfte att identifiera potentiella misstag. Vid data som genereras automatiskt är detta av mindre vikt (ibid.). Vad gäller loggfilerna samlades data in automatiskt, och den analyserades i övervägande del automatiskt. Detta minskade risken för manuella felinmatningar. Under delar av dataanalysen genomfördes dock en manuell felkontroll av hela datasetet avseende stavfel, plural av expertbegrepp, och skillnader i tempus. Detta gav ett mer inkluderande resultat, men introducerade också risken för felinmatning eftersom novisernas sökinmatningar manuellt korrigeras utefter expertbegreppen.

Bortfallet i studien resulterade i att 14 noviser deltog i hela testet istället för 24. Detta medförde en liten urvalsgrupp, vilket kan ha påverkat validiteten av studien negativt. Denscombe (2014) skriver att stora urvalsgrupper förser en god grund för generalisering i kvantitativa studier. Eftersom urvalsgruppen var liten så kan alltså möjligheterna för generalisering försämras. Dock, skriver Denscombe (ibid.) att perspektivet på detta i kvalitativ forskning skiljer sig från kvantitativ. I kvalitativ forskning fokuserar perspektivet mer på överförbarheten av upptäckterna (eng. *findings*) till andra instanser, snarare än den grad som upptäckterna existerar i andra instanser (ibid.). Överförbarheten av upptäckterna i denna studie, med avseende på bortfallet, blir således en uppgift för läsaren att bedöma.

Vidare är en utmaning vid validiteten av kvalitativ data är att en social miljö inte kan återskapas (Denscombe 2014). Det finns alltså ingen garanti att kvalitativ data är giltig, men andra åtgärder kan

tas för att öka validiteten. Denscombe (ibid.) skriver att kvalitativ data vanligen grundar sig i omfattande tidigare forskning. Resultaten stämmer i hög grad överens med tidigare forskning, och därmed ökar studiens validitet. Exempel på detta är att noviserna i denna studie upplevde svårigheter med källkritik och de angav även att deras domänkunskap ökade under sökprocessen.

## 5.8 Reliabilitet

Reliabilitet avser huruvida studier kan återskapas under olika tidpunkter och generera likadana resultat (Denscombe 2014). Den första enkäten behandlar novisernas domänkunskap. Eftersom elever på gymnasieskolor inte alltid kan förväntas ha lika domänkunskaper så kan inte detta resultatet återskapas. Syftet med den första enkäten var däremot att kartlägga användargruppen, snarare än att använda den som en filtrering för vilka noviser som fick delta i studien. Syftet med den andra enkäten kvantitativa del var däremot att generera resultat som var generaliserabara. Studien är förankrad i tidigare forskning och lämpliga metoder har valts utifrån detta. Även om till exempel sökuppgiften kan skiljas mellan studier, så ska resultaten (från den andra enkäten) som delvis besvarar frågeställningen vara möjliga att återskapa.

Datan som genererats av studiens loggfiler bedöms däremot vara svår att återskapa även med en likadan grupp noviser och sökuppgift. Detta på grund av att en novis kan variera sitt sökbeteende på flera olika sätt, både för till exempel sökinmatningarna och ordningen av dessa. Även sökresultaten som sökinmatningarna genererar kan förändras med korta tidsintervaller, och de kan uppdateras kontinuerligt. Detta innebär en utmaning för att återskapa resultaten från studiens tillämpning av datainsamlingsmetoden. Samtidigt är detta något som tidigare forskning behandlar, då till exempel novisers ökade vokabulär under sökprocessen påverkar studiers resultat (White, Dumais & Teevan 2009).

I kvalitativ forskning skiljer sig perspektivet på reliabilitet från i kvantitativ forskning. Fokuset skiftar från att kunna återskapa likadana resultat, till att återskapa studier som leder till likadana slutsatser (Denscombe 2014). För att kommunicera detta perspektiv på reliabiliteten av studien har metoder och verktyg redogjorts för utförligt. Forskare kan läsa om denna studie och själva bedöma om metoderna är tillförlitliga och slutsatserna resonliga (ibid.).

