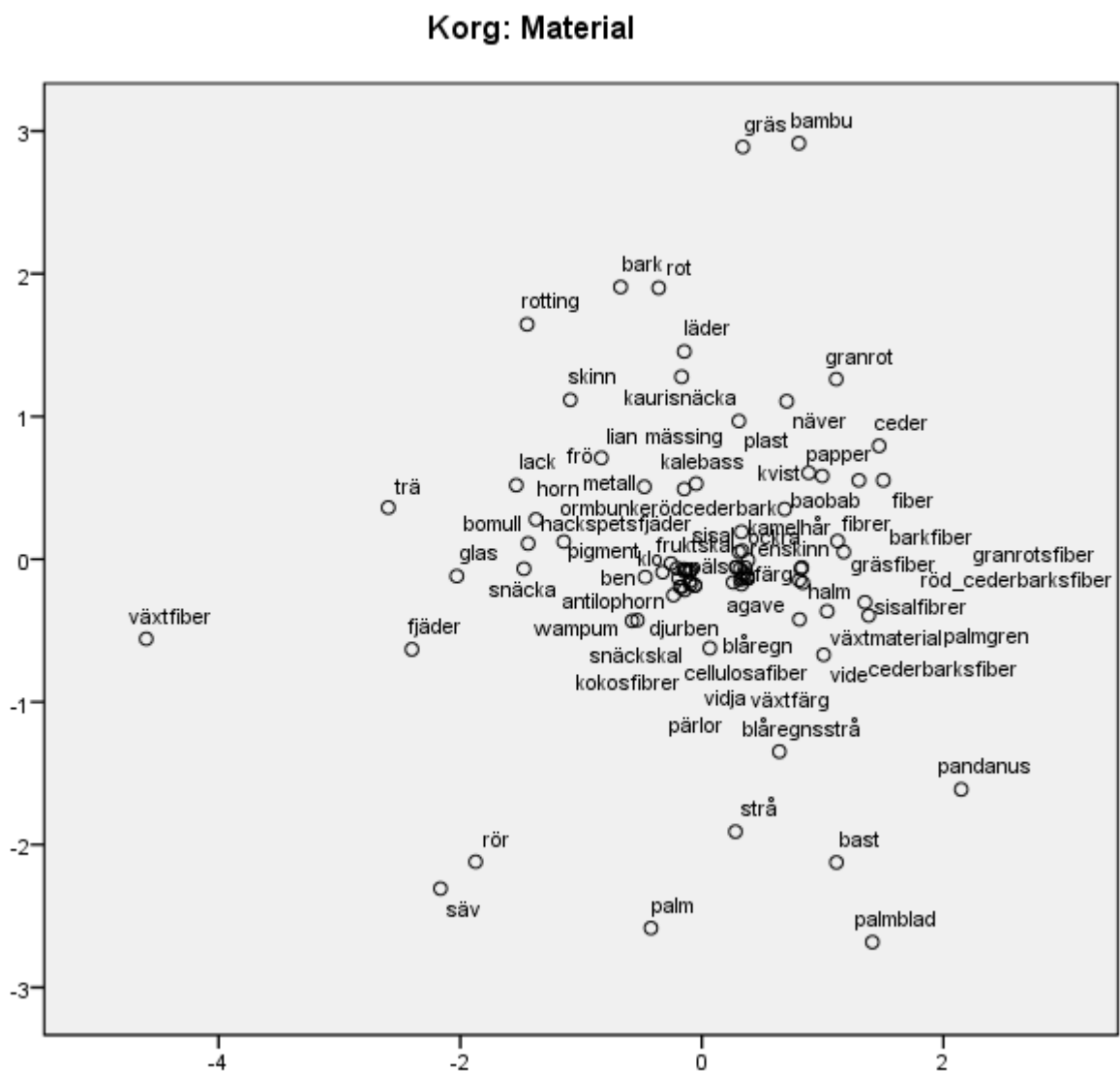
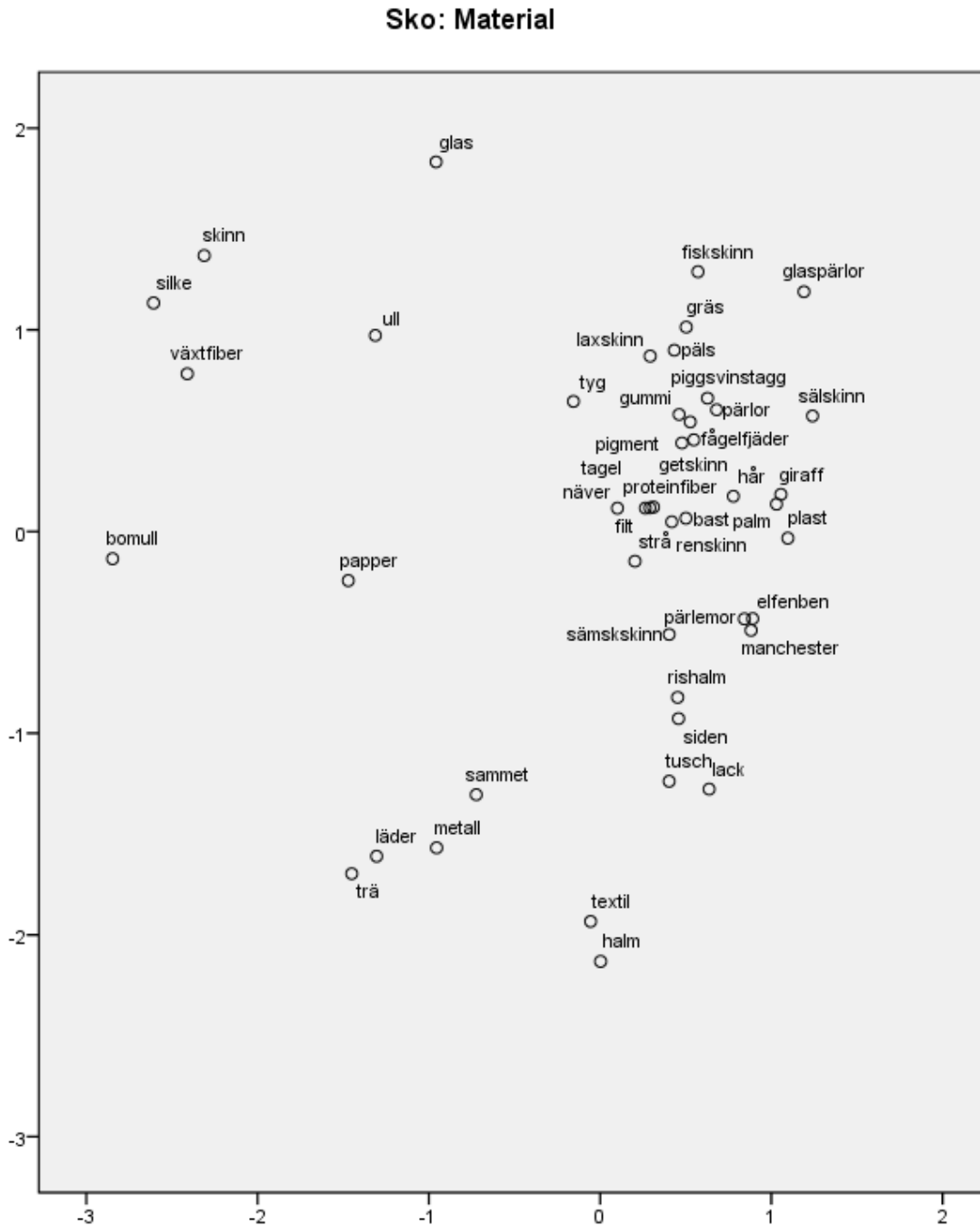


