

MONA Referensgruppsmöte 2024-12-06

Närvarande:

Susanna Bylin, Vetenskapsrådet
Susanne Dahllöf, MONA-supporten, SCB, ordförande
Olof Elfwering, MONA-supporten, SCB
Lisa Freed, MONA-supporten, SCB
Jonas Holm, MONA-supporten, SCB
Oscar Erixson, Uppsala universitet
Olle Hammar, Linnéuniversitetet, gäst
Martin Hällsten, Stockholms universitet
Freja Werke, sektionschef, mikrodataverksamheten, SCB
Jenny Lindberg, enhetschef Uppdrag och extern service SCB
Lucas Fischer Madsen, Göteborgs universitet
Helena Nilsson, Jönköpings universitet
Oscar Stålnacke, Umeå universitet

Frånvarande:

Martin Brandhagen, Vetenskapsrådet
Johan Holmberg, Umeå universitet
Roujman Shahbazian, Stockholms universitet
Torun Österberg, Göteborgs universitet

Inledning

Susanne hälsade alla välkomna. Mötet inleddes med en presentationsrunda.

1. Olle Hammar – Reproduktion och replikering med svenska registerdata

På förslag från Oscar Erixson var Olle Hammar inbjuden till referensgruppsmötet.

Reproduktion och replikering

Nationalekonomisk replikeringsstudie med svenska registerdata (med Anna Dreber, Institute for Replication, m fl)

Replikering: nya data

Reproduktion: samma data som i den ursprungliga studien

- Beräkningsreproduktion – data och kod tillgängliga
- Återskapandereproduktion – data och/eller kod saknas
- Robusthetsreproduktion – alternativa rimliga analytiska beslut på samma data

Olle visade bilder om replikering och reproduktion som en del av forskningsprocessen. Replikering är när man använder nya data medan reproduktion är samma data. Med reproduktion vill man testa att samma resultat ges från samma data.

Hur genomförs beräkningsreproduktion på svenska registerdata idag?

Ad hoc

- Tidigare vanligt att friskriva sig från kravet på reproduktion, nu svårare och svårare
- Ibland genomförs reproduktion av kollegor med tillgång till samma MONA-projekt
- Forskare visar eller filmar sina egna körningar
- Replikatör (t ex doktorand) ges tillfällig tillgång till IFAU-projekt

För att publicera forskningsartiklar kräver tidskrifter allt oftare beräkningsreproduktion. Utlämnande av kod, men inte data godkänns i allt lägre utsträckning än tidigare. Det är nu vanligt att någon i samma MONA-projekt, men som inte jobbat med den aktuella artikeln genomför reproduktionen.

Vad finns det för möjliga lösningar?

Söka etikprövning för separat reproduktions-/replikeringsprojekt

Koppla forskare till SCB med syfte att genomföra reproduktioner

Ge replikatör (annan forskare) tillfällig tillgång till projektet i MONA

Andra idéer?

SCB tog med sig frågorna nedan om vilka möjligheter och krav som finns för att ge en replikatör åtkomst till data och kod:

- förändras SCB:s syn på en replikatörs åtkomst till data när replikering/reproduktion är en del av forskningsprocessen?
- skulle SCB godkänna utlämnande till tidskrift, annat universitet eller doktorand på det egna universitetet om man i etikprövning eller forskningsplan tydliggjort att reproduktion ingår i studien?
- kan användarvillkoren ändras så att det blir tillåtet att visa sin skärm för en replikatör på en tidskrift eller på ett annat universitet?
- kan SCB genomföra reproduktioner?

2. På gång på SCB

Jenny Lindberg, enhetschef, SCB, presenterade nedanstående bilder:

På gång vid SCB

- SCB:s gd Joakim Stymne kommer att avgå vid årsskiftet och regeringen rekryterar nu en efterträdare.
- SCB har tagit fram och beslutat om en [AI-policy](#).
- En uppdaterad strategi för SCB ska vara klar till årsskiftet och samarbetet med SAMmyndigheterna ska utmynna i en ny gemensam målbild.
- I höst uppdaterar SCB sin visuella identitet och förändringarna rullas ut successivt över några månader



Budgetunderlag

Skälen för regeringens förslag

Tabell 6.10 Förändringar av anslagsnivån 2025–2027 för 1:9 Statistiska centralbyrån
Tusentals kronor

	2025	2026	2027
Anvisat 2024 ¹	624 174	624 174	624 174
Pris- och löneomräkning ²	19 964	34 177	46 437
Beslutade, föreslagna och aviserade åtgärder	43 100	48 551	50 455
varav BP25 ³	42 100	45 600	48 100
– Allmänna företagsregistret	21 100	21 100	21 100
– Effektiviserad datainsamling	20 000	23 500	26 000
– Medel för registerdata om integration	1 000	1 000	1 000
Överföring till/från andra anslag			
Förslagberäknat anslag	687 238	706 962	721 066

¹ Statens budget enligt förslagens beslut i december 2023 (bet. 2023/24 FUS10). Beslaget är således exklusive beslut om ändringar i statens budget.

