

SwedPop

Svenska befolkningsdatabaser
för forskning



Foto: Privat

En nationell forskningsinfrastruktur med data om
Sveriges historiska befolkning

SwedPop är en nationell forskningsinfrastruktur som finansieras till lika delar av Vetenskapsrådet (2023-2026) och av ett konsortium där följande parter ingår:

- Umeå universitet (medelsförvaltare)
- Göteborgs universitet
- Lunds universitet
- Riksarkivet
- Stockholms stadsarkiv

Följande databaser ingår i SwedPop:

POPUM-POPLINK (DDB)

Enheten för demografi och åldrandeforskning CEDAR, UmU.

Livsförloppsdata för personer från ca 100 församlingar i olika svenska regioner 1650-1960.
Folkräkningen 1950.

Göteborgs Befolkningspanel (GOPP)

Institutionen för ekonomi och samhälle, GU.

Fyraårspaneler för ett slumpvis urval individer (1%) observerade i Göteborg 1915-1967.

Skånes Ekonomisk-Demografiska Databas (SEDD)

Centrum för ekonomisk demografi, LU.

Livsförloppsdata för personer från ett antal skånska församlingar 1680-1968.

SweCens-databasen

Riksarkivet.

Folkräkningarna 1880-1940. Tvärsnitt vart tionde år för hela riket.

Rotemansarkivet

Stockholms stadsarkiv.

Livsförloppsdata för Stockholms befolkning 1878-1926.

Sveriges dödbok - Emiweb

Sveriges Släktforskarförbund.

Uppgifter om avlidna i Sverige 1830-2020.

1

Inledning

SwedPop är en nationell forskningsinfrastruktur som stöds av Vetenskapsrådet tillsammans med fem svenska universitet och arkivmyndigheter. Sedan 2016 har SwedPop haft stöd från Vetenskapsrådet. För perioden 2018-2022 beviljades infrastrukturen bidrag inom ramen för utlysningen Infrastruktur av nationellt intresse och i 2021 års utlysning för samma ändamål beviljades fortsatt stöd för perioden 2023-2026.

I SwedPop samordnas de viktigaste historiska befolkningsdatabaserna i Sverige till en gemensam resurs som gör det lättare för forskare att få tillgång till högkvalitativa svenska befolkningsdata för forskning. Ett gemensamt databasformat och samordnad kodning gör att den stora forskningspotentialen hos svenska befolkningsdata kan bättre tas till vara. Data från flera olika databaser kan användas tillsammans, till exempel för jämförande studier av hur människors liv, hälsa, sociala och ekonomiska levnadsvillkor har förändrats över tid. Våren 2021 lanserades ett uttagsverktyg på swedpop.se som gör det möjligt för forskare att kostnadsfritt ladda ned data från de ingående databaserna. Data som innehåller personuppgifter hanteras i enlighet med gällande lagstiftning med särskilt fokus på GDPR, OSL och Etikprövningslagen.

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, Lunds universitet, 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 1878-1926. 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 olika socioekonomiska miljöer och både stad och landsbygd finns representerade. De rikstäckande folkräkningarna är i detta sammanhang en viktig resurs. 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

Arbetet under bidragsperioden 2023-2026 syftar till att färdigställa och konsolidera SwedPop som en nationell infrastruktur och en värdefull och använd resurs för svensk forskning. Fram till och med 2026 kommer utvecklingen och driften av SwedPop att fokusera på följande övergripande områden:

- Slutföra harmonisering, standardisering och länkning av data.
- Utveckla uttagsverktyget på swedpop.se med nya funktioner och format.
- Ta fram ett avancerat och effektivt användarstöd, samt anordna utbildning i användning av infrastrukturen.
- Slutföra den datainsamling som inleddes under projektperioden 2018-2022.
- Förbättra tillgången till longitudinella inkomstdata på individnivå i de ingående databaserna.
- Komplettera data över döda med dödsorsaker under pandemiåren 1917-1920.
- Konsolidera SwedPop som nationellt kunskapscentrum för äldre befolkningsdata (före SCB-registren) på mikronivå.

Särskilda villkor för bidraget (2021-00183) samt budget och övergripande arbetsplan utifrån föreliggande projektansökan fastställdes av Vetenskapsrådet 2022-03-07.

Utgångspunkter

Verksamhetsplanen för SwedPop tar sin utgångspunkt i de särskilda bidragsvillkor och den budget och övergripande arbetsplan för bidragsperioden 2023-2026 som fastställts av Vetenskapsrådet (2021-00183)

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 året innan aktuellt verksamhetsår.

