Innehåll

[Dokumentation 2](#_Toc108515853)

[Dokumentanalys och förproduktion 2](#_Toc108515854)

[Bildfångst och efterbearbetning 3](#_Toc108515855)

# Dokumentation

Selecting, collecting, managing

Dokumentanalys och förproduktion   
Vid dokumentanalysen använde vi oss av de kriterier som föreslagits av Deutsche Forschungsgemeinschaft (2013).

|  |  |  |
| --- | --- | --- |
| Kriterium | Ja | Nej |
| Risk to ink; risk to paint layer / priming coat; ink / paint corrosion |  | x |
| Loose stitching |  | x |
| Binding or cover material too stiff / inflexible |  | x |
| Binding (partially) fractured |  | x |
| Cover material at joint (partially) fractured |  | x |
| Spine inlay too stiff; spine binding damaged |  | x |
| Cracks, tears or flaws |  | x |
| Delicate leather grain on spine |  | x |
| Risk to spine gilding |  | x |
| Very thick vellum volume with hollow spine |  | x |
| Cover (partially) fractured |  | x |
| Cover loose (front/back) |  | x |
| Closure straps too stiff / partially fractured |  | x |
| Inside margin too narrow |  | x |
| Book block corrugated |  | x |
| Leaves / signatures very stiff |  | x |
| Unavoidable loss of text |  | x |
| Original cannot be laid flat |  | x |
| Extreme format |  | x |

Utifrån dessa kriterier ansågs materialet lämpligt att utgå ifrån i det fortsatta arbetet. Däremot noterades vid dokumentanalysen att papperskvaliteten var väldigt tunn och att försiktighet måste iakttas vid hantering av materialet samt att materialet hade en stor genomblödning som måste beaktas vid bildfångsten. Vidare noterades i dokumentanalysen att materialet innehöll handskrivna noteringar i blyerts på två ställen i materialet. Under dokumentanalysen genomfördes också en diskussion kring vilka aspekter av materialet som vi önskade fånga i reproduktionen av materialet och önskvärda riktmärken för bildfångst noterades. Som stöd i detta arbete konsulterades Cornell University Library (2003), Deutsche Forschungsgemeinschaft (2013) och Terrass (2008).

|  |  |
| --- | --- |
| Riktmärken för bildfångst |  |
| Upplösning | 400 dpi |
| Bitdjup | 24-bit |

Eftersom senare årgångar av tidskriften innehåller färgfotografier så valde vi att använda bitdjup 24-bit som riktmärken för bildfångsten, även om detta avviker från den standard som anges för det materialet som vi arbetat med i detta projekt och resulterar i ett större filformat än nödvändigt utifrån det specifika materialets kvaliteter (jfr. Terrass, 2008). Syftet med detta är strävan efter att skapa ett semi-automatiserats arbetsflöde där alla årgångar skannas med samma inställningar för att undvika kostnaden av att undersöka varje årgång separat (jfr. Cornell University Library, 2003).

• Strategies

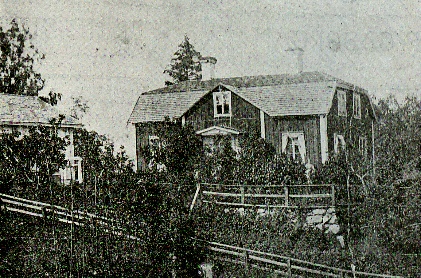
• Tools (selecting, acquiring,   
developing)

## Bildfångst och efterbearbetning

För bildfångst av materialet användes en bokskanner av märket Zeta comfort HX-4899 från Zeutschel som vi fick tillgång till via Nordiska museets bibliotek. Valet av denna skanner grundade sig på en strävan efter att skapa ett homogent och återupprepningsbart arbetsflöde för att senare årgångar längre fram ska kunna digitaliseras på ett tids- och kostnadseffektivt sätt. Vidare ansågs skanner väl lämpad för det valda materialet då dess inbyggda bokvagga bidrar till att stabilisera materialet och skydda bokryggen under bläddringen vid skanningen. Skannerns inbyggda linjaler var ytterligare en aspekt som talade för valet av skannern då detta hjälper till att skapa raka bilder som är en viktig aspekt för att få fram bildfiler som lämpar sig för OCR-behandling (ref på detta?). Vidare gav skannern möjlighet att direkt spara filerna i TIFF-format vilket är det filformat som rekommenderas för masterfiler vid digitalisering inom minnesinstitutioner (Deutsche Forschungsgemeinschaft, 2013; Cornell University Library, 2003).

