

## KIK-møde

**Dato:** 11. APRIL 2018, kl.11-14

**Sted:** Rigsarkivet Odense

**Deltagere:** Barbara Ana Revuelta-Eugercios (BR) referent, John Gregers Hvidkjær (JGH) begge fra Rigsarkivet, Odense, Svend Erik Christiansen (SEC) KIP-grundlægger mm, Jens V. Olsen (JVO) KIIP-udvikler, Hans Jørgen Marker (HJM) KIP-grundlægger mm, Hans Chr. Johansen (HCJ) forsker, Arne Christiansen (AC) i stedet af Jesper Skov DIS-Danmark, Jens Karl Rasmussen (JKR).

**Yderligere deltagere:** Allan Vestergaard (AV) og Katrine Tovgaard-Olsen (KET) fra Rigsarkivet og Nicolai Donitzsky (ND), Sophie Hirt) og Peter Goodwin fra Ancestry.

**Afbud:** Michael Dupont.

## Dagsorden

1. Godkendelse af dagsorden
2. Status fra Rigsarkivet Odense (og præsentation af Barbara Revuelta-Eugercios)
3. Præsentation af Ancestrys planer med og ansøgning om brug af data fra DDD v. Nikolai Donitzsky (Ancestry)
4. Diskussion af et revideret forslag til Kommissorium for KIK
5. Persondataforordningen fra 2018 og KIP – orientering
6. Status FT1901 og de øvrige folketællinger
7. Næste møde
8. Eventuelt

## Referat

### Ad 1

AV indledte mødet med at præsentere nye deltagere: Barbara Revuelta-Eugercios (BRE), som har overtaget Nanna Floor Clausens (NFC) funktioner og opgaver, og tre medarbejder fra virksomheden Ancestry. Barbara er blevet fastansat fra 1. februar. Hun er uddannet i historie og demografi, og kommer fra en forskerstilling på Københavns Universitet. Hun har bl.a. arbejdet med de danske folketællinger fra 1880 og 1885 for København. Hun overtager også sekretærfunktioner for KIK. De tre medarbejder fra Ancestry er blevet inviteret som opfølgning på seneste KIK-møde, hvor det blev noteret at KIK ønskede at få en nærmere præcisering af Ancestrys publiceringsplaner for folketællingsdata fra DDD, såfremt dette overdrages til Ancestry.

Dagsordenen blev godkendt.

## **Ad 2 Status fra Rigsarkivet Odense (og præsentation af Barbara Revuelta-Eugercios)**

BRE beskrev status for DDD siden sidste møde på forskellige områder:

### Indtastning:

Den største nyhed er at 1860 blev færdigindtastet i januar. Det er vores 8. færdigindtastede folketælling, med 1.610.013 poster. Rensning af data til forskning er begyndt men der er stadigvæk et stort arbejde foran os. I alt har DDD 20,119,739 poster pr. 10. April. I år er der blevet indtastet 397,363 poster indtil nu, hvilket er mindre (103,739) end sidste år på samme tid (I 2017 indtil 10/April var 501,102 indtastede). Indtil nu er og har 133 indtastere været i gang i år med KIP.

Et særligt, igangværende projekt er indtastningen af 1840-folketællingen fra den danske koloni i Frederiksnagore (i dag Serampore) i Indien. Frederiksnagore var en dansk handelsstation i Indien 1755-1845. Projektet er et samarbejde med Nationalmuseet i forbindelse med ”Serampore Initiativet”. Nationalmuseet v. Simon Rasten kontaktede først CS-portalen i forhold til indtastning af folketællinger efter billeder som de selv havde taget, men efterfølgende er KIIP programmet blevet valgt som det bedst egnede. NFC præsenterede projektet for nogle af indtasterne og det har vakt interesse. Billeder er fordelt på 12 kipnr., der hurtigt blev reserverede. Indtil nu er 10 ud af 12 indtastede. De 10 afleverede pakker svarer til 10,158 poster. Det er et fint eksempel på hvad KIP kan! Det var relativt nemt at overbevise frivillige om, at der var et godt projekt fordi Nationalmuseet har en webside hvor man kan læse alt om projekter og forstå hvorfor det er vigtigt for forskerne – og så er materialet selvfølgelig ekstra spændende p.g.a. den eksotiske baggrund.