Verksamhetsplanen är strukturerad utifrån de moduler som finns angivna i budget och arbetsplan för projektet. Verksamhetsplanen fastställdes 2022-09-22 av styrgruppen för SwedPop.

Moduler och 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.). För att göra det lättare att följa projektet i sin helhet har dessa moduler numrerats på samma sätt som under projekttiden 2018-2022. Några av rubrikerna har dock fått revideras något för att anpassas till projektläget under perioden 2023-2026. De fyra moduler där arbetet är helt avslutat (M6,M7,M11 och M14) finns för fullständighetens skull med i översikten, men de kommer inte att bli föremål för någon aktivitet under innevarande period. I texten finns för varje modul referenser till av VR fastställd projektbudget i formatet *M1.1*, inom parantes.

Moduler 1-15:

| | |
|-----|--|
| M1 | Organisation och ledning |
| M2 | Standardisering, kodning och datastruktur |
| M3 | Tillgängliggörande och användarstöd |
| M4 | Länkning, utvärdering och dokumentation |
| M5 | Utvidga GOPP-databasen med nya data |
| M6 | <i>Komplettering och länkning av data i Rotemansdatabasen (klart 2022)</i> |
| M7 | <i>Komplettering av dödsorsaker i Rotemansdatabasen (klart 2022)</i> |
| M8 | Komplettering av Rotemansdatabasen med ekonomiska data |
| M9 | Komplettering av SEDD, ekonomiska data Landskrona 1905-1947 |
| M10 | Komplettering av SEDD, befolkningsdata Landskrona 1882-1910 |
| M11 | <i>Komplettering av DDB, befolkningsdata Nysåtra församling (klart 2022)</i> |
| M12 | Komplettering av DDB, ekonomiska data Skellefteå 1881-1949 |
| M13 | Inkludering av 1940 års folkräkning i SweCens-databasen |
| M14 | <i>Inkludering av 1950 års folkräkning i SwedPop (klart 2022)</i> |
| M15 | Komplettering, kodning och länkning av Sveriges Dödbok |

Modulerna 1-4 är av övergripande karaktär och omfattar projektledning/organisation och administration (M1), standardisering, kodning och datastruktur (M2), utveckling och underhåll av uttagsverktyg samt användarstöd (M3) och länkning, utvärdering och dokumentation (M4). Dessa fyra moduler är starkt beroende av varandra och genomförandet av de planerade aktiviteterna inom dessa moduler är av avgörande betydelse för att SwedPop ska bli en välfungerande resurs för svensk forskning. Återstående moduler kan struktureras tematiskt i tre olika områden (I-III), vilka samtliga, var för sig och tillsammans, väsentligt kommer att bidra till att öka infrastrukturens vetenskapliga värde och göra det möjligt att fullt ut ta tillvara den stora forskningspotentialen hos svenska befolkningsdata.

Dessa tre områden (I-III) är:

- I. Addera och komplettera data om inkomst och socio-ekonomiska förhållanden och dödsorsaker (M8, M9, M12, M15)
- II. Strategisk komplettering av databaser i SwedPop med data från vetenskapligt relevanta geografiska områden eller tidsperioder (M5, M10, M15)
- III. Komplettering med kodade och länkade data från nationell folkräkning (M13)

Planerade delresultat under bidragsperioden beskrivs i den strategiska plan för SwedPop som utarbetas utifrån den övergripande arbetsplanen för bidragsperioden som fastställts av Vetenskapsrådet. Den strategiska planen fastställs formellt av styrgruppen under 2023.

Under 2023 är målsättningen att arbetet ska leda fram till följande delresultat:

| År | Nr | Delresultat | Modul |
|------|----|---|-------|
| 2023 | | | |
| | 1 | En första workshop för forskare/användare genomförs, inriktad på användning av SwedPop:s uttagsverktyg, statistiska metoder och analys genomförs. | M3 |
| | 2 | Inkomst- och taxeringsuppgifter för åren 1906, 1907, 1908, 1909, 1911, 1912, 1913, 1914 och 1916 har registreras och länkas till SEDD. | M9 |

| | | |
|---|---|-----|
| 3 | Församlingsböcker från Landskrona 1891-1892 har registrerats och länkats till SEDD. | M10 |
| 4 | Inkomst- och taxeringsuppgifter för åren 1881, 1885, 1889, 1893, 1897 och 1901 har registreras och länkas till DDB. | M12 |

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 2023.

