

928 LifeWatch Norge

Sluttrapport fra forprosjektet

NINA Rapport

Frank Hanssen, Tor G. Heggberget, Nils Valland, Wouter Koch,
Jannicke Moe, Jan Karud, Helge Sagen, Sjur Ringheim Lid og Dag
Endresen



NINAs publikasjoner

NINA Rapport

Dette er en elektronisk serie fra 2005 som erstatter de tidligere seriene NINA Fagrapport, NINA Oppdragsmelding og NINA Project Report. Normalt er dette NINAs rapportering til oppdragsgiver etter gjennomført forsknings-, overvåkings- eller utredningsarbeid. I tillegg vil serien favne mye av instituttets øvrige rapportering, for eksempel fra seminarer og konferanser, resultater av eget forsknings- og utredningsarbeid og litteraturstudier. NINA Rapport kan også utgis på annet språk når det er hensiktsmessig.

NINA Temahefte

Som navnet angir behandler temaheftene spesielle emner. Heftene utarbeides etter behov og serien favner svært vidt; fra systematiske bestemmelsesnøkler til informasjon om viktige problemstillinger i samfunnet. NINA Temahefte gis vanligvis en populærvitenskapelig form med mer vekt på illustrasjoner enn NINA Rapport.

NINA Fakta

Faktaarkene har som mål å gjøre NINAs forskningsresultater raskt og enkelt tilgjengelig for et større publikum. De sendes til presse, ideelle organisasjoner, naturforvaltningen på ulike nivå, politikere og andre spesielt interesserte. Faktaarkene gir en kort framstilling av noen av våre viktigste forskningstema.

Annen publisering

I tillegg til rapporteringen i NINAs egne serier publiserer instituttets ansatte en stor del av sine vitenskapelige resultater i internasjonale journaler, populærfaglige bøker og tidsskrifter.

LifeWatch Norge

Sluttrapport fra forprosjektet

Hanssen, Frank (NINA)

Heggberget, Tor G. (NINA)

Nils Valland (Artsdatabanken)

Wouter Koch (Artsdatabanken)

Jannicke Moe (NIVA)

Jan Karud (NIVA)

Helge Sagen (Havforskningsinstituttet)

Sjur Ringheim Lid (Havforskningsinstituttet)

Dag Endresen (Naturhistorisk Museum, UiO)



ARTSDATABANKEN



UiO : Naturhistorisk museum

LifeWatch Norge. Sluttrapport fra forprosjektet – NINA Rapport
[928. 45 s.]

Trondheim 15. Januar 2013

ISSN: 1504-3312

ISBN: 978-82-426-2532-8

RETTIGHETSHAVER

© Norsk institutt for naturforskning

Publikasjonen kan siteres fritt med kildeangivelse

TILGJENGELIGHET

[Åpen]

PUBLISERINGSTYPE

Digitalt dokument (pdf)

REDAKSJON

Frank Hanssen

KVALITETSSIKRET AV

Jiska Van Dijk og Roald Vang

ANSVARLIG SIGNATUR

Administrerende Direktør Norunn Myklebust (sign.)

OPPDRAGSGIVER(E)/BIDRAGSYTER(E)

Norsk Forskningsråd

KONTAKTPERSON(ER) HOS OPPDRAGSGIVER/BIDRAGSYTER

Ian Gjertz, Norsk Forskningsråd (NFR)

FORSIDEBILDE

Arnstein Staverløkk (NINA)

NØKKELOD

Artsdatabanken, Bioinformatikk, databaser, åpen deling av biodi-
versitetsdata, dataforvaltning, ESFRI, forskning, GBIF, infrastruk-
tur, LifeWatch, kapasitetsbygging, metadata, NMDC, standardise-
ring, åpen tilgang til biodiversitetsdata

KEY WORDS

NBIC, Bioinformatics, databases, open biodiversity data sharing,
data management, ESFRI, scientific research, GBIF, Infrastructure,
LifeWatch, capacity building, metadata, NMDC, open access to
biodiversity data, standardization

KONTAKTOPPLYSNINGER

NINA hovedkontor

Postboks 5685 Sluppen
7485 Trondheim
Telefon: 73 80 14 00

NINA Oslo

Gaustadalléen 21
0349 Oslo
Telefon: 73 80 14 00

NINA Tromsø

Framsenteret
9296 Tromsø
Telefon: 77 75 04 00

NINA Lillehammer

Fakkeltgården
2624 Lillehammer
Telefon: 73 80 14 00

www.nina.no

Sammendrag

LifeWatch Norge. Sluttrapport fra forprosjektet – NINA Rapport [928. 45 s.]

LifeWatch er et europeisk initiativ med langsiktig fokus på utvikling av vitenskapelig baserte e-løsninger for biodiversitets- og økosystemforskning. LifeWatch er en del av det europeiske veikartet for infrastrukturutvikling (ESFRI- European Strategy for Research Infrastructure), og kom inn på det norske veikartet for e-infrastrukturutvikling i 2010.

NINA ble i 2008 invitert av Forskningsrådet til å følge utviklingen av det europeiske LifeWatch prosjektet (finansiert gjennom EU sitt 7. Rammeprogram for forskning og infrastrukturbygging) og til å initiere et norsk LifeWatch forprosjekt. Forprosjektet ble formelt oppstartet i januar 2011 med finansiell støtte fra Forskningsrådet. Konsortiet har vært koordinert av NINA i tett samarbeid med Havforskningsinstituttet (HI), Artsdatabanken, Naturhistorisk Museum ved Universitetet i Oslo (norsk node i Global Biodiversity Information Facility) og Norsk institutt for vannforskning (NIVA).

Forprosjektets hovedformål har vært å utrede muligheter og utfordringer knyttet til realisering av en norsk e-infrastruktur for fri deling av biodiversitets- og miljødata på tvers av institusjons- og landegrenser. Forprosjektet LifeWatch Norge er tett integrert med LifeWatch Europa gjennom NINA sin deltagelse i LifeWatch Preparatory Phase (2008-2010) og observatørstatus i styringsgruppen for LifeWatch Europa. Med finansiering fra NordForsk har LifeWatch konsortiene i Norge, Sverige, Danmark og Finland, i samarbeid med Islands Naturhistoriske Museum nylig startet opp et nordisk LifeWatch forprosjekt for å utrede muligheten for å realisere en nordisk e-forskningsinfrastruktur. På sikt vil et nordisk tyngdepunkt innen LifeWatch være strategisk viktig for de nordiske forskningsmiljøene. En nordisk overbygning vil kunne danne en felles grenseflate mellom nasjonale prosjekter og det europeiske nivå.

Forprosjektet har kartlagt brukerbehov og metadata, samt foretatt en gjennomgang av de mest sentrale prinsipper, lover og retningslinjer knyttet til deling av offentlig finansierte miljø- og biodiversitetsdata. Videre er sentrale utfordringer, barrierer og rammevilkår for deling og tilgjengeliggjøring av offentlig finansierte forskningsdata drøftet.

Utgangspunktet er at det i dag eksisterer store mengder relevante miljø- og biodiversitetsdata som av ulike årsaker er lite, eller overhodet ikke tilgjengelig for samfunnet. Konsortiet kommer på bakgrunn av dette med relevante anbefalinger og tiltak for realisering av en nasjonal e-infrastruktur for deling av offentlig finansierte miljø- og biodiversitetsdata. Dette omfatter utvikling av kontraktmaler, strategier og handlingsplaner for bedre dataforvaltning, kapasitetsbygging, utvikling av datamobiliseringsverktøy, samt evaluering av gjeldende rammevilkår og virkemiddelbruk rettet mot instituttsektoren.