# 6 Slutsats

För denna studie var frågeställningen: ”Vad har en search trail genererad av en expert (hög domänkunskap) för effekt på novisers (låg domänkunskap) nytta för uppgifter på hög komplexitetsnivå?”. Slutsatsen är att en search trail genererad av en expert är delvis av nytta för noviser. Studien visar även att noviser upplever svårigheter med att bedöma om källor som expertens search trail leder till är trovärdiga. De upplever att expertens search trail är hjälpsam, men de är osäkra på huruvida de löst sökuppgiften. En fråga till en framtida studie är huruvida en experts search trail är av nytta för noviser om de inte är säkra på att de löst sökuppgiften? Vidare visar studien att novisers behov av expertens search trail minskar, på grund av ökad domänkunskap under sökprocessen. Expertens search trail är alltså av nytta under en begränsad tid. Framtida studier skulle kunna undersöka huruvida search trailen fungerar som en språngbräda för novisers domänkunskap och förmåga att söka självständigt. Studien visar även att noviser undviker begrepp de inte förstår, men upplever att de begrepp de använder är hjälpsamma. En utmaning för forskningsområdet blir således att utveckla tekniker som uppmuntrar noviser till att använda begrepp de inte förstår, eftersom de troligen senare har användning för dem, och kommer att förstå dem. Slutligen kan studien inte påvisa att en experts search trail är hjälpsam för uppgifter av hög komplexitet specifikt, och därmed är ett lämpligt förslag att i framtida forskning utforma en studie som analyserar återanvändningen av expertbegrepp som syftar på specifikt dimensionen i en sökuppgift.

# Referenser

Bagley, K. & Levkowitz, H. (2011). Roadmaps to success: Conceptual Mile Markers to accelerate exploratory search sessions. I *Emerging Technologies for a Smarter World (CEWIT)*. New York, USA 3 november 2011, ss. 94-98.  
DOI:10.1109/CEWIT.2011.6135877.

Byström, K. & Järvelin, K. (1995). Task Complexity Affects Information Seeking and Use. *Information Processing & Management*, 31(2), ss. 191-213.  
DOI:10.1016/0306-4573(95)80035-R

Capra, R., Arguello, J., O'Brien, H., Li, Y. & Choi, B. (2018). The Effects of Manipulating Task Determinability on Search Behaviors and Outcomes. I *SIGIR '18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. Ann Arbor, USA juli 2018, ss. 445-454.  
DOI:10.1145/3209978.3210047

Capra, R. & Arguello, J. (2019). Using Trails to Support Users with Tasks of Varying Scope. I *SIGIR'19: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Paris, Frankrike juli 2019, ss. 977-980.  
DOI:10.1145/3331184.3331290

Case, D. O. (2007). *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior*. 2. uppl., London: Elsevier, s. 80.

Charron-Bost, B., Mattern, F. & Tel, G. (1996). Synchronous, asynchronous, and causally ordered communication. *Distributed Computing*, 9, ss. 173-191.  
DOI:10.1007/s004460050018

Clarke, V., Braun, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology 2006*, 3(2), ss. 77-101.  
DOI:10.1191/1478088706qp063oa

Denscombe, M. (2014). *The Good Research Guide*. Berkshire: Open University Press, ss. 41-298.

Hendahewa, C. & Shah, C. (2017). Evaluating user search trails in exploratory search tasks. *Information Processing & Management*, 53(4), ss. 905-922.  
DOI:10.1016/j.ipm.2017.04.001

Hiemstra, D. (2009). Information Retrieval Models. I Göker, A. & Davies, J. (red.) *Information Retrieval: Searching in the 21st Century*. Hoboken: Wiley, ss. 1-19.  
DOI:10.1002/9780470033647

Hjerm, M., Lindgren, S. & Nilsson, M. (2014). *Introduktion till samhällsvetenskaplig analys*. 2. uppl. Malmö: Gleerups, ss. 110-116.

Huurde man H. C., Kamps, J. (2020). Designing Multistage Search Systems to Support the Information Seeking Process. I Fu, W. & van Oostendorp, H. (red.). *Understanding and Improving Information Search. Human–Computer Interaction Series*. Cham: Springer, ss. 113-137.  
DOI:10.1007/978-3-030-38825-6\_7.

Ingwersen, P. & Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context*. Dordrecht: Springer, ss. 1-85.

