



Riksarkivet

Nordisk Elarkivseminar 2011

Konverteringsprojekt P47

Magnus Geber

2011-05-27



P47 (Projekt 47)

Migrering av alla born-digital e-arkivfiler på gamla magnetband (9-kanals och 3480)

- Pågående projekt.
- Konvertering har i huvudsak inte utförts ännu
- Fysisk kopiering och recovery genomförts
- Inventeringsarbete genomförts



MÅL

- Migrera alla 9-kanals magnetband och 3480 bandkassetter.
- Kontrollera datafilerna logiskt och mot dokumentation för att upptäcka ej accepterade format
- Rätta till ej accepterade format genom olika typer av konverteringar.
- Upprätta ny dokumentation.
- Skapa nya lagringspaket.
- Registrera i Arkis2.



Bakgrund

- Riksarkivet har sedan 1970-talet tagit emot leveranser av Born digital-material.
- De tidigaste leveranserna ej tekniskt kontrollerade pga ingen fungerande utrustning före 1993
- Bevarande har varit eftersatt då nya leveranser tog det mesta av resurserna
- Viss migrering har skett genom åren
 - 1986 servicebyrå
 - SCB
 - 1996 fysisk kopiering till server



Förstudien

- En inventering av vilka arkivalier som var i behov av konvertering/migrering/kopiering.

Bedöma:

- Arbetets omfattning:
 - Vilka slags kopieringar och konverteringar som behöver göras
 - Tid- och resursåtgång för detta, räknat i personmånader och kalendertid.
- Behov av personal,
- Personalens typ av kompetens.



Projektets organisation

- 3 grupper
 - Intern personal och 1 projektanställd
- Inventeringsgruppen
 - Teknisk status
 - Metadata
- Konverteringsgruppen
 - Vilka konverteringar
 - Prioriteringar
- Arbetsprocessgruppen.
 - Sammanfatta
 - Skapa rutinbeskrivningar
- Parallell utveckling av verktyget KRAM



Inventering

- Antal filer: ca 14 725
- Antal filer med logiska problem: ca 1000
- Antal filer med teckenrepresentation som måste konverteras: ca 11000
- Fysiska problemband: 289 (av ca 12 000 levererade band). Avtal med företag i Norge för inläsning.



Fysisk kopiering, recovery

- Inläsningar problemband del 1 (2010):
- 1:a omgången skickade 100 band ex1
- 2:a omgången skickade 30 band ex2
- Förlorad information, ca 3 block i en fil.
- Inläsningar problemband del 2 (2011):
- 189 band ex1 (ej skickat)



Förarbete

- Välj ut vad som ska bearbetas utifrån prioriteringslistan.
- Förbered filer samt plocka fram metadata
- Skanna metadata på papper till fil (PDF/A-1b)
- Identifiera vilka konvertering som behöver göras genom analys av filer,
- Genomföra prekonverteringar
- Förbered körning i KRAM (kontrollramverket) genom att fylla i Excelark med metadata.



Konvertering

- Konvertering av teckenrepresentation, packade fält.
- Konvertering/utblankning av skräptecken.
- Utbrytning av repeterande grupper, posttyper (som finns i en och samma fil).
- Kontroll av att konverteringen gjorts korrekt.
- Korrigering av metadata (till exempel när nya filer måste skapas).
- Filnamngivning.