For å kunne realisere en operativ norsk LifeWatch e-infrastruktur må en tverrsektoriell forankring og organisering sikres. Vi anbefaler at nåværende LifeWatch-konsortium i samarbeid med Norges Forskningsråd arrangerer en interdepartemental LifeWatch konferanse innen Juni 2013 med relevante fagdepartement. Dette bør være første fase i utvidelsen av det nasjonale LifeWatch konsortiet, hvor hovedformålet bør være å klarlegge potensialet for merverdi og effektivisering for forskning og forvaltning.

Forskningsrådet ønsker å styrke sin innsats for innovasjon i offentlig sektor og har derfor nylig startet arbeidet med strategiutvikling hvor grunntanken er å styrke kunnskapsflyten mellom forskning, næringsliv/industri og utdanningssektoren. Vi mener at en nasjonal LifeWatch e-forskningsinfrastruktur for miljø- og biodiversitetsdata generelt kan bidra til økt kvalitet, innovasjon og effektivitet innen forskningsmiljøene, samt økt anvendelse og politikktutforming innen næringsliv og offentlig forvaltning.

Konkret hvordan en nasjonal LifeWatch e-infrastruktur kan etableres må løses med utgangspunkt i de behov som er avdekket i forprosjektet. Videre er det meget viktig at denne e-infrastrukturen samordnes godt med parallelle initiativer på nasjonalt, nordisk og internasjonalt nivå.

Hovedkomponentene i den nasjonale LifeWatch e-infrastrukturen tenkes bygd opp som en distribusjonsportal og en analyseportal. Distribusjonsportalen er beregnet for brukere som vil søke opp grunnlagsdata og koble disse med miljødata for videre analyser i egne verktøy. Analyseportalen er beregnet for brukere som ønsker tilrettelagt funksjonalitet for søk, filtrering, beregninger, presentasjon, skjermvisning og nedlasting. Alle typer data skal leses online fra dataeierens originaldatabaser via standardiserte WEB-protokoller og tjenester.

Det er på det nåværende tidspunkt også vanskelig å si noe om hvordan et fullskala LifeWatch-prosjekt bør organiseres. De sentrale støttefunksjonene i LifeWatch Europa kan her bistå med veiledning og rådgivning. I tillegg vil det nordiske samarbeidet og deling av erfaringer derfra være nyttige. Det vil være essensielt at forvaltningsinstitusjonene deltar aktivt i et framtidig LifeWatch konsortium, for på denne måten å sikre koblingen mellom forskning og forvaltning når det gjelder miljødata og tidsserier.

Konsortiet har som ambisjon å søke finansiering av et storskala LifeWatch-prosjekt fra Forskningsrådets nasjonale INFRASTRUKTUR-program. Neste utlysning vil sannsynligvis bli i oktober 2014. EU sitt nye rammeprogram for forskning og infrastrukturbygging, HORIZON 2020 (oppstart 2014) kan også være en aktuell finansieringskilde dersom det norske konsortiet søker sammen med en eller flere andre europeiske partnere som f. eks de nordiske landene.

Hanssen, Frank (frank.hanssen@nina.no)
Heggberget, Tor G. (tor.g.heggberget@nina.no)
Nils Valland (nils.valland@artsdatabanken.no)
Wouter Koch (wouter.koch@artsdatabanken.no)
Jannicke Moe (jannicke.moe@niva.no)
Jan Karud (jan.karud@niva.no)
Helge Sagen (helge.sagen@imr.no)
Sjur Ringheim Lid (sjur.ringheim.lid@imr.no)
Dag Endresen (dag.endresen@nhm.uio.no)

Abstract

LifeWatch Norway. Final report from the pilot project – NINA Report [928. 45 pp.]

LifeWatch is a pan-European initiative with a long-term focus on the development of scientifically based e-solutions for biodiversity- and ecosystem research. LifeWatch is a part of the European roadmap for infrastructure development (ESFRI- European Strategy for Research Infrastructure), and came into the Norwegian roadmap for infrastructure development in 2010.

NINA was in 2008 invited by the Research Council of Norway to follow the development in the European LifeWatch project (funded by the European FP7 program for research and e-infrastructure development), and to initiate a Norwegian pilot project. The Norwegian LifeWatch pilot project was formally kicked-off in January 2011 with financial support from the Research Council of Norway. The national consortium has been coordinated by NINA in close cooperation with the Norwegian Institute on Marine Research (IMR), The Norwegian Biodiversity Information Centre (NBIC), the Norwegian Institute for Freshwater Research (NIVA) and the Natural History Museum at the University of Oslo (which holds the Norwegian node in the Global Biodiversity Information Facility).

The aim of this pilot project has been to investigate the possibilities and challenges related to the realization of a Norwegian e-infrastructure for free sharing of biodiversity and environmental data across institutional and national borders. The Norwegian LifeWatch pilot project is tightly integrated with LifeWatch Europe through the NINA participation in the European LifeWatch Preparatory Phase (2008-2010), and through the NINA observatory seat in the LifeWatch Board meetings. With financial support from NordForsk, the LifeWatch consortiums of Norway, Sweden, Denmark and Finland, together with the Icelandic Museum of Natural History recently initiated a Nordic LifeWatch pilot project to investigate the possibilities of realizing a Nordic research e-infrastructure. A strong Nordic cooperation within LifeWatch Europe could have a long-term strategic importance for Nordic research institutions. Also, the Nordic LifeWatch level could represent a common interface between the different national levels and the European level.

The pilot project has mapped user needs and metadata, and considered the most relevant principles, laws and guidelines for sharing of public funded environmental and biodiversity data. Further on, main challenges, barriers, terms and conditions for sharing of public funded research data are considered.

Today large amounts of relevant biodiversity data are only partly accessible for the public society. The Consortium addresses relevant recommendations and actions for the realization of a national infrastructure for sharing of public funded environmental and biodiversity data. This includes the development of contract templates, strategies and action plans for improved data management, capacity building, development of data mobilization tools, and an evaluation of current terms and conditions towards the research institute sector in Norway.

An operative Norwegian LifeWatch infrastructure demands mandatory support from all relevant Ministries. We recommend that the current Norwegian LifeWatch Consortium in cooperation with the Research council of Norway organize a national inter-Ministerial LifeWatch conference within June 2013. This should be the first step towards extending the national LifeWatch Consortium. The main goal of this conference should be to investigate potential added values and how the research and management sector could benefit from these added values in terms of increased efficiency.

The Research Council of Norway wants to increase its effort in public innovation and has therefore started to develop strategies on how to improve the flow of knowledge between research, industry and the educational sector. A national LifeWatch e-infrastructure for environmental

and biodiversity data in general will contribute to increased quality, innovation and efficiency within these three sectors.

At this stage it is hard to be specific on how a future national LifeWatch e-infrastructure should be organized. The specific needs for such an e-infrastructure should be scoped to satisfy the needs identified in this pre-project. Further on it is very important that this e-infrastructure will be sufficiently aligned with existing parallel initiatives at national, Nordic and international levels.

The main components of a national LifeWatch e-infrastructure are established as a Distribution portal and an Analysis portal. The Distribution portal will be designed for users that want to search biodiversity data and link these with environmental data for further analysis in other tools. The Analysis portal will be designed for users that want WEB functionality for search, filtration, calculation, presentation, screen display and downloading. All involved data will be distributed online from the stakeholder's primary databases using standardized WEB protocols and services.

Concerning organization of a large scale LifeWatch project, the central support functions of LifeWatch Europe can guide and support a future national project. In addition, the Nordic LifeWatch cooperation and sharing of common experiences will be useful. It will be essential that the public management institutions actively participate in order to ensure the link between research and management when it comes to environmental data and long time series.

