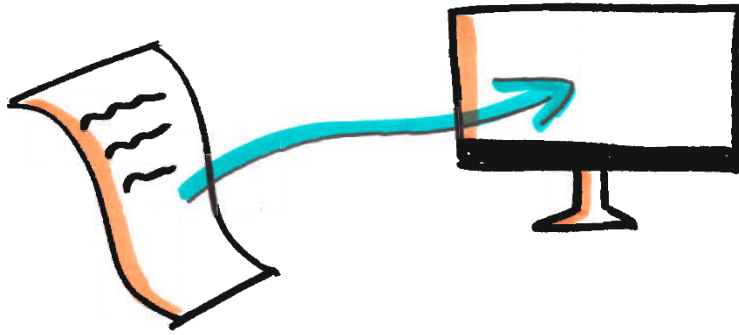


08.06.2021

NHMs digitaliseringsstrategi 2021-2025





Innhold

NHMs digitaliseringsstrategi 2021-2025.....	1
Samlingene som grunnmur.....	3
Hvorfor digitalisere?	3
Mål for digitaliseringsarbeidet og tiltak.....	4
Bakgrunn	6
Definisjon av digitalisering	6
Føringer	6
Internasjonalt samarbeid.....	7
Digitalisering ved NHM i dag og framover.....	7
Prioritering	7
Datastandarder	7
Digitaliseringsnivå	8
Tilgjengeliggjøring.....	9
Verktøy, metoder og prosesser	9
Status for digitaliseringen av samlingene ved NHM – hva gjenstår	11
Historikk	13
Digitalisering og databaseløsninger.....	13
Deling av data	14
Konklusjon.....	15

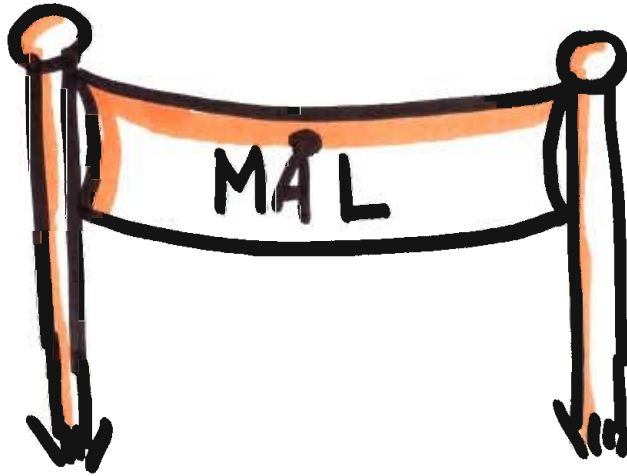
Samlingene som grunnmur

De naturhistoriske samlingene utgjør en vitenskapelig dokumentasjon innen systematikk, er et magasin for videre forskning, og et uerstattelig arkiv over norsk og internasjonal natur. Våre samlinger er derfor grunnmuren i museets virksomhet. NHM skal sikre, bevare, og gjøre samlingene tilgjengelig for forskning, formidling, undervisning og forvaltning.

Hvorfor digitalisere?

Verden står ovenfor enorme utfordringer knyttet til endringer i klima og tap av naturmangfold. Ett ledd i å møte disse utfordringene er å tilby kvalitetssikrede data slik at man kan forske på, undervise i, formidle og forvalte naturens mangfold. NHMs samlingsdata må derfor kunne brukes til både tradisjonelle, moderne og framtidige forsknings- og overvåkningsmetoder. Gjennom tilgjengeliggjøring av mer og flere typer data muliggjør digitaliserte samlinger slik forskning, undervisning og formidling, og dermed bærekraftig forvaltning av naturmangfold. NHMs viktigste mål for digitalisering er derfor bedre digital oversikt over og bedre tilgang til museumssamlingene.

Digitalisering av samlingene vil gjøre samlingsforvaltningen mer effektiv på mange områder. For det første legges det til rette for forskningsaktiviteter gjennom skreddersydde verktøy og rutiner for fangst av forskningsdata. For det andre muliggjør digitaliserte samlinger utvikling av digitale verktøy som støtter og effektiviserer arbeidsprosessene. Eksempler er digital dataflyt fra felt til samlingsdatabase og elektronisk behandling av låneforespørsler. Slike verktøy vil kunne spare de museumsansatte for mye arbeid og øke kvaliteten på samlingsdataene. For det tredje blir det lettere for museets ledelse å følge med på utviklingen av samlingsaktiviteten, og de kan ta avgjørelser om ressursbruk i samlingene på et bedre faktagrunnlag.



