

LINKÖPINGS UNIVERSITET
Institutionen för Datavetenskap

Studie av gränssnittsprototyp i projektet
Webbklustring
- användarupplevelsen

Namn Evelina Rennes
E-mail `evere305@student.liu.se`

Innehåll

1	Inledning	1
2	Bakgrund	1
2.1	Komplexitetssmåt	1
2.2	Beskrivning av systemet	1
3	Genomförande	2
3.1	Product Reaction Cards	2
3.2	Uppgifter	2
4	Resultat	3
4.1	Förvirrande	3
4.2	Ofärdigt	4
4.3	Användbart	5
4.4	Långsamt	5
5	Diskussion	6
5.1	Resultatdiskussion	6
6	Slutsats	7
	Referenser	7
A	Gränssnitt: Klusternivå	8
B	Gränssnitt: Sökning	9
C	Gränssnitt: Hovring över kluster	10
D	Gränssnitt: Dokumentnivå	11
E	Product Reaction Cards -ordpar	12

1 Inledning

Projektet Webbklustring som drivs på Linköpings universitet är ett webbaserat system vars syfte är att effektivisera informationssökning. Ett övergripande mål att bistå ovana användare eller personer med läs- och skrivsvårigheter med ett enkelt verktyg för att underlätta informationssökning på internet. Det system som utvecklas skall kunna jämföra hemsidor med samma innehåll och således utesluta dubletter, men också sammanfatta texter och avgöra hemsidors läsbarhet. Denna undersökning har ämnat undersöka användarnas upplevelse av den andra gränssnittsprototypen inom projektet och ge förslag på eventuella förbättringar.

2 Bakgrund

Projektet är delvis baserat på resultat från tidigare projekt: Webblättnäst och FriendlyReader. Webblättnäst rangordnar texter efter komplexitetsmått och FriendlyReader sammanfattar texter i syfte att göra dem mer lättlästa.

2.1 Komplexitetsmått

De komplexitetsmått som tillämpas i systemet är läsbarhetsindex (LIX) och ordvariationsindex (OVIX). LIX är ett i svenskan vedertaget mått som baseras på antalet långa ord (ord längre än sex bokstäver) och medelantalet ord per mening.

OVIX ser till hur många unika ord som finns i en text jämfört med totalt antal ord. En text som har stor ordvariation anses vara mer svårläst än en text med liten ordvariation, och till skillnad från LIX-värdet tas ingen hänsyn till ordens längd.

2.2 Beskrivning av systemet

Gränssnittet (se Appendix A) bestod av ett antal sektioner; en sökfunktion, ett rutnät som representerar sökresultaten visuellt samt en dokumentlista. Sökfunktionen fanns lokaliserat högst upp på sidan och består av en textruta och, till höger om rutan, ett reglage där användaren kontrollerar hur många sökresultat som ska visas.

Vid sökning sorteras dokumenten i kluster baserat på likhet och sökresultaten representeras på två sätt: som en hexagon i rutnätet samt i en klassisk dokumentlista baserat på relevans (se Appendix B). Dokumentets placering på rutnätet beror på vilken läsbarhet och relevans dokumentet har.

Så väl rutnätet som dokumentlistan går att sortera efter det komplexitetsmått som önskas användas (LIX eller OVIX) och dokumentlistan kunde utöver det sorteras efter relevans. Defaultvärdet för både rutnätet och dokumentlistan vid sökning är LIX.

När muspekaren förs över ett kluster visas en informationsruta innehållande en multidokumentsammanfattning av dokumenten i det specifika klustret samt dess olika komplexitetsvärden, återgivna i ord och siffror (se Appendix C).

Genom ett musklick på ett kluster öppnas det specifika klustret upp och visade

de unika dokumentens placering i samma rutnät som beskrivits ovan (se Appendix D). Genom att hålla muspekaren över en hexagon i rutnätet, motsvarande en hemsida, visas en informationsruta liknande den på klusternivå med komplexitetsmått samt en snippet. Genom att klicka på ett dokument i rutnätet öppnas hemsidan i fråga.

