



Sammanställning av remissvar FGS Databas (Alternativ 1, relationsdatabaser och register)

Ett utkast till FGS Databas (alternativ 1, relationsdatabaser och register) remitterades den 20 juni 2019. Remissen var öppen och kunde besvaras av specifikationens samtliga intressenter. Det vill säga av såväl statliga och kommunala myndigheter som av leverantörer av olika system och andra intresserade. Sista svarsdatum för remissen var den 30 september 2019. Remissen kommunicerades ut via Riksarkivets hemsida och riktades även direkt till de 579 mottagarna i adresslistan för nyhetsbrevet om FGS:er samt till 13 statliga myndigheter, 7 kommuner/regioner¹ och 1 enskild organisation som varit med i projektets referensgrupp.

Det har sammanlagt inkommit 8 svar på remissen. Av dessa är det 1 svar som har kommit från en privatperson, 1 svar som kommit från ett företag, 2 svar som kommit från regionala kommunala myndigheter och 4 svar som har kommit från statliga myndigheter.

De aktörer som på något sätt har besvarat remissen är följande:

Arbetsförmedlingen
ArkivIT
En privatperson
Polisen
Regionarkivet, Göteborg
Regionarkivet, Skåne
Skolverket
Stockholms universitet (SU)

Remissen omfattade det övergripande konceptet för FGS Databas, men först och främst förslaget till FGS Databas (Alternativ 1, relationsdatabaser och register)². Detta beskrevs i de enskilda dokument som utkastet till specifikation innehåller vilka var:

"Introduktion till förvaltningsgemensamma specifikationer (FGS), FGS Databas"

"FGS Databas (Alternativ 1, relationsdatabaser och register)"

"FGS Databas (Alternativ 1, relationsdatabaser och register), Tillägg"

Remissen innefattade *ej* FGS Databas (Alternativ 2, relationsdatabaser)³ som är en specifikation som i framtiden kommer att ingå i familjen FGS Databas.

Önskemålet var att remissvaren skulle fokusera på följande frågeställningar:

¹ En av dessa är egentligen ett samarbete mellan 10 olika regioner

² Som bygger på ADDML-formatet

³ Som bygger på SIARD-formatet



1. Är texten i dokumentet "*Introduktion till förvaltningsgemensamma specifikationer (FGS) FGS Databas*" begriplig, beskrivande och klargörande? Finns det delar som borde förtydligas eller som saknas?
2. Är de element som finns beskrivna i dokumentet "*FGS Databas (Alternativ 1, relationsdatabaser och register)*" tillräckliga för överföring av informationen i databaser? Vilka element saknas?
3. Är texten i dokumentet "*FGS Databas (Alternativ 1, relationsdatabaser och register)*" begriplig, beskrivande och klargörande? Finns det delar som borde förtydligas eller som saknas?
4. Är texten i dokumentet "*FGS Databas (Alternativ 1, relationsdatabaser och register), Tillägg*" begriplig, beskrivande och klargörande? Finns det delar som borde förtydligas eller som saknas?
5. Är de egendefinierade värdelistorna i dokumentet "*FGS Databas (Alternativ 1, relationsdatabaser och register) Tillägg*" relevanta och begripliga?

Remissen gick ut i ett gemensamt missiv tillsammans med remissen för en anpassad FGS för arkivleveranser, *Anpassning av FGS Databas (alternativ 1, relationsdatabaser och register)*⁴ för arkivleveranser, då anpassningen har ett beroende till *FGS Databas (alternativ 1, relationsdatabaser och register)*. I missivet poängterades att det var två separata remisser och att svar skulle skickas till två olika e-postadresser. Det är dock tydligt att de två remisserna skapat viss förvirring, och ett flertal av remissvaren har slagit ihop svaren på båda remisserna i en text. Detta får följden att vissa svar upprepas i de båda sammanställningarna.

Den följande sammanställningen av remissvaren går igenom och kommenterar de inkomna svaren på frågorna ovan. Därefter följer en sammanställning över övriga inkomna synpunkter på förslaget till *FGS Databas (Alternativ 1, relationsdatabaser och register)*.

1. Är texten i dokumentet "*Introduktion till förvaltningsgemensamma specifikationer (FGS) FGS Databas*" begriplig, beskrivande och klargörande? Finns det delar som borde förtydligas eller som saknas?

