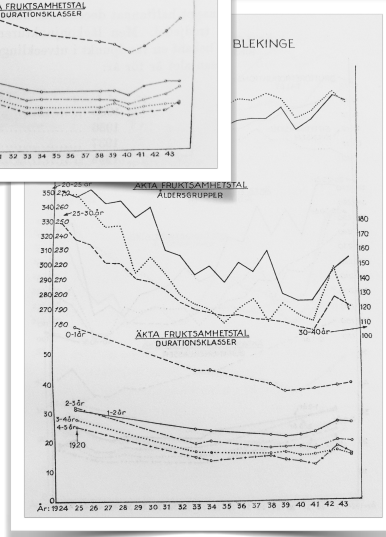
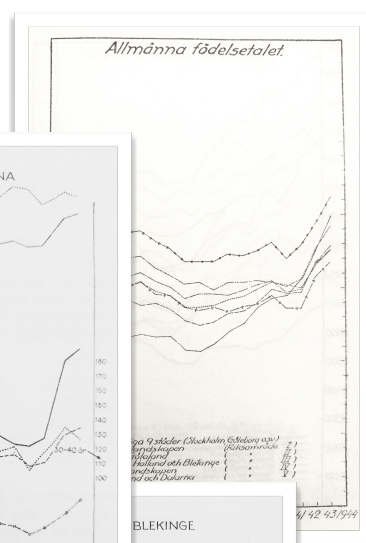
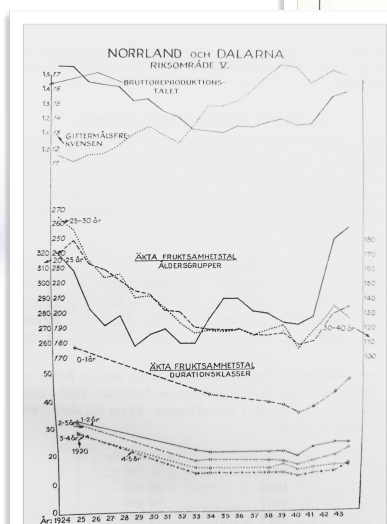


SwedPop

Svenska befolkningsdatabaser för forskning



RIKSARKIVET STOCKHOLMS STADSARKIV

UMEÅ
UNIVERSITETGÖTEBORGS
UNIVERSITETLUNDS
UNIVERSITET

SwedPop finansieras av Vetenskapsrådet 2018-2022 och stöds av ett konsortium bestående av följande parter:

- **Umeå universitet: Demografiska Databasen (DDB)** Enheten för demografi och åldrandeforskning CEDAR. Livsförloppsdata för personer från ca 100 församlingar i olika svenska regioner 1650-1960. Folkräkningen 1950. Koordinator för projektet.
- **Göteborgs universitet: Göteborgs Befolkningspanel (GOPP)** Institutionen för ekonomi och samhälle. Fyraårspaneler för ett slumpvis urval bosatta i Göteborg under något av panelåren 1915-1943.
- **Lunds universitet: Skånes Ekonomisk-Demografiska Databas (SEDD)** Centrum för ekonomisk demografi. Livsförloppsdata för personer från ett antal skånska församlingar 1680-1968.
- **Riksarkivet: SweCens-databasen.** Folkräkningarna 1880-1940. Tvärsnitt vart tionde år för hela riket. Sveriges Dödbok (via avtal med Sveriges Släktforskarförbund)
- **Stockholms stadsarkiv: Rotemansarkivet.** Livsförloppsdata för Stockholms befolkning 1877-1926.

1. Inledning

SwedPop: Svenska befolkningsdatabaser för forskning 2018-2022

Bakgrund

Under 2017 beviljades den nationella infrastrukturen SwedPop bidrag från Vetenskapsrådet inom ramen för utlysningen *Bidrag till infrastruktur av nationellt intresse*. Projektets syfte är att utveckla och samordna de viktigaste historiska befolkningsdatabaserna i Sverige till en gemensam resurs, som väsentligt kommer att öka, förbättra och förenkla forskares tillgång till högkvalitativa svenska befolkningsdata för forskning. Databaserna samordnas och tillgängliggörs i en ny infrastruktur, SwedPop, som bland annat kan användas för jämförande studier av hur människors liv, hälsa, sociala och ekonomiska levnadsvillkor formats och förändrats i samspel med samhällets utveckling under mer än 200 år. Genom länkning till andra forskningsregister möjliggörs även studier av flera aktuella medicinska forskningsfrågor. Våren 2021 lanserades ett egenutvecklat uttagsverktyg som gör det möjligt att kostnadsfritt ladda ned data kostnadsfritt från web-portalen swedpop.se. Data som innehåller personuppgifter hanteras inom ramen för gällande lagstiftning och lämnas ut efter föreskriven etik- och sekretessprövning.

