

Til: Styret for Naturhistorisk museum

Sakstype: Vedtak
Saksnr.: O-1
Journalnr.: 15/294
Møtedato: 26.mars 2015
Notatdato:
Saksansvarlig: Rindal, Lønnve, Rogstad

Sakstittel: Massedigitalisering

Videreføring av massedigitalisering er ett av NHMs satsningsområder som spilles inn til UiOs budsjett for 2016. Frist for innsendingen er fredag 20/3.

Vedlegg

Vedlegg til NHM innspill til store satsninger utenfor rammen - Budsjett UiO 2016



Massedigitalisering

NHM ønsker å videreføre suksessen med massedigitalisering av samlingene. Derfor søker NHM om 4 millioner per år over 5 år (til sammen 20 millioner). De vitenskapelige samlingene på Tøyen utgjør Norges største forskningsinfrastruktur i sitt slag, og ligger til grunn for vår forståelse av liv og livsprosesser. De viktigste vitenskapelige spørsmålene i vår tid dreier seg om livsvitenskap. Massedigitalisering av dette betydelige biologiske arkivet vil gjøre denne verdifulle kilden til kunnskap tilgjengelig, og bidra til at de riktige beslutningene for fremtiden tas. På denne måten imøtekommes målsetningene i Strategi 2020, som setter museet i en unik brobyggerrolle mellom samfunnet og universitetet.

Bakgrunn

Hensikten med vitenskapelige samlinger er at de skal brukes i forskning, men på grunn av manglende dokumentasjon har samlingene i liten grad vært tilgjengelige for andre enn spesialister ved NHM. Digitalisering vil føre til en tilgjengeliggjøring av samlingen for alle forskere internasjonalt og de vil kunne være et bidrag til å styrke livsvitenskapssatsingen og gi økt tverrfaglig samarbeid.

De vitenskapelige samlingene ved NHM består av ca. 6 millioner objekter. Dette er landets desidert største forskningsinfrastruktur, og en ressurs for forskere fra hele verden. For forskingen er det avgjørende at samlingene er riktig kuratert, forsvarlig oppbevart og digitalt tilgjengelig. Som en oppfølging av riksrevisjonens rapport (Dokument 39, 2002-2003) satte universitetet i gang REVITA prosjektet. Dette resulterte i store forbedringer innen kuratering og konservering av objektene. I de senere år har man også løst noen av riksrevisjonens alvorlige anmerkninger angående de fysiske oppbevaringsforholdene ved å innrede nye magasinlokaler på Økern.

Massedigitaliseringsprosjektet startet for å effektivisere digitaliseringsarbeidet. Til dette pionerarbeidet fikk NHM innvilget tilsammen 9 mill. kr. over tre år.

Hva har vi oppnådd

Vi er nå inne i det siste av disse tre årene. Da vi begynte prosessen i 2013 var markedet for massedigitalisering nesten ikke eksisterende. Naturhistorisk museum var helt i fronten internasjonalt innen digitalisering i industriell skala, og antakelig det første museum i verden som pakket ned og sendt samlingene utlands for digitalisering. Den første perioden har vært en suksess og satt Universitet i Oslo på kartet, som en av de mest progressive museumseieren i verden. Det har vært en glede å kunne representere UiO på konferanser og lignende og å legge frem dette prosjektet.

Sommeren 2013 startet NHM prosessen med å sende fra oss materiale til digitalisering. Vår første partner var Digitarium i Joensuu, Finland. Der ble utenlandsherbariet for karplanter (dvs. blomsterplantene og deres slektninger) avfotografert i alt ca. 230 000 objekter. Transkribering etikettinformasjon skulle så bli utført av en underleverandør (DigForsk). Kun 40 % av etikettene ble transkribert av DigForsk, og de nådde dessverre ikke det forventede antall.

Materialet av transkriberte data er i ferd med å legges ut på våre nettsider, se <http://www.gbif.no/datasets/e45c7d91-81c6-4455-86e3-2965a5739b1f/>. Kombinasjonen av tekst og bilde vil gjøre materialet vårt tilgjengelig på en helt ny måte for forskere over hele verden. Dette

vil redusere behovet for utlån og dermed frigjøre ressurser lokalt på NHM. Vi oppretter således et internt handlingsrom som brukes direkte inn mot forskningsstøtte.