Jansen, B. J. (2008). The methodology of search log analysis. I Jansen, B. J., Spink, A. & Taksa, I. (red.). *Handbook of Research on Web Log Analysis*. Hershey: IGI Global, ss. 100-123.  
DOI:10.4018/978-1-59904-974-8.ch006

Jenkins, C., Corritore, C. L. & Wiedenbeck, S. (2003). Patterns of information seeking on the Web: A qualitative Study of domain expertise and Web expertise. *IT & Society*, 1(3), ss. 64-89.

Johannesson, P. & Perjons, E. (2014). *An introduction to design science*. Stockholm: Springer International Publishing AG, ss. 56-57.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5), ss. 361-371.

Marchionini, G. (1996). Information Seeking in Electronic Environments. *Journal of Education for Library and Information Science*, 9(1), ss. 26-60.

DOI:10.2307/40324289

Pirolli, P. & Card, S. K. (1999). Information Foraging. *Psychological Review*, 106(4), ss. 643-675.  
DOI:10.1037/0033-295X.106.4.643

Sanderson, M. & Croft, B. M. (2012). The History of Information Retrieval Research. I *Proceedings of the IEEE v100 (Special Centennial Issue)*. 13 maj 2012, ss.1444-1451  
DOI:10.1109/JPROC.2012.2189916.

Shah, C. & Marchionini, G. (2009). Query Reuse in Exploratory Search Tasks. I *Proceedings of the Third Workshop on Human-Computer Interaction and Information Retrieval*. Washington, D.C., USA 23 oktober 2009, ss. 91-94.

Wang, A., Zhao, Y. & Chen, Y. (2020). Information Search Trail Recommendation Based on Markov Chain Model and Case-based Reasoning. *Data and Information Management*, 5(1), ss. 228-241.  
DOI:10.2478/dim-2020-0047

White, R. W., Dumais, S. T. & Teevan, J. (2009). Characterizing the influence of domain expertise on web search behavior. I *WSDM '09: Proceedings of the Second ACM International Conference on Web Search and Data Mining*. Barcelona, Spanien februari 2009, ss. 132-141.

Wildemuth, B., Freund, L. & Toms, E. G. (2014). Untangling search task complexity and difficulty in the context of interactive information retrieval studies. *Journal of Documentation*, 70(6), ss. 1118-1140.  
DOI:10.1108/JD-03-2014-0056

# Bilaga A – Medgivandeavtal för experten (SearchX)

## Samtycke till deltagande i användartest (experten)

Nedan ger du ditt samtycke till att delta i ett användartest där användarnas interaktion med systemet SearchX dokumenteras i form av loggfiler. Du kan när som helst under testet avbryta ditt deltagande, utan att delge anledning. Läs igenom detta avtal noga innan du skriver under längst ned på nästa sida. Kom ihåg att du behöver vara minst 18 år gammal för att delta i studien och ge ditt samtycke. Du har tillfrågats för deltagande i denna studie eftersom du anses tillhöra målgruppen expert.

Syftet med studien är att undersöka nyttan noviser har av en experts search trail för en sökuppgift av hög komplexitet. Noviser definierar vi som personer med låg kunskap inom en viss domän, och experter definierar vi som personer med hög kunskap inom en viss domän. En search trail är en visuell representation av tidigare användares sökinmatningar. En sökuppgift är problemet som användaren försöker lösa. Det kan till exempel vara en fråga användaren vill ha svar på. Testet genomförs i systemet SearchX, där du som expert placeras i en egen session med en sökuppgift. Efter att du genomfört sökuppgiften ombeds du skriva ned de begrepp som kan kategoriseras som expertbegrepp, och skicka dem till oss (som genomför studien).

Datan som samlas in är helt avgränsad till SearchX, och inkluderar till exempel sökinmatningar och klick. Data från andra sökmotorer dokumenteras alltså inte. Din IP-adress dokumenteras inte, utan du tilldelas ett autogenererat användar-id och grupp-id. Daten kommer att användas i en slutgiltig rapport för studien, men den är anonym och kan inte kopplas till specifika personer. Vi (som genomför studien) kommer vara tillgängliga för frågor under testet via samtal i Teams, men detta dokumenteras inte (det kommer alltså inte att spelas in). Syftet med vår digitala närvaro är enbart att fungera som stöd.

