

Digisams förslag till principer för verksamhet som omfattas av regeringens digitaliseringsstrategi 2011-15, Digit@lt kulturarv

Förslaget utgår från TOGAFs modell för arkitekturprinciper.¹ Det är utarbetat av Digisam (Johanna Berg, Sanja Halling och Rolf Källman) med konsultstöd av Peter Krantz. Det har vidare ventilerats med experterna i Digisams styrgrupp: Martin Bjersby (Riksarkivet), Torsten Hökby (RAÄ), Göran Konstenius (KB) och Sven Rentzhog (Centralmuseernas samarbetsråd/Nordiska museet).

Inledning

För att möjliggöra effektivt samarbete mellan ett stort antal aktörer krävs en samling överenskomna grundläggande principer. Principerna är tänkta att användas som stöd för inriktning av beslut och bidra till att samordna digital informationshantering på ett effektivt och ändamålsenligt sätt.

Eftersom principerna ska vara långsiktiga bör de inte innehålla referenser till specifika tekniker. Sättet att leva upp till principerna kommer att variera över tid. En förhoppning är dock att principerna leder till en diskussion om behov av mer specificerat stöd i form av t.ex. riktlinjer och anvisningar.

Principerna är indelade utifrån områdena styra, producera, använda och bevara.

---- Styra ----

P1. Digitaliseringsverksamhet ska styras av respektive institutions ledning

Beskrivning: Digitaliseringsverksamhet ska utgå från planer och strategier som beslutats av institutionen.

Motiv: Digitaliseringsverksamhet som inte har stöd från ledningen riskerar att inte få de resurser som behövs för att genomföra effektiv produktion eller för tillgängliggörande och långsiktig förvaltning av producerad eller insamlad data.

Konsekvenser:

- Varje institution behöver en fastställd strategi och plan och för digitalisering, tillgängliggörande och bevarande.
- Varje institution behöver fatta beslut om planerad digitaliseringsverksamhet.
- Planer och strategier formuleras utifrån nationella rekommendationer.

P2. Respektive institution ska ansvara för den metadata - analog såväl som digital - som rör institutionens samlingar, liksom för de objekt som är skapade digitalt

Beskrivning: Samlad information om såväl analoga som digitala samlingarna ska förvaltas på ett stabilt och ansvarsfullt sätt och kvalitetshöjning av metadata prioriteras.

Motiv: Högkvalitativ och uppdaterad metadata om förvaltade samlingar är en förutsättning för utvecklad användbarhet och ökad användning med större samhällsnytta som resultat.

Konsekvenser:

- Varje institution behöver avsätta resurser för en aktiv förvaltning av sina digitala samlingar.

P3. Digitaliseringsprojekt ska följas upp och utvärderas av respektive institution

Beskrivning: Beslutade digitaliseringsinsatser och projekt ska följas upp och utvärderas utifrån gemensamma kriterier. Resultat av uppföljningar och utvärderingar ska dokumenteras och delas med andra för erfarenhetsutbyte och gemensamt lärande.

Motiv: Genom uppföljning och erfarenhetsutbyte ges insyn i digitaliseringsprojekt och möjlighet att undvika misstag som andra har gjort. Det ger ett effektivare resursutnyttjande och bidrar till kompetensutveckling. Det bidrar även till spridning av goda exempel.

Konsekvenser:

- Varje institution behöver inför digitaliseringsinsatser planera för uppföljning och utvärdering.
- Varje institution behöver publicera resultatet av uppföljningar och utvärderingar, även av mindre framgångsrika projekt. Uppföljningar och utvärderingar görs utifrån nationella rekommendationer.

---- **Producera** ----

P4. Digitalisering ska genomföras så att resurser och kompetens används på ett effektivt och ändamålsenligt sätt

Beskrivning: I samband med planering och genomförande av digitaliseringsprojekt ska en avvägning göras av var och hur digitaliseringsinsatser kan genomföras på det mest effektiva och ändamålsenliga sättet utifrån fastställd kravspecifikation. Massdigitalisering bör alltid övervägas.