Appendix Bilaga till slutrapport:
Maskininlärningsbaserad indexering av digitaliserade museiartefakter.
Dnr 353-3849-2009, Riksantikvarieämbetet

A.1. MDS-diagram över deskriptorer i kategorin Korg : Material

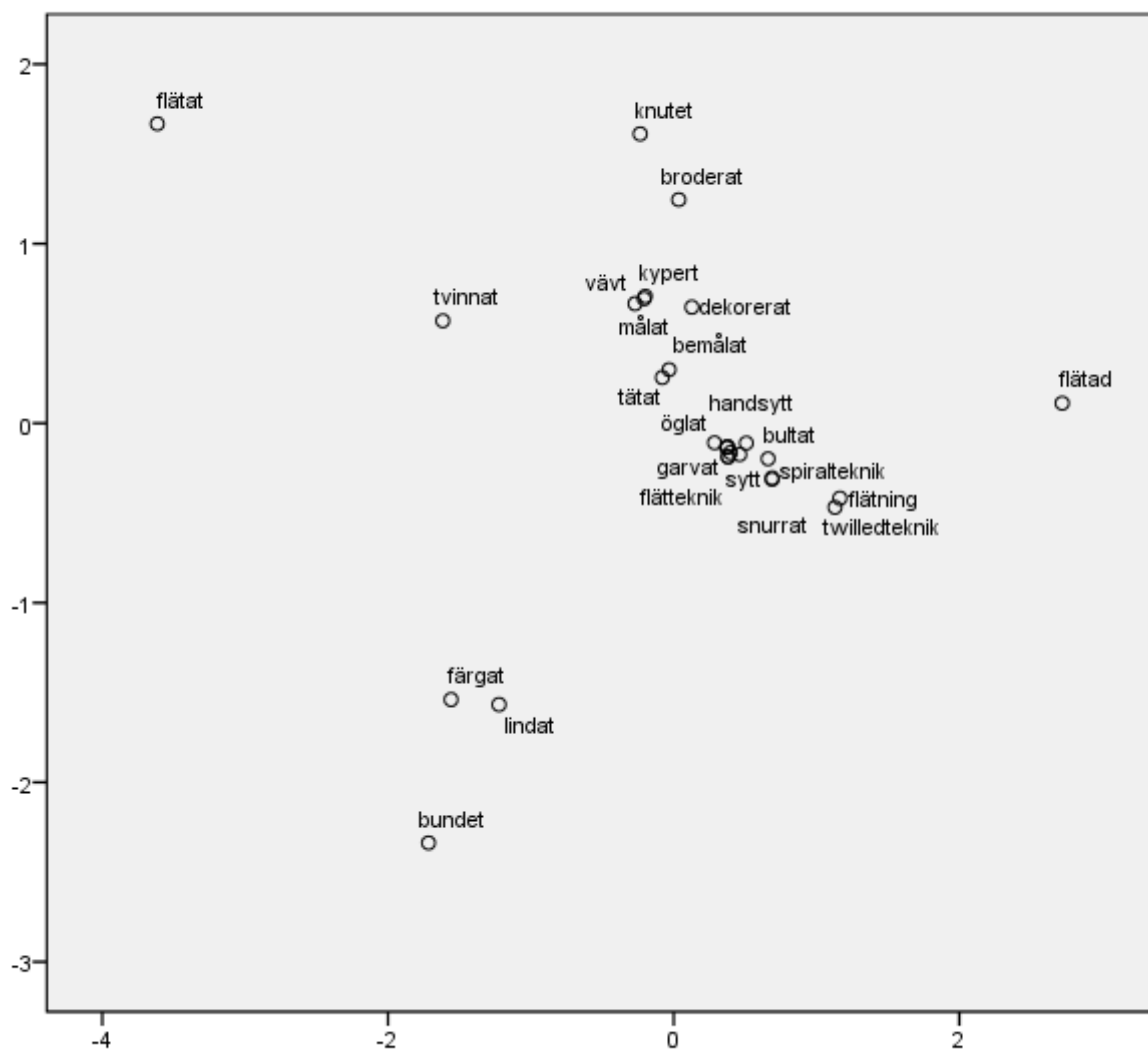


A.2. MDS-diagram över deskriptorer i kategorin Sko : Material



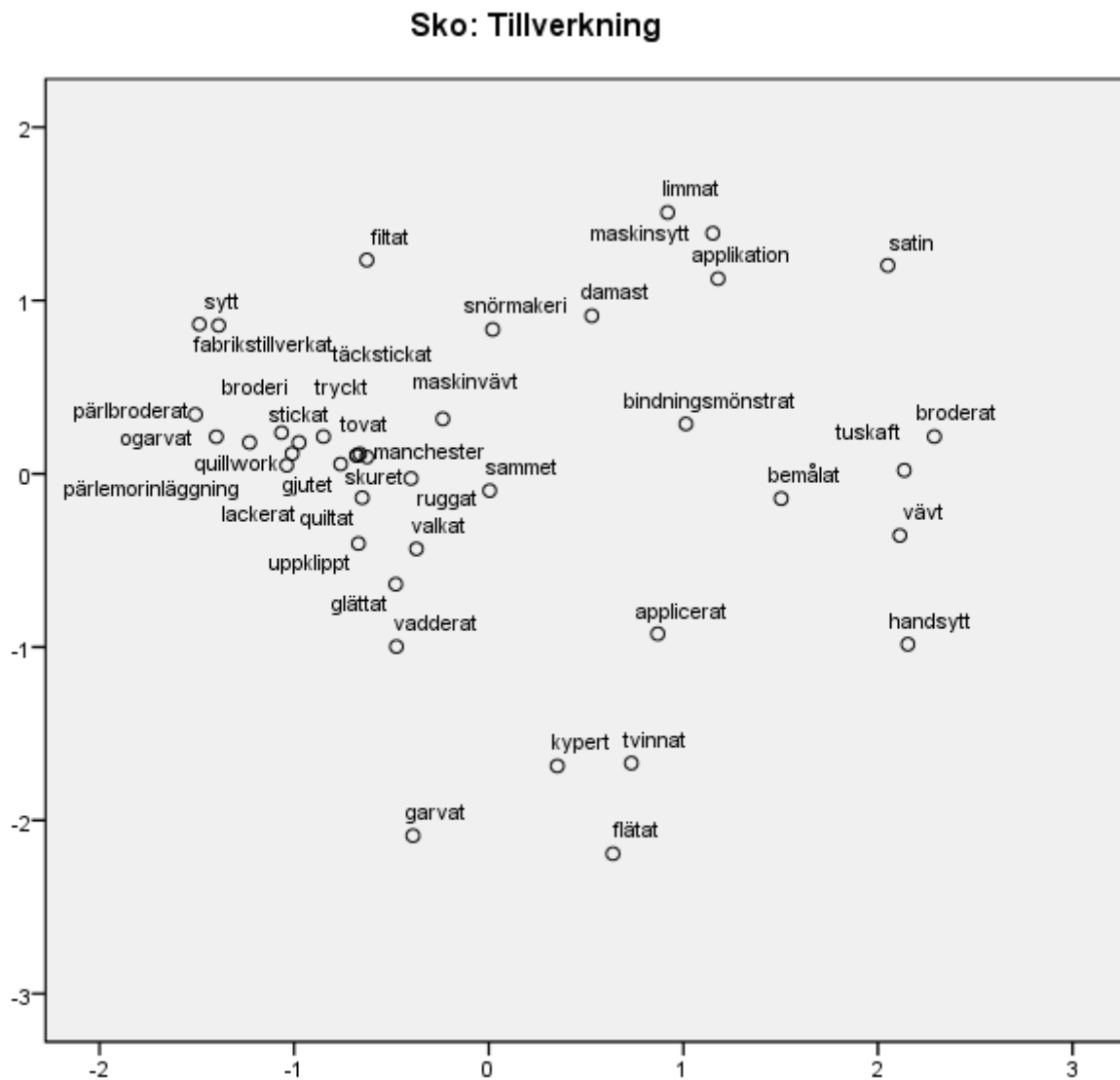
A.3. MDS-diagram över deskriptorer i kategorin Korg : Tillverkning