Kärnan i SwedPop utgörs av fem stora forskningsdatabaser: Databaserna POPUM och POPLINK vid Demografiska databasen (DDB), Umeå universitet, Skånes Ekonomisk-Demografiska Databas (SEDD) vid Centrum för ekonomisk demografi i Lund, Rotemansdatabasen vid Stockholms stadsarkiv, SweCens-databasen vid Riksarkivet och Gothenburg Population Panel (GOPP) vid Institutionen för ekonomi och samhälle, Göteborgs universitet. De storskaliga databaserna DDB och SEDD innehåller länkade individdata baserade på den svenska kyrkobokföringen i utvalda regioner, som gör det möjligt att följa individer, familjer och hushåll i sin sociala och ekonomiska kontext från sent 1600-tal till 1960-talet. Rotemansdatabasen innehåller motsvarande data från Stockholm 1877-1926. SweCens-databasen baseras på de svenska historiska folkräkningarna med information om hela den svenska befolkningen och GOPP är en panelstudie baserad på folkbokförings- och mantalsskrivningsdata från Göteborg under 1900-talet. Tillsammans erbjuder de fem databaserna sammanhängande information om den svenska befolkningen under nära tre sekler, med regionala nedslag i skilda delar av Sverige. Data omfattar såväl stad som landsbygd och olika socioekonomiska miljöer, vilket ökar möjligheterna att göra jämförande studier, till exempel av samhälls- och befolkningsutveckling i olika svenska regioner, i städer och på landsbygden. Även andra databaser kan knytas till SwedPop för att öka infrastrukturens vetenskapliga värde och täckningsgrad. En sådan databas är Sveriges Dödbok med information om avlidna svenskar, insamlad av Sveriges Släktforskarförbund, som inkluderas i SwedPop mot slutet av bidragsperioden 2018-2022. Databasen förvaltas av Riksarkivet, genom avtal med Sveriges Släktforskarförbund. Folkräkningen 1950 har även förvärvats för projektets räkning och förvaltas av CEDAR/DDB

Övergripande målsättningar för projektet

Några av databaserna i SwedPop är redan väletablerade i forskningsvärlden, medan andra inte har använts lika mycket för forskning. För att öka forskningspotentialen, för samtliga databaser och för infrastrukturen som helhet, kommer utvecklingen av SwedPop under bidragsperioden 2018-2022 att fokusera på följande övergripande områden:

- Samordning av av kodning av yrken, dödsorsaker och geografiska enheter i den ingående databaserna
- Strukturering av data enligt en gemensam databasstruktur med enhetlig nomenklatur och kodningssystem
- Tillgängliggörande av data för forskare från en gemensam portal, swedpop.se
- Länkning av individer mellan de olika databaserna
- Insamling av kompletterande data för att fylla luckor i tid och rum samt för att öka jämförbarheten mellan databaserna
- Etablering av SwedPop som nationellt centrum för äldre befolkningsdata (före SCB-registren) på mikronivå

Särskilda villkor för bidraget (2017-00666) samt budget och övergripande arbetsplan fastställdes av Vetenskapsrådet 2017-12-21.

I september 2021 beviljades SwedPop fortsatt stöd från Vetenskapsrådet för perioden 2023-2026. Särskilda villkor för bidraget (2021-00183) kommer att fastställas under hösten 2021.

Utgångspunkter

Verksamhetsplanen för SwedPop tar sin utgångspunkt i de särskilda bidragsvillkor och den budget och arbetsplan som har fastställts av Vetenskapsrådet (2017-00666) för bidragsperioden 2018-2022.

Verksamhetsplan för varje år utarbetas av ledningsgruppen för SwedPop. Planen fastställs därefter av styrgruppen för projektet och ska under bidragsperioden vara Vetenskapsrådet tillhanda senast den 15 november.

Verksamhetsplanen är strukturerad utifrån de moduler som finns angivna i budget och arbetsplan för projektet. Verksamhetsplanen fastställdes **2021-10-12** av styrgruppen för SwedPop.