Av de 8 inkomna remissvaren så har samtliga direkt eller indirekt besvarat frågan om texten i introduktionen är begriplig, beskrivande och klargörande.

Bland de inkomna remissvaren framgår det i 4 av svaren att man anser att texten är tillräckligt begriplig, beskrivande och klargörande. Ett svar lyfter exempelvis fram att texten är "*tydlig och föredömligt kortfattad*", och endast möjligen skulle önska att det inkluderades "*länkar till de officiella sidorna för ADDML och SIARD*".

I 2 av remissvaren framgår att det råder viss förvirring kring att introduktionsdokumentet tar upp både ADDML och SIARD. I det ena svaret framförs att det blir "*förvirrande när man nämner att alternativ 1 är om ADDML men trots detta tar upp SIARD i flera av avsnitten*". I det andra svaret påpekas att dokumentet "*ger en mycket bra bild av ADDML, men inte riktigt lika förklarande i förhållande till SIARD*", något som man menar skapar "*en obalans i dokumentet, där ADDML framstår som det primära alternativet*".

⁴ RA 21-2017/6701



Det är 4 remissvar som särskilt tar upp avsnittet om skillnader, för- och nackdelar med ADDML och SIARD. I ett av dessa svar skrivs att man önskar att skillnaden mellan Alternativ 1 (ADDML) och Alternativ 2 (SIARD) tydliggörs *”så att det inte finns någon tvekan om vilket format som är lämpligast i ett konkret fall”*. I ett annat framförs att det vore *”önskvärt med praktiska rekommendationer kring de två standardernas tekniska lämplighet”*. I samma svar önskar man även att skillnaderna åskådliggörs på ett mer tydligt sätt, exempelvis i tabellform, och menar att det blir svårt att överblicka när de två standardernas styrkor och svagheter blandas på det sätt det görs. I ytterligare ett av svaren menar man att avsnittet om skillnader, för- och nackdelar inte tar upp några nackdelar utan endast skillnader och fördelar. I det svaret påpekar man även att informationen om ADDML och SIARD förekommer på flera ställen i dokumentet, och man anser att det vore lättare att orientera sig kring de två standarderna om all text var samlad under en rubrik. Slutligen har ett av svaren specifikt kommenterat punkt 9 i listan och framfört att den är ganska missvisande då man anser att båda formaten klarar av att definiera databasschema.

I 1 av de inkomna remissvaren efterfrågas en tydligare begreppslista med kortare beskrivningar av t.ex. leveransöverenskommelse och sekventiell textfil: *”Vad läsaren är insatta i för begrepp beror på dess yrkesroll. Det kan vara bra [då] innebörden av många av begreppen inte är kända för läsaren”*. I samma remissvar anser man även att det behövs en bättre övergång till punktlistan i avsnitt 4.1 *omfattning och avgränsning* som förklarar vad syftet är. Man framhåller även att historik och tillämpning bör separeras.

I 1 av remissvaren tar man upp modellen i avsnitt 5.1.3 och lyfter fram vissa otydligheter i den, samt i den tillhörande texten. I svaret presenteras även ett förslag på en annan modell.

I 1 av remissvaren framförs att man anser att Riksarkivet borde vara *”ett stöd i att närmare förklara och utbilda i de standarder som FGS:en baseras på”* istället för att hänvisa till de organ som ger ut respektive standard.

Utöver det är det 1 remissvar som har besvarat frågan genom att påpeka en rad språkliga fel i dokumentet samt efterfrågat några exempel på typer av *”äldre databaser”* i avsnitt 4.1.1.

Kommentarer: Dokumentet kommer att ses över och tillmötesgå ovan uppräknade önskemål så långt det är möjligt. Särskilt kommer texten om ADDML och SIARD att ses över och förtydligas.

Länkar till de *”officiella”* sidorna för ADDML och SIARD kommer att läggas till.

Anledningen till att både ADDML och SIARD tas upp i dokumentet är för att det är meningen att introduktionsdokumentet ska gälla för båda de två alternativ som kommer att ingå i familjen FGS Databas: FGS Databas (Alternativ 1, relationsdatabaser och register) som bygger på ADDML-formatet och FGS Databas (Alternativ 2, relationsdatabaser) som bygger på SIARD-formatet. Det andra alternativet, med SIARD specifikationen, kommer att tas fram i ett senare skede. Introduktionen kommer således att uppdateras i samband med att SIARD-alternativet tas fram och då kommer också mer text om SIARD-formatet att läggas till.