Av de opprinnelige massedigitaliseringsmidlene har vi igjen ca. 4,2 mill. for 2015 som vi planlegger å bruke på resten av karplantesamling. Dette er vår desidert største botaniske samling på ca. 700 000 objekter. Ved utgangen av 2015 vil vi ha ca. **1 million** bilder og fått transkribert **alle** karplanteobjekter ved NHM. Dermed vil landets største samling av planter være mobilisert for forskning og forvaltning.

Veien videre

I dag er markedet for digitaliseringstjenester større, og den siste avtalen vi har inngått har derfor betydelig bedre vilkår enn den første. Prisen er nå ca. ¼ det vi betalte Digitalium. Vi inngikk en tjenestekjøpsavtale med nederlandske Picturae etter en anbudsprosess i 2014, og har nå en avtale som kan vare i opptil 10 år.

Vår ambisjon er å få avfotografert og transkribert (til et brukerstyrt minimumsnivå) resten av de botaniske samlingene gjennom tjenestekjøpsavtalen. Av de botaniske samlingene gjenstår:

- * Mosesamlingene ca. 230 000 objekter
- * Lavsamlingene ca. 320 000 objekter
- * Soppsamlingene ca. 350 000 objekter
- * Algesamlingene ca. 10 000 objekter

Resten av samlingene består av ca. 4 millioner zoologiske og geologiske objekter. Disse samlingene er ikke egnet til å bli sendt ut av landet på samme måten. Dette har vi tatt høyde for i planlegging av fasilitetene på Økern. Her blir lokalene utformet på en fleksibel måte som gjør det mulig å sette opp digitaliseringslinjer. Dette aktualiseres i første rekke i forbindelse med flyttingen av fossiler og mineraler fra Brøggers hus i forbindelse med den forestående renoveringen av museumsbygget på Tøyen. Der vil det også la seg gjøre å etablere digitaliseringslinjer for zoologiske samlinger.

Massedigitaliseringen og «Strategi 2020»

Universitetets største satsning i inneværende strategiperiode er «*en tverrfaglig satsing på livsvitenskap (Life Science) i utvidet forstand*» (s. 6 Strategi 2020). Museene er en del av livsvitenskapen og sitter på det som kan være en gullgrube for denne forskningen. UiOs egen strategi på Livsvitenskap trekker frem forskningen på DNA-strekkoder ved NHM som ett eksempel på spennende forskning innen livsvitenskap ved UiO (s.3 UiOs Strategi for Livsvitenskap). Dette er en forskning som utnytter den unike kombinasjonen av samlinger og et aktivt forskningsmiljø.

Strategi 2020 uttrykker klart at «*Museenes omfattende samlinger skal utnyttes bedre både i forskning og undervisning*» (s. 7 Strategi 2020). Dette er bare mulig hvis informasjonen fra samlingene er lett tilgjengelig, og massedigitalisering gjør dette mulig via web-innsynsløsninger som både gir tilgang til digitaliserte data, og til «rå-dataene» via bilder. Ved å gjøre samlingene tilgjengelig via webløsninger vil også samfunnet, spesielt igjennom forvaltningen, få tilgang til museets samlinger. Dermed vil man bidra til å gjøre museene til en «*brobygger mellom samfunnet og universitetet*» (s. 7 Strategi 2020).

NHM har sammen med Kulturhistorisk museum tatt ett initiativ til å anskaffe felles forskningsinfrastruktur for å mobilisere samlingen gjennom digitalisering. Denne tverrfaglige satsingen hadde aldri kommet i stand uten Massedigitaliseringsprosjektet og dets positive ringvirkninger.