Delresultat

Projektet är organiserat i moduler, där arbetet inom varje modul mynnar ut i ett antal mätbara delresultat (färdigställda verktyg, datainsamlingar, tillgängliggörande av data osv.). Planerade delresultat under bidragsperioden finns beskrivna i den strategiska planen för SwedPop fastställd av styrgruppen 2018-06-04. Under 2022 är målsättningen att arbetet ska leda fram till följande delresultat:

A	SEDD-databasen har kompletterats med inkomstdata för åren 1918, 1923 och 1928. **	9
B	Rotemansdatabasen har kompletterats med saknade data.	6
C	SEDD-databasen har kompletterats med befolkningsdata från Landskrona 1895-1904.	10
D	Information från Sveriges Dödbok har kodats, konverterats till IDS-format och länkats till databaserna i SwedPop.	15
E	Komplettering av GOPP-databasen med information om personer födda i eller inflyttade till Göteborg 1944-1961 och deras hushåll har påbörjats.	5
F	Komplettering av Rotemansdatabasen med uppgifter om inkomst har påbörjats.	8
G	POPLINK-databasen har kompletterats med inkomstdata för Lövångers församling.	12
H	Delar av 1940 års folkräkning har inregistrerats, kodats och länkats till övriga databaser i SwedPop.	13
	Kvarstår från 2021:	
*	GOPP-databasen har kompletterats med uppgifter om inkomst och yrke för personer observerade 1943 (Delmål 2021:a) kommer att skjuta över på 2022)	5
*	Ta fram gemensamma principer och metoder/programvara för båda länkingsstegen beskriva i modul 4 (Delmål 2019:c, kvarstår från 2019).	4

2. Moduler - aktiviteter och tidsplan

Modul 1

Organisation och ledning

Modul 1 har en stödjande roll i projektet. Här finns budget för projektledning, administration, resor och projektmöten. Inga specifika verksamhetsmål har utarbetats för 2022.

Modul 2

A. Harmonisering av kodning:

1. Harmonisering av variabler

- Dödsorsaker
- Yrke
- Församlingskodning

1a: Dödsorsaker

Nuläge

Det grundläggande arbetet med harmonisering av dödsorsakskodning av de ingående databaserna till SwedPop-standard är avslutat och kodschema samt rutiner för löpande kodning av nya dödsorsaker till har tagits fram. Kodschema och dokumentation av fastställda principer för kodning av dödsorsaker har under 2021 gjorts öppet tillgängliga på [swedpop.se](https://swedpop.se/data-description/) <https://swedpop.se/data-description/>. Till följd av covid-19-pandemin har ambitionen att utse en referensgrupp för kodning av nya dödsorsaker tillfälligt lagts på is.

Det faktiska arbetet med kodning av dödsorsaker till SwedPop-standard i de ingående databaserna påbörjades under samordningsprojektet 2016-2017 och har fortsatt under projektiden 2018-2021. Dödsorsaker i databaserna i Umeå, Lund och Rotemansdatabasn kodas enligt gemensamma principer. SweCens- och GOPP-databaserna innehåller inga dödsorsaker.

Aktiviteter under 2022

1. Utse en referensgrupp för kodning av nya dödsorsaker.

Tidsplan

Övergår till löpande verksamhet.

1b. Yrken

Nuläge

Det grundläggande arbetet med harmonisering av yrkeskodning till SwedPop-standard har slutförts. Ett kodschema för yrkeskodning, *Yrkeskodning i SwedPop* har tagits fram och rutiner för fastställande av nya yrkeskoder, när behov av sådana uppstår under det löpande arbetet, har tagits fram. Nu har standardiseringen av samtliga yrkestitlar avslutats och i april 2021 fastställde ledningsgruppen ytterligare en uppdaterad version av kodschemat v.3 där samtliga yrken är både kodade och standardiserade. Nya versioner av kodschemat kommer att behöva tas fram när samtliga yrken i Rotemannen, 1940 och 1950 års folkräkningar samt i Sveriges dödbok har blivit kodade och standardiserade. Implementeringen av den nya yrkeskodningen (genomförs av varje databas för sig), där ca 1/3 av strängarna har fått nya koder har lett till att kodningen blivit mera precis.

Kodschema och dokumentation av de principer som används för yrkeskodning till SwedPop-standard har under våren 2021 gjorts öppet tillgängliga på [swedpop.se](https://swedpop.se/data-description/) <https://swedpop.se/data-description/>.

Avsikten är att den löpande yrkeskodningen inom SwedPop i huvudsak ska kunna utföras maskinellt i framtiden.

Aktiviteter under 2022

1. HISCO-kodning och standardisering av återstående nya yrkessträngar i Rotemannen.

Tidsplan

Övergår till löpande verksamhet så snart standardisering är slutförd.