The Norwegian LifeWatch Consortium recommends applying for funding from the Research Council of Norway for a large scale LifeWatch project as soon as possible. The next Infrastructure call will probably be in October 2014. Also the upcoming European program for Research and Innovation, HORIZON 2020 (to be started in 2014), could be an actual source of funding if the Norwegian consortium applies together with one or several other European partners such as for instance the Nordic countries.

Hanssen, Frank (frank.hanssen@nina.no)
Heggberget, Tor G. (tor.g.heggberget@nina.no)
Nils Valland (nils.valland@artsdatabanken.no)
Wouter Koch (wouter.koch@artsdatabanken.no)
Jannicke Moe (jannicke.moe@niva.no)
Jan Karud (jan.karud@niva.no)
Helge Sagen (helge.sagen@imr.no)
Sjur Ringheim Lid (sjur.ringheim.lid@imr.no)
Dag Endresen (dag.endresen@nhm.uio.no)

Innhold

Sammendrag	3
Abstract	5
Innhold	7
Forord	9
1 Bakgrunn, formål og forhold til nasjonale og europeiske strategier for etablering av forskningsinfrastruktur	10
2 Kort historikk om Norges satsning på infrastruktur for deling av biodiversitets- og miljødata	13
3 Resultater fra forprosjektet	16
3.1 Forskningens og forvaltningens behov for bedre tilgang til biologiske data (arbeidspakke 2).....	16
3.1.1 Generelt om databehov.....	17
3.1.2 Om behov for ulike datatyper og kart.....	17
3.1.3 Om behov for data fra ulike økosystemer.....	18
3.1.4 Om behov for data fra ulike artsgrupper.....	18
3.1.5 Om behov for metadata (informasjon om eksisterende datasett).....	19
3.1.6 Om behov for verktøy og tjenester.....	20
3.1.7 Oppsummering av resultater.....	21
3.2 Kartlegging av metadata (arbeidspakke 2).....	21
3.3 Prinsipper, lover og retningslinjer for deling av miljø- og biodiversitetsdata (arbeidspakke 3).....	24
3.3.1 OECD- prinsippene.....	24
3.3.2 Viderebruksdirektivet.....	25
3.3.3 Offentlighetsloven.....	25
3.3.4 Geodataloven, Geodataforskriften og INSPIRE.....	25
3.3.5 Miljøinformasjonsloven.....	25
3.3.6 Naturoppsynsloven.....	25
4 Drøfting	27
4.1 Begrepene Åpne data og Åpen tilgang.....	27
4.2 Begrepet Maskinlesbart format.....	27
4.3 Deling av nyetablerte og historiske datasett.....	28
4.4 Utdringer og barrierer knyttet til deling og tilgjengeliggjøring av data.....	28
4.4.1 Teknologi, standarder og økonomiske rammevilkår.....	28
4.4.2 Bedriftskultur og holdninger.....	29
4.4.3 Behovet for akademisk merittering ved deling av data.....	30
4.4.4 Dataforvaltning, datastrategier og kontraktsforhold.....	31
4.5 Juridiske og administrative rammevilkår for deling av offentlig finansierte data.....	32
4.6 Norsk Lisens for Offentlige Data (NLOD).....	33
4.7 Nasjonal Infrastruktur for deling av offentlige miljø og biodiversitetsdata.....	34
5 Anbefalinger og tiltak	35
5.1 Utvikling av kontraktmaler.....	35
5.2 Utvikling av strategier og handlingsplaner.....	35
5.3 Kapasitetsbygging og utvikling av datamobiliseringsverktøy.....	35
5.4 Evaluering av rammevilkår og virkemidler rettet mot instituttsektoren.....	36

6 Utvidelse av det nasjonale LW- konsortiet	37
7 Realisering av et storskala LifeWatch- prosjekt	39
8 Referanser	42
9 Appendiks 1: Spørsmål stilt i brukerundersøkelsen	44

Forord

Forskning er grunnlaget for løsninger på mange av de kunnskapsutfordringer samfunnet vårt står overfor. Med effektiv tilgang til data og verktøy gjennom en avansert e-forskningsinfrastruktur frigjøres mer tid til forskning og mindre tid på tilrettelegging og sammenstilling av data fra ulike kilder. Norske myndigheter og Forskningsrådet har de senere år jobbet målrettet med strategiutvikling og finansieringsordninger rettet mot utvikling av avansert e-forskningsinfrastruktur for enkel tilgang til og bruk av offentlig finansierte forskningsdata.

Denne rapporten avdekker entydige behov for en mer effektiv tilgang til offentlig finansierte miljø- og biodiversitetsdata blant biologer innenfor forskning og naturforvaltning. Med dette som utgangspunkt har vi utformet relevante anbefalinger og tiltak for realisering av en norsk e-infrastruktur for fri deling av biodiversitets- og miljødata. Vi vil understreke at dette best kan skje i samordning med parallelle nasjonale initiativer som Norsk Marint Datasenter, Artsdatabanken og den norske noden i Global Biodiversity Information Facility ved Naturhistorisk Museum på Universitetet i Oslo.

Som en oppfølging av forprosjektet har det norske LifeWatch konsortiet som ambisjon å utvide konsortiet og søke finansiering av et storskala LifeWatch- prosjekt fra Forskningsrådets nasjonale INFRASTRUKTUR- program ved neste utlysning. EU sitt nye rammeprogram for forskning og innovasjon, HORIZON 2020 (oppstart 2014), kan også være en aktuell finansieringskilde dersom det norske konsortiet søker sammen med en eller flere andre europeiske partnere som f. eks de nordiske landene.

Til slutt vil vi rette en stor takk til alle som har deltatt i spørreundersøkelsen. Tusen takk også til alle dere som har bidratt med nyttige innspill, kommentarer og kvalitetssikring av denne slutt-rapporten.

Trondheim 15. Januar 2013

Hanssen, Frank (frank.hanssen@nina.no)
Heggberget, Tor G. (tor.g.heggberget@nina.no)
Nils Valland (nils.valland@artsdatabanken.no)
Wouter Koch (wouter.koch@artsdatabanken.no)
Jannicke Moe (jannicke.moe@niva.no)
Jan Karud (jan.karud@niva.no)
Helge Sagen (helge.sagen@imr.no)
Sjur Ringheim Lid (sjur.ringheim.lid@imr.no)
Dag Endresen (dag.endresen@nhm.uio.no)

1 Bakgrunn, formål og forhold til nasjonale og europeiske strategier for etablering av forskningsinfrastruktur

[Norges Nasjonale strategi for forskningsinfrastruktur \(2012 - 2017\)](#) har som hovedmålsetning at «Norske forskningsmiljøer skal ha tilgang til relevant og oppdatert infrastruktur som understøtter forskning av høy kvalitet, som i sin tur vil bidra til å møte samfunnets kunnskapsutfordringer». I perioden 2008 - 2012 ble dette fulgt opp med særskilte finansieringsordninger for forskningsinfrastruktur gjennom Kunnskapsdepartementet (KD) og [Norges Forskningsråd \(NFR\) sitt Infrastruktur-program](#). Medio 2012 har Norge investert vel en milliard kroner fordelt på 38 større infrastrukturprosjekter (Verktøy for forskning- del 1, Nasjonal strategi for forskningsinfrastruktur 2012- 2017. NFR 2012). Det er imidlertid et veldig stort gap mellom omsøkte midler og disponible midler for finansiering.