Motiv: Genom att använda existerande digitaliseringsutrustning utnyttjas gjorda investeringar på ett bättre sätt. Det leder till högre kvalitet, lägre kostnader och bättre utnyttjande av den kompetens som finns på respektive institution. Genom att koncentrera produktionen till färre och mer storskaliga produktionsanläggningar kan

samhällets totala kostnader för utrustning samt vidmakthållande och utveckling av kompetens optimeras.

Konsekvenser:

- Varje institution behöver undersöka vilka alternativ som finns inom och utom offentlig sektor innan egen utrustning anskaffas eller intern produktion initieras.
- Den institution som gjort investeringar i utrustning och kompetens behöver göra det så enkelt som möjligt att samutnyttja dessa resurser med andra institutioner.

P5. Öppna standarder och specifikationer lämpade för tillgängliggörande och långsiktigt bevarande ska användas

Beskrivning: Rekommenderade öppna standarder¹ ska användas i all digitaliseringsverksamhet.

Motiv: Användning av öppna standarder ger förutsättningar för långsiktigt bevarande, interoperabilitet² samt bred användning och återanvändning av digitaliserat material.

Konsekvenser:

- Varje institution liksom varje förvaltare av aggregeringstjänster behöver tillämpa rekommenderade standarder.
- Varje institution liksom varje förvaltare av aggregeringstjänster behöver i anskaffningsförfaranden säkerställa att utrustning och programvara kan hantera dessa standarder.

---- Använda ----

P6. Publikt tillgängliggörande ska vara en del av varje digitaliseringsinsats

Beskrivning: Det digitaliserade materialet ska tillgängliggöras så snart som möjligt och på ett sätt som möjliggör en så bred användning som möjligt.

Motiv: Digitaliserad information har ett stort värde för samhället. Ett snabbt tillgängliggörande gör det möjligt för aktörer utanför institutionen att använda den inom andra områden, samtidigt som slitage på handlingar och föremål minskas och medborgarna ges ökad tillgång oberoende av fysisk plats.

Konsekvenser:

¹ Med öppen standard avses här den definition som återfinns i European Interoperability Framework version 1.0 s 9.

² Interoperabilitet är förmågan hos olika system att fungera tillsammans och kommunicera med varandra. Exempel på detta kan vara att systemen kan använda samma protokoll eller läsa och skriva samma filformat. Interoperabilitet uppnås genom samordning och (internationella) standardiseringsprocesser.

- Planering av digitaliseringsinsatser behöver även innefatta en plan för tillgängliggörande av resultatet.
- Digitaliserat material bör tillgängliggöras löpande.

P7. Digitaliserade textresurser ska tillgängliggöras som maskinläsbar text

Beskrivning: För att nå sin fulla potential av nytta och användbarhet för många olika användargrupper behöver digitaliserad text tillgängliggöras på ett sådant sätt att den kan processas maskinellt.

Motiv: Textresurser som kan maskinläsas möjliggör bred och högkvalitativ användning för många grupper som har begränsad nytta av text som bara kan läsas av det mänskliga ögat. Här har många forskare samma grundläggande behov som icke svensktalande eller personer med funktionsnedsättning.

Konsekvenser:

- Varje institution behöver inför digitalisering av textresurser planera för hur informationsinnehållet kan göras maskinläsbart.

P8. Digitala filer och metadata ska vara så enkla som möjligt för andra att använda och återanvända

Beskrivning: Institutionen vinnlägger sig om att göra filer och metadata lätta att hitta och använda. De digitala objekten ska vara tydligt upphovsrättsligt licensierade/märkta.

Motiv: Nyttan med information uppstår när den används. Genom att göra det så enkelt som möjligt för så många som möjligt att använda resultatet av digitalisering skapas maximal nytta för samhället. Genom att tydliggöra rättigheter och använda öppna licenser bidrar institutionen till kunskapsspridning och nytt skapande.

Konsekvenser:

- Rättighetsmärkning behöver anges på ett maskinläsbart sätt.
- I oklara fall (orphan works) behöver institutionen göra en avvägning som balanserar eventuell kränkning av enskilda rättigheter mot det breda samhällsintresset av öppen licensiering.
- Institutionen behöver etablera en enkel kontaktväg för den som har synpunkter på tillgängliggörandet och en rutin för att eventuellt dölja eller avpublicera material.