Mål for digitaliseringsarbeidet og tiltak

Effektmål 1: Bedre digital oversikt over og tilgang til museumssamlingene

Resultatmål 1.1: Alle samlinger er digitalisert til det nivået deres kurateringsnivå¹ tilsier

Tiltak:

- Registrere alle NHMs samlinger i GBIF Registry of Scientific Collections²
- Skaffe oversikt over digitaliseringsstatus per delsamling
- Med utgangspunkt i Samlingsplanen 2030, beskrive til hvilket digitaliseringsnivå og hvordan hver delsamling skal digitaliseres
- Digitalisere delsamlinger til aktuelt nivå, i prioritert rekkefølge
- Utvikle rutiner for digitalisering som følger FAIR-prinsippene (se [Datastandarder](#))

Resultatmål 1.2: Alle digitaliserte samlingsdata er offentlig tilgjengelige via internett

Tiltak:

- Bedre kvaliteten på de data som NHM deler til Artskart³, GBIF⁴ og DISSCo⁵, f.eks. bedre georeferering, synonymisering av personnavn og øke bruken av UUID'er
- Videreutvikle og synliggjøre egne nettportaler
- Tilgjengeliggjøre samlinger som ikke deles i dag, f.eks. paleontologiske og geologiske data
- Tilgjengeliggjøre mer data fra samlingene det deles data fra i dag, f.eks. analyseresultater, revisjonshistorikk og koordinatrevisjoner

¹ Se revidert samlingsplan, ferdigstilles 2021

² <https://www.gbif.org/grscicoll>

³ <https://artskart.artsdatabanken.no/>

⁴ <https://www.gbif.org/>

⁵ <https://www.dissco.eu>

Effektmål 2: Mer effektiv samlingsdrift

Resultatmål 2.1: Flere arbeidsprosesser i samlingene er digitaliserte og (halv-)automatiserte

Tiltak:

- Kartlegge arbeidsprosesser som kan digitaliseres og/eller automatiseres (se MUSITs virksomhetsanalyse 2014)
- Lage verktøy for de prosesser der det er hensiktsmessig, f.eks. digital dataflyt fra felt til samlingsdatabaser
- Sikte mot samlingsdatabaser som støtter opp under arbeidsprosessene

Resultatmål 2.2: Alle samlingsdatabaser følger internasjonale standarder for digital informasjon

Tiltak:

- Avgjøre hvilke metadatastandarder NHM i dag skal følge for deling av data (for eksempel ABCD, DarwinCore)
- Innføre aktuelle standarder

Resultatmål 2.3: Sikkert driftsregime for samlingsdatabasene

Tiltak:

- Utarbeide rutiner for arkivering og langtidslagring/backup
- Lage migreringsstrategier for fremtidige overganger til nye databasesystemer
- Beslutte hvilke Collection Management Systems vi skal ta i bruk i fremtiden
- Gjennomføre øvelser for gjenoppretting av data fra backup

Effektmål 3: Økt tilrettelegging for forskning på samlinger

Resultatmål 3.1: Alle relevante forskningsdata er knyttet til sine samlingsobjekter

Tiltak:

- Kartlegge hvilke forskningsdata det må anskaffes eller utvikles løsninger for, som det ikke finnes eksisterende løsninger for fra før (f.eks. LIMS)
- Anskaffe og ta i bruk verktøy og system(er) for lagring av disse forskningsdataene
- Innføre rutiner for lagring og arkivering av forskningsdata

Resultatmål 3.2: Enklere å legge forskningsmateriale inn i samlingene

Tiltak:

- Tilrettelegge for at alle produserte forskningsdata blir digitalisert til høyeste nivå (MIDS nivå 3, se [Digitaliseringsnivå](#))

Resultatmål 3.3: Samlingene er mer tilgjengelige for forskning

Tiltak:

- Etablere rutiner for å kontinuerlig fange opp hvilke data forskerne ønsker digitalisert
- Utarbeide rutiner for å digitalisere data forskerne etterspør
- Tilrettelegge for kobling av våre data med eksterne databaser

Bakgrunn

Definisjon av digitalisering

Staten definerer digitalisering som «en samlebetegnelse for overgangen fra analoge, mekaniske og papirbaserte løsninger, prosesser og systemer, til elektroniske og digitale løsninger. Begrepet digitalisering rommer derfor også etablering av nye IT-systemer som opprettholder rutiner som utføres manuelt, selv om noen av de største besparelsene ved digitalisering ofte kommer når rutiner ikke bare blir digitale, men også kan automatiseres.»⁶ Denne definisjonen avviker fra det man tradisjonelt har oppfattet som digitalisering ved NHM, spesielt i forbindelse med samlinger, der fokuset har vært på avskrift av etikettinformasjon og innlegging av dette i en samlingsdatabase. En mer moderne definisjon omfatter også de gevinster man kan få ved digitale arbeidsprosesser.

Digitalisering ved NHM i 2020 defineres derfor som

- 1) Registrering av opplysninger om objekter i samlingene våre i digital form (inkludert ev. bilder av objektene)
- 2) Håndtering og utnytting av de resulterende digitale dataene ved hjelp av digitale verktøy (databaser o.a.)
- 3) Tilgjengeliggjøring av de digitale dataene for samfunnet
- 4) Digitalisering av manuelle/analoge arbeidsprosesser slik at de automatiseres eller effektiviseres

Føringer

Statens digitaliseringsstrategi er nedfelt i Regjeringens digitaliseringsstrategi⁷ for offentlig sektor 2019-2025 og de årlige digitaliseringsrundskrivene⁸, det siste fra 05.01.2021. Relevante anbefalinger for oss fra disse dokumentene er å sørge for at data skal kunne gjøres tilgjengelig i et langtidsperspektiv med opprettholdt integritet, autensitet, anvendbarhet og pålitelighet, og å legge til rette for gjenbruk og videre bruk av informasjon. Sagt på en annen måte betyr dette at vi må sørge for at det ikke kommer feil inn i dataene over tid, at de ikke endres, og at det er tydelig hvilken kontekst man kan bruke dem i.