Internasjonalt samarbeid er strategisk viktig for norsk forskning og samfunnsutvikling. Behovet for at Norge deltar aktivt i det internasjonale samarbeidet er understreket i Forskningsmeldingen Klima for forskning (2008 - 2009). Europeisk forskningspolitikk (både på nasjonalt og på felleseuropeisk nivå) preges av en stadig større oppmerksomhet omkring etablering av infrastruktur for forskning. [European Strategy Forum on Research Infrastructures \(ESFRI\)](#) arbeider aktivt for å harmonisere den europeiske politikktutviklingen for europeisk forskningsinfrastruktur. Norge deltar aktivt i ESFRI og KD har bedt NFR om å utarbeide beslutningsgrunnlag og fremme anbefalinger om norsk deltagelse i infrastrukturprosjekter som er tatt med på ESFRIs veikart.

Den første utgaven av [Norsk veikart for forskningsinfrastruktur](#) kom i 2010. Nåværende 2012-utgave av det norske veikartet underbygger anbefalingene gitt i Nasjonal strategi for forskningsinfrastruktur om å trappe opp investeringsvolumet og sørge for en langsiktig finansiering av forskningsinfrastruktur i årene som kommer.

Forskningsmeldingen [Klima for forskning \(2008- 2009\)](#) krever at det skal stilles strenge krav til hvilke forskningsinfrastrukturer som skal tas inn på veikartet, både når det gjelder kvalitet og samfunnsmessig betydning. Tre prosesser kan føre et prosjekt inn på det nasjonale veikartet: (1) Prosjekter inkludert etter søknadsbehandling innenfor den nasjonale satsningen på forskningsinfrastruktur, (2) Større prosjekter (>200 millioner) inkludert etter søknadsbehandling og departementalt vedtak (KD) i samråd med NFR, og (3) Prosjekter på ESFRIs veikart hvor Norge har inngått forpliktelser eller gitt politiske signaler om ønsket deltagelse. Prosjekter innenfor sistnevnte kategori er grundig evaluert av ESFRI og kan av NFR vurderes å ha stor strategisk betydning for norsk forskning.

[LifeWatch \(LW\) Preparatory Phase \(2008- 2010\)](#) kom inn på det norske veikartet i 2010. NFR inviterte i 2008 [Norsk institutt for naturforskning \(NINA\)](#) til å følge utviklingen i Preparatory Phase, og NINA har siden oppstarten deltatt aktivt i arbeidspakken Data Providers Platform og som observatør i LW Stakeholder Board. Parallelt med dette ble NINA også invitert til å initiere og koordinere et norsk LW forprosjekt.

Søknad om NFR-finansiering ble innvilget i november 2010 med 500.000 NOK. Forprosjektet Establishment of the LifeWatch Norwegian ble formelt oppstartet i januar 2011 med et konsortium bestående av [Havforskningsinstituttet \(HI\)](#), [Artsdatabanken](#), [Naturhistorisk Museum ved Universitetet i Oslo \(GBIF Norge\)](#), [Norsk institutt for vannforskning \(NIVA\)](#) og [Norsk institutt for naturforskning \(NINA\)](#). Forprosjektets hovedformål er å utrede muligheter og utfordringer knyttet til realisering av en norsk infrastruktur for fri deling av biodiversitets- og miljødata på tvers av institusjons- og landegrensener som en integrert del av det europeiske ESFRI-prosjektet LW.

Sammensetningen av det norske LW-konsortiet gjenspeiler en felles motivasjon om å bygge videre på eksisterende nasjonale initiativ (som f.eks. Artsdatabanken og den norske GBIF-noden ved Naturhistorisk Museum), samt å klargjøre gjeldende rammevilkår og behov innenfor offentlig/privat FOU-sektor, forvaltning, næringsliv og samfunnet for øvrig. Dersom forprosjektet videreføres vil relevante institusjoner og organisasjoner innenfor offentlig og privat sektor inviteres med i konsortiet.

Parallelt med det norske forprosjektet har det norske LW-konsortiet initiert et samarbeid med en rekke vitenskapelige institusjoner i Island, Sverige, Danmark og Finland om et [Nordisk LW forprosjekt](#). Selv om de nordiske landene befinner seg på ulike stadier i sin tilnærming til LW er motivasjonen for å innlede et nordisk samarbeidsprosjekt stor. For det første har de nordiske landene mange fellestrekk som vil kunne katalysere et samarbeid. For det andre pågår det allerede et omfattende samarbeid mellom flere nordiske land om etablering av infrastruktur for deling av biodiversitetsdata (f.eks. Artsdatabankene i Norge og Sverige, samt de nordiske nodene i [Global Biodiversity Information Facility \(GBIF\)](#) som også har sitt globale sekretariat i København. For det tredje er det ønskelig å utnytte synergier fra parallelle initiativ som f.eks. the Baltic Biodiversity Project som er et samarbeidsprosjekt mellom de Baltiske landene, Sverige, Finland og Danmark. For det fjerde har de nordiske landene relativt liten gjennomslagskraft i europeisk sammenheng hver for seg. Et nordisk tyngdepunkt innen LW vil være strategisk viktig for alle de nordiske forskningsmiljøene. En nordisk LW overbygning vil kunne danne en felles grenseflate mellom de nasjonale prosjektene og det europeiske LW. Representanter for de ulike institusjonene møttes 18. februar 2011 i Stockholm for å utrede muligheter for samarbeid og å lage en felles søknad om finansiering rettet til NordForsk. Møtet resulterte i opprettelsen av et Nordisk LW konsortium og en søknad til NordForsk som ble innvilget med 500.000 NOK den 10. februar 2012. Oppstartsmøtet ble avholdt 8-9 november 2012 på Naturhistorisk Riksmuseum i Stockholm. Det nordiske samarbeidet som koordineres av NINA tillegges stor vekt i LW Europa, Koordinator Eva Alonso ved LifeWatch Service Center, hadde følgende kommentar etter møtet:

Et nordisk LifeWatch samarbeid er svært gunstig for videre etablering og utvidelse av nasjonale LifeWatch infrastrukturer gjennom å bygge videre på felles standarder og mekanismer for deling av data og tjenester. Det nåværende samarbeidet mellom de nordiske landene om felles satsningsområder er en nøkkelfaktor for å styrke det vitenskapelige perspektivet innenfor forskningsinfrastruktur og for å kunne tilby nye vitenskapelige baserte tjenester for samfunnet.

Det europeiske LW gikk den 19. januar 2011 over i Construction Phase (2011 - 2016). Dette ble formalisert gjennom en samarbeidsavtale om etablering av en infrastruktur for forskning på biodiversitet og økosystemer, signert på myndighetsnivå av Italia, Spania, Ungarn, Nederland og Romania. Parallelt med dette har flere europeiske land undertegnet intensjonsavtaler med LW Europa om vilje til samarbeid. LW skal i Construction Phase organiseres som et juridisk uavhengig selskap basert på EUs forordning om et juridisk rammeverk for felles forskningsinfrastrukturer, [European Research Infrastructure Consortium \(ERIC\)](#) vedtatt 28. august 2009. ERIC-forordningen trådte formelt i kraft 28. august 2009 og skal gjøre det enklere å etablere og drifte forskningsinfrastrukturer på europeisk nivå gjennom felles konsortier for forskningsinfrastrukturer. Forordningen tilbyr et rammeverk for eierskap og drift av de felleseuropeiske forskningsinfrastrukturene. Hensikten med et slikt rammeverk er å gjøre det enklere å realisere ny forskningsinfrastruktur av felleseuropeisk interesse, primært basert på ESFRI sitt veikart for slik infrastruktur. ERIC-forordningen er ment å forenkle samarbeid mellom EU-medlemsland, assosierte land og tredjeland om forskningsinfrastruktur, og å bidra til å realisere ESFRI's veikart. En ERIC-infrastruktur vil på bakgrunn av forordningen bli et selvstendig rettssubjekt med full rettslig handlingsevne. Forordningen spesifiserer innholdet i en ERIC, og hvilke forutsetninger som må være oppfylt for at en infrastruktur skal kunne få en slik status. Dette omfatter organisasjon og vedtekter, søknadsprosedyre og klagemuligheter. En ERIC må ha minimum tre EU-land som medlemmer (eiere), og EU-land skal til sammen ha majoriteten av stemmer i styrende organer. Et assosiert land til EUs rammeprogrammer for forskning, teknologisk utvikling og

demonstrasjonsaktiviteter kan også være sete for en ERIC-infrastruktur. Kommisjonen skal ikke være medlem i en ERIC.