Se även: P5

P9. Digitalisering av material där skyddstiden löpt ut ska inte medföra att nya rättigheter uppstår

Beskrivning: Återgivning i nya medier av material med utslocknade rättigheter skapar inte nya verk, med nya skyddstider. Sådana resurser ska inte licensieras men rättighetsmärkas på ett maskinläsbart sätt.

Motiv: Nyttan med information uppstår när den används. Genom att tillhandahålla fria resurser utan villkor bidrar institutionen till kunskapsspridning, innovation och nytt skapande.

Konsekvenser:

- Rättighetsmärkning behöver anges på ett maskinläsbart sätt.
- Material med utslocknad upphovsrätt tillhandahålls för återanvändning utan restriktioner. Angivande av upphovsman/källa bör uppmuntras men kan inte krävas.

P10. Länknings mellan egen och andras information ska eftersträvas

Beskrivning: Information ska kontextualiseras på ett standardiserat sätt genom att kopplas till information utanför den egna institutionen.

Motiv: Kulturarvsinformation har naturliga kopplingar till många platser, händelser, tidsperioder, personer och objekt. Genom att koppla ihop digital information med andra källor ökar möjligheterna för återanvändning av materialet. När sammankopplingen sker på ett standardiserat sätt sänks tröskeln för användare av informationen och nya möjligheter öppnas för bl. a. forskning.

Konsekvenser:

- De digitala objekten behöver förses med persistenta identifierare i enlighet med gemensamma riktlinjer.
- Institutionen eftersträvar semantisk länkning till externa informationsresurser.
- Institutionen behöver göra det lätt för andra att semantiskt länka till information som förvaltas av den egna institutionen.

P11. Semantiska beskrivningar ska utformas på ett standardiserat sätt

Beskrivning: Metoder för semantisk³ beskrivning som används i informationen ska vara dokumenterade på ett standardiserat sätt och publicerade på webben med en

³ Medan världswebben är en standard för att presentera (sammanlänkade) *dokument* så går den semantiska webben längre och standardiserar även hur man uttrycker *data* och *kunskap*. Man har bl a standardiserat hur man refererar till objekt som man vill beskriva. De identifieras med unika identifierare, så kallade URIs, som tekniskt sett är webbadresser. Data och kunskap uttrycks sedan genom så kallade tripletter sammansatta av subjekt, predikat och objekt som var för sig identifieras på ett standardiserat sätt. Att beskriva semantik på ett standardiserat sätt innebär att redogöra för på vilket sätt man semantifierat sitt datainnehåll.

öppen licens. Etablerade vokabulärer ska återanvändas för att främja en gradvis framväxande semantisk interoperabilitet.

Motiv: Med semantiska beskrivningar blir det lättare för användare att integrera informationen i egna lösningar. Utan sådan beskrivning riskerar man att användare drar felaktiga slutsatser av informationen. Genom att beskriva semantik på ett standardiserat sätt möjliggörs återanvändning och oklar tolkning av begrepp och definitioner minimeras.

Konsekvenser:

- För att främja interoperabilitet på flera nivåer bör redan etablerade semantiska beskrivningar för tvärssektoriell kulturarvsdata återanvändas.

P12. Samarbete runt informationen ska möjliggöras

Beskrivning: Institutionen ska uppmuntra allmänheten att bidra med kunskap till den gemensamma informationsbanken och göra det enkelt för alla att bidra med information.

Motiv: Mycket kunskap om samlingarna finns utanför institutionen som därför bör göra det så lätt som möjligt för externa krafter att bidra. Användningen ökar om kvaliteten höjs och det finns många som kan rätta, komplettera och länka information hos olika förvaltare.

Konsekvenser:

- Plattformer för användargenererat innehåll behöver användas.
- Rutiner och verktyg behöver implementeras så att värdefull information kan tas emot och relateras till institutionens egen metadata i samlingshanteringssystem eller motsvarande.