3 Genomförande

Deltagarna bestod av 10 studenter utan läs- eller skrivsvårigheter från Linköpings universitet, varav 6 var män och 4 var kvinnor. Åldern på deltagarna varierade mellan 20 och 27 ($M = 23$, $SD = 2,6$) och samtliga deltagare ansåg sig ha god datorvana.

Innan testet blev deltagarna deltagarna informerade om att de var fullständigt anonyma, att de hade möjlighet att avbryta testet när de ville samt att resultat inte skulle kunna spåras tillbaka till dem.

3.1 Product Reaction Cards

Product Reaction Cards syftar till de 118 kort som utvecklades av Benedek och Miner från Microsoft (2002) men i denna studie användes en anpassad version av dessa kort.

De ord som användes togs fram genom en fokusgruppsdiskussion av personer som har varierande koppling till projektet Webbklustring. Samtliga deltagare i diskussionen var bekanta med systemet. Den uppsättning av kort som användes bestod av 20 olika ord med antingen en positiv eller negativ klang. Samtliga ord återfinns i Appendix E.

Försöksdeltagaren fick efter testsessionen välja ut ett godtyckligt antal kort som de ansåg beskriva systemet bäst. Ur dessa fick de sedan ranka de tre ord som de ansåg var mest representativa från 1 till 3 där 1 var mest representativt och 3 minst representativt. De fick sedan motivera valet av kort verbalt och intervjun spelades in med en ljudupptagningsenhet.

3.2 Uppgifter

Uppgifterna utformades i syfte att motsvara systemets samtliga funktioner och togs fram av personer som arbetar med projektet Webbklustring.

De uppgifter som användes var:

- Hitta den mest läsbara hemsidan om Helan och Halvan enligt LIX.
- Hitta det kluster som är det mest relevanta.
- Hitta den hemsida är den mest relevanta.
- Hitta det kluster som är det mest lättlästa.
- Vilken sammanfattning tycker du är bäst och varför?
- Hitta en sida som beskriver hur du lagar en punktering.

- Hur många olika typer av svar på frågan hur du lagar en punktering hittar du?
- Hur många olika kluster hittar du?

4 Resultat

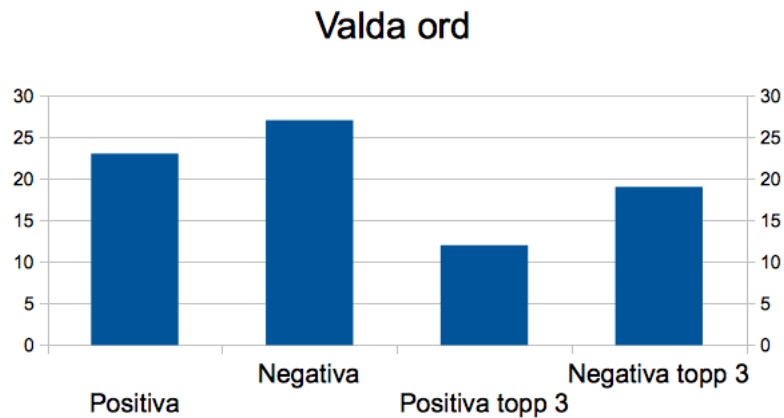
Av de 20 ord som deltagarna fick välja mellan för att beskriva systemet blev 17 ord valda. De ord som aldrig valdes var *Värdelöst*, *Färdigt* samt *Motarbetande*.

Ord	Antal totalt	Antal relevanta
Förvirrande	8	7
Ofärdigt	5	3
Användbart	5	3
Långsamt	4	4

Tabell 1: Antal personer som valde det specifika ordet för att beskriva systemet samt antal personer som valde ordet till de tre ord som bäst beskrev systemet.