### Kontakt:

Vid frågor angående ditt deltagande kan du vända dig till oss som genomför studien.

Nedan finner du våra namn och vår gemensamma mejladress:

Stella Millwood  
Adam Gillström  
[examensarbetevt21@gmail.com](mailto:examensarbetevt21@gmail.com)

**Samtycke:**

Jag har tagit del av informationen kring studien och är medveten om hur den kommer att gå till.

Jag har fått tillfälle att få mina frågor angående studien besvarade innan den påbörjas och vet vem jag ska vända mig till med frågor.

Jag deltar i denna studie helt frivilligt och har blivit informerad om varför jag har blivit tillfrågad och vad syftet med deltagandet är.

Jag är medveten om att jag när som helst under studiens gång kan avbryta mitt deltagande utan att jag behöver förklara varför.

Jag är medveten att mitt deltagande dokumenteras i form av loggfiler och ska bearbetas och presenteras anonymt, i forskningssyfte.

\* Required

Ger du ditt samtycke till att delta i studien och till att dina anonyma uppgifter dokumenteras och bearbetas i forskningssyfte? \*

Ja

Nej

**Submit**

# Bilaga B – Medgivandeavtal för noviser (SearchX)

## Samtycke till deltagande i användartest (noviser)

Nedan ger du ditt samtycke till att delta i ett användartest där användarnas interaktion med systemet SearchX dokumenteras i form av loggfiler. Du kan när som helst under testet avbryta ditt deltagande, utan att delge anledning. Läs igenom detta avtal noga innan du skriver under längst ned på nästa sida. Kom ihåg att du behöver vara minst 18 år gammal för att delta i studien och ge ditt samtycke. Du har tillfrågats för deltagande i denna studie eftersom du anses tillhöra målgruppen novis.

Syftet med studien är att undersöka nyttan noviser har av en experts search trail för en sökuppgift av hög komplexitet. Noviser definierar vi som personer med låg kunskap inom en viss domän, och experter definierar vi som personer med hög kunskap inom en viss domän. En search trail är en visuell representation av tidigare användares sökinmatningar. En sökuppgift är problemet som användaren försöker lösa. Det kan till exempel vara en fråga användaren vill ha svar på. Testet genomförs i systemet SearchX, där du som elev placeras i en egen session med lärarens sökinmatningar från ett tidigare test.

Datan som samlas in är helt avgränsad till SearchX, och inkluderar till exempel sökinmatningar och klick. Data från andra sökmotorer dokumenteras alltså inte. Din IP-adress dokumenteras inte, utan du tilldelas ett autogenererat användar-id och grupp-id. Daten kommer att användas i en slutgiltig rapport för studien, men den är anonym och kan inte kopplas till specifika personer. Vi (som genomför studien) kommer vara tillgängliga för frågor under testet via samtal i Teams, men detta dokumenteras inte (det kommer alltså inte att spelas in). Syftet med vår digitala närvaro är enbart att fungera som stöd.

### Kontakt:

Vid frågor angående ditt deltagande kan du vända dig till oss som genomför studien.  
Nedan finner du våra namn och vår gemensamma mejladress:

Stella Millwood  
Adam Gillström  
[examensarbetevt21@gmail.com](mailto:examensarbetevt21@gmail.com)

Samtycke:

Jag har tagit del av informationen kring studien och är medveten om hur den kommer att gå till.

Jag har fått tillfälle att få mina frågor angående studien besvarade innan den påbörjas och vet vem jag ska vända mig till med frågor.

Jag deltar i denna studie helt frivilligt och har blivit informerad om varför jag har blivit tillfrågad och vad syftet med deltagandet är.

Jag är medveten om att jag när som helst under studiens gång kan avbryta mitt deltagande utan att jag behöver förklara varför.

Jag är medveten att mitt deltagande dokumenteras i form av loggfiler och ska bearbetas och presenteras anonymt, i forskningssyfte.

