

## UPPDRAG:

# Bevara Moderna museets samling digitalt i 1 000 år

Arkivbyrån Archiwwwe sammanställde och bearbetade Moderna Museets samling för digitalt bevarande i 1 000 år. Matias Vangsnes, grundare och VD berättar om utmaningar och insikter från projektet.

Text: Matias Vangsnes Bild: artboard.studio  
matias.vangsnes@archiwwwe.se

**I DEN DIGITALA** arkivvärlden har vi lärt oss betydelsen av att löpande konvertera föråldrade filformat och samtidigt flytta över arkiven till nya databärare. På så sätt bevaras arkiven över tid. De flesta arkivförvaltare brukar ägna sig åt detta var 5–10:e år. Den långa intervallen beror främst på att filformaten förändras och att dagens databärare inte har längre garanterad hållbarhet.

Men om det fanns en bärare som höll längre tid? Och om vi enbart behövde bekymra oss om filformatens överlevnad, hur skulle vi göra då? Jo, filerna skulle behöva vara maximalt tillgängliga, det vill säga så öppna som möjligt, men framförallt innehålla dokumentation om själva filformatet.

I projektet med att långtidsbevara Moderna museets samling digitalt föreställde vi oss att mjukvaran för att öppna filerna inte kommer att existera i framtiden. Därför bevarade vi också mjukvarans källkod i

arkivet, så att det i framtiden finns möjlighet att bygga en ny mjukvara som klarar att läsa filerna. Bildfilerna sparades i formatet PNG med tillhörande dokumentation om filformatet och mjukvaran som visar bilderna. Vi lade också till instruktioner om vilka algoritmer som används för komprimering och hur återinläsning av filerna går till. All metadata och den övergripande instruktionen om arkivet sparade vi i rena ASCII textfiler. För att förenkla eventuell framtida transkription formaterades all text till versaler.

Att bädda in mjukvaran i koden är något som Googles Internet evangelist Vint Cerf har belyst med projektet *Digital Vellum* där man ville spara hela kedjan av beroenden för att öppna filer (operativsystem, programvara) i en behållare tillsammans med filernas innehåll. Det var högtflygande planer som fortfarande idag mest är teorier, och komplext att lösa tekniskt.

### ATT KOMMUNICERA MED FRAMTIDEN

För att eftervärlden på bästa sätt skall kunna tolka informationen i ett arkiv med tusentals avbildade objekt krävs övergripande information om arkivet och metadata om varje enskilt objekt. Det är idealiskt om arkivförvaltarna också med hundra års intervall styrker autenticiteten och vid behov adderar kommentarer och information som hjälper framtida tolkning. Den tusen år gamla makimonon "Trä och sten" av Su Shi från Songdynastin är ett bra exempel på just omsorgsfullt omhändertagande av ett objekt över lång tid. Varje enskild förvaltare har inte bara bidragit till exemplarisk proveniens utan också lämnat kommentarer, vilket skapar ett mervärde i sig.

### ENTROPI ÄR DIGITALA DATABÄRARES FIENDE

De tillgängliga lagringsmedierna som vi hittills har haft för att lagra digital information, såsom hårddisk, magnetband, optiska media, flashminne med flera, har alla ett gemensamt

tillkortakommande: en i arkivsammanhang kort livslängd. Det är areadensiteten, det vill säga mängden data per ytenhet, som avgör hur mycket som går att lagra i mediet. Och när behovet av att lagra mer information ökar försöker tillverkarna tillmötesgå detta genom att öka densiteten. Lagringsmedierna idag har extremt hög densitet och är därför extra känsliga för den mikroskopiska nedbrytning som konstant påverkar all materia i universum.

Exempelvis har en tusen år gammal runsten med informationen "Björn, Finnvid's son, lät resa denna sten till minne av sig själv" en låg densitet, och håller därför över lång tid.

Areadensitet mäts i mättenheten Tbps (Terrabyte per kvadrattum). Om man räknar runstensens antal tecken per yta kan de jämföras med nutida databärare.