De mest valda orden var *Förvirrande* (8), *Ofärdigt* (5) samt *Användbart* (5). Av de ord som deltagarna rankade som mest relevanta valdes *Förvirrande* (7), *Långsamt* (4), *Användbart* (3) samt *Ofärdigt* (3).

Det var övervägande negativa ord som valdes ut för att beskriva systemet.



Figur 1: Antalet ord med positiv respektive negativ laddning som valdes av försöksdeltagarna.

4.1 Förvirrande

Flertalet deltagare fann systemet något förvirrande. Anledningen till denna förvirring verkade framför allt vara problem med den terminologi som användes vid testningen. Här har användaren problem med att förstå vad ordet kluster syftar till:

P01: För det första förstod ja inte va kluster va först. Det är för visso inom parentes på förklaring. Sen så tyckte ja de var förvirrande att man va tvungen att gå in i alla de olika grupperna liksom. För att komma till svaren.

Begreppen läsbarhet och relevans var också problematiska:

P05: Förvirrande kändes som de asså man visste inte riktigt va de va dom här olika asså va man ville mena va som menades med de här olika asså relevans och läsbarhet. Speciellt inte läsbarhet. Relevans känns ju tydligt för de har man ju hört om innan.

P09: Även lite förvirrande de va läsbarhetsmått som ja inte har koll på alls va de är för nåt överhuvudtaget. (...) förvirrande va ba dom här mått som ja ska välja olika läsbarhets va de nu va så vet jag inte va de är för nåt å det kan ju förvirra lite. När ska jag söka me link lix link nå jag kommer inte ihåg va de hette men när ska jag använda de olika och jag vet inte va de är för skillnad på dom eller nånting så såna saker är ju förvirrande för mig som inte vet va de är.

Ett fåtal av de deltagare som beskrev systemet som förvirrande grundade sin åsikt i det grafiska sättet att representera sökresultaten, vilket indikeras i följande motiveringar:

P10: Man kan ju lätt ta till sig det och hur upplägget är men samtidigt så blir de förvirrande för att jag tror att jag förutsäger att det är på ett sätt å då blir det förvirrande när det sen inte är de på det sättet. Till exempel med a men bara grafiskt knapparna vilken knapp jag kan trycka på för att komma tillbaka ehh det tog ett tag innan jag fatta att ja kunde trycka på plupparna på själva plupparna så jag antog att jag kunde trycka och sen så vart jag förvirrad när ja skulle gå tillbaka så.

P06: De var lite förvirrande att till exempel de här me teman å eller olika typer ehh de va lite förvirrande.

En deltagare beskrev designen som förvirrande, men framhöll även att det kunde vara hjälpsamt att representera sökresultat grafiskt.

P07: Jämför man me typ google som sökmotor så känner ja väl att de ä jobbigt att få den här va ska man säga den här grafiska layouten. Just att man ser såhär kluster istället för att se direkta sökresultat. Men på samma gång så kan de till viss del vara hjälpsamt att se asså väldigt tydligt att amen dom här resultaten finns och framförallt med ehh relevans och läsbarhetsaxeln så är det ju på ett sätt en fördel gentemot typ google som sökmotor.

4.2 Ofärdigt

De deltagare som beskrev systemet som ofärdigt beskrev ett antal faktorer som påverkade upplevelsen av systemet. Systemet beskrevs som långsamt och som att det hängde upp sig vid körningarna:

P05: Asså ofärdigt känns som att man kan göra nånting bra av de för det kändes som om att de är en bra grej att ha grafer å så. Och därför kändes det som att det var ett ofärdigt system. Likadant att de va långsamt och sådär. Å hängde upp sej.

Behovet av utförligare instruktioner uttrycktes:

P06: Ofärdigt. Det kändes att det va nånting som fattades. De skulle behöva lite mer instruktioner i början. Lite mer förklaring. Men man lär sig efter man har klickat lite.