1c. Församlingskodning

Nuläge

Arbetet med harmonisering av församlingskodning har gått vidare under 2021 under ledning av en särskild arbetsgrupp. En gemensamt kodschema för kodning av församlingar till SwedPop-standard fastställts, baserat på Riksarkivets register TORA Topographic Register at the Swedish National Archives. Då församlingsgränser inte sällan förändras över tid kan nya versioner av kodschemat behöva tas fram när nya data läggs till i databasen. Den första uppdateringen kommer att behöva göras redan innevarande år, när folkräkningen 1950 är kodad och upparbetad.

Dokumentation och kodschema har under våren 2021 gjorts öppet tillgängliga på swedpop.se, <https://swedpop.se/data-description/>.

Implementering och omkodning av den geografiska kodningen utförs av varje databas för sig. Harmoniseringen av församlingskodning är därmed slutförd.

Aktiviteter under 2022

1. Uppdatera fastställt kodschema med koder för svenska församlingar 1950.

Tidsplan

Fortsatt harmonisering av församlingskodning sker vid behov, när nya, tidigare okodade geografiska områden läggs till, enligt de principer som fastställs.

B. IDS-konvertering av databaserna:

Nuläge

Arbetet med IDS-konvertering av databaser i SwedPop startade redan under samordningsprojektet. Efter en omfattande probleminventering och analys kunde gemensamma definitioner och principer för konverteringen till IDS 4.0 fastställas under 2020. En detaljerad dokumentation inklusive metadata avseende konvertering av data till SwedPop:s gemensamma databasformat har under våren 2021 gjorts öppet tillgänglig på swedpop.se, <https://swedpop.se/ids-structure/>

Varje part ansvarar för att konvertera sina data till fastställt format, samt för att ta fram erforderlig systemdokumentation för konverteringen. Inför den första datareleasen på swedpop.se i juni 2021 genomfördes noggranna tester för att säkerställa fullständig harmonisering mellan de ingående databaserna.

Aktiviteter under 2022

Skapa IDS-versioner av Rotemannen, GOPP och Folkräkningen 1950.

Tidsplan

Avsikten är att IDS-arbetet ska vara avslutat under 2022.

Modul 3

A.Utveckling av ett uttagsverktyg på webportalen

Nuläge

Arbetet startade 2018 och fortsatte under 2019 med framtagande av en detaljerad kravspecifikation för uttagsverktyget. Systemutvecklingen har ägt rum vid DDB och SEDD och Riksarkivet har deltagit i planeringen av det pågående arbetet. Utvecklingen av uttagsverktyget visade sig ta mer tid i anspråk än beräknat, då det tidigt stod klart att den underliggande databasstrukturen behövde kompletteras för att nå önskvärd funktionalitet. En reviderad projektplan fastställdes av ledningsgruppen 2019-10-29 och arbetet med utveckling av verktyget fortsatte under 2020-2021. En testversion togs fram 2020, som vidareutvecklades efter användartester. I juni 2021 kunde uttagsverktyget lanseras på swedpop.se där forskare nu kan ladda ned harmoniserade data från DDB, SEDD och SweCens (folkräkningen 1880), som inte utgör personuppgifter. Data från övriga databaser kommer att tillgängliggöras så snart kompletteringar, kodning och harmonisering är klar. Frågor om juridik, etik och integritetsskydd i samband med tillgängliggörande diskuterades vid en workshop under september 2020. Dessa frågor kommer att fortsätta att bearbetas i samråd med parternas jurister och dataskyddsombud. Ambitionen är att det under 2022 ska finnas tydliga riktlinjer och rutiner utarbetade för utlämnande av data som innehåller personuppgifter.

Aktiviteter under 2022

1. I dialog med juridisk expertis hos parterna utarbeta tydliga riktlinjer och rutiner för utlämnande av data som innehåller personuppgifter.

Tidsplan:

Arbetet ska vara slutfört under 2022.

B. Utveckling av webbportal

Nuläge

Under 2021 har swedpop.se omarbetats och uppdaterats. Uttagsverktyg och data har publicerats och metadata har tillgängliggjorts tillsammans med dokumentation och information om databaser och data. Publicerade data har försetts med versionshantering och PID kopplat till varje version. Under 2022 kommer webbplatsen att utvecklas vidare med fokus på användarstöd. Webbplatsens nyhetssida kommer också att vidareutvecklas för marknadsföring av den nya infrastrukturen.

Aktiviteter under 2022

1. Vidareutveckling av resurser för användarstöd på swedpop.se
2. Fortsatt marknadsföring av swedpop.se i strategiska kanaler.

Tidsplan:

Arbetet fortgår under 2022.