Korg: Tillverkning

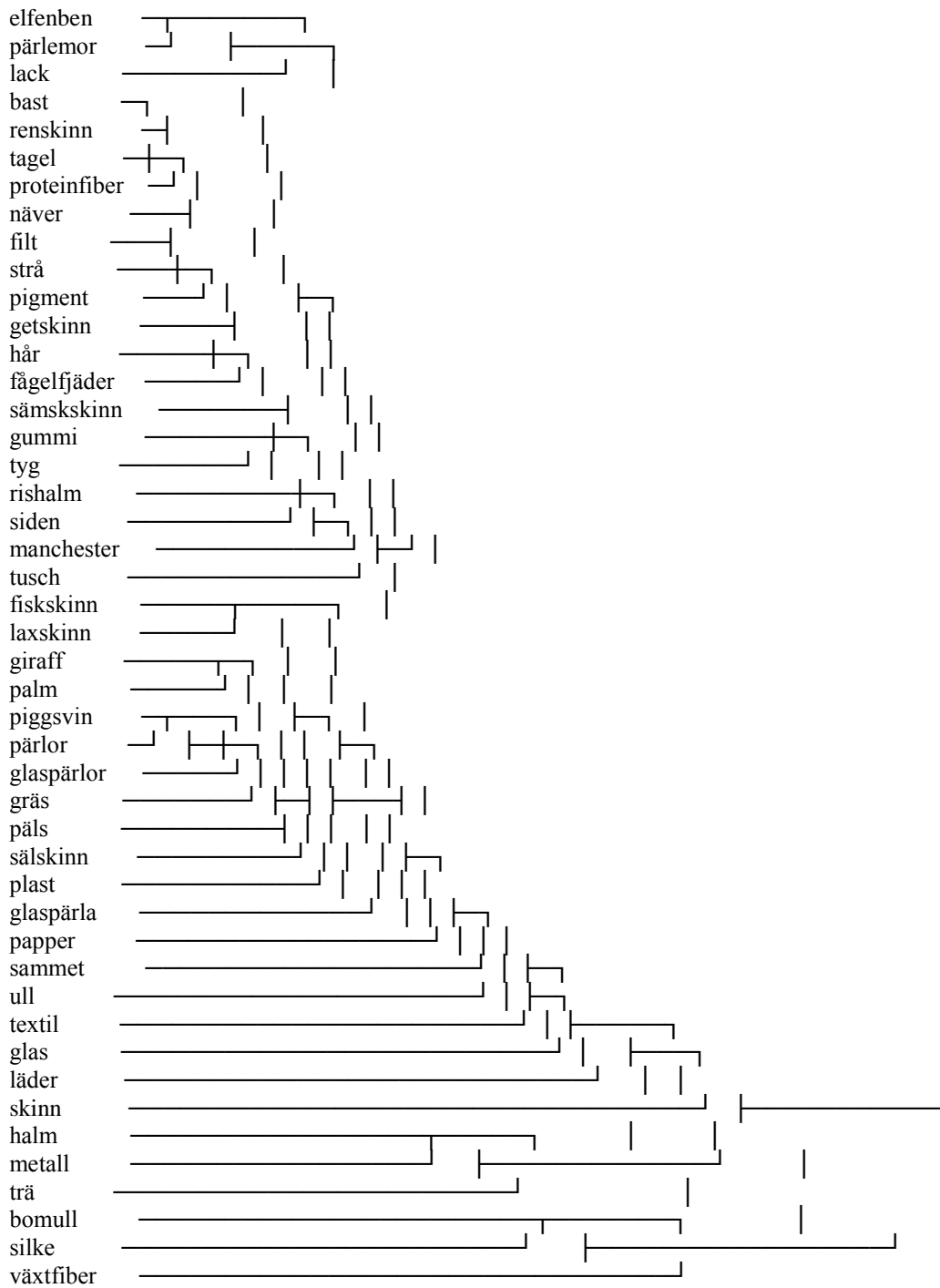


A

A.4. MDS-diagram över deskriptorer i kategorin Sko : Tillverkning



B. Dendrogram över deskriptorer i kategorin Sko : Material



C. Entropibaserade mått på semantiska relationer

Vi ger här en formell beskrivning av de mått på semantiska relationer som har använts i denna studie. Låt X vara en slumpvariabel som kan anta n olika tillstånd. Uttryckt i kommunikationstermer kan X vara ett alfabet och de n tillstånden utgörs av dess bokstäver; ett annat exempel är att låta X vara en vokabulär och de n tillstånden motsvaras av de ord som ryms i vokabulären. Vi låter ett enskilt tillstånd betecknas x och sannolikheten för att detta tillstånd skall observeras skrivs $P(x)$. Entropin för x har definitionen