Et vesentlig punkt er at en ERIC skal ha status på linje med en internasjonal organisasjon med hensyn til direktiver om merverdiavgift, skatt og offentlige anskaffelser. Dette innebærer at en ERIC skal ha fritak for merverdiavgift og omfattende unntak fra skatteplikten, og at den har unntak fra direktiv om offentlige anskaffelser. Det understrekes i forordningen at også infrastruktur med en annen juridisk plattform enn ERIC vil kunne anerkjennes som et bidrag til realiseringen av ESFRIs veikart. Det er med andre ord ikke obligatorisk å bruke ERIC. Andre måter å gjøre dette på må baseres på eksisterende nasjonal, internasjonal eller europeisk lovgiving. Det er imidlertid en sentral begrunnelse for forordningen at den skal bidra til vesentlig forenkling hva gjelder etablering og drift av felles infrastruktur for flere europeiske land. Det forventes at de fleste av prosjektene på ESFRIs veikart vil bygge på ERIC- forordningen.

Hvordan forholder Norge seg til ERIC? EU har et eksplisitt krav om at alle vedtak i en ERIC skal være basert på flertallsvedtak av EU-medlemsland. Dette innebærer at assosierte land ikke får samme status som EU-medlemsland innenfor en ERIC. Dette er ikke akseptabelt for Norge og KD har derfor vært i forhandlinger med EU om å sikre Norge likeverdig partnerstatus gjennom en tilpasning av EØS-avtalen. Dette kunne ikke aksepteres av EU-kommisjonen som i stedet vil endre ERIC-forordningen slik at medlemsland og assosierte land i styringsstrukturen likestilles. En slik endring må imidlertid godkjennes av EU Council og EU-parlamentet. Dette arbeidet er forventet ferdigstilt sommeren 2013. Dersom Norge kan akseptere endringene i ERIC- forordningen vil denne bli innlemmet i Lov om ansvarlige selskaper og kommandittselskaper (Selskapsloven). Spania leder de juridiske prosessene som skal ligge til grunn for den endelige konstitueringen av LW ERIC, men på grunn av strukturelle omstillinger i den spanske regjering er disse prosessene blitt noe forsinket.

2 Kort historikk om Norges satsning på infrastruktur for deling av biodiversitets- og miljødata

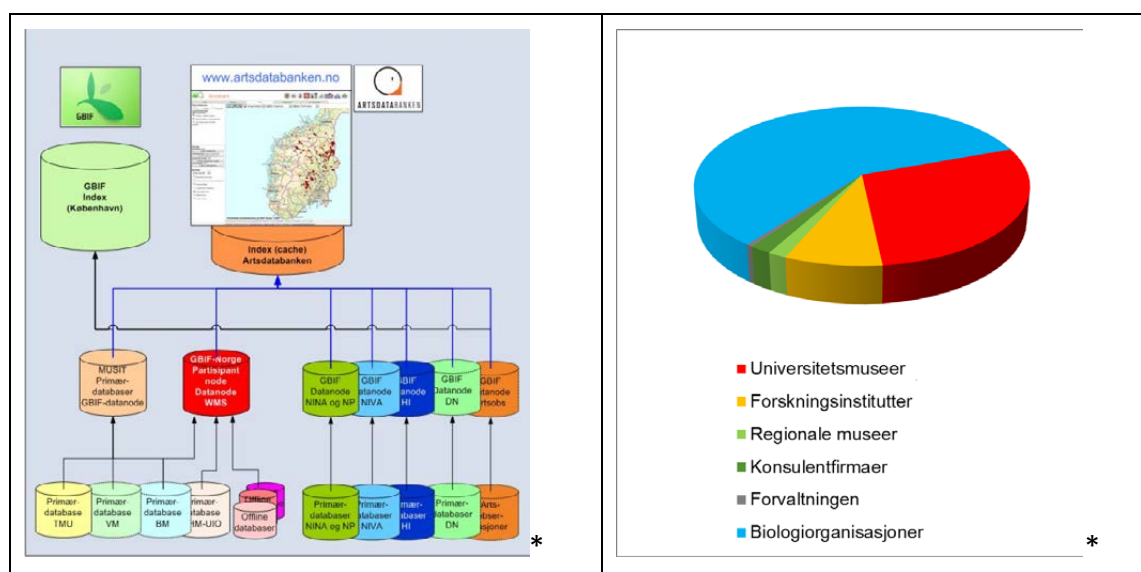
[Stortingsmelding nr. 42 \(2000-2001\)](#) redegjorde for tiltak knyttet til styrking av en kunnskapsbasert forvaltning, bl.a. ved at det skulle etableres et Nasjonalt program for kartlegging og overvåking av biologisk mangfold, herunder etablering av en nasjonal Artsdatabank. Målet med dette programmet var å gi informasjon om:

- Stedfesting og verdiklassifisering av viktige områder for biologisk mangfold
- Endringer i biologisk mangfold over tid
- Årsakene til endringene og forslag til tiltak
- Oppfølging av tiltak

I 2002 ble det utformet et program for datadeling og dataflyt i «Nasjonalt program for kartlegging og overvåking av biologisk mangfold». Miljøforvaltningen og miljøforskningsinstitusjoner har brukt disse prinsippene for dataetablering uten at det ble laget en nasjonal datadelingsinfrastruktur. [Artsdatabanken](#) ble etablert i 2004 for:

- Generelt å styrke arbeidet med truede og sårbare arter, naturtyper og gener
- Å få større faglig tyngde i dette arbeidet
- Å skape større uavhengighet i forhold til forvaltningen
- At resultater skal få større og bredere autoritet og legitimitet
- Å gjøre informasjon om Norges biologiske mangfold lett forståelig og lett tilgjengelig

Artsdatabanken har siden 2007 hatt en formidlingstjeneste av stedfestede artsdata i [Artskart](#). For tiden høstes over 14 millioner poster fra over 100 primærdatabaser i 32 kildeinstitusjoner, herunder de fleste norske forskningsinstitusjoner. Det er stedfestede artsdata for mer enn 27 000 arter. Dataflyten foregår ved at dataeierne/-forvalterne gjør sine stedfestede artsdata tilgjengelig i et internasjonalt standardisert format (Darwin Core) som høstes daglig av Artsdatabanken og dataene blir søkbare i Artskart. Brukeren kan søke informasjon om geografiske områder eller arter og artsgrupper. Brukeren kan velge data etter hvilken statuskategori artene har (etter Norsk rødliste for arter 2010 eller Fremmedartsstatus 2012) Alle dataene er nedlastbare og kan brukes fritt, men kildene skal synliggjøres. Alle dataene i Artskart distribueres også internasjonalt gjennom dataportalen til GBIF.



Figur 1: Dataflyt inn til Artskart

Artskart- portalen har om lag 10 000 unike brukere pr. måned og webtjenestene til Artskart er også hyppig brukt. Både eksterne portaler som [Miljøstatus for Norge](#), [Kilden](#) og andre viser data fra Artskart. I skognæringen anvendes dataene direkte i forvaltningsapplikasjoner.

