

Beskrivning av SND

Detta är en beskrivning av vad SND är och vad SND gör och inte gör. Det är ett levande dokument som ska uppdateras kontinuerligt allt eftersom SND utvecklas och utifrån de synpunkter som framkommer om vad som saknas eller bör förtydligas i dokumentet.

Vad är SND

[Svensk Nationell Datatjänst](#) (SND) är en nationell infrastruktur för forskningsdata som i dagsläget tar emot forskningsdata direkt från forskare eller forskargrupper och gör dem tillgängliga via SND:s forskningsdatakatalog. För att möta kraven om öppen tillgång till forskningsdata ([Forskningsprop 2020](#)) utvecklas verksamheten till en distribuerad storskalig modell i samarbete med ett flertal av Sveriges lärosäten och ett antal forskande institutioner och myndigheter. I ett nästa steg kommer intresset hos regionerna och det privata näringslivet att undersökas.

SND har fått bidrag från Vetenskapsrådet för att ge det svenska forskarsamhället ett samordnat system för att beskriva, hitta och dela forskningsdata, både nationellt och internationellt. Finansieringsperioden är på fem år med start 1 januari 2018 och en ny ansökan för perioden 2023-2026 lämnas in i februari 2021.

SND är ett konsortium som består av nio lärosäten med kontor vid Göteborgs universitet. I SND:s distribuerade nätverksmodell ingår för närvarande 35 av landets lärosäten och forskande myndigheter inklusive de nio ägarlärosätena. Syftet med nätverket är att stödja de svenska forskarna i arbetet med att tillgängliggöra forskningsdata. Centralt i detta arbete är att varje lärosäte har en lokal stödfunktion, Data Access Units (DAU), som arbetar för att forskningsdata ska bli så tillgängliga som möjligt med så bra dokumentationsnivå som möjligt. SND:s arbete förändras i hög grad från att direkt ge råd och stöd till forskare till att vara support gentemot DAU i deras arbete med att ge råd och stöd till lärosätets forskare. Några DAU:er är gott igång med arbetet medan andra är i etableringsfasen.

En motsvarighet till SND:s distribuerade modell finns inte i sin helhet i något annat land, utan vi ska tillsammans utveckla ett nytt sätt att organisera arbetet med att göra forskningsdata tillgängliga. Detta innebär att vägen framåt inte blir helt spikrak och att omtag ibland kan behöva göras. Det är ett samarbete mellan många aktörer där samtligas bidrag och engagemang behövs för att modellen ska bli framgångsrik. SND:s utveckling ingår också i ett globalt sammanhang, och kan bl.a. ses som en del i uppbyggnaden av the European Open Science Cloud ([EOSC](#)) där SND är medlem (via Göteborgs universitet).

SND-konsortiet

[Konsortiet](#) består av Göteborgs universitet, Chalmers tekniska högskola, Karolinska Institutet, Kungliga tekniska högskolan, Lunds universitet, Stockholms universitet, Sveriges lantbruksuniversitet, Umeå universitet och Uppsala universitet. Göteborgs universitet är värduniversitet för SND, vilket

innebär att verksamhetens kontor finns i Göteborg och att styrgruppens ordförande är från Göteborgs universitet. Konsortiet fattar beslut som gäller SND genom en årlig stämma och en styrgrupp. Styrgruppen följer upp att den strategiska planen, de årliga verksamhetsplanerna och budgeterna följs.

Vid lärosätena i konsortiet finns expertis i form av så kallade domänspecialister (DoSp). Domänspecialisterna är experter inom olika vetenskapliga ämnesområden eller områden rörande olika former av forskningsdata. De fungerar som en nationell resurs för forskare och DAU:er och som en lokal resurs på sina respektive lärosäten. De samarbetar även med SND-kontoret i relevanta arbetsuppgifter.

Som en sammanhållande grupp inom konsortiet finns dessutom konsortie-DAU-rådet som bildades hösten 2019. Rådet består av en representant från varje konsortiedlem samt en representant från SND-kontoret.

SND-nätverket

Nätverket utgörs i dagsläget av de 35 lärosäten och forskande institutioner och myndigheter som alla etablerar en DAU-funktion. De 26 nätverksmedlemmarna som inte ingår i konsortiet har genom en avsiktsförklaring (Letter of intent) gått med i samarbetet och åtagit sig att skapa förutsättningar för etablering av en DAU. Denna avsiktsförklaring ersätts under 2020/2021 av ett formellt samverkansavtal för att klargöra ansvar och behövda resurser.