Modul 2

Standardisering och dokumentation

Nuläge

Under bidragsperioden 2018-2022 har ett grundläggande arbete utförts med harmonisering och standardisering av kodning och datastruktur. Detta har resulterat i genomarbetade kodscheman för kodning av yrkesuppgifter, dödsorsaker och geografiska områden till etablerad internationell och nationell kodstandard. En svensk standard för den internationella HISCO-kodningen för yrken och för den nya kodstandarden för historiska dödsorsaker, ICD10H, har tagits fram. En svensk standard har etablerats för kodning av geografiska områden baserad på Riksarkivets TORA-system och svenska standarddefinitioner har utarbetats för det internationella dataformatet IDS - Intermediate Data Structure. Det omfattande arbetet med kodning och standardisering har varit helt avgörande för att uppnå projektets syfte: att tillgängliggöra harmoniserade historiska befolkningsdata data av hög kvalitet som en resurs för svensk forskning. Tillgången till standardiserade och dokumenterade kodscheman gynnar också samordning och jämförbarhet i ett större nationellt perspektiv, genom att etablerade genomarbetade standards för centrala variabler nu finns att tillgå när nya databasprojekt skapas. En basal version av kodscheman och grundläggande principer för kodning finns redan nu publicerad på swedpop.se, men denna information kommer under bidragsperioden att ersättas av en fullständig och utförlig dokumentation av utarbetade standards.

Inom ramen för M2 kommer det omfattande arbetet med harmonisering och standardisering att slutföras. Det innebär bland annat följande:

- Det nyligen utarbetade internationella klassifikationssystemet ICD10H, där SwedPop deltagit i utvecklingsarbetet kommer fullt ut att implementeras i SwedPop. Systemet är utformat för att hantera äldre och inte sällan ospecifika dödsorsaker, med bibehållet forskningsvärde och ökad kontextuell förståelse (M2.1).

- En vidareutveckling av den geografiska standardkodningen görs, till att omfatta fler administrativa enheter (härad, tingslag, fögderier, pastorat, kontrakt och stift). Detta kommer att möjliggöra mer detaljerade spatiala analyser av SwedPop-data (M2.2).
- En fullständig och utförlig dokumentation på svenska och engelska av utarbetade svenska standarddefinitioner för kodning av yrken (SwedPop-HISCO), dödsorsaker (SwedPop-ICD 10H), geografiska och administrativa enheter (SwedPop-TORA) och databasstruktur (SwedPop-IDS) kommer att utarbetas och publiceras, till ledning för både forskare och framtida databasprojekt (M2.1).
- För att upprätthålla kodstandard och harmonisering av data kommer en Clearing Committee att inrättas som en permanent del av organisationen (M2.3-M2.5).
- Inom denna modul finns också budget för att lägga in metadata för SwedPop i RUT (M2.6-M2.10).

Aktiviteter under 2023

1. Dokumentation av SwedPop-standards påbörjas. Dokumentationen läggs upp löpande på swedpop.se med målsättning att samtlig dokumentation ska finnas tillgänglig under 2025.
2. Vidareutveckling av SwedPop-TORA-kodning med fler administrativa enheter påbörjas.

Tidsplan

Standardisering av geografiska variabler ska vara avslutad under 2024. En detaljerad dokumentation ska vara publicerad under 2025. Metadata för SwedPop ska ha lagts in i RUT senast under 2026.

Modul 3

Tillgängliggörande och användarstöd

A.Support swedpop.se

Nuläge

Under bidragsperioden 2018-2022 har webbplatsen swedpop.se byggts upp och utvecklats. Webbplatsen är en viktig nod i SwedPop, som kräver löpande underhåll och teknisk uppdatering (M3.1). Där får forskare tillgång till data, dokumentation och information och under innevarande bidragsperiod kommer den också att få en viktig roll när det gäller användarstöd. Därtill kommer serverkapaciteten att utökas, med tanke på såväl ökade datamängder som ett ökat antal användare (M3.8).

Aktiviteter under 2023

Löpande underhåll av webbplatsen swedpop.se.

Tidsplan:

Pågående, under hela projektiden.

B. Vidareutveckling och underhåll av uttagsverktyget på swedpop.se

Nuläge

I juni 2021 lanserades ett uttagsverktyg på swedpop.se där forskare kan ladda ned harmoniserade SwedPop-data i IDS-format, för närvarande från DDB, SEDD och SweCens. Data från övriga databaser kommer att tillgängliggöras så snart kompletteringar, kodning och harmonisering är klar. Under bidragsperioden kommer uttagsverktyget att vidareutvecklas för att möjliggöra output i fler dataformat (M3.2). Fokus ligger på format som är kompatibla med de vanligaste analysprogrammen. Uttagsverktyget kräver också löpande underhåll och tekniska uppdateringar (M3.3).