\* Required

Vänligen klistra in ditt grupplD. (Du kan hitta det längst upp i instruktionerna) \*

Your answer

Ger du ditt samtycke till att delta i studien och till att dina anonyma uppgifter dokumenteras och bearbetas i forskningssyfte? \*

Ja

Nej

Submit

# Bilaga C – Medgivandeavtal för experten (E-postkontakt)

## **Samtycke till datainsamling i form av mejlkontakt**

Nedan ger du ditt samtycke till att svara på ett antal frågor via e-postkontakt. Syftet med detta moment är att säkerställa att du har den nivå av domänexpertis som studien förutsätter. Du kan när som helst avbryta ditt deltagande, utan att delge anledning. Läs igenom detta avtal noga innan du skriver under längst ned på denna sida. Dina svar presenteras i slutrapporten för studien anonymt.

### **Kontakt**

Vid frågor angående ditt deltagande kan du vända dig till oss som genomför studien.

Nedan finner du våra namn och vår gemensamma mejladress:

Stella Millwood  
Adam Gillström  
examensarbetevt21@gmail.com

### **Medgivande**

- Jag har tagit del av informationen kring studien och är medveten om hur den kommer att gå till.
- Jag har fått tillfälle att få mina frågor angående studien besvarade innan den påbörjas och vet vem jag ska vända mig till med frågor.
- Jag deltar i denna studie helt frivilligt och har blivit informerad om varför jag har blivit tillfrågad och vad syftet med deltagandet är.
- Jag är medveten om att jag när som helst under studiens gång kan avbryta mitt deltagande utan att jag behöver förklara varför.
- Jag är medveten att mitt deltagande dokumenteras i form av loggfiler och ska bearbetas och presenteras anonymt, i forskningssyfte.

Stockholm den \_\_\_\_ / \_\_\_\_ 2021

### **Deltagare:**

\_\_\_\_\_  
\_\_\_\_\_  
Namnteckning/Namnförtydligande

### **Testledare:**

\_\_\_\_\_  
\_\_\_\_\_  
Namnteckning/Namnförtydligande  
\_\_\_\_\_  
\_\_\_\_\_  
Namnteckning/Namnförtydligande

# Bilaga D – Medgivandeavtal för den första enkäten (noviser)

## **Samtyckesblankett för Enkät-undersökning alt. Elektroniska val**

Stockholms universitet är personuppgiftsansvarig för den behandling av personuppgifter som sker i Survey and Report.

Den lagliga grunden för behandlingen är att du gett samtycke till behandlingen i någon enkät.

För att projektet ska kunna utföras kommer ansvarig för studien och ansvarig för projektet ha tillgång till personuppgifterna. Uppgifterna kommer att behandlas så att inte obehöriga kan ta del av dem.

Om detta material bedöms ha ett bestående värde enligt de riktlinjer som anges i 6-8 §§ i RA-FS 1999:1 kommer det att bevaras för framtiden.

Enligt EU:s dataskyddsförordning samt nationell kompletterande lagstiftning har du rätt att:

återkalla ditt samtycke utan att det påverkar lagligheten av behandling som skett i enlighet med samtycket innan det återkallades,

begära tillgång till dina personuppgifter,

få dina personuppgifter rättade,

få dina personuppgifter raderade,

få behandlingen av dina personuppgifter begränsad.

Under vissa omständigheter medger dataskyddsförordningen samt kompletterande nationell lagstiftning undantag för dessa rättigheter, som kan komma att tillämpas.

Om du vill åberopa någon av dessa rättigheter Kontakta Dataskyddsombudet vid Stockholms universitet ([dso@su.se](mailto:dso@su.se)).

Mer information om detta finns på Integritetsskyddsmyndigheten (tidigare Datainspektionen) hemsida. <https://www.imy.se/>

Jag har läst och accepterar villkoren

Jag nekar samtycke

Jag samtycker

# Bilaga E – Medgivandeavtal för den andra enkäten (noviser)

## **Samtyckesblankett för Enkät-undersökning alt. Elektroniska val**

Stockholms universitet är personuppgiftsansvarig för den behandling av personuppgifter som sker i Survey and Report.

Den lagliga grunden för behandlingen är att du gett samtycke till behandlingen i någon enkät.

För att projektet ska kunna utföras kommer ansvarig för studien och ansvarig för projektet ha tillgång till personuppgifterna.

Uppgifterna kommer att behandlas så att inte obehöriga kan ta del av dem.