Modul 4

Utveckling av verktyg och metoder för att länka samman databaserna i SwedPop

Nuläge

Länkningsarbetet kan delas in i två steg:

1. Länkning av samtliga SwedPop-databaser till Sveriges Dödbok.
2. Länkning av SEDD, DDB, Rotemannen och GOPP till folkräkningarna i SweCens.

Under första delen av bidragsperioden har arbetet fokuserat på att testa och analysera metoder för att länka folkräkningar och kyrkboksdata (SEDD, DDB) till Sveriges dödbok. Testlänkningsarbetet har utförts och analyserats och förberedande arbete pågår inför länkning av databaserna till den nya uppdaterade versionen av Sveriges dödbok (se modul 15) med harmoniserad församlingskodning enligt SwedPop-standard. Tester har också visat på behovet av standardisering av förnamn och efternamn inom samtliga SwedPop databaser.

Under hösten 2021 genomfördes en workshop med fokus på länkningens andra fas: länkningen av samtliga databaser till folkräkningarna i SweCens samt sammanlänkning av folkräkningarna till varandra. Där diskuterades möjligheten att använda sig av en ny länkningsmetod som utvecklats vid MPC i Minneapolis, baserad på machine learning. Metodens lämplighet för länkning mellan folkräkningar och kyrkboksdata i en svensk kontext kommer att testas och utvärderas.

Gemensamma principer och metoder/programvara kommer att tas fram för båda dessa länkningssteg. Därmed kommer delmål 2019:c att behöva omformuleras något. Det

kommer också att behöva senareläggas till 2022, då länkningen kan genomföras först när flertalet av de kompletteringar som pågår inom projektets moduler är avslutade (modul 6-8, 15).

Aktiviteter under 2022

1. Ta fram gemensamma principer och metoder/programvara för båda länkingsstegen (delmål 2019:c).
2. Länkning av samtliga databaser till Sveriges dödbok.
3. Länkning mellan folkräkningarna 1880-1890-1900-1910.
4. Länkning av samtliga databaser till folkräkningarna i SweCens.

Tidsplan:

Arbetet ska vara avslutat under 2022.

Modul 5

Utvidga GOPP-databasen med nya data

Nuläge

GOPP-databasen är en panelstudie som idag omfattar ett 1%-sample av befolkningen i Göteborg 1915-1943. I denna modul kommer databasen att utvidgas, genom att de personer som kan observeras 1943 följs fram till 1967 och att nya urvalspersoner födda eller inflyttade efter 1943 tillförs och följs till 1967. Under 2021 har arbetet i modulen fortskridit enligt fastställd tidsplan, med den modifieringen att registrering av uppgifter från digitalt tillgängliga källor (församlingsböcker och personakter) har prioriterats medan registrering från papperskällor skjuts upp på grund av pandemin. Detta gör att delmål 2021:a inte har kunnat nås under året och kommer att kvarstå under 2022. Alla urvalspersoner har länkats till databaserna Sveriges dödbok och Sveriges befolkning 1990, och för cirka 84 % har uppgifter registrerats från personakterna. Excerpering ur församlingsböckerna har gjorts för 34% av urvalet och pågår.

Aktiviteter under 2022

1. Registreringen av uppgifter från personakterna färdigställs och registreringen från församlingsböckerna fortsätter.
2. Registrering av uppgifter från inkomstlängderna om yrke och inkomst fortsätter.
3. Färdigställande av komplettering av uppgifter om inkomst och yrke för personer observerade 1943 (delmål 2021:a).

4. Komplettering av GOPP-databasen med information om personer födda i eller inflyttade till Göteborg 1944-1961 och deras hushåll har påbörjats (delmål 2022:e).

Tidsplan

Datainsamlingen kommer att pågå hela bidragsperioden 2018-2022 och kommer att avslutas under nästa bidragsperiod 2023-2026.

Modul 6

Komplettering och länkning av data i Rotemansdatabasen

Nuläge

När databasen Rotemannen byggdes upp registrerades ca 40% av alla poster med en förenklad metod som innebar att viss information utelämnades. För att öka databasens forskningspotential har denna information behövt kompletteras och länkas. Förstudie, validering och dokumentation utfördes under 2018 och kompletteringsarbetet startade 2019. Det fortsatte under 2020-2021 och kommer att vara avslutat under 2022, försenat i förhållande till tidigare tidsplan. Förseningen är föranledd av Stadsarkivets flytt till nya lokaler 2019, vilken ledde till tidsmässiga omprioriteringar mellan modul 6-8. Under 2020 har arbetet försenats ytterligare till följd av personalbrist och förändrade prioriteringar under den pågående Corona-pandemin. Under 2021 har delar av standardiseringsarbetet samordnats med modul 15, så förseningar avseende kodning av yrke och geografi har kunnat arbetas in.