Innan hela materialet skannades genomfördes en provskanning på delar av materialet. Vid detta tillfälle upptäcktes att den valda skannern inte klarade av att generera en upplösning högre än 300 dpi, trots att vi fått information om att den skulle klara av att skanna i upp till 600 dpi. Vidare framkom vid provskanningen att det inte fanns möjlighet för oss att kalibrera skannern. Vi valde att ändå använda denna skanner då det var den bästa utrustning vi hade att tillgå för bildfångst av materialet.

Under provskanningen undersöktes också vilken färg på papper som gavs bäst resultat i syfte att minska genomblödningen i materialet. Vi valde att använda ett grått papper eftersom det gav en viss minskning i genomblödning men inte, utifrån en subjektiv bedömning, påverkade färgåtergivningen av materialet lika mycket som en svart bakgrund, detta var framför allt tydligt i fråga om fotografier där en svart bakgrund gav mycket starkare kontraster i färgåtergivningen än vad som fanns i originalmaterialet. En viktig anmärkning här är att det är det är det tryckta materialet vi eftersträvat att skapa en reproduktion av och att en aspekt av detta är att fånga såväl den gulnade färgen på materialet samt den bleknande kvaliteten av de tryckta fotografierna i materialet även om detta innebär att kvaliteten på fotografiernas detaljer blev sämre.

 En bild som visar träd, utomhus, gammal, hus

Automatiskt genererad beskrivning

Figur 1 Ljusgrått papper använt bakom materialet Figur 2 Svart papper använt bakom materialet

Det är dock viktigt att lägga märke till att jämförelsen av färgåtergivningen mot originalet är just en *subjektiv* bedömning. Bilderna skannades inte i en miljö som följer ISO 3666 – Viewing Conditions for Graphic Technology and Photography och det finns flera aspekter, exempelvis dagsljus insläpp som påverkat resultatet. Vidare har vi inte haft tillgång till en standardiserad färgkarta att använda som referens vid bildfångsten och vare sig skannern eller de datorer som använts för att granska materialet har kalibrerats med avseende på färgåtergivning (jfr. Terrass, 2008).

Under provskanningen provade vi att använda oss av tyngder för att hålla sidorna på plats och få dem att ligga slätt eftersom sidorna, framför allt på vänster uppslag, kom att bölja uppåt på grund av höjden från det papper vi använde som bakgrund. På grund av papperets tunna kvalitet och den risk för slitage på materialet som vi upplevde vid användandet av tyngder valde vi dock att skanna materialet utan tyngder även om detta innebar att motivet blev mer snett utan tyngder. För en diskussion om att OCR:n ändå blev bra i avsnittet om OCR, borde vi köra de olika sidorna med och utan tyngder i OCR och se om det blir en markant skillnad?

En bild som visar text

Automatiskt genererad beskrivning En bild som visar text

Automatiskt genererad beskrivning

Efter skanningen beskars alla filer i GIMP. Här lämnades en tunn kant av bakgrunden runt varje bild för att visa att hela originalmaterialet har skannats (Deutsche Forschungsgemeinschaft, 2013). Efter redigering av inbäddad metadata och filnamngivning gjordes kopior av masterfilerna i JPEG-format för användning som presentationsmaterial.