---- Bevara ----

P13. Varje institution ska ha en strategi för bevarande av digital information

Beskrivning: Vid digitalt långtidsbevarande måste, förutom tekniskt bevarande, även bevarande av äkthet och kontext beaktas för att informationen i framtiden ska kunna återsökas och förstås.

Motiv: Digitalt bevarande innebär interoperabilitet⁴ över tid, dvs. att informationen är långsiktigt användbar och tillgänglig. För att digital information ska vara tillgänglig

⁴ Interoperabilitet är förmågan hos olika system att fungera tillsammans och kommunicera med varandra. Exempel på detta kan vara att systemen kan använda samma protokoll eller läsa och skriva samma filformat. Över tid handlar det om att inte fastna i (tekniska) lösningar som blir omöjliga att läsa/tolka i nya tekniska system.

och användbar i framtiden krävs användning av lämpliga standarder, paketering av information och tekniska åtgärder.

Konsekvenser:

- Information som genereras genom digitalisering behöver utformas så att kontexten säkras så att informationen kan återsökas och förstås i framtiden.

Se även: P5.

P14. Aktuella digitaliseringsinsatser ska beskrivas på institutionens webbplats

Beskrivning: Varje institution ska förteckna pågående och planerade digitaliseringsprojekt på sin webbplats. För varje projekt ska en kort beskrivning ges av omfattning, tidsplan och eventuell tillhörande dokumentation.

Motiv: Ett stort antal digitaliseringsprojekt planeras och genomförs parallellt. För att undvika dubbelarbete och underlätta samverkan bör information om projekt publiceras öppet för insyn.

Konsekvenser:

- Organisationen behöver inrätta rutiner för att publicera information om planerade, pågående och avslutade digitaliseringsprojekt.

ⁱ För vidare info se <http://pubs.opengroup.org/architecture/togaf9-doc/arch/chap23.html>

Utgångspunkterna för TOGAF är av Peter Krantz formulerade som följer:

Krav på principer

För att principer ska bli användbara brukar man ställa följande krav på dem. Principer bör vara:

1. Begripliga - så att missförstånd undviks
2. Robusta - för att kunna användas som underlag för beslut även i svåra situationer.
3. Kompletta - för att kunna vara stödjande i samtliga typer av situationer som kan uppkomma.
4. Konsekventa - så att principer inte motsäger varandra eller är för detaljerade.
5. Beständiga - de bör vara formulerade med tillräcklig flexibilitet för att kunna hantera omvärldsförändringar utan att de löpande behöver göras om (t.ex. undvika för specifika tekniska detaljer).

Mall för att formulera principer

Förslagsvis dokumenterar man en princip med följande delar:

1. "Viljetrytning". En mening om vad man vill.
2. Motiv till principen. 1-2 stycken som beskriver principen kortfattat och varför den behövs.
3. Möjliga konsekvenser av principen i punktform. (i klarspråk, även "jobbiga" konsekvenser)

För ett exempel se princip #16:

http://pubs.opengroup.org/architecture/togaf9-doc/arch/chap23.html#tag_23_06_03_01

Mall**Name**

<Name of Principle>

Statement

The Statement should succinctly and unambiguously communicate the fundamental rule. For the most part, the principles statements for managing information are similar from one organization to the next. It is vital that the principles statement be unambiguous.

Rationale

The Rationale should highlight the business benefits of adhering to the principle, using business terminology. Point to the similarity of information and technology principles to the principles governing business operations. Also describe the relationship to other principles, and the intentions regarding a balanced interpretation. Describe situations where one principle would be given precedence or carry more weight than another for making a decision.

Implications

The Implications should highlight the requirements, both for the business and IT, for carrying out the principle – in terms of resources, costs, and activities/tasks. It will often be apparent that current systems, standards, or practices would be incongruent with the principle upon adoption. The impact to the business and consequences of adopting a principle should be clearly stated. The reader should readily discern the answer to: "How does this affect me?" It is important not to oversimplify, trivialize, or judge the merit of the impact. Some of the implications will be identified as potential impacts only, and may be speculative rather than fully analyzed.