Aktiviteter under 2023

1. Vidareutveckling av uttagsverktyget startar.

Tidsplan:

Nya outputformat beräknas vara tillgängliga under 2026.

C. Användarstöd

Nuläge

Efter den första uppbyggnadsperioden 2018-2022 befinner sig SwedPop i en fas när det är dags att börja flytta fokus från databasuppbyggnad, till tillgängliggörande och användarstöd. En koordinator för forskarstöd (deltid) kommer att knytas till infrastrukturen, ansvarig för forskarkontakter och tillgängliggörande av data som innehåller personuppgifter (M3.5). En permanent Data Access Committee kommer att inrättas, med uppdrag att koordinera och genomföra utlämningsprövning av data i enlighet med OSL och GDPR (M3.4). För att stimulera användning av infrastrukturen och öka kunskapen om SwedPop-data kommer två workshops riktade till forskare att anordnas 2023 och 2025. För att stötta forskare i användningen av uttagsverktyget kommer video-tutorials att produceras, som ska finnas tillgängliga på swedpop.se tillsammans med andra resurser för användarstöd (M3.6).

Aktiviteter under 2023

1. Genomföra en första workshop för forskare/användare av SwedPop (Delmål 2023:1).

Tidsplan:

En workshop genomförs under andra halvan av 2023.

D. Tillgängliggörande av data

Nuläge

I juni 2021 gjordes den första releasen av SwedPop-data på swedpop.se. Den första versionen SwedPop version 1.0 (<https://doi.org/10.48524/SWEDPOP-V1-0>) omfattade longitudinella befolkningsdata från DDB och SEDD till och med 1910 samt och folkräkningsdata (1880) från Riksarkivet tillsammans med tillhörande dokumentation. I juni 2022 publicerades SwedPop version 2.0 (<https://doi.org/10.48524/SWEDPOP-V2-0>) där data utökats med folkräkningarna

1890-1910 samt data från Sveriges Dödbok 1832-2020. Avsikten är att genomföra en ny datarelease per år, med nya och upparbetade data. I nästa release kommer data från Rotemannen att tillgängliggöras för forskning på swedpop.se. Data kommer även att utökas med länkar mellan de olika databaserna.

Aktiviteter under 2023

1. Genomföra en ny datarelease, SwedPop version 3.0, med länkar och tillhörande dokumentation under september 2023.

Tidsplan:

Nya datareleaser kommer att genomföras löpande under bidragsperioden.

Modul 4

Länkning utvärdering och dokumentation

Nuläge

Länkningsarbetet kan delas in i tre steg:

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

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. Testlänknings 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 2022, när flertalet av de kompletteringar som pågår av databaserna är avslutade och harmoniseringen är slutförd kommer länkningens andra fas att genomföras enligt utarbetade och testade metoder. Samtliga länkar kommer att publiceras på swedpop.se tillsammans med den länkningskod som använts. Under bidragsperioden 2023-2026 kommer aktiviteterna inom M4, förutom länkning av tillkommande data, att fokusera på en utvärdering och dokumentation av den länkning som gjordes 2022 (M4.1-4.4). Syftet att utvärdera kvaliteten på de länkar som har skapats, undersöka länkningsgrad, potentiell bias och andelen true/false positives. Dokumentationen kommer att publiceras på swedpop.se och vara utformad så att forskare kan förstå och värdera de överväganden som har gjorts.

Aktiviteter under 2023

1. Analysera och utvärdering av kvalitet och länkningsmetoder.
2. Analys och utvärdering av länkningsgrad, potentiell bias mm.
3. Påbörja dokumentation av länkning.

4. Löpande arbete med länkning när nya data läggs till.

Tidsplan:

Pågående arbete under hela tidsperioden, med fokus på utvärdering och förbättring under de två första åren.