P10: Ja först tänker ja att de är lite ofärdigt de hör ihop med att det är förutsägbart systemet. (...) Så då känns det lite ofärdigt att de kanske lite mer tydlighet grafiskt sett men annars så fungerade ju programmet väldi bra

upplägget var bra.

4.3 Användbart

Flera deltagare ansåg att systemet var eller skulle kunna vara användbart, vilket indikeras i följande motiveringar:

P04: Det var ju väldigt annorlunda mot en sökmotor som ja brukar ju använda google. Så det är väldigt annorlunda och liksom använda. Och sen tycker ja det är ju liksom nånting som du ska utveckla så det märks ju att de inte är färdigt. Det kan säkert bli bra och därför tog ja användbart också fö jag tror att det kan det kan säkert eh fungera och va lättförståeligt liksom när det blir färdigt men ja tyckte de va lite förvirrande.

En person uttryckte ett specifikt användningsområde för systemet inom läraryrket:

P06: Användbart. Ja ser absolut användningen för det här den här typen. Jag kan tänka mej att använda de själv som lärare. Just de är jättebra de här med eh att de sorteras på lix-värde eftersom de ja de behöver ju ja veta för att ge rätt texter till mina elever.

Även grupperingen av sökresultat samt den grafiska layouten uttrycktes vara användbara:

P09: Vi börjar me användbart för de verkar som att systemet kan kunna va när de ä fulländat. (...) Mm asså användbart jag ser ett syfte för de. På det sättet är det användbart. Man kan ja gillar det här me att man kan gruppera saker ja vet inte hur dom grupperas men det känns användbart att man kan gruppera saker och att man kan visuellt se ehh olika grupper och hur den relevans som dom har.

P10: Absolut användbart för dom som behöver lite mer grafiskt se vad en sökning innebär och hur sökträffar hänger ihop å så. Kanske för unga människor eller äldre. Så de va absolut förståeligt.

4.4 Långsamt

Utöver de mest valda orden valdes även ordet *Långsamt* att beskriva systemet. Det som framför allt klagades på var den långa väntetiden från att deltagaren hade tryckt på *Sök* till att resultaten dök upp, samt att systemet ibland hakade upp sig.

P05: Aa asså långsamt kändes det som det gick fö att ja de haka ju upp sig. Det hände ju inte så mycke så rätt segt så. Ändå fick man inte så många svar typ tretti.

P06: Långsamt va att vi fick eller du fick starta om sidan ett par gånger.

P07: Lite lätt långsamt kanske. Den kan vi sätta som tredje istället så gör vi. Just för de tog sån tid att för att få fram sökresultaten.

5 Diskussion

Den metod som användes i denna studie togs fram med inspiration av de 118 reaktionskort framtagna av Benedek och Miner från Microsoft. Det fanns dock flertalet signifikanta skillnader mellan metoderna, vilket innebär att den version av reaktionskortet som användes i denna studie inte är en erkänd eller etablerad metod. För att hantera de problem som detta kan innebära togs ordparen fram med hjälp av en fokusgruppsdiskussion, och förhoppningen är att detta innebar en högre reliabilitet än om de hade varit godtyckligt valda.

Ett problem som kan uppstå vid användning av reaktionskort är att ord kan ha olika betydelser för olika personer. Detta hanterades genom en inspelad intervju, där deltagarna fick motivera sitt ordval och ge en utförligare beskrivning av ordets betydelse.

5.1 Resultatdiskussion

Det var övervägande negativa ord som valdes för att beskriva systemet. Detta indikerar att upplevelsen av systemet inte var tillfredsställande och att förbättringar måste göras.

De mest valda orden var *Förvirrande* (8), *Ofärdigt* (5) samt *Användbart* (5). Av de ord som deltagarna rankade som mest relevanta valdes *Förvirrande* (7), *Långsamt* (4), *Användbart* (3) samt *Ofärdigt* (3).