Nasjonal strategi for tilgjengeliggjøring og deling av forskningsdata sier at «Forskning som skjer ved bruk av offentlige midler, skal være til det beste for alle. Derfor er det viktig at også dataene bak forskningsresultatene er tilgjengelige for flest mulig: for andre forskere, men også for forvaltningen og næringslivet»⁹. Naturhistorisk museums strategi 2030 følger statens føringer; våre samlinger «skal være digitalisert og digitalt tilgjengelige for forskning og allmennheten».

Oppsummert sier føringene NHM må forholde seg til at man skal digitalisere samlings- og forskningsdata med høy pålitelighet og høy kvalitet, som lagres sikkert for lang tid, og tilgjengeliggjøres for omverdenen.

⁶ <https://www.regjeringen.no/no/tema/statlig-forvaltning/ikt-politikk/digitaliseringen-i-offentlig-sektor/id2340245/>

⁷ <https://www.regjeringen.no/no/dokumenter/en-digital-offentlig-sektor/id2653874/>

⁸ <https://www.regjeringen.no/no/dokumenter/digitaliseringsrundskrivet/id2826781/>

⁹ <https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-tilgjengeliggjoring-og-deling-av-forskningsdata/id2582412/>

Internasjonalt samarbeid

Det er flere europeiske samarbeidsprosjekter som omhandler samlinger og digitalisering. Det største av disse er DiSSCo, «Distributed System of Scientific Collections»¹⁰, som skal være en europeisk infrastruktur for naturhistoriske samlinger. Målet er å endre det fragmenterte landskapet av viktige vitenskapelige ressurser til én sømløs samling med enhetlig tilgang til data og tjenester for alle brukerne. Et av de viktigste bidragene fra DiSSCo vil være berikelsen av samlingsobjektene med flere datatyper fra andre kilder, f.eks. gendata. NHM representerer de norske museumssamlingene som partner i DiSSCo Prepare (ESFRI preparatory phase).

Et annet viktig satsingsområde er «European Loans and Visits System»¹¹ (ELViS) som utvikles av SYNTHESYS+ sammen med DiSSCo, og skal gjøre tilgangen til samlinger enklere, raskere og bedre.

Det avsluttende EU-prosjektet ICEDIG¹², «Innovation and consolidation for large scale digitisation of natural heritage», produserte standarder og arbeidsflyter for storskala digitalisering. Det har blitt produsert flere gode leveranser her som NHM kan bruke i sitt digitaliseringsarbeid.

Det amerikanske prosjektet iDigBio¹³ har som mål å gjøre samlingene i USA digitaliserte og tilgjengelige på nett. NHM har hatt kunne nytte erfaringene og den kunnskapen dette prosjektet har produsert gjennom flere år.

Digitalisering ved NHM i dag og framover

Prioritering

Delsamlingene som skal digitaliseres de neste fem årene har et gap mellom ønsket digitaliseringsnivå og faktisk digitaliseringsnivå i dag. Ønsket digitaliseringsnivå følger av REVITA planen og Samlingsplanen 2030. Ønsket digitaliseringsnivå tar utgangspunkt i NHMs minimumsstandard for digitalisering som i store trekk følger MIDS (se [Digitaliseringsnivå](#)).

Tilvekst skal prioriteres over etterslep slik at vi ikke bygger opp nye restanser. Tilveksten skal digitaliseres fortløpende, og senest 12 måneder etter inntekt.

Datastandarder

For å kunne utnytte potensialet som ligger i dataene tilknyttet samlinger må de foreligge i en form som muliggjør utveksling mellom datatilbydere og brukere.

Darwin Core er en datastandard for biologisk materiale og er en utvidelse av den EU-anbefalte standarden Dublin Core. Standarden er utviklet for å forenkle deling av informasjon om biologisk mangfold ved å tilby identifikatorer, «merkelapper» (labels) og definisjoner. Darwin Core er først og fremst basert på taksa og deres forekomst i naturen dokumentert av observasjoner, objekter og prøver, og data relaterte til disse objektene. Darwin Core brukes av for eksempel GBIF og Artskart og utvikles og

¹⁰ <https://www.dissco.eu>

¹¹ <https://elvis.dissco.eu>

¹² <https://icedig.eu>

¹³ <https://www.idigbio.org>

vedlikeholdes av Biodiversity Information Standards (TDWG), et internasjonalt konsortium for datastandarder i forbindelse med naturhistoriske samlinger.