Som afslutning på samarbejdet med NAPP (North Atlantic Population Project), afleverede NFC en forskningskopi af FT1845 og FT1885 folketællinger til dem kort før sin afgang.

### Frivillige:

Vi har fortsat en gang om måneden såkaldt ’tastecafé’, hvor frivillige indtastere kan komme og taste eller diskutere problemstillinger etc. Fremmødet er ikke så stort, som vi havde håbet, men vi fortsætter indtil videre og håber på, at flere efterhånden vil møde op.

### Udlevering og formidling:

Der har været 4 udleveringer af data til opgaveskrivninger og forskningsprojekter i 2018. BRE fortalte om sit arbejde med at øge forskningsbrugen af DDD: Digital History Methods Lab. Baggrunden for det nye initiativ er behovet for at træne historikere i de forskningsmetoder der er relevante når man bruger DDD. Digital History Methods Lab er en serie af workshops og seminarer på Københavns Universitet, der tager disse emner op. Der er foreløbigt holdt 4 seminarer på Københavns Universitet og der er planer om at udvide til SDU og Aarhus Universitet.

Der er med samme formål lavet noget ændringer på DDD websiden for at synliggøre forskningsmuligheder, herunder en ny fane kaldet ”Forskning”, hvor der findes oplysninger om hvordan man kan få udleveret data. Derudover er der også planer om en lille side om ”citizen science” side hvor vi vil præsentere projekter som har brug for hjælp af frivillige (som Nationalmuseets Serampore-projekt).

### Linking:

Sidste år var der to projekter i *SDU Supercomputer challenge* som brugte DDD data til at linke personer gennem folketællinger. Det ene hold vandt konkurrencens tredje plads. Vi har fået deres resultater og nu arbejder vi med at finde ud af hvor gode de er. En af deltagerne har været i praktik hos DDD og her arbejdet på en videnskabelig artikel.

#### Forskning med DDD:

Der har været nogle aktiviteter med direkte brug (og planer om brug) af DDD data til forskning: BRE og KET holdt således foredrag om arbejdet med DDD på konferencen *Digital Humanities in the North*-konference i Helsinki i marts og BRE agter at søge om penge til et mindre projekt hos Kulturministeriets forskningspulje. Der fortsættes også med at søge om penge til det store Link Lives projekt. En ny ansøgning blev sendt til Innovationsfonden i februar, i samarbejde med Københavns Universitet (Anne Løkke) og Københavns Stadsarkiv.

Det er Rigsarkivets plan, at DDD skal indgå i crowdsourcing-portalen. For at det kan lade sig gøre skal portalen udvikles betydeligt. Det betyder, at DDD fortsætter på den nuværende platform i sandsynligvis mindst 1 til 1½ år endnu. Vi arbejder nu på at skabe et redskab til at importere data til cs-portalen (gennem en API), som er en forudsætning for at udfase DDD.

### **Ad 3 Præsentation af Ancestrys planer med og ansøgning om brug af data fra DDD v. Nikolai Donitzsky (ND) (Ancestry)**

AV sammenfattede Ancestrys projekt og introducerede ND: Ancestry agter at filme kirkebøger for 1812 til 1892 og derudover at afskrive kirkebøgerne for perioden 1812 til 1914. Ancestry filmer fra originalkilderne (kontraministerialbøgerne). Til brug for kvalitetssikring af deres indtastninger har de allerede fået en kopi af visse data fra DDD. Ancestry har tilbudt at overdrage Rigsarkivet en kopi af de afskrevne kirkebøger, mod at modtage kopi af de afskrevne folketællinger fra Dansk Demografisk Database.

ND præsenterede sig selv og hans hold. Sophie Hirt er projektleder og koordinerer forskellige teams og workflow fra London og Peter Goodwin er ansvarlig for affotografering og scanning. Begge arbejder fra London. ND er ansvarlig for hele projektet.