Kritisk analys

Slutsats

Sammantaget bör projektet förstås som en förstudie för en digitalisering av tidskriften Hälsingerunor i större skala. Detta baseras främst på att den tekniska utrustning som användes för bildfångst i detta projekt inte gav adekvata resultat eftersom vi saknade möjlighet att kalibrera den innan skanning. Detta ledde till att själva skanningen tog längre tid än vad som hade behövts då flera skanningar av samma sida behövde genomföras för att få ett tillfredställande resultat och färgåtergivningen på framför allt annonserna varierar kraftigt. Då även arbetsstationen där skannern är placerad inte är optimal för bildfångst (dagsljusinsläpp, sviktande trägolv, passerande biblioteksbesökare och dylikt rekommenderas val av annan teknisk utrustning och arbetsstation vid ett fortsatt arbete med att digitalisera tidskriften.

• OCR and other technologies  
• Text encoding and other   
optimization, e.g. indexing  
• Metadata

Vid provskanningen framkom att den valda skannern inte automatiskt sparade datum då bilderna skannades samt skannerns märke och modell. Dessa element lades därför in manuellt, för detta arbete använde vi programmet EXIF Pilot 64. Varför anser vi att det är viktigt att dessa sparas med bildfilerna? För resterande inbäddad metadata i bildfilerna använde vi Adobe Bridge.

|  |  |  |  |
| --- | --- | --- | --- |
| IPCT | Definition | Anteckningar | Exempel |
| Creator | Contains the name of the photographer, but in cases where the photographer  should not be identified the name of a company or organisation may be  appropriate. |  |  |
| Description | A textual description, including captions, of the image. | Används endast för artiklar med bilder. Används för att beskriva bilder på exempelvis fotografier på objekt och konstverk i artiklarna. |  |
| Description writer | Identifier or the name of the person(s) involved in writing, editing or correcting the Description, Alt Text (Accessibility), or Extended Description (Accessibility) of the image. |  |  |
| Copyright notice | Contains any necessary copyright notice for claiming the intellectual property for this photograph and should identify the current owner of the copyright for the photograph. Other entities like the creator of the photograph may be added in the corresponding field. Notes on usage rights should be provided in "Rights usage terms". | Unknown används i de fall då de scannade artikalarna undertecknats med psedonym eller saknar angiven författare. Används även då artikelförfattare är känd men fotografi eller bild i artikeln saknas upphovsägare. Public domain används då upphosvrätten för den scannade artikeln gått ut. Copyrighted används då den scannade artikeln är upphovsrättskyddad |
| Name of featured organisations | Name of the organisation or company which is featured in the image. | Används endast för annonssidor |  |
| Artwork or Object | A structured datatype for details about artwork or an object in an image. Includes a Copyright Notice, a Creator, a Date Created, a Source, a Source Inventory Number and a Title. | Används för att ange vilken årsbok av Hälsingerunor som bilder är scannade från. Använd inte för fotografier av exempelvis konstverk i artiklarna här används istället *Description* |
| Type of source | The type of the source of this digital image |  |  |
| Licensor | A person or company that should be contacted to obtain a licence for using the item or who has licensed the item. | Används endast för scannade artiklar där upphovsrätten är okänd. Används inte för artiklar i public domain. |  |
| Licensor: name |  |  |  |
| Licensor: identifier |  | Använd id från https://libris.kb.se/katalogisering/ | https://libris.kb.se/fcrttcbz55x1520#it |
| Copyright owner | Owner or owners of the copyright in the licensed image. | Används endast för scannade artiklar som är upphovsrättsskyddade. Använd id från från https://libris.kb.se/katalogisering/ |  |

Filnamngivningen framgår av tabellen nedan och har hämtat inspiration från Kungliga bibliotekets filnamngivning för inbundna böcker så som det presenterats i Andrea Davids Kronlunds föreläsning ” Digital imaging: 3 Image Capture” den 10 februari 2021.