Om detta material bedöms ha ett bestående värde enligt de riktlinjer som anges i 6-8 §§ i RA-FS 1999:1 kommer det att bevaras för framtiden.

Enligt EU:s dataskyddsförordning samt nationell kompletterande lagstiftning har du rätt att:

återkalla ditt samtycke utan att det påverkar lagligheten av behandling som skett i enlighet med samtycket innan det återkallades,

begära tillgång till dina personuppgifter,

få dina personuppgifter rättade,

få dina personuppgifter raderade,

få behandlingen av dina personuppgifter begränsad.

Under vissa omständigheter medger dataskyddsförordningen samt kompletterande nationell lagstiftning undantag för dessa rättigheter, som kan komma att tillämpas.

Om du vill åberopa någon av dessa rättigheter Kontakta Dataskyddsombudet vid Stockholms universitet ([dso@su.se](mailto:dso@su.se)).

Mer information om detta finns på Integritetsskyddsmyndigheten (tidigare Datainspektionen) hemsida. <https://www.imy.se/>

Jag har läst och accepterar villkoren

Jag nekar samtycke

Jag samtycker

# Bilaga F – Kategorier och teman

Teman/Kategorier	Reviderade teman	Definierade teman
Sökuppgiften var utmanande	Låg domänkunskap med svårigheter för källkritik och förståelse för sökuppgift	Samband mellan låg förståelse för sökuppgiften och låg domänkunskap
Positiv inställning till sökresultat	Search trail var hjälpsam och vägledande	Noviser upplevde att search trailen vägledde dem i början av sökprocessen, och efter ökad domänkunskap minskade behovet av att använda den
Hjälpsamt att använda search trail	Domänkunskap ökar under sökprocessen	Expertens search trail var hjälpsam, men noviser upplevde svårigheter med källkritik
Källkritik	Undvek att ta hjälp av search trail	Noviser undvek begrepp de inte förstod, men upplevde att begreppen de använde var hjälpsamma
Låg kunskap och erfarenhet i domänen		
Search trail var vägledande		
Domänkunskap ökar under sökprocessen		
Förståelse för expertbegrepp ökar under sökprocessen		
Delvis positiv till search trails hjälpsamhet		
Sökstrategi baserad på vana		
Undvek okända expertbegrepp		

# Bilaga G – Expertens search trail

## ⌚ Recent queries

x

- ⌚ digital skärm färgsystem
- ⌚ ljusskärm
- ⌚ cmyk tryck
- ⌚ färgcirkel
- ⌚ färglära
- ⌚ ittens färglära
- ⌚ rgb cmyk skillnad
- ⌚ rgb vs cmyk
- ⌚ which color system to use
- ⌚ Färgsystem tillämpningar
- ⌚ LAB colors photoshop
- ⌚ LAB colors
- ⌚ cielab
- ⌚ Lab color system
- ⌚ NCS color system uses
- ⌚ NCS color system
- ⌚ additiv färgblandning
- ⌚ subtraktiv färgblandning
- ⌚ cmyk färgsystem
- ⌚ rgb färgsystem
- ⌚ rgb
- ⌚ färgsystem

# Bilaga H – Gränssnittet i SearchX

The screenshot shows the SearchX search interface. At the top left is the "SearchX" logo. To its right is a search input field with a magnifying glass icon. Above the input field, there are two lines of text: "User id: ff1be697ea50114467c047522ca5d4" and "Group id: 50b0014e74e587c8b1ae1ced882ef5 ?". A "Reset group" button is located next to the group ID. On the right side of the interface, there are two panels: "Recent queries" (with a list of terms like "digital skärm färgsystem", "ljusskärm", etc.) and "Saved documents" (which is currently empty). Below these panels, a message reads "Your search results will appear here :)".

User id: ff1be697ea50114467c047522ca5d4  
Group id: 50b0014e74e587c8b1ae1ced882ef5 ? [Reset group](#)

Recent queries

- digital skärm färgsystem
- ljusskärm
- cmyk tryck
- färgcirkel
- fäglära
- ittens fäglära
- ... och annan aktuellt

Saved documents

Your search results will appear here :)

About [SearchX](#).