ND forklarede at Ancestry bruger indtastere der er bosiddende i Østen, men som kan læse gotisk.

HCJ spurgte om de indtaster feltet ”erhverv” og Sophie Hirts, svarede at de ikke gør det fordi det netop i dette felt er svært at holde et højt kvalitetsniveau. KET spurgte om kvalitetsniveauet generelt, og DN svarede at de beregner kvalitet i procent af korrekte ord (ikke tegn). Han refererer til de tyske eksempler hvor det varierer for eksempel fra 93% i Hamburg til 98,3% i Lübeck. I Norge har det været op imod 98%.

DN forklarede fortsat at Ancestry agter at stille det danske materiale, herunder også data fra DDD såfremt de kunne råde over dette, til rådighed frit for alle, uden at opkræve betaling for brugen.

AV spurgte hvorfor Ancestry er interesserede i det her projekt og ND svarede at det er en del af deres store projekt i Norden (der er i forvejen arbejde i gang i Sverige, Norge og Finland men Danmark mangler).

Til sidst viste Peter Goodwin slides med status for affotograferingen i Danmark og ND viste hvordan brugere vil få adgang til materialet, fra søgning til de original kilder. Alle brugere i verden vil have adgang til de danske folketællinger gennem deres webside. Og brugere vil være til stand til at bidrage med fejlrettelser.

Da delegationen fra Ancestry havde forladt mødet, fortsatte diskussionen.

HCJ spurgte til fremtiden for KIP, efter at arbejdet med afskrift af kirkebøger var afsluttet. AV svarede at man kan forestille sig en række projekter, og at Rigsarkivet f.eks. håbede, at indtasterne kunne være interesserede i at "linke" personer. Efter en lille samtale om udfordringer i kobling, blev det besluttet at emnet tages op igen på næste KIK møde med et oplæg om kobling erfaring og Link Lives projektet.

Der var enighed blandt komiteens medlemmer om, at Ancestry havde virket overbevisende i forhold til at demonstrere, at de var i stand til at levere affotograferinger og afskrivninger i høj kvalitet, og at Ancestrys tilbud om at give gratis adgang til danske ressourcer og til at udveksle afskrifter af kirkebøger med data fra DDD var attraktivt. KIK gav derfor sin støtte til, at Rigsarkivet kontakter de enkelte indtastere med henblik på at opnå disses tilladelse til at overdrage kopier af data til Ancestry.

#### **Ad 4 Diskussion af et tilrettet forslag til Kommissorium for KIK**

KET fremlagde en ny version af udkast til kommissoriet, hvori de ret få foreslåede ændringer fra forrige møde var indarbejdet. Der fulgte en kort diskussion om Michael Duponts forslag om, at Samfundet for Dansk Genealogi og Personallhistorie også får en repræsentant med i Kildeindtastningskomiteen. Forslagen blev accepteret. Rigsarkivets direktion skal nu godkende den endelige tekst, hvorefter denne tænkes publiceret på DDD's hjemmeside.

#### **Ad 5 Den kommende Persondataforordning**

AV oplyste at Rigsarkivet er i gang med at udarbejde en politik for håndtering af data i DDD, som ventes at være færdig kort efter 1. Maj.

#### **Ad 6 Status for FT1901 og de øvrige folketællinger**

BRE oplyste at 1901 i alt omfatter 2.4449,540 poster og der er til dato indtastet 65,68%. Der mangler dele af Jylland og hele Staden København. Resten har følgende dækningsgrad: 1890: 53.5%; 1855: 43,54%; 1870: 34,74%, 1930: omkring 20%

Den nye prioriterede folketælling er 1901. Det drøftedes om ikke det var bedre, at fokusere på de folketællinger som ikke er omfattet af det nye persondatadirektiv.

#### **Ad 7 Næste møde**

Onsdag den 9. oktober 2018 kl. 10.15 i Odense

#### **Ad 8 Eventuelt**

SEC har nogle forslag til forbedresøgemuligheder i forhold til forskellige stavmåder (ø, ö).