$$H(x) = -\log_2 P(x)$$

Enligt ovanstående definition följer att entropin för ett tillstånd är stor om sannolikheten för att tillståndet skall observeras är låg (och vice versa). Vi kan med andra ord säga att entropin (informationsmängden) för tillståndet står i omvänt förhållande till chansen att tillståndet observeras. På motsvarande sätt är självinformation för två samtidiga tillstånd x och y

$$H(x, y) = -\log_2 P(x, y)$$

Den *gemensamma informationen* (eng. mutual information) mellan x och y definieras

$$I(x, y) = H(x) + H(y) - H(x, y)$$

och kvantifierar mängden information som tillstånden uttrycker om varandra. Om x och y exempelvis utgörs av observerade beskrivningsord i en databas kommer måttet $I(x, y)$ ange hur stor information deskriptorerna ger om varandra. Ju större värde på måttet, desto starkare är den semantiska associationen mellan deskriptorerna. Denna typ av mått kallar vi likhetsmått eftersom det för en given händelse x maximeras då händelsen jämförs med sig själv. Formellt: $I(x, x) \geq I(x, y)$ för alla händelser x och y .

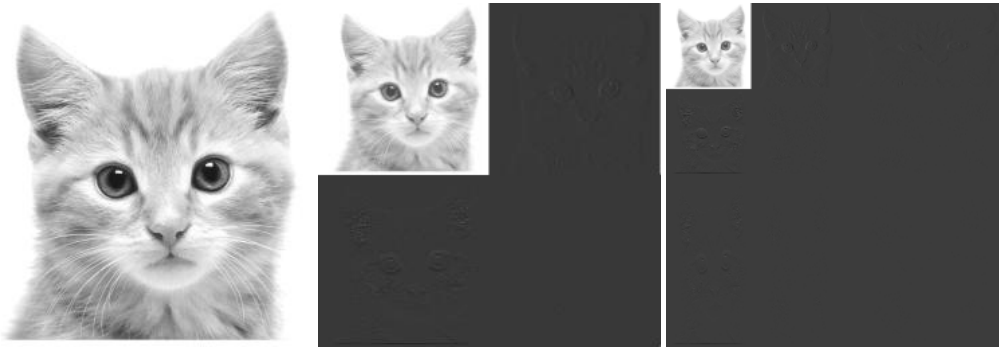
För att generera visuella kartläggningar av deskriptorernas relationer är det däremot önskvärt att transformera $I(x, y)$ till ett mått på *olikhet* snarare än *likhet*. Det är därvidlag lämpligt att skapa ett mått som uppfyller kriterierna för ett geometriskt avståndsmått (eng. *metric*). Ett sådant mått erhålles enkelt genom följande definition:

$$d(x, y) = H(x, y) - I(x, y) = 2H(x, y) - H(x) - H(y)$$

Detta avståndsmått har använts för att visualisera semantiska relationer i 2-dimensionella kartor med hjälp av multidimensionell skalning (se beskrivning ovan).

D. Exempel på Haar-transformation

I figuren nedan ser vi ett exempel på de två första transformationstegen i Haar-transformationen av en bild. Bilden delas successivt upp i information som approximerar de ursprungliga bildpunkterna respektive information om kontrasten mellan närliggande bildpunkter (vilket syns i bilderna nedan som en mörk relief). Informationen om skillnaden mellan bildpunkter är mycket användbar för inläring av klassificerade bildmönster.



E. Omarbetning av bildvektorerna

Bildvektorerna är den allra största delen av datamängden och den del av webb-implementationen som har menligast inverkan på prestanda, främst inladdningstid. En särskiljande egenskap hos klientskript som JavaScript är att koden läses in som ren text. den förkompileras inte (moderna webbläsare har däremot en sk. *just-in-time*-kompilering). Den rena filstorleken ska skickas till klienten och kan utgöra en stor prestandabegränsning.

Bildvektorerna lagras i en JavaScript-struktur, en associativ array av nyckel- och värde par där nycklarna är id för föremålet och värdet är en vektor med de 3072 värdena enligt nedan.

Form för varje vektorbeskrivning i JavaScript-form:

1002920 : [-0.5,-1.25,-1,-0.75,-1,0,-1,0,-1.25,-0.75,-0.75,-0.75,0,0.25,0,-1.25,-1.75,-2,-2.5,-2...

Bildvektorerna har för det första uttryckts kortare genom säkra förändringar av mättnoggrannhet (icke påverkande avslutande nollor i decimaltal). Även i denna form har bildvektorerna utgjort den klart största datamängden som läses in vid sidinladdning. Flera olika sätt att minska denna storlek har övervägts.

- Det vore möjligt att räkna ut alla resultat av vektorjämförelser (med cosinusmättet respektive euklidiskt avstånd) för par av föremål på förhand. Eftersom sökgränssnittet dock tillåter att ett godtyckligt antal bilder används för att formulera varje bildsökfråga (centroiden av dessa tas fram) blir de möjliga kombinationerna av bildvektorer mycket stort, även om det som behöver lagras enbart skulle vara rangordningen mellan föremålen i resultatet.
- *Skalarprodukterna* för parvisa jämförelser (vilka är återkommande värden i algoritmerna, se nedan) skulle kunna tas fram på förhand och laddas in. Detta innebär kortare beräkningstid för frågor och kortare inladdningstid.
- En omformulering av elementens datavärden i variabler, *komprimering*, (varje värde lagras i variabel med kort namn och slås upp vid beräkningar).