Universitetsmuseene hadde i perioden 1998-2006 et stort prosjekt for datamobilisering, [Museumsprosjektet](#), med et digitaliseringsprogram for vitenskapelige samlinger. Som del av dette prosjektet ble det også utviklet samlingsdatabaser for museene. Dette er fra 2007 videreført av universitetsmuseenes felles IT-infrastruktur (MUSIT) og omfatter også datadeling.

[GBIF](#) ble etablert i 2001 etter forslag fra Organization for Economic Co-operation and Development ([OECD](#)) Global Science Forum i 1999. Danmark ble valgt som vertskap for GBIF sekretariatet, og dette har vært lokalisert i København siden 2001. Norge ble medlem av GBIF i 2004, representert av [Forskningsrådet](#). Den norske [GBIF noden](#) ble etablert i 2005 og er lokalisert ved Naturhistorisk museum ved Universitetet i Oslo ([NHM-UiO](#)). Den norske GBIF-noden er ansvarlig for koordinering av norske biodiversitets-datasett som publiseres til GBIF portalen. Noden tar imot datasett fra dataeiere som ønsker assistanse med datapubliseringsen og hjelper dataeiere som ønsker å publisere sine egne datasett med teknisk implementering. Norge bidrar i dag med totalt 74 datasett fra 22 dataeiere med over 11 millioner dataposter. Alle datasett som publiseres til GBIF portalen publiseres også til [Artskart](#) som er en meget populær tjeneste fra [Artsdatabanken](#).

Forskningsrådet fikk i perioden 2003 - 2004 gjennomført et omfattende arbeid for å kartlegge og vurdere viktigheten av å ta bedre vare på lange tidsserier for miljøovervåking og forskning ([Lange tidsserier for miljøovervåking og forskning, Rapport 1, 2 og 3, NFR 2004](#)). Tre arbeidsgrupper leverte fagrapporter for henholdsvis klimadataserier (koordinert av Meteorologisk institutt), terrestriske og limniske dataserier (koordinert av Direktoratet for naturforvaltning) samt marine dataserier (koordinert av Havforskningsinstituttet). Arbeidsgruppene identifiserte spesielt viktige, utvalgte tidsserier og kom med klare anbefalinger og tiltak for å sikre videre finansiering for drift, vedlikehold og tilgjengeliggjøring av tidsseriene for forskning og forvaltning.

Forskningsrådet leverte i 2007 [et forslag til en nasjonal strategi for forskningsinfrastruktur](#) for Forskningsrådets arbeid med databaser (inkludert registre og tidsserier) og samlinger som anses å ha stor samfunnsmessig verdi. I strategidokumentets innledning oppsummeres strategiarbeidet slik:

«Norges forskningsråds arbeid med en nasjonal strategi for forskningsinfrastruktur må sees på bakgrunn av nasjonale og internasjonale aktiviteter på området. For Norge er det viktig at man satser tilstrekkelig på infrastruktur til at forskningsmiljøene har mulighet til å arbeide i den internasjonale forskningsfronten, og at nasjonale prioriteringer er etablert slik at disse kan legges til grunn for deltagelse i den markante EU-satsingen innen feltet. Forskningsrådets ledelse og styrende organer har derfor engasjert seg i dette som en viktig arbeidsoppgave i 2007. Målsetningen for dette arbeidet er å utarbeide en nasjonal strategi (veikart) for forskningsinfrastruktur samt å etablere en egnet beslutningsstruktur for behandling av de ulike kategorier forskningsinfrastruktur som gir effektiv utnyttelse av ressursene».

Den 16. februar 2012 la Forskningsrådet frem [en oppdatert nasjonal strategi for forskningsinfrastruktur og et nytt veikart over strategisk viktige infrastrukturer \(det første veikartet ble lansert i 2010\)](#). Nasjonal strategi for forskningsinfrastruktur «Verktøy for forskning del 1» er en oppdatering av strategien med samme navn fra 2007. Mange av anbefalingene fra forrige strategidokument er fulgt opp og realisert. Blant annet sto det sentralt i den forrige infrastrukturstrategien å få på plass den finansieringsordning vi i dag kjenner som Nasjonal satsing på forskningsinfrastruktur.

Den oppdaterte strategien peker på hvordan strukturen for finansiering av forskningsinfrastruktur skal utvikles videre. Den rommer også klare råd til departementene og forskningsinstitusjonene, samt en oversikt over hva Forskningsrådet selv vil gjøre fremover.

Strategien og veikartet er et tydelig forskningspolitisk råd fra Forskningsrådet til bevilgende myndigheter om fortsatt å prioritere infrastruktur i form av anlegg, utstyr og databaser. Samtidig inneholder strategien en klar anbefaling til departementene om å bruke veikartet når de planlegger sine investeringer, og å nyttiggjøre seg det grundige arbeidet og de betydelige ressursene som er nedlagt i å kvalitetssikre prosjektene på veikartet.

3 Resultater fra forprosjektet

Forprosjektet ble organisert i tre arbeidspakker:

- Arbeidspakke 1: Kartlegge forskningens og forvaltningens behov for bedre tilgang til biologiske data.
- Arbeidspakke 2: Kartlegge metadata om eksisterende og relevante biologiske data.
- Arbeidspakke 3: Avklare prinsipper, lover og retningslinjer for deling av biologiske data

3.1 Forskningens og forvaltningens behov for bedre tilgang til biologiske data (arbeidspakke 2)

Brukerundersøkelsen ble organisert som en nettbasert spørreundersøkelse rettet mot biologer innenfor skoler, høyskoler, universiteter, forskningsinstitutter, miljø- og naturforvaltning, konsultantselskaper, naturfaglige organisasjoner og media. Med «Biologiske data» menes i denne sammenheng observasjoner av individer og bestander i naturen som er identifisert til art eller høyere nivå taksonomiske nivåer, og som har tids- og stedsangivelse (begrenset til Norge og Svalbard).

Invitasjon til å delta på brukerundersøkelsen ble sendt til 170 respondenter per epost. I tillegg ble undersøkelsen sendt til alle medlemmer av Norsk Biologforening (ca. 900 mottakere). Vi oppfordret samtlige representanter til å videreformidle vår invitasjon ut til sine faglige nettverk. Responsen på spørreundersøkelsen var noe høyere enn forventet. 115 av 170 respondenter deltok i undersøkelsen. Selv om dette svarmaterialet ikke er stort nok for statistiske sammenligninger, mener vi at det kan gi en god indikasjon på databehovet hos biologer i Norge.

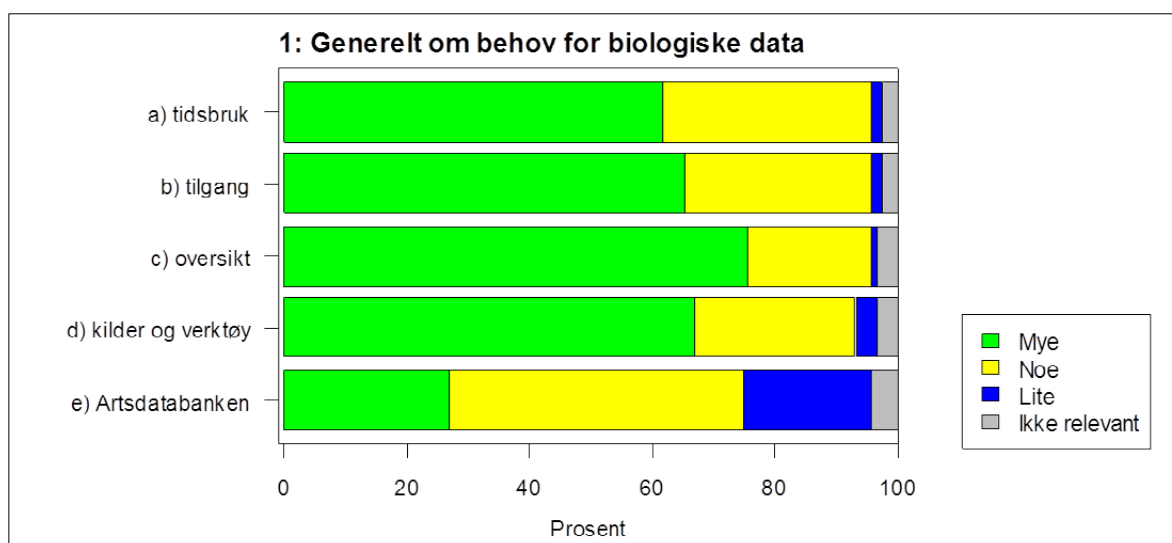
Respondentene var relativt jevnt fordelt på yrkesgrupper: Forskning (32 %), naturforvaltning og naturvern (26 %), undervisning (17 %) og andre yrkesgrupper (22 %). Den sistnevnte gruppen inkluderer bl.a. kartlegging av biologisk mangfold, arealforvaltning, rådgivning/konsulentvirksomhet, fylkesmannen, skadedyrbekjempelse, artsbestemmelse og GIS/Geodata. 3 % oppga ikke tilhørighet til noen spesiell yrkesgruppe.