SND i relation till DAU

SND-kontorets roll mot DAU:erna är två: som stödfunktion på olika sätt i uppbyggnaden av DAU och som kravställare för hur forskningsdata som publiceras i den nationella forskningsdatakatalogen ska beskrivas. I detta arbete ingår att ta fram en kravbeskrivning (vs 2, 2019-03-14) för att säkerställa att data blir så [FAIR](#) (Findable, Accessible, Interoperable, Reusable) som möjligt. Som hjälp i arbetet finns också en DAU-handbok i form av en wiki med beskrivning av hur kraven ska uppfyllas och hur datahantering bör bedrivas under hela forskningsprocessen, alltså även vad de som arbetar i forskningsprojekten bör tänka på och ta hänsyn till vid hantering av forskningsdata. Vidare utvecklas funktioner på SND:s webbsida som underlättar för DAU att granska och bearbeta metadata. Bland dessa funktioner är DORIS (DataORganiserings-och-InformationSystem) som lanserades våren 2020. DORIS är ett sammanhållet system där forskare och DAU kan arbeta för att göra forskningsdata tillgängliga. Systemet utvecklas kontinuerligt och målet är att skapa ett verktyg som går att använda i hela processen, från beskrivning och delning/tillgängliggörande av forskningsdata till beställning av data för återanvändning eller validering.

En viktig aspekt i arbetet är att SND är ett certifierat repositorium ([CoreTrustSeal](#)) för forskningsdata. Detta innebär en kvalitetssäkring av arbetsprocessen med forskningsdata från "ax till limpa", från dataset som beskrivs, bearbetas och lagras säkert till att slutligen kunna beställas eller laddas ner. I det distribuerade arbetet är målsättningen att varje lärosäte och myndighet på samma sätt, i samarbete med SND, arbetar för att uppfylla kraven för certifiering. Att forskningsdata finns tillgängliga från ett certifierat repositorium är ofta ett krav från tidskrifter och anslagsgivare.

Brukaravgift

De 26 nätverksmedlemmar som inte ingår i konsortiet betalar en brukaravgift till SND. Stämman beslutar om brukaravgiften varje år, men den kan, som beskrivet i ansökan till VR, inte bli högre än 100 000 kronor. Brukaravgiften avser att främst täcka kostnader för möten och fortbildningsträffar, besök av SND för stöd i DAU-verksamheten, tillgång till tjänster som DORIS (för bl.a. redigering av metadata via inloggning), framtagande av DAU-handbok och best practice-dokument samt medlemskap i DataCite.

Utbildning

För att stödja och utbilda forskare och DAU-medarbetare i datahantering finns centrala resurser framtagna av SND. BAS Online är en webbutbildning som i första hand är riktad till medarbetarna vid DAU:erna, men stora delar av den är också användbar för forskare och andra som vill lära sig mer om datahantering. För forskare finns dessutom på SND:s hemsida uppdaterade informationssidor om datahantering på både svenska och engelska. Det finns även utbildningspaket riktade direkt till forskare och ytterligare planeras. Som påbyggnad till BAS Online finns en återkommande kurs anordnad av Högskolan i Borås och SND: "[Forskningsdata: tillgänglighet, hantering och samverkan](#)".

Anpassning till framtidens SND

Tidigare har SND fokuserat på områdena samhällsvetenskap, humaniora samt medicin och hälsa. Den nya satsningen innebär ingen begränsning av forskningsdata när det gäller vetenskapsområde eller typ av forskningsdata. Vid beskrivning eller sökning av forskningsdata kräver denna breddning en anpassning av SND:s tjänster till nya metadatastandarder, nya metadataprofiler och ny terminologi. Detta är ett arbete som pågår och där behöver vi göra prioriteringar. I dagsläget finns i tillägg till en generell metadataprofil fem anpassade: samhällsvetenskap, medicin och hälsovetenskap, arkeologi och historia, språkresurser samt miljö-, klimat- och geovetenskaper. Det planeras för fler ämnesanpassade metadataprofiler.

Anpassningen innebär också en satsning på att samarbeta med andra aktörer som beskriver eller lagrar forskningsdata, som t.ex. specifika portaler och lagringsplattformar som Bolin center, Språkbanken och DiVA. Genom integrationen med DiVA kan forskarna koppla artiklar och andra publikationer till sina forskningsdata. Diskussioner pågår också med RUT som är Vetenskapsrådets verktyg för att söka datavariabler i Sveriges många register. SND kommer att samarbeta med dem kring beskrivning av forskningsdata med personnummer.

"Mina sidor" – portal till SND:s tjänster

"Mina sidor" täcker de tjänster som går att nå genom inloggningen hos SND. Centralt här är [DORIS](#) där forskare enkelt kan få en översikt över sina befintliga databeskrivningar, skapa nya och återuppta pågående databeskrivningar. Varje DAU har dessutom en översikt över de databeskrivningar som har lärosätet som huvudman. DAU:en kan granska dessa och skicka dem vidare för publicering i SND:s nationella forskningsdatakatalog.

På Mina sidor anpassas också information och tjänster efter vem som loggar in. En forskare får, förutom åtkomst till DORIS, information om vilket lokalt stöd som finns för datahantering på deras

eget lärosäte samt länkar till relevanta sidor om datahantering. En DAU-medarbetare som loggar in når istället exempelvis DAU-handboken och aktuella utbildningsresurser.

Inloggning till tjänsten sker genom universitetskonton kopplade till EduGAIN, sociala medier som Facebook och Google, eller genom att ett lokalt konto upprättas hos SND.