Att systemet upplevdes som något förvirrande är tydligt då 8 av 10 personer valde att låta ordet representera systemet och 7 av 10 valde det som ett av de mest relevanta orden. Då deltagarna fick motivera sitt ordval blev det tydligt att det framförallt var problem med terminologin som ställde till problem. Ord som kluster, läsbarhet och relevans tycktes vara otydliga, vilket försämrade upplevelsen av systemet.

En ytterligare orsak till rådande förvirring ansågs bero på den grafiska representationen av sökresultaten. Det var inte självklart hur systemet fungerade, till exempel att det var möjligt att öppna kluster för att se de enskilda dokumenten. Att systemet inte inbjuder till korrekt användning är tydligt och bör åtgärdas.

Hälften av alla deltagare upplevde systemet som ofärdigt. Det som framför allt motiverade detta ordval var rent tekniska problem, så som att sökningsprocessen tog lång tid och att servern kraschade vid sökning. Ingen deltagare tyckte använda ordet ofärdigt ur ett designmässigt perspektiv, utan samtliga motiveringar behandlade systemets prestation. Ett behov av tydligare instruktioner uttrycktes.

Trots att den grafiska representationen inte upplevdes intuitiv uttryckte flertalet deltagare att de såg ett syfte i att representera sökresultat grafiskt. Hälften av alla deltagare valde ordet *Användbart* att representera systemet och 3 av 10 deltagare valde det som ett av de mest relevanta orden. De målgrupper som deltagarna trodde skulle gagnas av en grafisk representation av sökresultat var unga, gamla och personer som av någon anledning behöver en tydligare bild av hur sökresultat hänger ihop. En person upplevde att denne i sitt läraryrke kunde ha användning för verktyget vid framtagning av texter av rätt svårighetsgrad till sina elever.

Att systemet ansågs vara långsamt grundades även det i vissa prestationspro-

blem, till exempel tiden det tog att få fram sökresultaten samt de sporadiska krascher som skedde i samband med sökning.

Utifrån de problem som uppstod vid dessa test är det tydligt att systemet behöver förtydligas. Då det är svårt att ytterligare förenkla den terminologi som är implementerad i systemet behövs förklaringar till existerande termer tillföras. Detta skulle kunna göras genom att ha ett hjälpavsnitt som mer ingående beskrev de olika komplexitetsmått samt även mer ingående hur systemet fungerar. Då dokumentklustring fortfarande verkar vara ett relativt okänt koncept för gemene man vore det fördelaktigt för användaren att få mer ingående beskrivningar vid behov.

Flertalet deltagare fann den grafiska representationen av sökresultaten något förvirrande, samtidigt som många fann det användbart. Detta motiverar att den grafiska representationsformen görs valbar. Genom att låta sökresultaten först presenteras på det sätt som användarna är vana vid från vanliga webbsökningar och göra grafiken valfri elimineras den eventuellt störande faktorn som deltagarna uttryckte.

De tekniska problem som uppstod i samband med testningen är till stor del redan åtgärdade.