En annen og mer omfattende standard som TDWG utvikler er ABCD standarden (Access to Biological Collection Databases). Denne inneholder også felter for geologiske og paleontologiske samlingsdata i tillegg ABCD-efg (Extended for Geosciences)¹⁴.

NHM følger EUs anbefalinger for hvordan data skal lagres og formidles¹⁵. For å sikre langtidsbevaring av data må filene være på formater som sikrer at de kan leses av datamaskiner i fremtiden, og museet må ha rutiner for sikkerhetskopiering (backup) og langtidslagring.

Forsknings- og samlingsdata skal forvaltes etter FAIR-prinsippene¹⁶. Dette er i tråd med regjeringens strategi, Forskningsrådets policy og EUs krav. For å sikre at forskningsdataene kan deles og gjenbrukes må både data og metadata være søkbare (Findable), tilgjengelige (Accessible), mulige å integrere med andre data (Interoperable) og gjenbrukbare (Reusable). FAIR-prinsippene skal være en rettesnor for digitaliseringsarbeidet ved NHM, og skal gjenspeiles i NHMs rutiner for digitalisering.

Digitaliseringsnivå

CETAF, "The Consortium of European Taxonomic Facilities" i samarbeid med TDWG, utarbeider en ny standard for digitaliseringsnivå av naturhistoriske samlinger; MIDS (Minimum Information about a Digital Specimen)¹⁷, som opererer med tre nivåer av data:

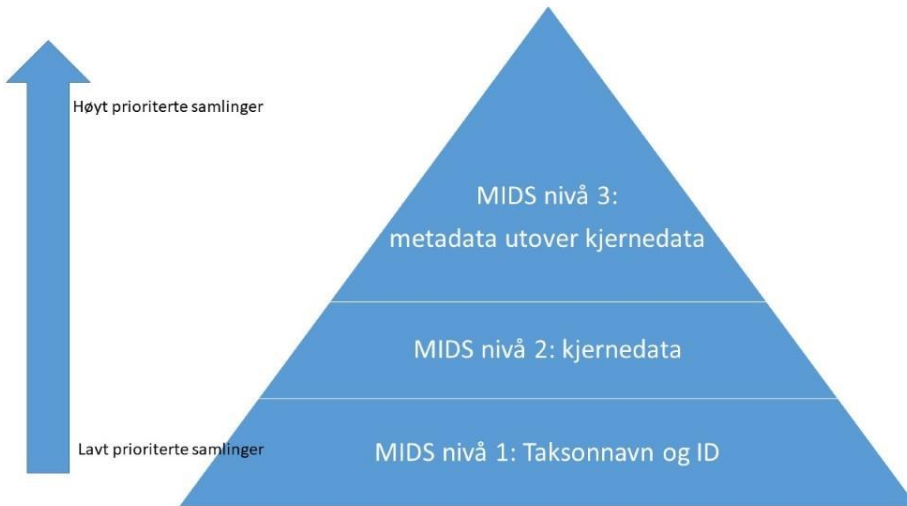
1. Grunnleggende data, typisk taksonnavn og en «persistent identifier», og teknisk info rundt opprettelsen av den digitale posten.
 2. Etikettinformasjon; inkludert det vi kaller kjernedata: **Hva** er det vi har, **Hvor** er det samlet inn, **Når** er det samlet inn / er det fra, **Hvem** har samlet det inn.
 3. Utvidet informasjon, for eksempel bestemmelsehistorikk, analyseresultater og andre annoteringer
- Bilder av objekter kommer inn under nivå 1 eller 2.

¹⁴ <http://www.geocase.eu/efg>

¹⁵ <https://www.succeed-project.eu/outputs>

¹⁶ <https://www.go-fair.org/fair-principles/>

¹⁷ <https://www.tdwg.org/community/cd/mids/>



Tilgjengeliggjøring

NHMs grunnleggende prinsipp er at alle data skal deles så sant det er mulig. Unntak fra dette prinsippet kan begrunnes i at det koster for mye ressurser å dele dataene, tilstanden til dataene er av for dårlig kvalitet, det pågår forskning på dataene, dataene er konfidensielle eller personvern-relaterte, eller dataene avslører sårbare funnlokalteter.

Data som holdes tilbake midlertidig skal ha en tidsfrist for tilgjengeliggjøring. Om mulig skal overordnede data om de aktuelle objektene tilgjengeliggjøres.

NHMs data skal deles i alle relevante kanaler og spesielt skal NHMs samlingsportal¹⁸ videreutvikles. NHM har via sin egen portal bedre kontroll på dataene som deles fra NHMs samlinger: vi kan dele mer enn i andre portaler, vi unngår eventuelle filtre andre aktører legger på dataene, og vi unngår i mange tilfeller duplisering av data i samme portal. I tillegg promoterer museet sterkere gjennom sin egen portal.

Verktøy, metoder og prosesser

Databaser

NHM skal ha samlingsforvaltningssystemer (Collection Management Systems - CMS) som dekker NHMs behov for digital samlingsforvaltning og tilgjengeliggjøring av data, støtter arbeidsprosessene våre og sikrer datakvalitet og –integritet, uavhengig av enkeltpersoner. Dataene skal være enkle å migrere og det skal foreligge strategi for migrering til neste løsning. Antall CMSer skal reduseres så langt det lar seg gjøre, og CMSene skal være verktøy for god forskningsinfrastruktur.