Lagring

SND har *inte* fått bidrag för att ansvara för lagring av forskningsdata. Det pågår dock ett tätt samarbete med SUNET, som är en avdelning vid VR, där SND verkar för att integrera SUNET:s tjänst för distribuerad lagring, [SUNET STaaS](#), med SND:s system för databeskrivningar, DORIS, och [SND:s tjänster för persistenta identifierare](#) via DataCite. Liknande diskussioner förs även med SNIC rörande deras lagringstjänster.

Att flytta forskningsdata mellan olika lagringsutrymmen är en kostsam process, därför bör data så långt som möjligt stanna kvar på samma plats under hela projektiden fram till arkivering. Det är också enklast utifrån ett juridiskt perspektiv, framförallt när det kommer till data med personuppgifter eller annan känslig information.

Det finns ett behov av tjänster för lagring under projektets gång, arkivering och tillgängliggörande och SUNET STaaS är en plattform där sådana tjänster kan byggas.

SND har, tillsammans med SUNET, tagit fram en API-specifikation för kommunikation mellan de lagringsplattformar där data ligger lagrat och DORIS för att möjliggöra tillgängliggörande av data i en distribuerad lagringsmiljö. Detta API är inte lösningsspecifikt för SUNET STaaS, utan kan anpassas till andra lokala lagringslösningar på lärosätena.

SUNET STaaS kan även användas som arbetsytta för forskare till andra tjänster, som t.ex. e-arkiv. SND arbetar även för att DORIS i framtiden ska ha stöd för de arkivspecifika metadata som krävs ur ett e-arkivperspektiv.

SUNET STaaS innebär en kostnad, och offert kan erhållas från SUNET.

SND nu och SND i framtiden

I dagsläget tar SND emot forskningsdata direkt från forskare eller forskargrupper och efter genomgång av datamaterialet görs det tillgängligt via SND:s forskningsdatakatalog. Både servicen att dela forskningsdata via SND och att beställa data är kostnadsfri.

I den distribuerade modellen går metadata via DAU till den nationella forskningsdatakatalogen och forskningsdata lagras på organisationens egna yta/ytor, (läs mer om lagring ovan). Möjligheten att göra forskningsdata tillgängliga direkt via SND:s repositorium kommer dock att finnas kvar för organisationer (och eventuellt enskilda forskare) som inte har tillgång till en DAU, men detta kommer då innebära en kostnad för dessa organisationer.

Då inte all data kommer att kunna ligga direkt nerladdningsbara i forskningsdatakatalogen tar SND också fram stöd för en beställningsfunktion för data. Öppen tillgång till forskningsdata är grundprincipen, så det kommer inte att innebära någon avgift att beställa data via SND.

Övergångsperioden till den distribuerade modellen med DAU kommer sträcka sig över en längre tid och innebära parallella processer för SND. Detta blir nödvändigt då utvecklingen av DAU-funktionen går olika snabbt runtom hos nätverksmedlemmarna. Under övergångsperioden kan arbetsflödet för hantering av data och metadata ske på de olika sätt som beskrivs nedan:

1. Innan lärosätets DAU är på plats

Enligt SND:s traditionella process från forskare till SND via DORIS. Forskningsdata lagras hos SND i väntan på att SUNET-lösningen eller en annan lagringslösning blir klar. Dessa data kommer senare att överföras till lärosätets lagringsyta. Observera att dessa data inte får innehålla personuppgifter.

2. När DAU är på plats men innan lärosätets lagring är klar

Inkomna databeskrivningar går via DORIS till DAU för kontroll av data och metadata. Kravbeskrivningen och DAU-handboken är vägledande, men det praktiska arbetet med dokumentation och kontroll av forskningsdata sker i dialog med SND till en början. Data kommer att lagras hos SND, eller under begränsad tid hos SNIC, i väntan på att SUNET-lösningen eller en annan lagringslösning blir klar. Dessa data kommer senare att överföras till lärosätets lagringsyta. Observera att dessa data inte får innehålla personuppgifter.

3. När DAU och lärosätets lagring är på plats

SUNET-lösningen eller annan lokal lagringslösning är klar att användas. De lärosäten och myndigheter som har en etablerad DAU kan ladda upp metadata till forskningsdatakatalogen och forskningsdata till den lärosätesspecifika ytan hos SUNET (enligt avtal tecknat med SUNET) eller annat datalager. Med SUNET:s lagring, eller annan lokal lagringslösning, på plats får även forskningsdata med personuppgifter hanteras och göras tillgängliga efter prövning.

Tidsperspektiv

Det är flera faktorer som påverkar när samarbetet mellan DAU och SND kring forskningsdata kan ske i större skala. Varje lärosäte kan välja att aktivt arbeta med frågan genom att avsätta resurser och ta fram en policy för öppen tillgång till forskningsdata och arbetet med forskningsdata. Tidskrifters krav på publicering av data aktualiserar arbetet alltmer, medan arbetet med att införa öppen tillgång till forskningsdata och lärosätenas forskningsdatapolicys kommer generera ett större flöde kommande år.