För att försöka minska filstorleken för bildvektorerna, har här utvecklats en komprimeringslösning med variabler för värden. Värdena som förekommer i vektorrepresentationerna är nämligen i hög grad återkommande. I JavaScript är det skillnad mellan versaler och gemener som identifierare – *x1* är inte detsamma som *X1* (till skillnad från i t.ex. Visual Basic). Därför blir det möjligt att skapa ett stort antal variabler enbart med bara två tecken. Lösningen innebär att ersätta vektorvärden med variabler, vilkas namn inte får förekomma på annat sätt i runtime-miljön. Att representera värdet "-0,25" med t.ex. *Wx* innebär insparande av tre tecken för varje förekomst i bildvektorerna.

Denna representation ger en klart minskad storlek för vektorrepresentationerna. Det ska nämnas att variablerna dock först ska initieras med värdena en gång först. Nedanstående kodexempel visar hur samma inledning av första bildvektorn (för ett föremål av typen *korg* med id 1002920).

Ursprunglig form:

1002920:[-0.5,-1.25,-1,-0.75,-1,0,-1,0,-1.25,-0.75,-0.75,-0.75,0,0.25,0,-1.25,-1.75,-2,-2.5,-2,-0.5,-1.25,-2.25,-1.75,-0.25,0,-0.75,-1,0,-0.5,-1,-1,-0.5,-1.5,-0.25,-0.75,-0.25,-2.25,-0.25,-1,-1.75,-1,-0.75,0,15.75 ...

Motsvarande representation med globala variabler:

1002920:[c,h,o,d,o,e,o,e,h,d,d,e,b,e,h,i,ä,q,ä,c,h,l,i,a,e,d,o,e,c,o,o,c,j,a,d,a,l,a,o,i,o,d,e,aF ...

F. Konvertering av dataformatet för textbeskrivning av föremålen

De drygt 900 föremålen och deras beskrivningar har erhållits i ett XML-format, från Carlotta-databasen. Det är fråga om en detaljerad men i stycken överflödigt struktur. För att kunna användas praktiskt i webb-implementationen har data omformulerats till ett platssparande JavaScript-format, ett överordnat objekt (associativ array) som i sin tur består av ett objekt (associativ array) per föremålspost.

Detta förarbete gör att textdata blir enklare att behandla programtekniskt, och söka i, samt reducerar filstorleken för den textmässiga beskrivningen av föremålen till mindre än en femtedel av ursprungsformatet (från ca 4308 kB till ca 785 kB).

De nedanstående tabellerna visar representationerna av en enskild post i det ursprungliga XML-formatet samt därefter i det platsbesparande JavaScript-formatet.

```
<object id="1002848" register="2" registerName="Objekt" id-term="2" visible-term="72">
<term term-id="2" spe="OBJIDN" objspe="Inventarienummer" is-id-term="true">
<value>
<objtxt>1887.08.2489</objtxt>
</value>
</term>
<term term-id="72" spe="OBJNAM" objspe="Sakord">
<value>
<objtxt>sko</objtxt>
</value>
<value>
<objtxt>skodon</objtxt>
</value>
<value>
<objtxt>sandal</objtxt>
</value>
</term>
<term term-id="116" spe="OBJNAM" objspe="Sakord engelska">
<value>
<objtxt>shoe</objtxt>
</value>
<value>
<objtxt>sandal</objtxt>
</value>
</term>
<term term-id="117" spe="OBJNAM" objspe="Lokalt namn">
<value>
<objtxt>geta</objtxt>
</value>
</term>
<term term-id="143" spe="OBJCDE" objspe="OCM">
<value>
<objtxt>291</objtxt>
</value>
</term>
<term term-id="144" spe="OBJTME" objspe="Tidpunkt" objasc="tillverkning">
<value>
<objtxt>1800-tal</objtxt>
</value>
</term>
<term term-id="156" spe="FOLCDE" objspe="OWC" objasc="ursprung">
<value>
```

```

<objtxt>AB1</objtxt>
</value>
</term>
<term term-id="157" spe="GEONAM" objspe="Världsdel">
<value>
<objtxt>Asien</objtxt>
</value>
</term>
<term term-id="159" spe="GEONAM" objspe="Ort" objasc="brukare">
<value>
<objtxt>Tokyo</objtxt>
</value>
</term>
<term term-id="170" spe="GEONAM" objspe="Land" objasc="ursprung">
<value>
<objtxt>Japan</objtxt>
</value>
</term>
<term term-id="180" spe="OBJTXT" objspe="Förvärvsomständigheter" objasc="i fält">
<value>
<objtxt>Delsamling 1887.8.2489-2867 är inköpt i Tokyo och består av sandaler, skor, handdukar, ätpinnar,
symboliska svärd, många rituella och religiösa föremål och figurer, klädedräkter mm. (PH 960506)</objtxt>
</value>
</term>
<term term-id="181" spe="PERNAM" objspe="Personnamn" objasc="förvärvat från">
<value>
<objtxt>Stolpe, Knut Hjalmar</objtxt>
</value>
</term>
<term term-id="184" spe="OBJTME" objspe="Datum" objasc="förvärv till museet">
<value>
<objtxt>1887</objtxt>
</value>
</term>
<term term-id="187" spe="MATNAM" objspe="Material" objasc="tillverkning">
<value>
<objtxt>lack</objtxt>
</value>
<value>
<objtxt>sammet</objtxt>
</value>
<value>
<objtxt>metall</objtxt>
</value>
<value>
<objtxt>halm</objtxt>
</value>
<value>
<objtxt>tusch</objtxt>
</value>
</term>
<term term-id="189" spe="OBJMEA" objspe="Längd">
<value>
<objtxt>21,4 cm</objtxt>
</value>
</term>
<term term-id="190" spe="OBJMEA" objspe="Höjd">
<value>
<objtxt>H 8,5 cm</objtxt>
</value>