Digitaliseringsmetoder

I dag digitaliseres data i massedigitaliseringsprosjekter og som del av daglig drift av samlingene. Massedigitalisering gjøres eksternt eller internt på dugnad. Begge varianter krever både en del forarbeid

¹⁸ <https://samlingsportal.nhm.uio.no/museum/nhm>

(logistikk) og etterarbeid som kvalitetssikring og endelig ferdigstilling som bl.a. større import av data. Dette vil vi fortsette med framover.

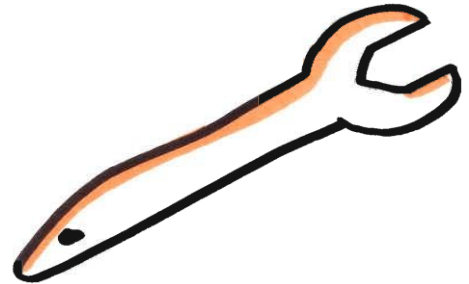
Andre metoder å vurdere er Crowdsourcing og Optical Character Recognition (OCR)-skanning. Det er gjort noen mindre prosjekter på Crowdsourcing ved NHM, med "Østafrikanske lav" og "karplanter fra Telemark" gjort av Telemark Botaniske Forening, og "fossiler fra polarsamlingen" utført på en WeDigBio arrangement. Crowdsourcing har potensiale, men man trenger et stabilt verktøy, samt et system for god tilbakemelding til publikum. Skal NHM utvikle dette til en ressurs for museet må vi ha en til to ansatte som har oppfølging av publikum og Crowdsourcing -prosjektet som en av sine hovedoppgaver.

OCR-skanning og tolking av etiketter består i at en programvare gjenkjenner bilder av bokstaver og tolker dem som tekst. Dette er i liten grad utprøvd ved NHM så langt. Utfordringen har vært at samlingene består av svært mange håndskrevne etiketter. For disse er ikke OCR-skanning god nok i dag, men det skjer en hurtig utvikling på feltet ved bruk av kunstig intelligens (AI) så om noen år er det kanskje mulig å lese disse etikettene også¹⁹. OCR egner seg imidlertid godt for maskinskrevne kartotek kort, spesielt de med mye tekst.

Digitale verktøy for arbeidsprosesser

Hittil har det ved NHM blitt utviklet software for eksempel for håndtering av bilder ved fotografering av objekter, og prøvetaking til DNA-strekkoding. Andre prosesser det er relevant å utvikle digitale verktøy for er:

- sjekklister i forbindelse med rutiner (feltarbeid, lån, tilgjengeliggjøring, tilstandsvurdering osv.)
- e-læring av nye rutiner
- dataflyt fra innkøst/felt til database, for interne og eksterne leverandører av materiale
- registrering av tilhørende dokumentasjon/metadata i database (foto, GIS-data, kart, tegninger, analyseresultater osv.)
- behandling av tilbakemeldinger angående NHMs samlingsdata fra omverdenen



Verktøyene som skal anskaffes eller utvikles er for eksempel Laboratory Information Management Systems (LIMS), Electronic Lab Notebook (ELN), Collection Management Systems (CMS) med moduler for analyseresultater og konservering, og applikasjoner knyttet til den eksisterende nettportalen.

Dataflyt mellom databaser og portaler skal ha høyt fokus: data til og fra f.eks. Samlingsportalen, Datadumper, Bionomina.net, Artskart, og mellom egne databaser. Hjelpemidler for dataflyt vil være for eksempel Wikidata, ORCID og Dataset DOI, og taksonregistre som ANB²⁰, CoL²¹ osv.