² Pris- och löneomräkningen baseras på anvisade medel 2024. Övriga förändringskomponenter redovisas i löpande priser och inkluderar därmed en pris- och löneomräkning. Pris- och löneomräkningen för 2025–2027 är preliminär.

³ Exklusive pris- och löneomräkningen.



Företagsregistret som SCB tidigare har fått ta ut avgifter för, kommer att bli fritt tillgängligt. SCB har därför fått medel som ersätter intäktsförlusten. SCB har också fått extra medel för att effektivisera datainsamlingen.

Utredningar/beslut kopplat till utlämnanden för forskning

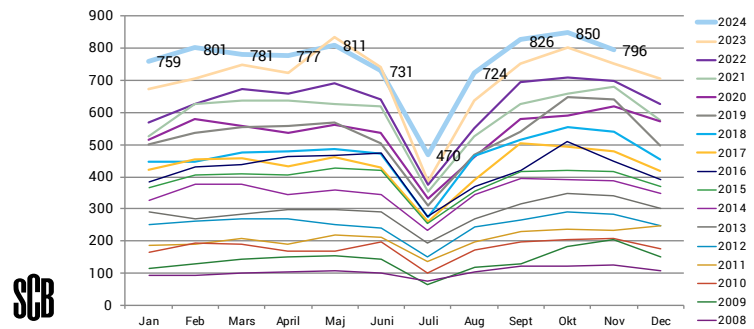
- Lagen (2024:1146) om vissa forskningsdatabaser (utfärdad 28 november 2024)
- Remiss - Ny lag om forskningsetiska krav på och etikprövning av forskning som avser människor
- Remiss - Det Europiska hälsodataområdet – sekundäranvändning och ansvarsfördelning mellan myndigheter
- Remiss - Ett nytt regelverk för hälsodataregister (SOU 2024:57)
- [Uppdrag till Vetenskapsrådet att lämna förslag om samordning och organisatorisk förändring av infrastruktur för forskning - Regeringen.se](#)



Susanna Bylin kompletterade med att VR kommer få en roll om Lagen om vissa forskningsdatabaser. Lärosätena ska skicka ansökningar till departementet, som skickar dessa vidare till VR.

3. Information från SCB om MONA

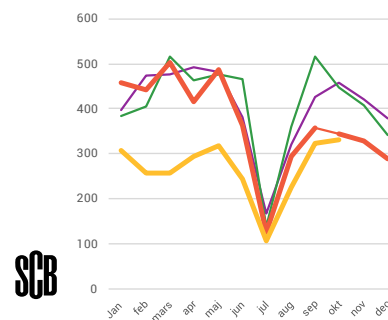
Antal unika inloggade användare, per månad, 2008 – nov 2024



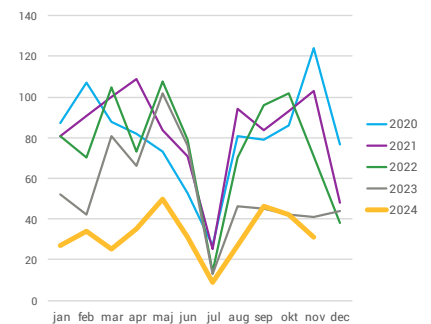
Susanne visade diagram över den ökande användningen av MONA. Nu finns fler än 2 300 användarkonton, Ca 1 590 av dem har använts under 2024. Antalet unika användare per månad överstiger 800 vissa månader.