Varje tecken (1 byte) på runstenen upptar cirka 4 kvadrattum vilket blir 0,0000000000025 Tbps, jämfört med ett modernt flashminne på 2,8 Tbps. Runstenen rymmer 64 tecken medan flashminnet klarar 100 biljoner. Båda bryts ner lika snabbt, men runstenen med sin låga densitet överlever i 10 000 år medan flashminnet endast håller cirka 10 år på grund av den höga densiteten.

### BEPRÖVAD TEKNIK MED NY TEKNOLOGI

Norska företaget Piql AS har tagit fram en typ av film som i vetenskapliga tester (Norner AS i Norge och Image Permanence Institute i Rochester, USA) visar att vid goda lagringsförhållanden kan informationen på filmen bevaras i minst tusen år. Filmen har låg areadensitet jämfört med andra tillgängliga databärare idag: en filmkapsel har en lagringskapacitet på 500 GB. Filmen är oföränderlig vilket innebär att det går att skriva och sedan läsa men inte ändra informationen. Dessa egenskaper lämpar sig för digitalt långtidsbe-



Den tusen år gamla makimonon "Trä och sten" av Su Shi från Songdynastin, källa: Wikipedia, public domain.



Pablo Picasso, La source från 1921, en utav de cirka 80 000 digitaliserade objekt i Moderna museets samling som nu bevaras för eftervärlden. Foto: Prallan Allsten, Moderna Museet.

varande med inga eller mycket långa intervaller för arkivvård.

### RUTOR PÅ FILM I 1 000 ÅR

Mycket kan hända mänskligheten under tusen år. Informationen skrivs på filmen i form av så kallade rutkoder (QR-kod); tvådimensionella koder för optisk avläsning. En fördel med QR-koden är att den kan läsas även om filmen är något skadad. QR-kod är en beprövad teknik som visar sig vara extra robust då filmen har

en extremt hög upplösning. För att skriva och läsa filmen har Piql tagit fram både en egen filmskrivare och en film läsare. Instruktioner om själva filmen har skrivits direkt i text på filmens intro och kan läsas direkt med till exempel ett förstoringsglas. Piql har även tagit fram en molnlösning för tillgång till innehållet på filmen.

### 80 000 DIGITALISERADE OBJKT

Uppdraget för Moderna Museet innebar att samla in och



Filmformat som kan bevara digital information i minst 1 000 år. Förutom QR-kod går det att skriva läsbar text och bild direkt på filmen. Källa: Piql AS

bearbeta 80 000 objekt från museets samling via den webbaserade bild databasen "Sök i samlingen". Det första vi gjorde var att indexera samtliga sökvägar till varje objekt. När vi hade definierat sökvägarna programmerades ett skript (i kommandotolken BASH) som laddade hem objektens bilder och tillhörande metadata. Därefter programmerades skript för att bearbeta och formatera metadata. Slutligen adderade vi kontrollsummor (MD5) till filerna för att möjliggöra kontroll av arkivet i framtiden.

"Det är areadensiteten, det vill säga mängden data per ytenhet, som avgör hur mycket som går att lagra i mediet."

#### METADATA SPELAR NYCKELROLL

Moderna Museet öppnade 1958. Om vi hypotetiskt öppnar det digitala bildarkivet om 1 000 år kanske inte avsändaren existerar längre, sannolikt inte heller de digitala verktyg vi har idag. Så när vi begrundar de digitaliserade objekten är det viktigt att skicka med information om vad objekten förställer, vem som är upphovsman, när och var verken uppkommit och när de avbildades.