Aktiviteter under 2022

1. Rotemansdatabasen har kompletterats med saknade data (delmål 2022:b).
2. Kontextinformation om familj och hushåll kodas enligt SwedPop-standard utifrån kompletteringsinformationen.

Tidsplan

Kodning av yrken och geografisk information kommer att vara genomförd under 2021. Övrig kodning görs i takt med fortsatt komplettering.

Modul 7

Komplettering av dödsorsaker i Rotemansdatabasen

Nuläge

För att Rotemansdatabasens information om dödsorsaker ska bli fullständig har en kompletterande inregistrering behövt ske av uppgifter i Stadsläkarens arkiv. Under 2019

påbörjades inskanningen av samtliga dödsorsaksintyg och länkning av tidigare olänkade poster startade. En analys har också gjorts för att identifiera var kompletteringar behöver ske. Nyregistrering av dödsorsaker är nu genomförd men harmonisering, länkning och omkodning enligt SwedPop-standard återstår. Arbetet skulle ha varit avslutat 2020, men kommer att bli försenat. Förseningen är liksom i M6 föranledd av Stadsarkivets flytt till nya lokaler 2019, samt av personalbrist och förändrade prioriteringar under den pågående Corona-pandemin. Under 2021 avslutas registrering och länkning av dödsorsaksintyg (delmål 2020:e).

Aktiviteter under 2022

1. Standardisering och kodning av dödsorsaker till SwedPop-standard, i samarbete med DDB.

Tidsplan

Arbetet beräknas vara slutfört vid halvårsskiftet 2022.

Modul 8

Komplettera Rotemansdatabasen med ekonomiska data

Nuläge

För att öka den komparativa forskningspotentialen i SwedPop ska Rotemansdatabasen kompletteras med ekonomiska data. Under 2019 och 2020 har ett förberedande arbete utförts med 1901 års 1901 års huvudförteckning (den del av mantalsuppgifterna som listar samtliga lägenheter med information om innehavare, rumsantal, hyra, mm), som finns tillgänglig men i råformat och endast partiellt länkad. Avsikten var att därtill registrera in ekonomiska data från taxeringslängder för åren 1901 och 1910, men dessa planer fick revideras när projekttiden reducerade från 8 till 5 år. Under 2021-2022 kommer arbetet därför att fokuseras på förberedande aktiviteter, skanning, systemutveckling och förberedelser för länkning av ekonomiska data. Resurser för transkribering finns budgeterade för sista halvåret 2022. En fortsatt inregistrering av mer omfattande uppgifter om inkomst, skatt, hyra och fastighetsinnehav från mantalslängder under en längre tidsperiod 1880-1910, samt länkning kommer att ske under projektperioden 2023-2026.

Aktiviteter under 2022

1. Fortsatt validering av registret för 1901 års huvudförteckning.
2. Fortsatt verifiering av tidigare utförd länkning av 1901 års huvudförteckning.
3. Komplettering av Rotemansdatabasen med uppgifter om inkomst har påbörjats (delmål 2022:f)

Tidsplan

Kompletteringsarbetet kommer att pågå fram till 2022.

Modul 9

Komplettering av SEDD med ekonomiska data för Landskrona 1905-1947

Nuläge

Arbetet inom denna modul avser en komplettering av ekonomiska data i form av inkomst- och taxeringslängder för tidsperioden 1905-1947, för att öka databasens forskningsvärde. Motsvarande data för tidsperioden 1948-1968 finns sedan tidigare inkluderade i databasen och används för forskning. Till och med 2020 har inkomstlängderna för åren 1906, 1910, 1915, 1920, 1925, 1930, 1933, 1935, 1938, 1940, 1942 och 1944-47 registrerats och länkats till databasen. Under 2019 justerades arbetsplanen för modulen av arbetsekonomiska skäl något i förhållande till projektets långsiktiga strategiska plan: Istället för att registrera alla år i ordning, så registreras längderna först med tio års mellanrum, därefter inkluderas data för mellanliggande år.

Aktiviteter under 2022

1. Inkomst- och taxeringsuppgifter för åren 1918, 1923 och 1928 kommer att registreras och länkas till SEDD.
2. Yrkesuppgifter från längderna kodas.
3. SEDD-databasen har kompletterats med inkomstdata för åren 1918, 1923 och 1928 (delmål 2022:a).

Tidplan

Arbetet kommer att pågå under hela bidragsperioden 2018-2022.