Aktuellt 2024

Besvarade mejl, 2021- nov 2024

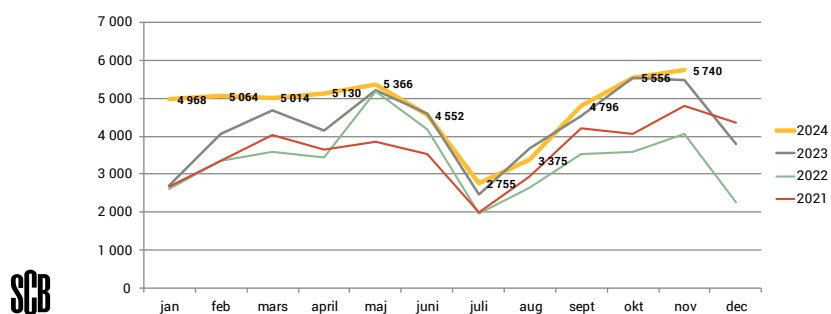


Besvarade samtal, 2021- nov 2024



Det nya inloggningssättet som infördes under våren 2023 har inneburit att färre användare behöver kontakta MONA-supporten för hjälp.

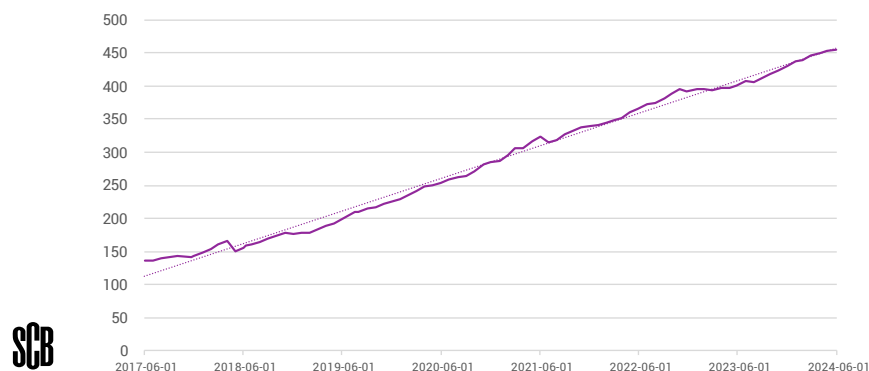
Antal batchjobb, 2021 – nov 2024



Fortsatt hög användning av batch-miljön under 2024 framgår av diagrammet som visades.

Susanne påminde om testet där man med manuella rutiner kan få projektdata överförda till BIANCA-klustret på Uppmax vid Uppsala universitet, där det finns betydligt bättre prestanda. Flera personer har utsetts för test, men än har ingen testat. Om någon i referensgruppen vill testa, kan man anmäla det.

Lagring i TB, juni 2017- maj 2024



Diagrammet över lagring i projektmappar visar en fortsatt ökad lagring. Det är viktigt att var och en löpande raderar data man inte behöver, det är inte bara en fråga om kostnader utan också att mängden data påverkar återställningstiden vid ett haveri.

Aktuellt 2024

- **Metadata**
 - mikrometadata.scb.se
 - > Lansering – uppdatering
 - > Synpunkter
 - RegiStream
- **En extra batchserver**
- MQ02a och MQ02b har utgått och ersatts av ny SQLserver under hösten



Mikrometadata.scb.se innehåller metadata om alla register vid SCB. Sidan lanserades under hösten. RegiStream - ett Stata-paket för att matcha kolumnnamn mot metadata finns nu tillgängligt i MONA. En extra batchserver har lagts till och SQL-servern har bytts ut. Ingen hade upplevt några större problem vid bytet till ny SQL-server.

Gruppen menade att det är bra att ytterligare en batchserver lagts till, men att det är en flaskhals med gränsen på 150GB i minne. Många av jobben som läggs där borde ha kunnat läggas på andra servrar. Alla förstår att det kan vara svårt att välja rätt, men finns det något sätt som kan hjälpa användaren att välja rätt server? MONA-supporten ser samma sak. Användare vet inte hur mycket minne ett jobb kommer ta, och då säkrar man. MONA-supporten ska jobba med att se över hela batch-lösningen, kanske i slutet av nästa år.

Martin Hällsten påpekade att när man jobbar väldigt mycket med BatchClient så har servern "hoppat runt" i gränssnittet, vilket då gör att jobbet läggs på en annan server än man tänkt sig.

Aktuellt 2024

Programvaror

- Nytt avtal om SPSSlicenser är klart
- Nytt avtal om SASlicenser ska upphandlas
- RTools installerades i MONA under hösten
- Uppdatering av Rpaket två ggr per år?
- Uppdatering av Python och paket en gång om året?
- Problem med Jupyter Notebook
- Nytt avtal om e-legitimationer: behövs fler än BankID och TOTP?