Modul 5

Utvidga GOPP-databasen med nya data

Nuläge

GOPP-databasen innehåller individuella paneldata med demografisk information och uppgifter om yrke, inkomst och hushållsförhållanden för ett slumpvis urval (1%) av Göteborgs befolkning 1915-1967. Information om urvalspersoner 1915-1943 (sample 1) samlades in inom ramen för ett forskningsprojekt finansierat av FAS (FORTE) och har ställts till SwedPops förfogande av Göteborgs universitet. Under bidragsperioden 2018-2022 utvidgades databasen genom ett nytt urval av personer som var bosatta i Göteborg någon gång 1947-1967 (sample 2). Aktiviteterna inom M5 syftar till att komplettera GOPP-databasen med uppgifter om hushållsförhållanden och inkomst och yrke för åren 1947, 1952, 1955, 1959, 1963 och 1967 för båda dessa urvalsgrupper (M5.2-M5.4). Arbetet har påbörjats under föregående projekt och kommer att slutföras inom ramen för föreliggande projekt. Därefter kommer samtliga individdata från 1915-1967 att länkas samman till en databas i SwedPop-standard och tillgängliggöras. Detta arbete kommer att vara avslutat under 2026.

Aktiviteter under 2023

1. Slutföra excerpering av inkomst och yrke för sample II.
2. Fortsätta excerpering av hushållsuppgifter för sample II.

Tidsplan

Datainsamlingen kommer att pågå hela bidragsperioden 2023-2026 och kommer att avslutas under 2026.

Modul 6

Komplettering och länkning av data i Rotemansdatabasen

Nuläge

Arbetet inom denna modul slutfördes under 2022.

Modul 7

Komplettering av dödsorsaker i Rotemansdatabasen

Nuläge

Arbetet inom denna modul slutfördes under 2022.

Modul 8

Komplettering av Rotemansdatabasen med ekonomiska data

Nuläge

För att öka den komparativa forskningspotentialen i SwedPop kompletteras Rotemansdatabasen med ekonomiska data. Under 2019-2022 har ett förberedande arbete utförts med skanning, systemutveckling och förberedelser för inregistrering av information om inkomst, skatt, hyra och fastighetsinnehav från mantalslängder. Inom ramen för M8 kommer Rotemansdatabasen att kompletteras med ekonomiska data på individnivå för Stockholms befolkning åren 1880, 1890, 1900, 1910 och 1920 (M8.1-M8.5). Arbetet beräknas vara avslutat 2026.

Aktiviteter under 2023

1. Utarbeta och driftsätta en webb-app för inregistrering och verifiering av ekonomiska data.
2. Inregistrering av data från 1880.

Tidsplan

Kompletteringsarbetet kommer att pågå fram till 2026.

Modul 9

Komplettering av SEDD, ekonomiska data 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. Under bidragsperioden 2018-2022 har databasen kompletterats med information från 37 inkomstår: vart femte år mellan 1906 till 1968. Under bidragsperioden 2023-2025 kommer data från återstående mellanliggande 27 inkomstår med fokus på tidsperioden 1906-1947 att adderas till databasen och tillgängliggöras för forskning (M9.1-M9.7). Arbetet beräknas sträcka sig över åren 2022-2025 och vara avslutat i och med 2025 års utgång.

Aktiviteter under 2023

1. Inkomst- och taxeringsuppgifter för perioden 1906-1914, samt 1916 registreras och länkas till SEDD.
2. Yrkesuppgifter från längderna kodas.
3. SEDD har kompletterats med ekonomiska data för åren 1906-1914 och 1916.

Tidplan

Kompletteringsarbetet beräknas vara avslutat under 2025.

Modul 10

Komplettering av SEDD, befolkningsdata Landskrona, 1882-1910

Nuläge

Befolkningsdata för Landskrona 1910-1968, och från omkringliggande församlingar 1813-1968 har inkluderats i SEDD inom ramen för tidigare projekt. Under bidragsperioden 2018-2022 har information från församlingsböcker, register över födda, döda, in- och utflyttade har registrerats för perioden 1905-1910, samt faders yrke för inflyttade registrerats för perioden 1905-1947. Aktiviteterna inom M10 har som mål att avsluta inregistreringen av befolkningsdata från Landskrona genom att fokusera på perioden 1882-1894. Med detta kan kompletta befolkningsdata för Landskrona för tidsperioden 1882-1968 tillgängliggöras för forskning (*M10.1-M10.7*).

Aktiviteter under 2023

Församlingsböcker från Landskrona 1891-1892 inregistreras och länkas till SEDD.

Tidplan

Arbetet beräknas pågå till och med 2025.

Modul 11

Komplettering av DDB med befolkningsdata för Nysätra församling

Nuläge

Arbetet inom denna modul slutfördes under 2022.