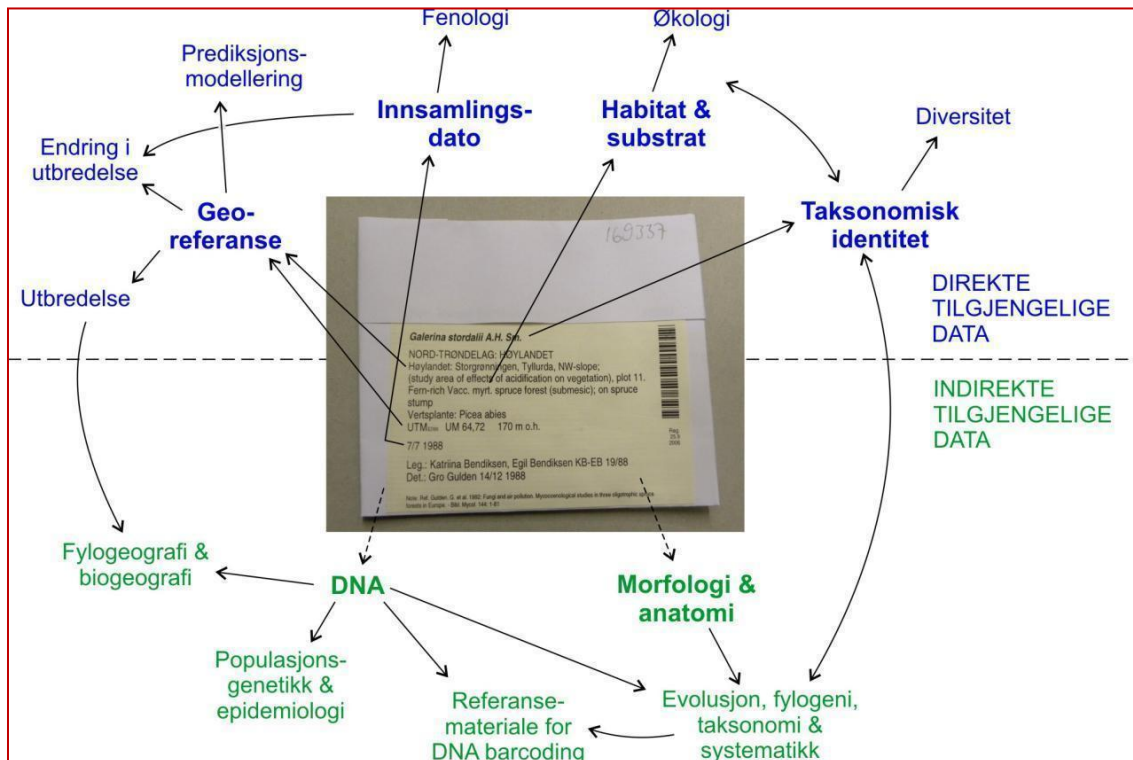
Vi søker derfor om en 5-års periode med massedigitalisering med 4 millioner NOK per år.

Budsjett for massedigitaliseringen

2016	2017	2018	2019	2020
4 000 000 NOK	4 000 000 NOK	4 000 000 NOK	4 000 000 NOK	4 000 000 NOK

Sammenhengen mellom digitalisering, forskning og livsvitenskap

Med å digitalisere samlingene kan man forstå flere ting, men det grunnleggende er å fange data assosiert med ett museum objekt og oppbevare disse dataene digitalt i en database e.l.



Figuren viser hvilke data som er tilgjengelig fra et objekt og hvordan dette kan utnyttes i forskning.

I starten var det viktig å skrive av det som stod på etiketten, ord for ord. Dette er det vi kaller direkte tilgjengelige data og kan gi grunnlag for mange forskningsområder:

- **Georeferanse** (koordinatfestete lokalitetsdata) kan gi grunnlag for forskning innen **endringer i utbredelse** f.eks. i forbindelse med klimastudier, **prediksjonsmodellering**
- **Innsamlingsdato**, gir oss mulighet til å studere fenologi, dvs. når på året organismer opptrer og hvordan dette endrer seg i respons på f. eks. klimavariabel
- **Habitat & Substrat** er viktige for å forstå økosystemer og hvordan menneskelig aktivitet kan påvirke naturen.

- **Taksonomisk identitet** er basis for alle studier av organismer, og en database med mye taksonomisk informasjon vil være en gullgrube for diversitetsstudier.

Denne type data er selvfølgelig fremdeles viktig, men utviklingen har også ført med seg andre behov. Dette er det vi kaller indirekte data, som kan fås fra objektene ved nærmere undersøkelser. Det er nesten uendelig mange forskjellige typer indirekte data som kan digitaliseres, men de viktigste er:

- **Morfologi & anatomi som** gir grunnlag for sammenliknende studier inne evolusjon, fylogeni, taksonomi og systematikk. Kjerneområdene i klassisk museal forskning.
- **DNA** har åpnet portene til en helt ny verden med muligheter innen biologisk forskning. Museene sitter på enorme samlinger av biologisk materiale, som hvis det blir mobilisert og tilgjengelige gjennom digitaliseringsprosesser kan gi opphav til banebrytende forskning.
- **Kjemiske analyser** kan gi svar på sammensetning av mineraler, forståelse av prosesser i organismer og deres celler.

Disse dataene kan gi opphav til **tidsserieanalyser** der man ser på endringer over tid. Dette gir oss muligheten til å se på tverrfaglige problemstillinger som effekter av menneskelig påvirkning på økosystemer, klimaendringer og responser på disse osv.