SCB har ett nytt avtal om SPSS som därmed kommer att kunna fortsätta användas i MONA. Nytt avtal om SAS-licenser ska upphandlas under 2025. RTools är tillgängligt och R-paket uppdateras två gånger per år. Python planeras uppdateras en gång per år. En säkerhetsincident med Jupyter Notebook har nu blockerats för näst intill samtliga användare och kommer troligtvis att avvecklas. Julia, en programvara som fanns i MONA under en kort period, har efterfrågats. Efterfrågan är i nuläget ändå för liten för att beslut ska fattas om installation. Varje programvara måste förvaltas, därför måste MONA-supporten vara restriktiv i vilka som läggs till.

Aktuellt 2024

- **Införa rollen projektadministratör**
 - En eller flera, per projekt
 - Registrera nya och ta bort befintliga användares behörighet från projekt
 - Ta hand om material från tidigare användares MyDocuments
 - Registrera slutdatum på projekt och användare
 - Följa lagringstillväxten och annan användning
- **Utveckla en ny MONAWeb**
 - Tillgänglighetskrav
 - SCB:s grafiska profil
 - På sv och eng
 - Mobilanpassad



MONA-supporten arbetar med den nya MONAWebben, förutom en allmän uppdatering av design och tillgänglighet (till exempel anpassning för mobil och personer med funktionsvariationer) tillkommer ny funktionalitet för användare, handläggare, MONA-supporten samt för den nya rollen projektadministratörer. Olof berättade hur långt man kommit i arbetet.

4. Synpunkter från referensgruppen

- Martin Hällsten, SU
 - Jag tycker MONA fungerar som vanligt, ganska bra.
 - Det var något som hände efter en uppdatering, att batchservern inte fungerade som den skulle.
 - Detta problem är nu löst, det hängde ihop med den nya servern.
 - Replikering blir en svår fråga, speciellt när man inte kan lämna över till någon. Kan vara en filstruktur som ska köras i en speciell ordning, det kan bli väldigt komplext.
- Oscar Erixson, UU
 - Folk är generellt sett nöjda.
 - Ibland kan det ta tid att lägga upp jobb i batchmiljön
- Helena Nilsson, JU
 - Bra presentation om replikering, som också stämmer med diskussionerna på hemmaplan.
 - Annars är alla väldigt nöjda och glada.
 - Kan man lägga in det som en budgetpost när man söker om pengar till ett forskningsprojekt, att ha med att det ska vara med pengar för replikering för att kunna exempelvis köpa den tjänsten av SCB?
 - Susanne svarade att det är svårt juridiskt om SCB ska ta del av datamaterialet, på samma sätt som om någon annan ska ta del av det.
- Lucas Fischer Madsen, GU
 - Alla är nöjda.
 - Stor eloge till supporten som är lätta att kontakta och snabba att hjälpa.
 - Intresserad av att testa projektadmin-rollen
- Oscar Stålnacke, UmU
 - Saker rullar på bra.
 - Hur länge kan ett projekt ligga i MONA, har hört om 10 år?
 - Susanne svarade att ett projekt som avslutas innebär att ingen kommer åt projektet. Vi har inte satt någon borte gräns, men vi ser gärna att ett åtminstone teoretiskt datum för gallring finns.
- Olle Hammar

- Bra idé att få in replikering i etikprövning respektive forskningsplanen.
- Största problemet är köerna i batchen.
- Susanna Bylin, VR
 - Susanna visade bilder om regeringsuppdraget om digitala infrastrukturer, utredaren Pär Weihed ska slutredovisa uppdraget senast den 15 juni. Viktiga hållpunkter blir en hearing i januari och en andra hearing i mars. Första hearing-tillfället är 14 januari där Susanne deltar från SCB.



Regeringen har gett Vetenskapsrådet i uppdrag att lämna förslag om *hur svensk e-infrastruktur för forskning kan utvecklas genom samordning och organisatorisk förändring*

Susanna beskrev att en e-infrastruktur är en teknisk grundstruktur som har tjänster för exempelvis bearbetning. SUNET, RUT, SND, NAISS och MONA.

5. Övriga frågor

- Olle Hammar erbjöds och tackade ja till att ingå i referensgruppen och bjuds in till kommande möten.
- SCB har fått i uppdrag av Vetenskapsrådet att framställa två av de tre dataset som ingick i utredningen om offentliga mikrodata för utbildningssyfte. Dataseten kommer att levereras första halvåret 2025.

Nästa möte

Nästa möte till våren. Susanne skickar kallelse.

Susanne tackade alla för deras medverkan.