```

```

</term>
<term term-id="191" spe="OBJMEA" objspe="Bredd">
<value>
<objtxt>10 cm</objtxt>
</value>
</term>
<term term-id="211" spe="OBJTME" objspe="Datum" objasc="dokumentation">
<value>
<objtxt>1995-04-05</objtxt>
</value>
<value>
<objtxt>2003-03-04</objtxt>
</value>
</term>
<term term-id="212" spe="PERNAM" objspe="Personnamn" objasc="dokumentation">
<value>
<objtxt>Holmberg, Petra</objtxt>
</value>
<value>
<objtxt>Svensson, Irene</objtxt>
</value>
</term>
<term term-id="214" spe="OBJTME" objspe="Datum" objasc="uppdatering">
<value>
<objtxt>2003-03-04</objtxt>
</value>
<value>
<objtxt>2009-11-04</objtxt>
</value>
<value>
<objtxt>2009-11-04</objtxt>
</value>
</term>
<term term-id="215" spe="PERNAM" objspe="Personnamn" objasc="uppdaterare">
<value>
<objtxt>Svensson, Irene</objtxt>
</value>
<value>
<objtxt>Dahlström, Maria</objtxt>
</value>
<value>
<objtxt>Dahlström, Maria</objtxt>
</value>
</term>
<term term-id="219" spe="OBJNAM" objspe="Ursprungsdatabas">
<value>
<objtxt>Etno</objtxt>
</value>
</term>
</object>

```

Ursprungsformatet (XML) har omformats till nedanstående kompakta kod-format utan någon egentlig dataförlust.

```

1002848:{"Inventarienummer":"1887.08.2489","Sakord":"sko,skodon,sandal","Sakord
engelska":"shoe,sandal","Lokalt namn":"geta","OCM":"291","Tidpunkt (tillverkning)":"1800-

```

```
tal", "OWC (ursprung)": "AB1", "Världsdel": "Asien", "Ort (brukare)": "Tokyo", "Land  
(ursprung)": "Japan", "Förvärvsomständigheter (i fält)": "Delsamling 1887.8.2489-2867 är inköpt i  
Tokyo och består av sandaler, skor, handdukar, ätpinnar, symboliska svärd, många rituella och  
religiösa föremål och figurer, klädedräkter mm. (PH 960506)", "Personnamn (förvärvat från)": "Stolpe,  
Knut Hjalmar", "Datum (förvärv till museet)": "1887", "Material  
(tillverkning)": "lack,sammet,metall,halm,tusch", "Längd": "21,4 cm", "Höjd": "H 8,5 cm", "Bredd": "10  
cm", "Datum (dokumentation)": "1995-04-05,2003-03-04", "Personnamn (dokumentation)": "Holmberg,  
Petra,Svensson, Irene", "Datum (uppdatering)": "2003-03-04,2009-11-04,2009-11-04", "Personnamn  
(uppdaterare)": "Svensson, Irene,Dahlström, Maria,Dahlström,  
Maria", "Ursprungsdatabas": "Etno", "id": "1002848" },
```

XML för databehandling har efter ett stort genomslag kring millennieskiftet ofta blivit kritiserat för sin stora luftighet och emellanåt överflödiga struktur, vilket lätt sänker prestandan. Förespråkare för XML-formatet anger ofta egenskapen ”human readable” som en styrka och förklaring. Det finns dock många konkurrerande format (bl.a. JSON)¹ idag.

¹ JSON står för *JavaScript Object Notation* och är relativt likt den formalism som här används.

G. Algoritmer för bildsökning och återkoppling

De algoritmer som använts är för bildsökning Cosinus-likhet eller euklidiskt avstånd samt för återkopplingsökning (relevans-feedback) Rocchio-algoritmen vilka visas nedan.

De två likhetsmått används vid bildsökning då en genomsnittlig vektor (sökfrågevektor) beräknas utifrån de föremål som användaren väljer. Därefter jämförs och rangordnas samtliga föremål mot denna frågevektor. Resultatet visas direkt och beräkningen och sorteringen kan ske nästan momentant.