Behovet av en begreppslista för samtliga FGS:er kommer att tas upp med FGS-funktionen.

Modellen över flatFiles, i avsnitt 5.1.3, kommer att göras om och uppdateras för att tydliggöra hur flatFiles



förhåller sig till det som den överordnar. Även den tillhörande texten ses över.

Angående önskemålet om att Riksarkivet ska vara ett större stöd i att förklara och utbilda i standarder så är FGS-dokumenterna inte menade att vara direkt utbildande, utan här får hänvisning ske till eventuella vägledningar och andra dokument. Behovet noteras och frågan förs vidare inom Riksarkivet.

Avsnitt 4.1.1 kommer att ses över och förtydligas.

2. Är de element som finns beskrivna i dokumentet ”FGS Databas (Alternativ 1, relationsdatabaser och register)” tillräckliga för överföring av informationen i databaser? Vilka element saknas?

Av de 8 inkomna remissvaren så är det 6 stycken som direkt eller indirekt besvarat frågan om de element som finns beskrivna i specifikationen är tillräckliga för överföring av information i databaser.

Bland de inkomna remissvaren är det 1 som svarat att man haft svårt att avgöra om det saknas element, och 2 som svarat att elementen verkar vara tillräckliga.

De övriga svaren som inkommit består mer av generella kommentarer kring texten och vissa av de existerande elementen.

I 1 svar påpekas att tecknet *é* inte tas upp i sektionen om teckenöversättning, och det poängteras att om det tyska *ü* tas upp, borde även *é* inkluderas. I samma svar framförs även ett önskemål om att lägga till boolean som ett tillåtet värde i elementet *datatyp*.

I 1 svar framförs en önskan om ett tätare samarbete mellan Sverige, Norge och Finland kring de element i ADDML som saknas i grundstrukturen. Man anser att vissa *additionalElements* som är obligatoriska i specifikationen borde omvandlas till riktiga element i schemat, samt att möjligheten till att använda så kallade generiska element borde begränsas.

I detta svar efterfrågas även möjligheten att ange SQL-script och XML i elementet *filtyp*.

I 1 svar framförs att man tycker det är oklart vad *addmlgenerator* ska användas till och undrar om det är systemet som driver databasen eller systemet som databasen är kopplad till som ska beskrivas. Om *addmlgenerator* inte är menat att avse bakomliggande databas och källsystem utan vad som exporterade databasen så efterfrågar man även ett element som beskriver källsystemet.

När det gäller *addmlgenerator* så tycker man att det underliggande elementet *value* borde struktureras så att namnet på programvaran och versionen av programvaran kan presenteras separat. ”*Namn och version kan vara sådant som vill presenteras separat. Det är exempelvis möjligt att dåliga namnkonventioner går ut över återsökning av handlingar efter system*”. Som förslag framförs att man separerar innehållet i namn och version och använder ett attribut för att urskilja innehållet. Dessutom ifrågasätts användandet av *datetime* i *addmlgenereradtid*.



I svaret tas även upp strukturen för primära och främmande nycklar, som man anser går ned en nivå längre än vad som behövs. För enkelhetens skull föreslår man därför att elementet *key* ska tas bort och att endast *primaryKey* och *foreignKey* används.

Det frågas vidare, i svaret, om elementet *flatFileDefinition* inte borde innehålla filnamn istället för posttypsnamn; ”Utgår man från elementet *<flatFileDefinitionReference>*, som enligt namnet är en referens till en *<flatFileDefinition>* och har filnamn som attribut, så borde *<flatFileDefinition>* också ha filnamn som attribut om man ska kunna referera till den”.

De tar även upp att främmande nycklar inte nödvändigtvis måste vara primärnycklar i andra tabeller; ”Att anta att det alltid är fallet innebär en risk att förlora information. Man borde också ta hänsyn till att främmande nycklar inte alltid har samma namn som den primära nyckeln den pekar på. I en tabell med namnet *user* kan exempelvis den primära nyckeln heta *id* och i en annan tabell kan den främmande nyckeln heta *user_id*. Vi föreslår istället att man inom elementet *foreignKey* refererar till den främmande nyckeln i tabellen i fråga och en referens till nyckeln i den andra tabellen”.