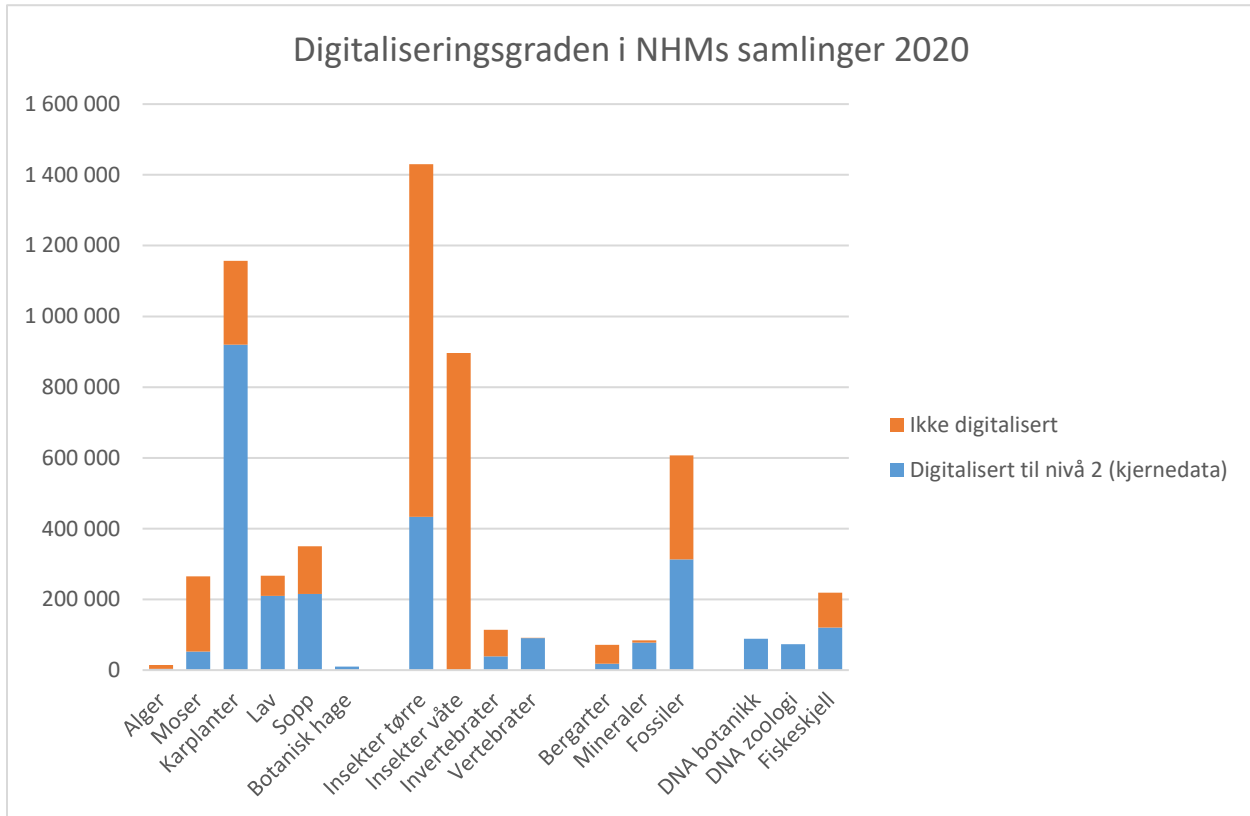
¹⁹ https://icedig.eu/sites/default/files/deliverable_d4.1_icedig_methods_for_automated_text_digitisation.pdf

²⁰ <http://www2.artsdatabanken.no/artsnavn/Contentpages/Sok.aspx>

²¹ <http://www.catalogueoflife.org>

Status for digitaliseringen av samlingene ved NHM – hva gjenstår

NHM har ca. 5,7 mill. objekter i sine samlinger, hvor ca. 46 % er digitalisert til minst nivå 2 (se [Digitaliseringsnivå](#)), og ca. 22 % av objektene er avfotografert (Tabell 1). En detaljert statistikk over samlingene og digitaliseringen føres årlig og legges ut på museets Wiki²². Digitaliseringsgraden varierer fra samling til samling.



²² <https://wiki.uio.no/nhm/skf/best-practices/index.php/Samlingstall>

Tabell 1. Antall objekter digitalisert (til MIDS nivå 2 – kjernedata, og i database), på nett (GBIF eller nettportal) og med foto på nett i NHMs ulike samlinger.

	Antall objekter	Andel digitalisert	På nett	Fotografert
BOTANISKE OG MYKOLOGISKE SAMLINGER				
Alger	16 839	35%	35%	0
Moser	265 208	20%	20%	0
Karplanter	1 156 730	80%	75%	63%
Lav	266 708	79%	78%	40%
Sopp	350 331	61%	61%	1%
Botanisk hage	9 718	100%	100%	24%
ZOOLOGISKE SAMLINGER				
Insekter tørre	1 430 489	30%	24%	0
Insekter våte	896 800	0%	0%	0
Evertebrater	113 695	34%	0%	0,4%
Vertebrater	91 560	99%	99%	17%
GEOLOGISKE OG PALEONTOLOGISKE SAMLINGER				
Bergarter	71 500	25%	0%	0
Mineraler	84 223	93%	0%	0
Fossiler	606 935	52%	0%	1%
DNA-BANKEN				
Botaniske samlinger	89 010	100%	100%	Ikke aktuelt
Zoologiske samlinger	73 024	100%	100%	Ikke aktuelt
Fiskeskjell	219 000	45%	45%	Ikke aktuelt
OPPSUMMERT				
Botaniske samlinger	2 063 057	68%	66%	36%
Zoologiske samlinger	2 532 544	22%	19%	1%
Geo og pal samlinger	762 658	54%	0	1%
DNA-banken	381 034	74%	74%	Ikke aktuelt

Historikk

Digitalisering og databaseløsninger

NHM startet tidlig med digitalisering av sine samlinger og allerede på 80-tallet var det flere initiativ i gang. Først ut var de paleontologiske samlingene som under Hans Arne Nakrems ledelse begynte innleggelsen av «typesamlingen» i en dBase III+ database i 1984. Fuglesamlingen fikk digitalisert den norske skinnsamlingen i perioden 1984 til 1985, dataene ble registrert av Vidar Bakken fra kartotek kort i en Access-database i et system fra Norsk Data. Deretter fulgte entomologene gjennom det som ble kjent som LepArb, et initiativ med utspring i Norsk Entomologisk forening for å registrere forekomsten av norske sommerfugler i 1986. Dette arbeidet gikk inn i NHMs entomologi-databaser ved årtusenskiftet. Botanikerne startet arbeidet med digitaliseringen i 1990 for karplantene, og lav kom deretter i 1994.