Modul 10

Komplettering av SEDD med befolkningsdata för Landskrona 1895-1910

Nuläge

Befolkningsdata för Landskrona 1910-1968, och från omkringliggande församlingar 1813-1968 har inkluderats i SEDD inom ramen för tidigare projekt. Under perioden 2018-2021 har information från församlingsböcker samt register över födda, döda, in- och utflyttade har registrerats för perioden 1905-1909 och länkats till övriga data i databasen. Databasen har även kompletterats med information om faders yrke för individer i Landskrona födda 1905-1947 (delmål 2021:e).

Aktiviteter under 2022

1. Kontroll av registreringen mot födelse-, död- , giftermåls-, och flyttningslängder.
2. Kodning av tillkomna yrken och dödsorsaker.
3. SEDD-databasen har kompletterats med befolkningsdata från Landskrona 1895-1904 (delmål 2022:c).

Tidplan

Arbete kommer att pågå under tidsperioden 2020-2022.

Modul 11

Komplettering av DDB med befolkningsdata för Nysätra församling

Nuläge

Sedan 2008 har befolkningsdatabasen POPLINK varit under uppbyggnad vid DDB i Umeå, inom ramen för tidigare projekt. Avsikten är att databasen på sikt ska omfatta befolkningen i Västerbotten 1700-1950. Arbetet med att inkludera befolkningsdata från Nysätra som startade 2019 har fortsatt under 2020. Ända fram våren 2020 har arbetet följt upprättad tidplan, men har inte kunnat slutföras som planerat till följd av den pågående Corona-pandemin och övergången till hemarbete. Förseningen avser de avslutande länkningsmomenten som måste utföras i den skyddade datormiljön på arbetsplatsen. Arbetet är inte tidskritiskt i förhållande till övriga moduler har därför pausats därför under 2020 och stora delar av 2021. Det kommer att slutföras under hösten 2021. Data kommer att tillgängliggöras i SwedPop:s uttagsverktyg under våren 2022 och i databasen POPLINK under hösten 2022.

Aktiviteter under 2022

Arbetet inom modulen är avslutat.

Tidplan

Arbetet avslutades under 2021.

Modul 12

Komplettering av DDB med ekonomiska data för Lövångers församling

Nuläge

För att öka den komparativa forskningspotentialen i SwedPop kompletteras DDB:s befolkningsdatabaser med ekonomiska data. Datainsamlingen startade 2021 och har löpt enligt upprättad tidplan.

Aktiviteter under 2022

1. Inregistrering och länkning av inkomst- och taxeringslängder från Lövångers färdigställs.
2. POPLINK-databasen har kompletterats med inkomstdata för Lövångers församling (delmål 2022:g).

Tidplan

Datainsamlingen avslutas under 2022.

Modul 13

Inkludering av 1940 års folkräkning i SweCens-databasen

Nuläge

Arbetet startade under 2018. Vissa arbetsmoment tidigare lades, vilket gjorde att arbetet vid 2019 års ingång låg något före fastställd projektplan. Ett nytt formulär i inregistreringssystemet PERSONA avsett för folkräkningar ta har tagits fram och vid 2020 års ingång hade de första församlingarna hunnit bli komplett inregistrerade. I samarbete med Sveriges Släktforskarförbund startade under hösten 2019 ett storskaligt citizen-science-projekt där frivilliga släktforskare deltar i inregistreringsarbetet. Under 2020-2021 har registreringsarbetet skalats upp i takt med att allt flera frivilliga släktforskare anslutit sig till projektet. Funktioner och formuläret i inregistreringsprogrammet och tillhörande manualer har också förbättrats under året, för att effektivisera inregistreringsarbetet. Datainsamlingen följer hittills upprättad plan, men förhoppningen är att samarbetet med Sveriges Släktforskarförbund ska bidra till att fler församlingar än planerat kan inkluderas under projektiden. Arbetet med volontärer har gjort att den ursprungliga tanken med att färdigställa folkräkningen län för län har behövt omprövas. En del i citizen-science-strategin är att volontärerna själva väljer vilka län de vill arbeta med, vilket gör att arbetet blir mer splittrat. I nuläget är prognosen att Jämtland och Värmland ligger närmast till att bli klara under bidragsperioden. Detta gör att delmål 2022:h delvis måste omformuleras.

Aktiviteter under 2022

1. Fortsatt inregistrering av data i 1940 års folkräkning.
2. Rekrytering av ytterligare registrerare i Sveriges Släktforskarförbunds regi.
3. Undersöka om det webbaserade gränssnittet för uppföljning kan förfinas, för att stimulera frivilliga registrerare.
4. Delar av 1940 års folkräkning har inregistrerats, kodats och länkats till övriga databaser i SwedPop ([delmål 2022:h](#)).