Kommentarer: Dokumentet kommer att ses över och tillmötesgå ovan uppräknade önskemål så långt det är möjligt inom ramen för FGS-konceptet och de aktuella framtagna mallarna som används för FGS dokument.

När det kommer till förslagen som rör t.ex. *additionalElements* och hur elementen är strukturerade, så kan inte Riksarkivet ändra i grundstrukturen för ADDML. Sådana typer av större grundläggande förändringar av strukturen måste diskuteras med Norge och Finland för att potentiellt leda till förändringar av standarden i framtiden.

I ADDML hanteras alltid databasuttaget som en textfil. Elementet *filtyp* är därför endast till för att ange om formatet på filen är fast fält- och postlängd, eller teckenseparerat med fältavskiljartecken. Andra filtyper hanteras inte i ADDML, utan hanteras i METS via MIME-typ.

Elementet *addmlgenerator* är avsett för att ange det program som har genererat ADDML-filen, oavsett om det är ett speciellt program eller en modul i ursprungssystemet. Viss metadata om källsystem finns i FGS Paketstruktur, t.ex. namn och version i separata element, och ytterligare metadata anges i den allmänna systemdokumentationen. Att dela elementet *addmlgenerator* och skapa ett separat element för version ser vi därför inget behov av.

Addmlgenereradtid är ett exempel på ett element där datatyp inte kan valideras genom schemat, se vidare i avsnitt 4.3 i FGS Databas (Alternativ 1, relationsdatabaser och register) Tillägg. Tidpunkten i elementet avser när ADDML-filen skapas och därför bör det inte vara något problem att generera det i formen *datetime*.

Med den nuvarande specifikationen går det bra att ange främmande nycklar som inte är primärnycklar i andra tabeller. En felaktig formulering i texten kommer att korrigeras. Nycklarna behöver heller inte ha samma namn, vilket redan framgår av specifikationen.

3. Är texten i dokumentet ”FGS Databas (Alternativ 1, relationsdatabaser och register)” begriplig, beskrivande och klargörande? Finns det delar som borde förtydligas eller som saknas?



Av de 8 inkomna remissvaren så är det 6 stycken som direkt eller indirekt besvarat frågan om texten i dokumentet "FGS Databas (Alternativ 1, relationsdatabaser och register)" är begriplig, beskrivande och klagörande.

Bland de inkomna remissvaren framgår det i 3 av svaren att man anser att texten är tillräckligt begriplig. I ett av dessa svar tillägger man dock att det är förutsätter att "*läsaren har vissa förkunskaper/ erfarenheter inom området*". I ett svar noteras även att informationen om tillåtna tecken för filnamn (avsnitt 3) skiljer sig från motsvarande information i FGS Paketstruktur (avsnitt 3.1.1). Ett av svaren framför att det är oklart i vilken omfattning FGS Databas relaterar till den ursprungliga standarden ADDML och dess versionshantering.

I 1 av remissvaren svarar man att det man anser att det blir "*rörigt med hänvisningar som finns utspridda. Det blir att läsandet hackas upp eftersom man hoppar fram och bak i texten*". Slutligen är det 1 remissvar som uttrycker att det tydligare bör framgå i texten vad som är skall-krav och vad som är bör-krav. Man anser även att det uppstår osäkerhet då man uppfattat det som att "förkortningar ska nyttjas" i elementet för tekniskt fältnamn, men att det är osäkert då det inte är obligatoriskt med fältnamn i klartext. Något man även anser påverkar elementet *främmande nycklar*. Utöver det framgår att man anser att förklaringen till elementet begreppsinnehåll är otydlig.

I övrigt har det endast påpekats en rad språkliga fel i dokumentet.

Kommentarer: Dokumentet kommer att ses över och tillmötesgå ovan uppräknade önskemål så långt det är möjligt inom ramen för FGS-konceptet och de aktuella framtagna mallarna som används för FGS dokument.

Vad gäller den motstridiga informationen i FGS Databas och i FGS Paketstruktur angående tillåtna tecken, kommer det uppdateras så att specifikationen följer det som är angivet i FGS Paketstruktur och alltså tillåter användandet av bindestreck i namn på filer och filkataloger.