För att arkivet skulle bli så konsekvent som möjligt valde vi att reducera objektens tillhörande metadata till de som är obligatoriska för systemet. I metadatafilerna ingår verkens upphovsmän, datum, storlek och beskrivningar etc. tillsammans med metainformation om bilderna. Vi använde mjukvara baserad på öppen källkod till att extrahera information från varje enskild bild som vi sedan adderade till textfilen med metadata.

```

1 #####
2 ARCHIVE META: SNAPSHOT OF THE MODERNA MUSEUM COLLECTION
3 #####
4 ARCHIVE NAME: SIS.MODERNAMUSEET.SE
5 ARCHIVE SNAPSHOT DATE: 2020-02-20
6 ORGANISATION NAME: MODERNA MUSEET
7 SERVICE BUREAU: ARCHIVUME SWEDEN AB
8 THE NAME OF THE SYSTEM: THE MUSEUM SYSTEM (TMS)
9 ORGANISATION LOCATION: SKEPPSHOLMEN, STOCKHOLM, SWEDEN
10 ORGANISATION POSITION: 59°19'33.9"N 18°05'04.9"E
11 SOURCE TRANSFER PROTOCOL: HTTP (TLS)
12 #####
13 # ARTIFACT META: RENDITION NUMBER: NM 6360
14 #####
15 ARTIST: PABLO PICASSO
16 CATEGORY: PAINTINGS
17 DATE: 1920S
18 COUNTRY: SPAIN (300)
19 ARTIST: PABLO PICASSO
20 1881 - 1973
21 SPAIN LA SOURCE, 1921
22 KALLAN
23 THE SPRING
24 OIL ON CANVAS
25 64 X 90 CM
26 PAINTINGS
27 DONATION 1970 FRÅN GRACE OCH PHILIP SANDBLOM
28 NM 6360 THIS ARTWORK IS NOT ON DISPLAY
29 #####
30 SAVE
31 - + ABOUT THE WORK
32 IN THE 1921 PAINTING THE SOURCE, WE SEE A WOMAN RECLINING IN AN
33 ARCADIAN LANDSCAPE. ARCADIA IS THE NAME OF A REGION IN GREECE, BUT ALSO
34 THE MYTHOLOGICAL HOME OF THE GOD PAN AND THE NYMPHS. THE WOMAN DEPICTED
35 IN THIS PAINTING IS PICASSO'S FIRST WIFE, OIRA KHOKHILOVA. AT THE TIME
36 #####
37 Pain Text Tab Width: 2 Lin 14, Col 43 INS

```

Exempel: utdrag ur ett objekts tillhörande META-data. Källa: Archiwvwe

#### ERFARENHETER OCH INSIKTER

Vilka erfarenheter och insikter kan vi dela med oss från projektet? Tidigt insåg vi att lagring på film inte är vidare tillämpningsbart på arkiv där arkivvården kräver korta intervall. Det krävs också öppnare filformat och mer dokumentation av tillhörande mjukvara, samt instruktioner för att undvika att mjukvaran bli föråldrad över tid. För riktigt långa intervall är däremot det nya filmediet perfekt.

En annan insikt är hur viktig förberedelsefasen är. Innan vi startade ställde vi en rad frågor in till projektet som vi eller arkivförvaltaren behövde svara på:

- Vilken terminologi och vilka benämningar skall vi använda vid dokumentation?
- Är databasen konsekvent och allt innehåll relevant?
- Skall vi begränsa omfånget?
- Har bildfilerna tillräcklig upplösning?
- Innehåller bildfilerna textinformation (EXIF)?
- Vilken identifierare skall vi använda för objekten?
- Vilken struktur är lättast att navigera?
- Innehåller databasens utdata etiketter för att identifiera fält?

När frågorna fått svar hade vi en tydlig bild av hur vi skulle gå tillväga och därefter föll stegen i processen naturligt på plats.

Antalet filformat var begränsat till en handfull öppna format. Bearbetningen av arkivet tog ungefär 80 timmar. Om underlaget hade varit mer komplext, med fler och dessutom leverantörsspecifika filformat hade det tagit betydligt längre tid.

Slutsatsen är att tekniken finns och går att appliceras på digitala arkiv som innehåller öppna filformat och inte måste konverteras eller flyttas över till nya databärare. Det är också över längre tid som film blir kostnads-effektivt.

Filmens stora fördelar jämfört med traditionella databärare är framförallt hållbarheten. Det finns säkert många organisationer som skulle ha stor nytta av att kunna lagra arkiv utan att vara beroende av hårdvaror, elektricitet och internet. Det gäller exempelvis sekretessbelagd information som förvaltaren måste lagra i hundra år och samtidigt vill begränsa tillgången till. Filmerna kan då lagras på ett säkert ställe i hundra år precis som vi gör idag med fysiska pappersarkiv, men till en bråkdel av kvadratmeterpriset. ✕