Tidplan

Datainsamling påbörjas 2019 och kommer att pågå hela bidragsperioden, till och med 2022. Datainsamlingen kommer att fortsätta under bidragsperioden 2023-2026.

Modul 14

Inkludering av 1950 års folkräkning i SwedPop

Nuläge

Under 2020 kunde äntligen arbetet med att inkludera 1950 års folkräkning starta, kraftigt försenat till följd av omständigheter utanför projektets kontroll. Då det under 2019 stod klart att en ny överenskommelse mellan Riksarkivet och Arkiv Digital, som digitaliserat folkräkningen i genealogiskt syfte, inte skulle kunna nås förbereddes en upphandling av data för projektets räkning. Innan dess undersöktes även möjligheten att digitalisera data i egen regi, men källmaterialet bedömdes inte hålla en sådan kvalitet att OCR-läsning skulle vara möjlig. Tanken på en digitalisering med manuella metoder uteslöt helt, då en sådan varken skulle rymmas inom projektets tids- och budgetramar. Styrgruppen för SwedPop gav 2019-10-16 medelsförvaltaren mandat att gå vidare med en upphandling med målsättning att projektet skulle få tillgång till data innan 2019 års utgång. Upphandlingen genomfördes i enlighet med gällande lagstiftning, med Arkiv Digital som enda anbudsgivare och tilldelningsbeslut fattades i december 2019. Beslutet överklagades till Förvaltningsrätten av en annan aktör, vilket ytterligare fördröjde processen. Överklagan avslogs emellertid 2020-03-31 och 2020-04-15 kunde avtal slutligen träffas mellan medelsförvaltaren Umeå Universitet och Arkiv Digital om inköp av 1950 års folkräkning. Arbetet med kodning och upparbetning startade vid halvårsskiftet 2020 och kommer att vara avslutat under hösten 2021 ([delmål 2020:g](#)). Inköpet kunde genomföras inom given budgetram, men krävde en viss omfördelning av resurser. Därtill har Umeå universitet och Riksarkivet gått in med ökad medfinansiering för att täcka kostnader för kodning och upparbetning.

Aktiviteter under 2022

Arbetet inom modulen är avslutat.

Tidplan

Arbetet avslutades under 2021.

Modul 15

Sveriges dödbok

Nuläge

Sveriges dödbok 1860-2017 är en databas som innehåller information om alla individer som har avlidit i Sverige från 1860 och fram till idag. Genom ett avtal med Sveriges Släktforskarförbund har SwedPop för forskningsändamål fått tillgång till dessa data för att inkludera dem som en del i SwedPop. Databasen är en väsentlig resurs för länknings mellan olika material i SwedPop och, då poster för avlidna efter 1947 innehåller personnummer som gör det möjligt att koppla SwedPop-data till andra externa databaser, exempelvis från SCB eller Skatteverket. Huvuddelen av databasen har skapats av frivilliga släktforskare, med målet att sökbarheten för genealogisk och kvalitativ forskning ska vara så bra som möjligt. Detta innebär att råmaterialet har vissa brister i standardisering och kodstruktur vid användning för kvantitativ forskning och för länkning till forskningsdatabaser. Dessutom saknas ibland fullständiga uppgifter om födelseförsamling, födelse- och dödsdatum. Sveriges Släktforskarförbund har under 2020, på uppdrag av SwedPop, arbetat med att identifiera och minimera dessa kvalitetsbrister. Arbetet har omfattat nyregistrering av födelseuppgifter, validering och korrigering av namnuppgifter, genus och födelseort, kontroll av personnummer och kontroll av att myndighetsuppgifter om dödsfall stämmer, att personerna verkligen är döda. Det senare har betydelse för om uppgiften kan klassificeras som personuppgift eller inte enligt GDPR. Under SwedPop-projektet har flera av de ingående databaserna provlänkats mot Sveriges dödbok, inom ramen för modul 4, och dessa erfarenheter är till stor nytta i det pågående förbättringsarbetet. Arbetet fortsätter och avslutas under 2021 med rättning och komplettering av uppgifter enligt ovan, harmonisering och kodning av namn, orter och yrken till SwedPop-standard samt komplettering med döda 2018-2020. Kodning till IDS-format utförs därefter av Riksarkivet.

Aktiviteter under 2022

1. Länkning till folkräkningar och övriga databaser i SwedPop.
2. Information från Sveriges Dödbok har kodats, konverterats till IDS-format och länkats till databaserna i SwedPop (delmål 2022:d).

Tidplan

Arbetet kommer att pågå under tidsperioden 2020-2022.