FGS Databas (Alternativ 1, relationsdatabaser och register) är en svensk version av den ursprungliga nordiska ADDML standarden. Versionen genomförs via de generiska elementen. Detta kommer att tydliggöras i texten.

När det gäller elementet för tekniskt fältnamn så är det inget krav att använda förkortningar, men det är ofta så de ursprungliga fältnamnen ser ut och det är det man vill få med. Med anledning av att det tekniska fältnamnet från ursprungssystemet ofta är kryptiskt så har ett frivilligt element lagts till för att kunna skriva in namnet i klartext. Detta kommer att förtydligas.

4. Är texten i dokumentet "FGS Databas (Alternativ 1, relationsdatabaser och register), Tillägg" begriplig, beskrivande och klagörande? Finns det delar som borde förtydligas eller som saknas?

Av de 8 inkomna remissvaren så är det 7 stycken som direkt eller indirekt besvarat frågan om texten i tilläggsdokumentet är begriplig, beskrivande och klagörande.

I 3 av de inkomna remissvaren framgår det att man tycker att texten är tillräckligt begriplig. I ett av dessa svar efterfrågas dock "*lite mer förklarande text i avsnitt 2.4 om utökade metadatabeskrivningar*".



Det är 1 som har besvarat frågan genom att påpeka ett språkligt fel i dokumentet.

I de resterande 3 remissvaren diskuteras kapitel 3 som handlar om BLOB:ar och CLOB:ar. I ett av svaren framförs att det är *"bra att ni ger upplysningar rörande CLOB/BLOB"*, i ett annat anser man att det är *"otydligt hur BLOB ska hanteras med hjälp av ADDML då det inte finns någon särskild datatyp för BLOB"*. I det tredje svaret lyfts önskemål om ett angivet gränsvärde för vart gränsen mellan "långa textfält" och CLOB:ar går, samt en vägledning från Riksarkivet.

Kommentarer: Dokumentet kommer att ses över och texten i ovan nämnda stycken kommer att förtydligas.

Vad gäller datatypen för CLOB/BLOB:ar så blir det string, då det är en länk som hänvisar till filen som sedan redovisas i FGS Paketstruktur som fil. Angivning av CLOB/BLOB i vissa metadatafält kommer att övervägas.

Gränsvärde för var gränsen mellan "långa textfält" och CLOB:ar går är svårt att ange generellt utan får avgöras i aktuella fall.

5. Är de egendefinierade värdelistorna i dokumentet "FGS Databas (Alternativ 1, relationsdatabaser och register) Tillägg" relevanta och begripliga?

Av de 8 inkomna remissvaren så är det 4 stycken som direkt eller indirekt besvarat frågan om de egendefinierade värdelistorna är relevanta och begripliga.

I 2 av de inkomna remissvaren framgår att man anser att värdelistorna är relevanta och begripliga, i det ena svaret tillägger man dock att det vore *"mer användarvänligt om denna information var en integrerad del av dokumentet FGS Databas (Alternativ 1, relationsdatabaser och register)"*.

I 1 av de inkomna remissvaren har man svarat att man haft *"svårt att avgöra om de egendefinierade värdelistorna är relevanta och begripliga"*.

Slutligen har 1 svarat att man ej har några synpunkter angående de egendefinierade värdelistorna.

Kommentarer: Dokumenten följer ramverket för FGS-konceptet och de aktuella framtagna mallarna som används för FGS dokument. Uppdelningen inom en FGS blir i normalfallet således i tre delar; introduktion, specifikation och tillägg. Synpunkten att integrera värdelistorna i specifikationen förs vidare till FGS-funktionen.

6. Övriga synpunkter och önskemål angående FGS Databas (Alternativ 1, relationsdatabaser och register)?

När det gäller övriga synpunkter som framförts i remissvaren så är det 2 stycken som i sitt svar påpekat att det blir otydligt med så många dokument. I det ena svaret påpekas att när dokumentation är *"uppdelad i flera olika dokument minskar generellt användarvänligheten. Det vore mer fördelaktigt ur användarsynpunkt att sammanställa de olika delarna till ett enhetligt och heltäckande dokument"*. I det andra svaret föreslås att innehållet i texten om anpassningen för arkivleverans borde inkluderas i själva FGS Databas, samt att *"bilagshänvisning borde inkluderas i samma dokument för att undvika för många dokument att hålla reda på"*.