I 1994 startet «Universitetsmuseenes naturhistoriske dokumentasjonsprosjekt» (UNADOK) med digitalisering av samlingene ved hjelp av arbeidssøkende og med midlertidige ansatte som prosjektmedarbeidere og kvalitetskontroll.

I 1998 avløste «Museumsprosjektet» (Muspro), som dekket kulturhistorie og naturhistorie, UNADOK, men mye etter samme modell som UNADOK. Muspro var et felles IT-prosjekt for de norske universitetsmuseene, og hadde som hovedmål å gi en samlet og rasjonell tilgang til informasjon i universitetsmuseene. NHM meldte seg ut av Muspro-samarbeidet ved årsskiftet 2003/2004, og KHM overtok NHMs fagkonsulentmidler. Muspro ble avsluttet i 2006, men NHM hadde fremdeles mye ugjort på digitaliseringsfronten. Muspro ble erstattet av organisasjonen MUSIT hvis hovedoppgave var å utvikle og drifte CMS for universitets museene i Norge.

Et stort og viktig skritt for museene kom med Riksrevisjonens undersøkelse av bevaring og sikring av samlingene ved fem statlige museer i 2002/2003²³. Dette resulterte i REVITA-prosjektet med formål om å revitalisere samlingene ved de norske universitetsmuseene. NHM fulgte opp Riksrevisjonen og laget en Revitaplan for 2004-2012.

I 2005 ble det ved NHM undersøkt mulige programvareløsninger for samlingsdatabaser ved Naturhistorisk Museum. Konklusjonen ble en databaseløsning utviklet av USIT (MUSIT). I NHMs samlingsplan av 2011²⁴ presiseres det at antall forskjellige databaseløsninger skal reduseres, og flest mulig av samlingene samles på felles plattform. Dette medførte at man i MUSIT målrettet gikk inn for å utvikle baser for nye fagområder. Arbeidet ble aldri fullført og antallet databaser fra 2011 til 2021 ble ikke redusert i særlig grad. Det var også en ganske stor motstand lokalt for å ta i bruk nye databaser.

I 2015 startet prosjektet ny IT-arkitektur for MUSIT. Den nye versjonen skulle være nettbasert og bestå av flere moduler; bl.a. analysemodul, magasinmodul, konserveringsmodul mm. Versjonen skulle gjelde for kultur og natur. I 2020 var ovenfornevnte moduler utviklet, og en objektmodul for marine evertebrater var 80% ferdig.

²³ https://www.stortinget.no/globalassets/pdf/dokumentserien/2002-2003/dok_3_9_2002_2003.pdf

²⁴ <https://www.nhm.uio.no/forskning/publikasjoner/nhm-rapporter/nhm-rapport-018-2011.pdf>

I 2019 ble MUSIT-organisasjonen på bestilling fra eierne evaluert av et eksternt firma. Rapporten avdekket kritikkverdige forhold på både bestiller- (museene) og tilbydersiden (USIT). Som en følge av de store utfordringene som ble avdekket ble MUSIT i sin daværende form avvirket i februar 2021. Nye avtaler som avløste MUSIT-samarbeidet består av 1) Videreutvikling av MUSIT (ny arkitektur) for kulturhistorie ved USIT, og 2) drift av de gamle MUSIT-basene for naturhistorie i ca. tre år, inntil ny løsning er på plass.

MUSIT som databaseløsning må erstattes for naturhistoriske samlinger ved alle Universitetsmuseene. Der de andre museene ønsker å benytte samme CMS og samarbeide om dette, har NHM stilt seg noe avventende. Vitenskapsmuseet/NTNU har i april 2021 satt i gang et prosjekt som skal vurdere og teste en håndfull eksisterende databaseløsninger (kommersielle) og innen slutten av året anbefale en løsning. NHM støtter prosjektet og deltar i prosjektgruppa, styringsgruppe og brukergruppe.

Digitalisering av data har generelt stort fokus i samlingsplanen 2011. I prinsippet er målsettingen i planen at alle samlinger skal digitaliseres på et eller annet nivå (fra oversikt til mer detaljert). Det pekes på et stort etterslep av digitalisering av data som krever 243 årsverk for å avvikle. Siden den gang har etterslepet blitt betydelig redusert gjennom massedigitaliseringsprosjekter, hovedsakelig på botanikk.

UiO sentralt bevilget NHM 9 millioner øremerket til massedigitalisering i 2013 fordelt på 3 mill. i perioden 2013-2015.