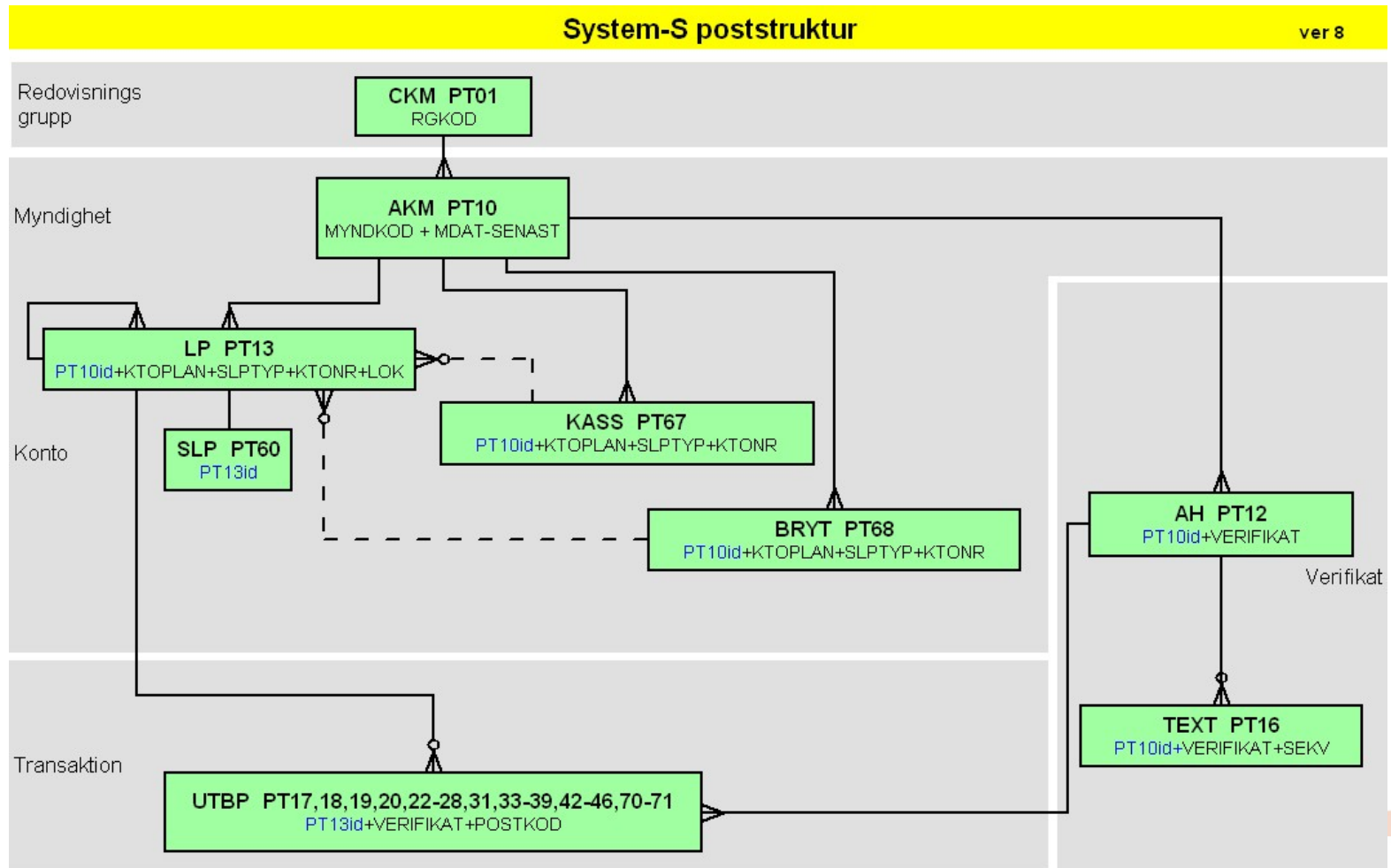


KRAM (Kontrollramverket)

- Verktyg för kontroll och viss konvertering av data och metadata (för närmare information se separat föreläsning).
- Packetering i AIP.
- Registrerar i ARKIS
- Samt arkivläggning i RADAR (EssArc), på disk (Online) och på band 2 ex (offline).



Exempel: System S



Förklaring

POSTNAMN PT**POSTKOD**
Fält som bedöms bilda unikt ID ingående ärvda fält

System S, åtgärder

- KRAM

- Konvertering av Packed Decimal
- Uppdelning av posttyper till olika filer
- Logisk koppling genom nya nycklar
- Även räknare på filnivå

- Preprocess

- Borttagning av kontroll och skräptecken
- Nya posttyper PD/text
- Förlängning med blanktecken
- Nya posttyper olika generationer postbeskrivningar
- Nya posttyper, intern hierarki



Erfarenheter hittills

- Det går att rädda data!
- Krävs mycket tid och resurser för att driva igenom ett sådant här projekt.
- Finns en vilja inom organisationen att genomföra projektet men det skall inte synas i plånboken.
- Önskan om dedicerad personal har inte infriats helt och hållet, fått en person på 100% i 6 månader.
- Fått medel för inläsningar av problemband, samt för inköp av skannner.
- Svårt att förutse allt