I 2 av remissvaren framförs en önskan om förtydligande gällande när FGS Databas ska användas och när andra FGS:er är att föredra: *"Är tanken att en tematisk FGS alltid ska användas före FGS Databas eller är det ett val myndigheten kan göra själv?"*.

Det är ett remissvar som uttrycker att innehållet i specifikationen kommer ge *"möjlighet att påvisa betydelsen av genomtänkta arkivexporter för yrkesgrupper som inte är vana att ta hänsyn till arkivaspekten"* och uttrycker att det är en välbalanserad specifikation, *"som utan att överarbeta sitt innehåll redogör för ett område där standardisering och rådgivning är mycket behövligt"*. Man framför att de oklarheter som finns inte beror på formuleringar utan på att innehållet ska kommuniceras till disparata yrkesgrupper med vitt skilda förkunskaper. Men just för att specifikationen ska kunna förstås av läsare från olika yrkesgrupper önskar man att Riksarkivet kunde konkretisera *"hur anpassningarna är tänkta att fungera i praktiken, till exempel i form av exempelfall från ett fiktivt system"*.

I 1 av de inkomna remissvaren belyser man att ADDML i stort sett är ett okänt format och att man ser en risk att myndigheter saknar förmåga att leverera utifrån den här specifikationen då det saknas verktyg för ADDML. Man efterfrågar även att Riksarkivet delar med sig av sina verktyg för skapande och återställande av paket enligt ADDML specifikationen.

I samma remissvar tar man även upp skillnaderna mellan ADDML och SIARD och undrar om det stämmer *"att ADDML fungerar för enkla/äldre typer av databaser med en eller ett fåtal tabeller, äldre hierarkiska databaser, register, datauttag och rapporter i tabellform?"*. Man nämner i samband med detta att det redan finns andra format för dataset t.ex. HDF och nämner även W3Cs rekommendation för tabulärt data. Det uttrycks att ADDML inte är helt kompatibelt med SIARD och att eventuell migrering från ADDML till SIARD kan medföra svårigheter. *"Det är viktigt att FGS Databas fungerar vid migrering av databaser. Det är önskvärt att leveranser från mellanarkiv till slutarkiv kan ske på effektivt och säkert sätt."*

I övrigt kan lyftas fram att det är 4 av remissvaren som uttrycker att man är mycket positiva till att det kommer en FGS för databaser, varav två särskilt nämner att de ser fram emot SIARD specifikationen.

Kommentarer: Dokumenten följer ramverket för FGS-konceptet och de aktuella framtagna mallarna som används för FGS dokument. Uppdelningen inom en FGS blir i normalfallet således i tre delar; introduktion, specifikation och tillägg.

Angående frågan om när denna FGS ska användas och när andra tematiska FGS:er ska användas, så är det något som ska diskuteras vidare med FGS-funktionen.

När det gäller praktisk användning och exempelfall, så är de framtagna dokumenten till FGS Databas inte avsedda att vara en vägledning och det är inget som rymts inom projektets uppdrag. Däremot noteras behovet av ytterligare vägledning och frågan förs vidare inom Riksarkivet.

När det gäller verktyg för skapande av paket så finns RALF (Riksarkivets leveransförberedelse-verktyg) för nedladdning från Riksarkivets hemsida, men detta verktyg har dock inga funktioner för editering och återställande. Riksarkivet har för närvarande inga möjligheter att, för externt bruk, paketera mjukvara för



återställande.

ADDML fungerar för enkla/äldre typer av databaser, äldre hierarkiska databaser, register samt datauttag och rapporter i tabellform ur relationsdatabaser.

Vad gäller migrering från ADDML till SIARD så går det bra att föra över datainnehåll och grundläggande struktur från ADDML till SIARD via import i en databashanterare. Däremot hanterar inte ADDML automatiskt funktioner från ursprungssystemet som t.ex. views och users.

När det gäller andra format än ADDML och SIARD så har inte sådana ingått i projektets uppdrag att titta på. Men för senare versioner av FGS Databas kan det i princip övervägas.

Riksarkivet
Avdelningen för bevarande och digital infrastruktur