NHM hadde først kontrakt med Digitalium, Finland, om fotografering av herbarieark fra karplante-generalherbariet. Karplanter var første gruppe ut fordi disse er montert på herbarieark, som er enklere å fotografere enn objekter i andre samlinger. På grunnlag av bildene ble etiketter transkribert av underleverandøren Digforsk i Kautokeino. I neste runde med massedigitalisering kjøpte NHM tjenester fra Picturae, Nederland; her ble karplanter-nordisk samling sendt. I tredje runde ble kapselsamlingene av mose og lav sendt til fotografering hos Picturae. For Picturae-prosjektene ble deler av etikettinformasjonen transkribert av en underleverandør av Picturae, i Alembo i Surinam.

Intern massedigitalisering har skjedd i mindre skala og er blitt utført for fugl, sopp og paleontologi på dugnadsbasis. Slike dugnader krever en del personalressurser i planlegging og organisering, foruten deltakelse i selve dugnaden.

Deling av data

Den eldste samlingsportalen fra NHM var for lavherbariet (NLD, Norsk LavDatabase) hvor man kunne søke i lavsamlingens database. Her ble det etter hvert hentet inn data om all lav som var samlet på norsk jord. Einar Timdal samlet inn datasett fra flere steder i verden. Den ble lagt ut på nett første gang i 1997. Dette ble senere kopiert for sopp, moser og karplanter (delvis) for norsk materiale.

De paleontologiske databasene var søkbare på nettet mellom 2000 og 2012. Hans Arne Nakrem lagde også søkbare zoo-databaser; fuglesamlinga og lakselus (Gyrodactylus), som også fungerte ut 2012. Grunnen til at dette stanset var at de nye versjonene av Windows ikke ville bruke FileMaker v5, som var den siste versjonen som hadde innebygd web-server.

GBIF-noden laget og driftet en samlingsportal for MUSIT-data fra NHM (og de andre universitetsmuseene) i noen år i første halvdel av 2010-tallet. I 2020 er en ny samlingsportal²⁵ i drift, hvor målet er å vise data fra alle digitaliserte samlinger, både biologisk og geologiske.

De levende samlingene i Hagen har blitt delt gjennom plattformen Garden Explorer²⁶ siden 2008, og Fuglesamlingen har blitt delt gjennom Collection Explorer²⁷ siden 2014. DNA-banken har delt sine data gjennom Global Genome Biodiversity Network (GGBN)²⁸ siden ca. 2016, og plante DNAet også i Collection Explorer siden 2015. Data om NHMs vertebratsamlinger har blitt delt i VertNet²⁹ siden ca. 2017.

I 2005 ble det foretatt en evaluering av universitetsmuseenes digitaliseringsarbeid³⁰. Her ble det anbefalt bl.a. samarbeid med Artsdatabanken og GBIF-noden.

Global Biodiversity Information Facility (GBIF)³¹ ble etablert i 2001 og Danmark ble valgt som vertsland for sekretariatet. Norge ble medlem i 2004 og den norske noden av GBIF ble opprettet ved NHM i 2005. GBIF Norge har ansvaret for koordineringen av norske biodiversitetsdatasett publisert til den globale GBIF-portal. GBIF Norge samarbeider tett med Artsdatabanken i Trondheim. Alle occurrence-datapunkter med geografiske koordinater fra lokaliteter i Norge publisert i det internasjonale GBIF-nettverket er inkludert i Artskart³², og alle observasjoner publisert i Artsobservasjoner publiseres i GBIF-portal.

Artsdatabanken drifter Artsnavnebasen³³, Artskart og Artsobservasjoner³⁴. Artskart har eksistert siden 2007 og er en tjeneste som viser funn og observasjoner av arter på kart. Informasjonen kommer fra frivillige rapportører i Norge (via Artsobservasjoner) og ulike biologiske fagmiljøer, herunder NHM.

NHM deler sine datadumper fra MUSIT og Corema med GBIF, som deretter sender dataene til Artskart.

Konklusjon

NHM har gjennom sitt langvarige digitaliseringsarbeid oppnådd gode resultater både når det gjelder kvalitet og kvantitet, og dette danner et godt grunnlag for videre digitaliseringsarbeid. Fremover vil man konsentrere seg om å digitalisere alle samlinger opp til ønsket nivå og ha et sterkt fokus på å ta ut de gevinstene digitaliserte arbeidsprosesser vil gi, både i samlingsarbeid og som støtte til NHMs forskning.

²⁵<https://samlingsportal.nhm.uio.no/museum/nhm>

²⁶<https://nhmbgoslo.gardenexplorer.org/>

²⁷<https://collectionexplorer.org/>

²⁸http://www.ggbn.org/ggbn_portal/

²⁹<http://vertnet.org/>

³⁰https://evalueringsportalen.no/evaluering/evaluering-av-universitetsmuseenes-digitaliseringsarbeid/Evaluering_av_universitetsmuseenes_digitaliseringsarbeid.pdf/@@inline

³¹<https://www.gbif.org/>

³²<https://artskart.artsdatabanken.no>

³³<http://www2.artsdatabanken.no/artsnavn/Contentpages/Hjem.aspx>

³⁴<https://www.artsobservasjoner.no>