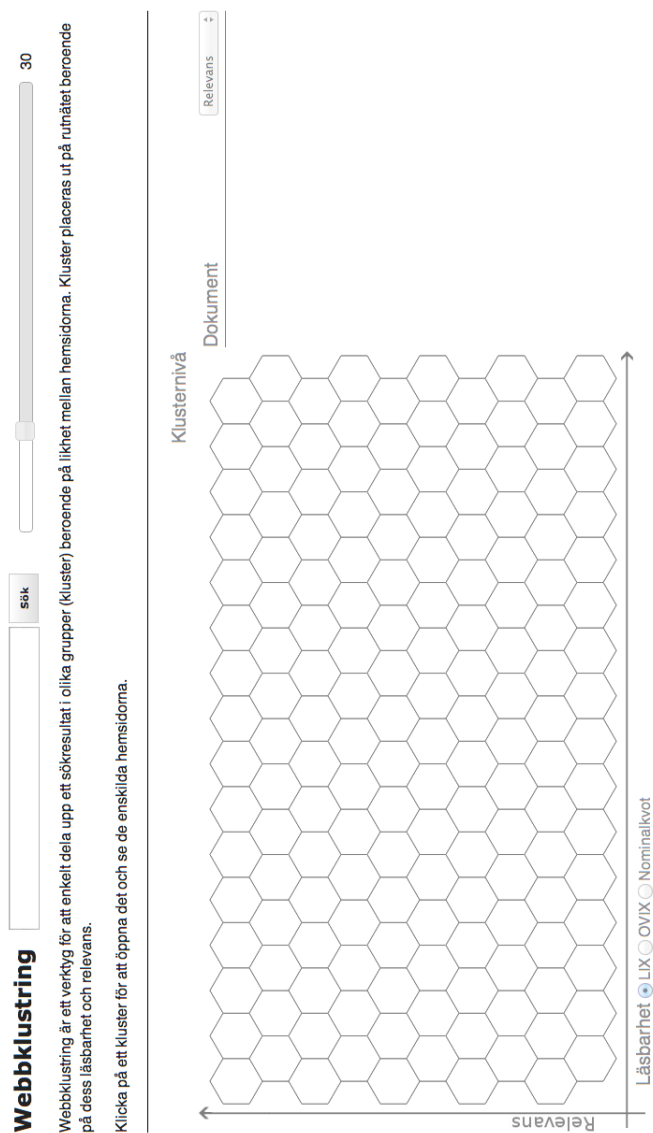
6 Slutsats

Syftet med denna studie var att undersöka användarnas upplevelser av prototyp 2 i projektet Webbklustring. För detta användes en version av Product Reaction Cards där deltagarna fick välja ut ett godtyckligt antal ord att beskriva systemet. De ord som valdes mest var *Förvirrande*, *Ofärdigt*, *Användbart* och *Långsamt*. Utifrån dessa ord, samt medföljande motiveringar, gavs förslagen att 1) inkludera ett hjälpavsnitt i systemet, samt 2) göra den grafiska representationen valbar.

Referenser

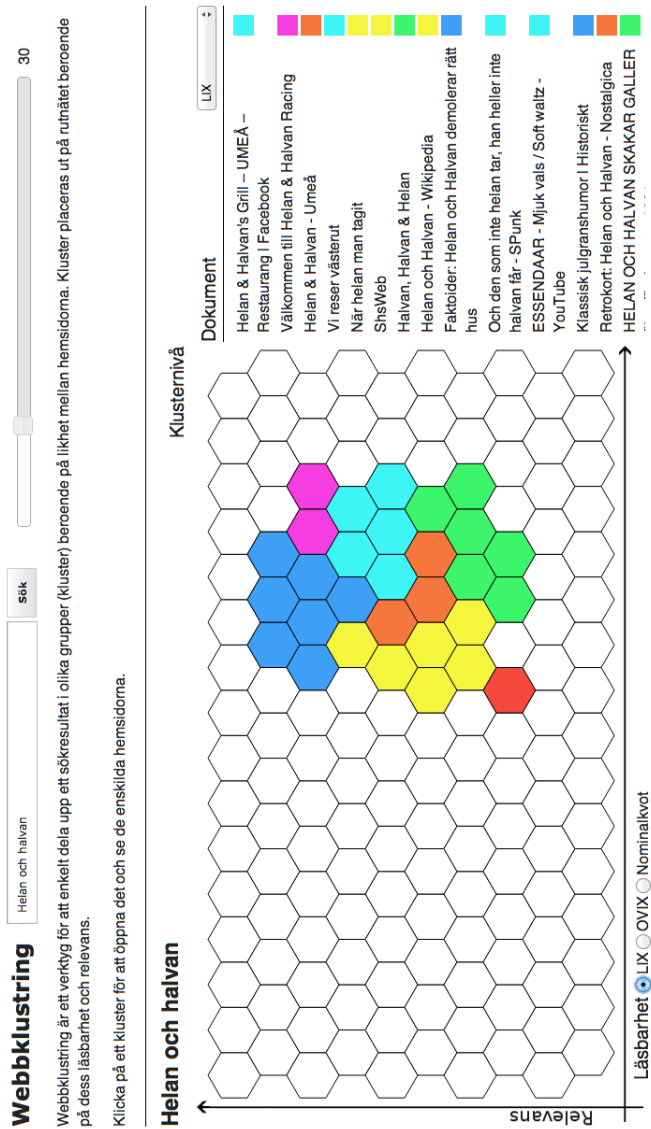
- Benedek, J. & Miner, T. (2002). *Measuring desirability: New methods for evaluating desirability in a usability lab setting*.
- Friendlyreader*. (u. å.). Hämtad 2013-12-29, från <http://www.ida.liu.se/projects/friendlyreader/>
- Webbklustring*. (u. å.). Hämtad 2013-12-29, från <http://www.ida.liu.se/projects/webbklustring/>
- Webblättläst*. (u. å.). Hämtad 2013-12-29, från <http://www.ida.liu.se/arnjo/webblattlast/>