# Bilaga I – Den första enkäten

Vänligen klista in ditt gruppID. (Du kan hitta det längst upp i instruktionerna)

Är du minst 18 år gammal?

- Ja
- Nej

Är du flytande i svenska i tal och skrift?

- Ja
- Nej

Har du studerat eftergymnasiala studier i bild och design?

- Ja
- Nej

Hur länge söker du information inom domänen bild och design, via en internetbaserad sökmotor?

- Mindre än 1 timme per vecka
- 1 till 5 timmar per vecka
- 6 till 10 timmar per vecka
- Mer än 10 timmar per vecka

När du söker information för en skoluppgift: hur många källor i genomsnitt kollar du på innan du är nöjd?

- 1
- 2-3
- 4-5
- Fler än 5

Vilka av följande källor brukar du använda när du söker information för en skoluppgift?

- Läromedlet för kursen (t.ex. kursboken)
- Nyhetsartiklar
- Wikipedia
- Forum så som Reddit eller Flashback
- Youtube
- Tidsskriftartiklar
- Publicerade rapporter
- Facklitteratur
- Annat:

Vad gör att du litar på informationen?

- Du litar på allt som kommer från ursprungskällan
- Du tycker att informationen stämmer med det du redan vet
- Du kollar om källan använder referenser
- Du använder källorna som en Wikipedia- eller nyhetsartikel refererar till
- Du har hittat källan genom en sökmotor för vetenskaplig litteratur (t.ex. Google Scholar eller Diva)
- Du ser att källan är "peer-reviewed"
- Du läser på om författarens bakgrund och motiv
- Du har inga kriterier för att avgöra om en källa är pålitlig
- Annat:

[Skicka nu](#)

# Bilaga J – Den andra enkäten

Detta är den andra enkäten i studien. Nedan presenteras en lista med förklaringar till begrepp vi använder i frågorna.

Novis = Person med låg kunskap inom en viss domän.

Expert = Person med hög kunskap inom en viss domän.

Sökinmatningar = De ord eller fraser som matas in i SearchX. För denna studie används expertens tidigare sökinmatningar.

Sökuppgift = Problemet som användaren försöker lösa. Det kan till exempel vara en fråga användaren vill ha svar på.

Sökprocess = I denna studie är sökprocessen det som sker från att du börjar försöka lösa sökuppgiften i SearchX fram till att du upplever att du är klar med att försöka lösa sökuppgiften.

**Vänligen klistica in ditt gruppID. (Du kan hitta det längst upp i instruktionerna)**

**På en skala 1-5, hur svår tyckte du att sökuppgiften var att lösa?**

- 1
- 2
- 3
- 4
- 5

Vänligen motivera ditt svar.

/

**På en skala 1-5, hur säker är du på att du lyckades lösa sökuppgiften?**

- 1
- 2
- 3
- 4
- 5

Vänligen motivera ditt svar.

/

**På en skala 1-5, hur hjälpsamma var expertens sökinmatningar för att lösa sökuppgiften?**

- 1
- 2
- 3
- 4
- 5

Vänligen motivera ditt svar.

/

**Återanvände du några begrepp från expertens sökinmatningar?**

- Ja  
 Nej

**Vad var anledningen? Vänligen beskriv problemet bakom varför du valde att återanvända begreppen.**

---

/

**Varför tror du att problemet uppstod?**

---

/

**Hur påverkade begreppsanhäntningen din förmåga att söka relevant information?**

---

/

**Var det några begrepp i expertens sökinmatningar som du inte förstod?**

- Ja  
 Nej

**Vänligen beskriv hur du hanterade detta.**

---

/

**När under sökprocessen hade du mest nytta av expertens sökinmatningar?**

- I början av sökprocessen  
 I mitten av sökprocessen  
 I slutet av sökprocessen

**Vänligen motivera ditt svar.**

---

/

**När under sökprocessen hade du minst nytta av expertens sökinmatningar?**

- I början av sökprocessen  
 I mitten av sökprocessen  
 I slutet av sökprocessen

**Vänligen motivera ditt svar.**

---

/

**Tycker du att detta har varit ett användbart sätt att söka information? Alltså, att ha en experts sökinmatningar tillgängliga? Vänligen motivera ditt svar.**

---

/