De fleste respondentene har bakgrunn innenfor biologi (74 %) eller andre naturfag (13 %). Litt mer enn halvparten jobber i offentlig sektor (57 %), en fjerdedel (24 %) i instituttsektoren og 16 % i privat sektor. Det er flest besvarelser fra fylker som har store forskningsmiljøer: Oslo (20 %), Sør- Trøndelag (17 %) og Rogaland (17 %) samt Akershus (10 %). 12 andre fylker var representert hver med mindre enn 5 % av besvarelsene.

Det var noe overraskende at forskningsmiljøene i og rundt Bergen (Hordaland) ikke ga flere besvarelser. Dette kan skyldes at disse har deltatt i lignende forundersøkelser i tilknytning til etableringen av Norsk Marint Datasenter i regi av Havforskningsinstituttet.

Brukerundersøkelsen var delt inn i 6 hovedtema: Databehov generelt, behov for ulike data typer og kart, behov for data fra ulike økosystemer, behov for data fra ulike artsgrupper, behov for metadata og behov for verktøy og tjenester. Spørsmålene er gjengitt i Appendiks 1.

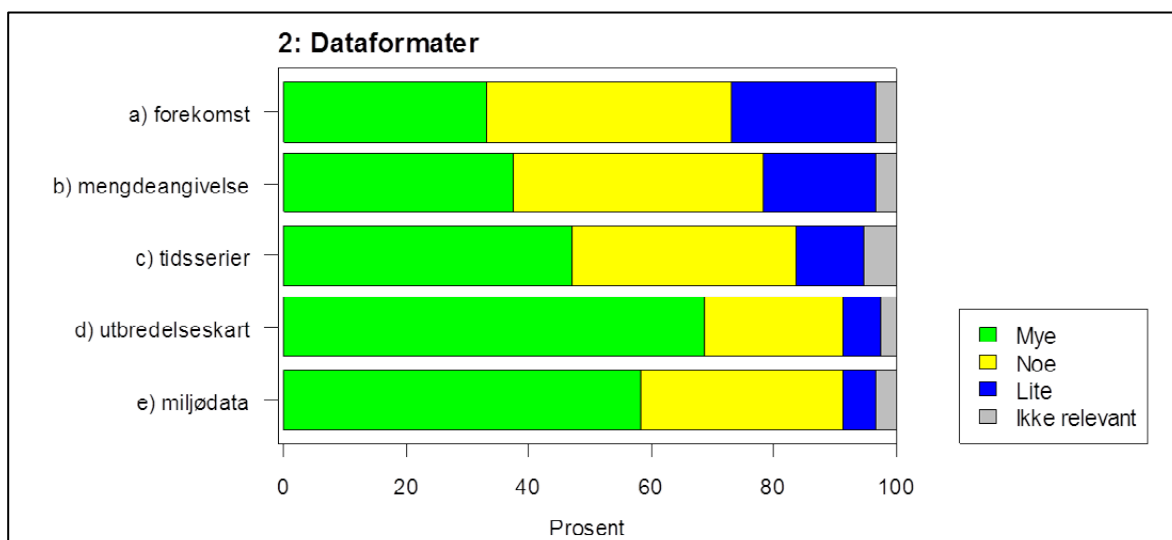
3.1.1 Generelt om databehov



Figur 2: Generelt om behov for biologiske data

Respondentene uttrykker generelt et stort behov for bedre dataoversikter og enklere tilgang til biologiske data. 63 % bruker biologiske data mye i sitt arbeid. 67 % mener de vil ha nytte av bedre tilgang til biologiske data. 78 % mener de vil ha nytte av en bedre oversikt over hva som fins av biologiske data. 69 % sier de er interessert i å ta i bruk nye kilder og verktøy for bruk av biologiske data. 75 % har benyttet seg av Artsdatabankens tjenester.

3.1.2 Om behov for ulike datatyper og kart

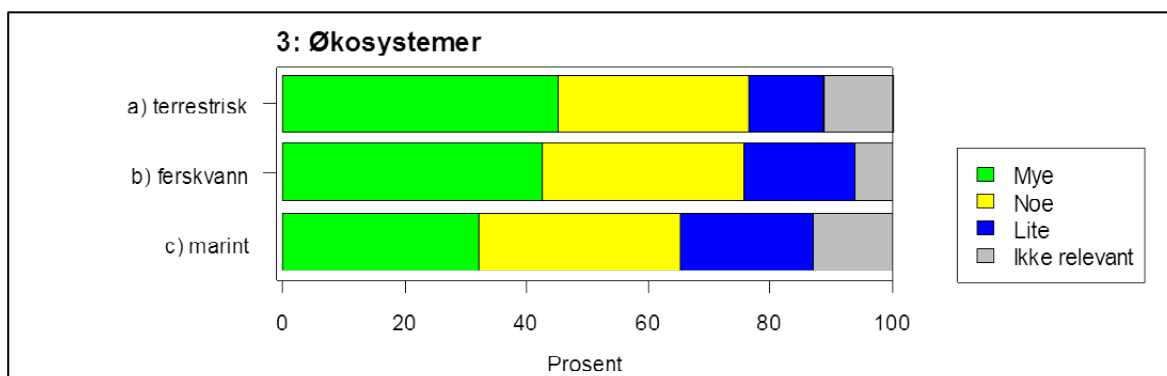


Figur 3: Om behov for ulike datatyper og kart

Respondentene ble bedt om å spesifisere behovet for henholdsvis forekomstdata, data om mengdeangivelse, tidsseriedata, miljødata og utbredelseskart. Det er klart størst behov for utbredelseskart for utvalgte arter (70 %), og dernest for miljødata som kan kobles til de biologiske dataene (60 %). Interessen for utbredelseskart var størst blant naturforvaltere og lignende yrker. Det var kun 48 % av respondentene som uttrykte stort behov for tidsseriedata (lange tidsserier) og tid- og stedfestede artsobservasjoner, av disse var den største andelen forskere

(46 %). Det hadde liten betydning om artsobservasjonene var oppgitt med mengdeangivelse, f.eks. som antall per km² eller bare som forekomst.