A Gränssnitt: Klusternivå



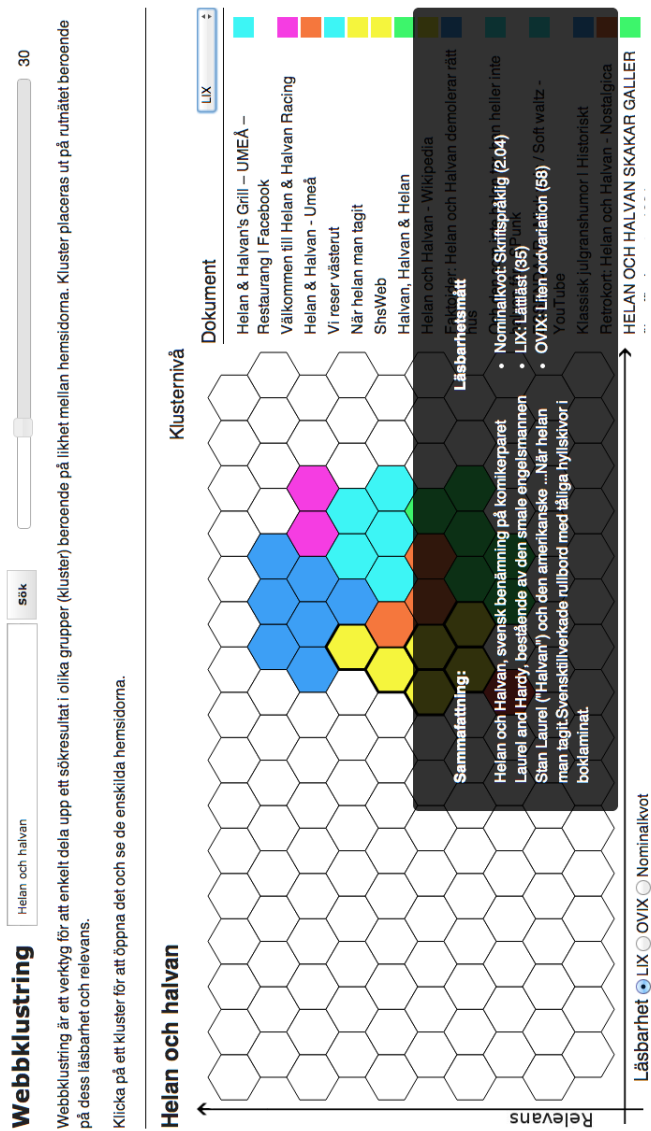
Figur 2: Startside med presentationstext. Det enda som är möjligt att göra i detta skede är att skriva i sökrutan och dra i reglaget för antal sökresultat (på bilden inställt på defaultvärde 30). Komplexitetsmättet nominalkvot tillkom efter användartestningens genomförande.

