

10. juni, 2013

DeiC Sekretariat
Danmarks Tekniske Universitet
Anker Engelunds Vej 1, Bygning 101A
2800 Kgs. Lyngby

Interesstillkendegivelse vedr. etablering og drift af national HPC-facilitet

Tilkendegivelsen sker med henvisning til det udsendte notat "Organisering af HPC i Danmark fra 2013".

Statsbiblioteket har pt. et driftscenter koncentreret om *indsamling* og *bevaring* af dansk kulturarv. Vi er her specialiseret om en del af den digitale kulturarv, som rummer meget store datamængder og som vokser i stadig acceleration: Pr. dags dato rummer arkivet mere end 1,5 Petabyte fordelt på fx Radio/TV, høstninger af det danske internet, reklamefilm, musik og digitaliserede aviser.

Der har de seneste år været stigende interesse for at få *adgang* til den digitaliserede kulturarv. Lovgivningen rummer mulighed for at give adgang på stedet til forskningsformål, men interessen omfatter naturligt nok netværksadgang, og adgang til at bruge kulturarvsdata i forskningsprojekter under anvendelse af moderne IT-værktøjer. Som illustration for interessen kan nævnes, at Statsbiblioteket pt. deltager i 3 forsknings- og forskningsinfrastrukturprojekter med baggrund i den digitaliserede kulturarv:

- LARM (KU med 9 partnere; bevilling 25 mio. kr.)
- COSOUND (DTU Informatik med over 10 partnere; bevilling 17,5 mio. kr.)
- DIGHUMLAB (Humaniora ved alle DK universiteter; bevilling 30 mio. kr.)

Samtlige projekter er afhængige af adgangen til den digitale kulturarv på Statsbiblioteket. Især anvendelsen i COSOUND bruger analyseværktøjer, som har karakter af HPC.

Adgangen til kulturarven er belagt med skrappe restriktioner, både i forhold til kommercielle rettigheder og med hensyn til persondatalovgivningen.

Som forskningsbibliotek er Statsbiblioteket endvidere involveret i arbejdet med forskningsdata illustreret ved deltagelsen i DeiCs initiativ om Strategi for Datamanagement.

Som bidrag til en national HPC-facilitet vil Statsbiblioteket kunne tilbyde:

1. Workspace for forskere, som ønsker at arbejde med værktøjer, der opererer direkte på Kulturarven. LARM, COSOUND og DIGHUMLAB illustrerer, at der interesse herfor, og forhandlingsforløb med rettighedshavere og myndigheder illustrerer, at en sådan service bedst kan afvikles under Statsbibliotekets driftskontrol. Alternativet vil være kopiering af data til en HPC facilitet, som kan være omkostningskrævende i form af betaling til Copydan for eksemplar fremstilling, eller direkte umulig pga. persondatalovgivning.
2. Datamanagement for forskningsdata som udvikles eller udarbejdes i tilknytning til kulturarven. Selvom disse data ikke vil være pålagt de samme restriktioner som de egentlige kulturarvsdata, vil det være naturligt, at de organiseres i direkte sammenhæng. Muligheden efterspørges fx af partnerne i LARM-projektet og Internet Arkiv forskerne i DIGHUMLAB.
3. Vejledning og rådgivning til forskere, som arbejder med algoritmer for datamining i såvel kulturarv som forskningsdata i almindelighed, herunder i brug af permanente identifikatorer og andre metoder til håndtering af videnskabelige data, inkl. langtidsbevaring.
4. Specialudvikling af forskningsværktøjer især til datamining i kulturarven, men også generelt for videnskabelige data.

Da vi betragter os selv som forskellig fra den traditionelle HPC leverandør, har vi ikke udfoldet denne interessetilkendegivelse til et egentligt forslag med økonomi for etablering og drift, men vil blot udtrykke, at vi ser store synergimuligheder mellem det, som er vores institutionelle ansvar omkring kulturarv og universitetsbibliotek, og det vi kan se som attraktive tilbud til dansk forskning omkring HPC i bred skala.

Vi foreslår, at vi drøfter tilkendegivelsen på vores møde den 21. juni.

Såfremt DeiC ønsker et mere udbygget forslag inden den pågældende dato, imødeser vi gerne en henvendelse herom.

Venlig hilsen

Arne Sørensen

IT-direktør