Cosinus-måttet:

Låt \mathbf{d} och \mathbf{q} vara en dokumentvektor respektive en sökfrågevektor i ett n -dimensionellt vektorrum. Cosinusmålet för likhet mellan dessa vektorer definieras då:

$$\cos \theta = \frac{\mathbf{d} \cdot \mathbf{q}}{\|\mathbf{d}\| \times \|\mathbf{q}\|} = \frac{\sum_{i=1}^n d_i q_i}{\sqrt{\sum_{i=1}^n d_i^2} \times \sqrt{\sum_{i=1}^n q_i^2}}$$

Euklidiskt avstånd:

Det euklidiska avståndet mellan två vektorer \mathbf{d} och \mathbf{q} definieras som längden hos vektorn $\mathbf{d} - \mathbf{q}$, dvs

$$\delta(\mathbf{d}, \mathbf{q}) = \|\mathbf{d} - \mathbf{q}\| = \sqrt{\sum_{i=1}^n (d_i - q_i)^2}$$

Rocchio-algoritmen för relevansfeedback:

Låt $\mathbf{q}^{(t)}$ vara sökfrågevektorn vid iteration t , \mathbf{C}_r centroiden för vektorerna för de kända relevanta dokumenten, samt \mathbf{C}_n centroiden för vektorerna för de kända icke-relevanta dokumenten. Den uppdaterade sökfrågevektorn $\mathbf{q}^{(t+1)}$ beräknas då

$$\mathbf{q}^{(t+1)} = \alpha \mathbf{q}^{(t)} + \beta \mathbf{C}_r - \gamma \mathbf{C}_n$$

där α , β och γ är koefficienter som används för att finjustera anpassningen av sökfrågevektorn.

H. Hjälptext i webbgränssnittet

Den nedanstående hjälptexten återfinns i webb-gränssnittet och är en guide för praktisk användning.

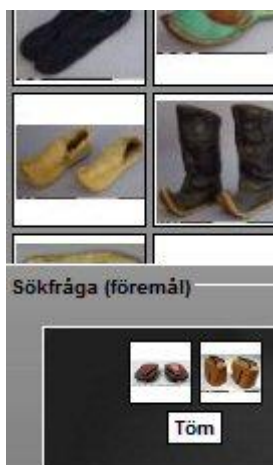
Sökning med bilder

Genom bildsökning (*query-by-example*) är det möjligt att välja en eller flera exempelbilder, och söka efter liknande bilder. De bilder som väljs utgör alltså "sökfrågan".

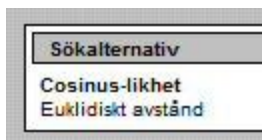
Söka på bildlikhet

Sök efter liknande bilder på följande sätt.

1. Klicka på fliken **Sökning med bilder** i vänsterdelen av sidan.
2. Klicka på den eller de bilder som ska vara sökfråga.



3. Välj sökalgoritm. Det är denna algoritm som avgör likhet mellan de vektorer som representerar bilderna.



4. Klicka på knappen **Sök**.
5. Resultatet visas i vänstervyn.

Observera att samtliga bilder finns med i sökresultatet, i bildsökning. Det betyder att en sökning kan ses som en speciell sortering av bilderna.

Återkopplingssökning med bilder

Det är möjligt att efter en sökning återupprepa sökning (relevans-feedback). Det sker genom att relevanta bilder i resultatet i mittvyn markeras. Därefter sker en återsökning och "sökfrågan" har då modifierats så att de nya markeringarna beaktas.

Gör en återkopplingssökning med bilder på följande sätt.

1. Inled med att göra en vanlig bildsökning.
2. Markera de föremål som ska ses som relevanta. De får en röd ram och dyker upp i **Sökfråga** i vänstervyn.



3. Klicka på knappen **Sök** för att genomföra återkopplingssökning.

Algoritmen för sökning heter *Rocchio*-algoritmen.

Textsökning efter föremål

Det är möjligt att söka efter föremål genom den textinformation som varje post har kopplad till sig. Den text som representerar varje post (hämtad från *Carlotta*-databasen) syns i höger vyn när en post markeras.

Inled en textsökning på följande sätt.

- Välj fliken **Sökning med text** i vänstervyn.

Det är möjligt att söka med ett enda villkor eller genom att kombinera en sökfråga genom att ange flera villkor.

Sök med ett villkor på följande sätt.

1. Fyll i en söksträng till vänster.
2. Ange vilka sökvillkor som ska gälla (om du väljer **Sök i visst fält** så kommer du att behöva ange fält).
3. Klicka på knappen **Sök med enbart detta kriterium**.
4. Resultatet visas i mittvyn.

Sök med flera villkor på följande sätt.

1. Ange sökvillkor genom att skriva en söksträng och ange sökalternativ.
2. Klicka på knappen **Lägg till detta kriterium**.
3. De poster som matchar villkoret visas i mittvyn. Nederst visas även kombinationen av kriterier som anges.
 - Lägg om så önskas till fler sökkriterier på samma sätt.

Resultatet kan slutligen visas i resultatläge genom att välja knappen **Visa föremålen i resultatvy** nederst i mittvyn. Resultatet visas i mittvyn.