B Gränssnitt: Sökning



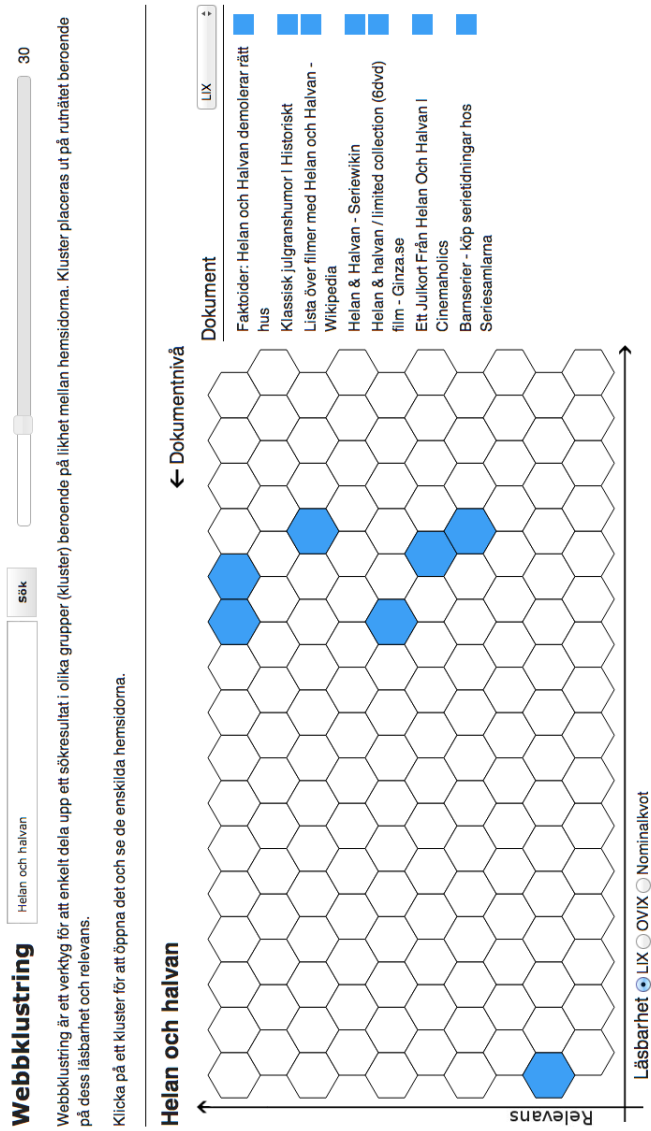
Figur 3: Det utvärderade gränssnittet på klusternivå efter en sökning på "Helan och Halvan". Komplexitetsmättet nominalkvot tillkom efter användartestningens genomförande.

C Gränssnitt: Hovring över kluster



Figur 4: Det utvärderade gränssnittet på klusternivå efter en sökning på ”Helan och Halvan”. Informationen som visas dyker upp då användaren håller muspekaren över ett specifikt kluster. Multidokumentsammanfattning samt komplexitetsmätt.

D Gränssnitt: Dokumentnivå



Figur 5: Det utvärderade gränssnittet på dokumentnivå efter en sökning på "Heelan och halvan".

E Product Reaction Cards -ordpar

Positiv	Negativ
Hjälpsamt	Motarbetande
Snabbt	Långsamt
Intuitivt	Svårförståeligt
Användbart	Oanvändbart
Värdefullt	Värdelöst
Förutsägbart	Oförutsägbart
Tillfredsställande	Otillfredsställande
Motiverande	Irriterande
Färdigt	Ofärdigt
Tydligt	Förvirrande

Tabell 2: De ord som användes i testet. De togs fram i ordpar men presenterades enskilt.