|  |  |
| --- | --- |
| **Filnamn** | **Förklaring i klartext** |
| Helsingerunor 1921-2914fp | Frampärmen |
| Helsingerunor 1921-2914fpi | Frampärmens insida |
| Helsingerunor 1921-1924,3 | Försättsblad, osignerat |
| Helsingerunor 1921-1924,4 | Försättsblad, osignerat |
| 1921.1,5 | Första annonsen i 1921 års första volym |
| 1921.1,6 | Andra annonsen i 1921 års första volym |
| 1921.1,7,1 | sida 1 i 1921 års första volym |
| 1921.1,8,2 | sida 2 i 1921 års första volym |
| Etc. fram till |  |
| 1921.1,23 | Tredje annonsen i 1921 års första volym |
| 1921.1,23 | Fjärde annonsen i 1921 års första volym |

|  |  |  |  |
| --- | --- | --- | --- |
| IPCT | Definition\* | Anteckningar | Exempel |
| Creator | Contains the name of the photographer, but in cases where the photographer  should not be identified the name of a company or organisation may be  appropriate. |  |  |
| Description | A textual description, including captions, of the image. | Används endast för artiklar med bilder. Används för att beskriva bilder på exempelvis fotografier på objekt och konstverk i artiklarna. |  |
| Description writer | Identifier or the name of the person(s) involved in writing, editing or correcting the Description, Alt Text (Accessibility), or Extended Description (Accessibility) of the image. |  |  |
| Date created | Designates the date and optionally the time the content of the image was created rather than the date of the creation of the digital representation | ält Datum då artiklarna scannades. Har använts i enligt med instruktionerna i hjälptexten som anger att det är datum då bilden togs som ska föras in som värde och hur IPCT standarden har använts i exempelbeskrivningar för objekt. se exempelvis http://www.iptc.org/std/photometadata/documentation/userguide/#\_a\_heritage\_artwork\_image\_by\_an\_agency\_photographer | |
| Copyright notice | Contains any necessary copyright notice for claiming the intellectual property for this photograph and should identify the current owner of the copyright for the photograph. Other entities like the creator of the photograph may be added in the corresponding field. Notes on usage rights should be provided in "Rights usage terms". | Unknown används i de fall då de scannade artikalarna undertecknats med psedonym eller saknar angiven författare. Används även då artikelförfattare är känd men fotografi eller bild i artikeln saknas upphovsägare. Public domain används då upphosvrätten för den scannade artikeln gått ut. Copyrighted används då den scannade artikeln är upphovsrättskyddad | |
| Name of featured organisations | Name of the organisation or company which is featured in the image. | Används endast för annonssidor |  |
| Artwork or Object | A structured datatype for details about artwork or an object in an image. Includes a Copyright Notice, a Creator, a Date Created, a Source, a Source Inventory Number and a Title. | Används för att ange vilken årsbok av Hälsingerunor som bilder är scannade från. Använd inte för fotografier av exempelvis konstverk i artiklarna här används istället *Description* | |
| Artwork or Object: Title |  |  |  |
| Artwork or Object: Creator |  |  |  |
| Artwork or Object: Source |  |  |  |
| Artwork or Object: Source Inventory number | |  |  |
| Type of source | The type of the source of this digital image |  |  |
| Licensor | A person or company that should be contacted to obtain a licence for using the item or who has licensed the item. | Används endast för scannade artiklar där upphovsrätten är okänd. Används inte för artiklar i public domain. |  |
| Licensor: name |  |  |  |
| Licensor: identifier |  | Använd id från https://libris.kb.se/katalogisering/ | ält https://libris.kb.se/fcrttcbz55x1520#it |
| Licensor: Phone number 1 |  |  |  |
| Licenso: Email adress |  |  |  |
| Licensor: Web adress |  |  |  |
| Copyright owner | Owner or owners of the copyright in the licensed image. | ält Används endast för scannade artiklar som är upphovsrättsskyddade. Använd id från från https://libris.kb.se/katalogisering/ |  |
|  |  |  |  |
| ält \* för att undvika idiosynkrata översättningar har beskrivningarna av termer hämtats från: https://iptc.org/std/photometadata/specification/IPTC-PhotoMetadata#iptc-core-schema-1-3-specifications och https://iptc.org/std/photometadata/specification/IPTC-PhotoMetadata#iptc-extension-schema-1-6-specifications | | |  |