Modul 12

Komplettering av DDB, ekonomiska data Skellefteå, 1881-1949

Nuläge

Arbetet inom M12 avser en komplettering av ekonomiska data i form av inkomst- och taxeringslängder för tidsperioden 1881-1949, för att öka databasens forskningsvärde. Under bidragsperioden 2018-2022 gjordes ett förberedande arbete med metodutveckling och datainsamling för Lövångers församling. Under bidragsperioden 2022-2026 kommer ekonomiska data på individnivå att samlas inför hela Skellefteåregionen, ett område som genomgick en snabb industrialisering och samhällsomvandling under tidigt 1900-tal. Data samlas in för vart fjärde år under perioden 1881-1949. Datainsamling, länkning, standardisering och kodning kommer att pågå under hela bidragsperioden och avsikten är att data ska kunna tillgängliggöras för forskning under 2026 (*M12.1-M12.5*).

Aktiviteter under 2023

1. Inscanning av källor.
2. Inregistrering och länkning av inkomst- och taxeringslängder från Skellefteå för åren 1881, 1885, 1889, 1893, 1897 och 1901.

Tidplan

Datainsamlingen beräknas vara avslutad under 2026.

Modul 13

Inkludering av 1940 års folkräkning i SweCens-databasen

Nuläge

Arbetet med att inkludera 1940 års folkräkning i SweCens-databasen startade 2018. I samarbete med Sveriges Släktforskarförbund startade under hösten 2019 ett storskaligt citizen-science-projekt där frivilliga släktforskare deltar i inregistreringsarbetet. Under 2020-2022 har registreringsarbetet skalats upp i takt med att allt flera frivilliga släktforskare anslutit sig till projektet. Detta samarbete kommer att fortsätta även under bidragsperioden 2023-2026. Projektet bistår med handledning, koordination och kvalitetskontroller av insamlade data, samt den efterföljande validering, kodning och länkning som krävs för att kunna tillgängliggöra data för forskning (M13.1-M13.3).

Tekniken för Handwritten Text Recognition (HTR) har utvecklats snabbt de senaste åren. Under våren 2023 genomförs en förstudie för att utreda förutsättningarna att använda Handwritten Text Recognition (HTR) för att snabba upp registreringen av 1940 folkräkning.

Aktiviteter under 2023

1. Fortsatt inregistrering av data i 1940 års folkräkning.
2. Rekrytering av ytterligare registrerare i Sveriges Släktforskarförbunds regi.
3. Genomföra en förstudie om förutsättningarna för Handwritten Text Recognition (HTR) av folkräkningen 1940.

Tidplan

Datainsamling kommer att pågå hela bidragsperioden 2023-2026.

Modul 14

Inkludering av 1950 års folkräkning i SwedPop

Nuläge

Arbetet inom denna modul slutfördes under 2022.

Modul 15

Komplettering, kodning och länkning av Sveriges Dödbok

Nuläge

Sveriges dödbok (SDB) 1860-2017 är en databas som innehåller information om alla individer som har avlidit i Sverige från 1860 och fram till 2017. Genom ett avtal med Sveriges Släktforskarförbund har SwedPop fått tillgång till SDB för att inkludera databasen som en del i SwedPop. Under bidragsperioden 2018-2022 har en SwedPop-harmoniserad version av SDB tagits fram och länkats till övriga SwedPop-databaser. Tillgång till namn, datum och födelseplats gör SDB till en värdefull resurs för länkning mellan de olika SwedPop-databaserna. Under bidragsperioden 2023-2026 kommer huvuddelen av arbetet i M15 att ägnas åt registrering, validering, harmonisering och kodning av en uppdaterad version av SDB, utvidgad med data för åren 1815-1859 och successivt även för 2020-2025 (*M15.3*). Datainsamling för 1815-1829 har påbörjats, finansierad av medel utanför SwedPop. Komplettering, validering och upparbetning utförs av Sveriges släktforskarförbund och Stockholms stadsarkiv i samverkan,. Viss kodning utförs också av CEDAR, Umeå universitet och Riksarkivet. Därtill kommer en kompletterande datainsamling att göras med dödsorsaker för åren 1917-1920, perioden omedelbart före och under utbrotten av pandemisk influensa (spanska sjukan) 1918-1920, som är av stort forskningsintresse (*M15.2*).

Aktiviteter under 2023

1. Upparbetning och kodning av döda 1830-1859.
2. Datainsamling av dödsorsaker 1917-1920 har påbörjats.
3. Inregistrering av data om döda 1815-1829 pågår.

Tidplan

Arbetet kommer att pågå under tidsperioden 2023-2026.