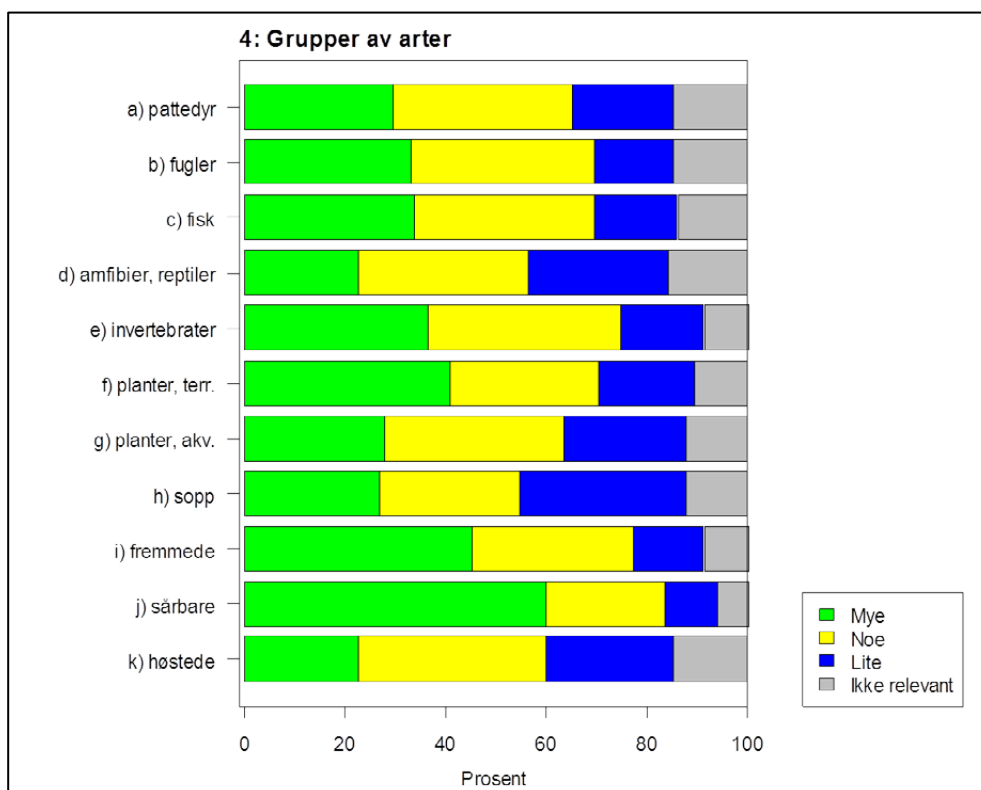
3.1.3 Om behov for data fra ulike økosystemer



Figur 4: Om behov for data fra ulike økosystemer

Respondentene uttrykte størst interesse for data fra terrestre økosystemer (48 %) og fra ferskvannsystemer inkludert våtmark (44 %). Det var relativt liten interesse for data om marine økosystemer (34 %), dette kan skyldes at marinbiologer av en eller annen grunn ser ut til å være underrepresentert i denne undersøkelsen (blant annet få besvarelser fra Hordaland som har store marine forskningsmiljøer i norsk målestokk). En annen forklaring kan være at marinbiologiske data allerede er bedre organisert og lettere tilgjengelig enn data fra terrestre og ferskvannsystemer, slik at behovet for å få bedre tilgang til marine data faktisk er lavere for marinbiologer.

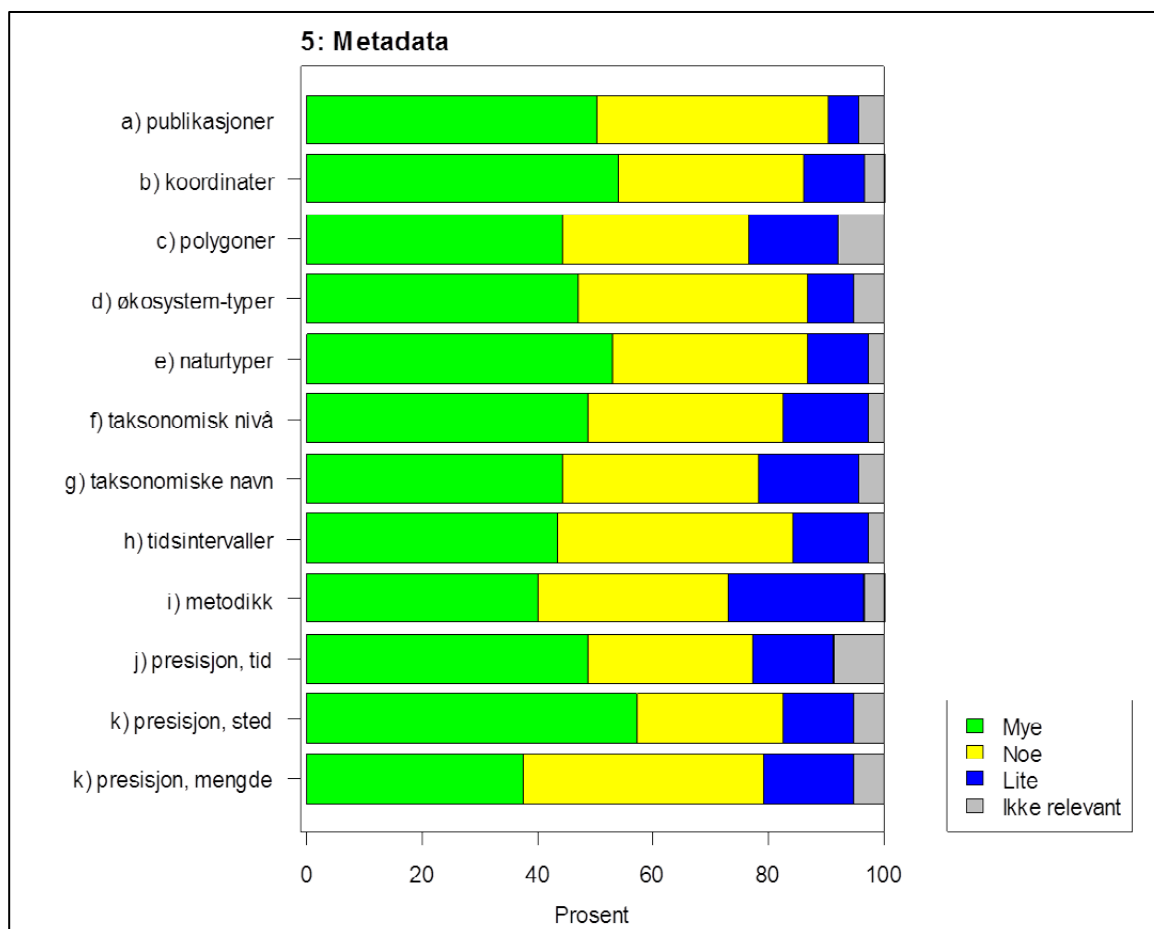
3.1.4 Om behov for data fra ulike artsgrupper



Figur 5: Om behov for data fra ulike artsgrupper

De ulike taksonomiske artsgruppene var omtrent like mye etterspurt blant respondentene. For hver av de terrestre/akvatiske artsgruppene pattedyr, fugler, fisk, invertebrater og planter uttrykt stort behov (80 %). Det var litt mindre behov for artsgruppene sopp, amfibier og reptiler (60 %). Andre artsgrupper som respondentene spesielt trakk fram var rovfugl og insekter. Behovet for mikrobielle artsgrupper ble også nevnt. Dette viser at det generelt er et stort behov for data fra de fleste taksonomiske grupper. Det ble ellers kommentert at det generelt er et større behov for data om utvalgte arter og naturtyper enn data om taksonomiske grupper. Videre var det spesielt stort behov for data om truede/sårbare arter (86 %) og introduserte/invaderende arter (81 %). Det var generelt noe mindre behov for data om høstbare arter (65 %).

3.1.5 Om behov for metadata (informasjon om eksisterende datasett)



Figur 6: Om behovet for metadata

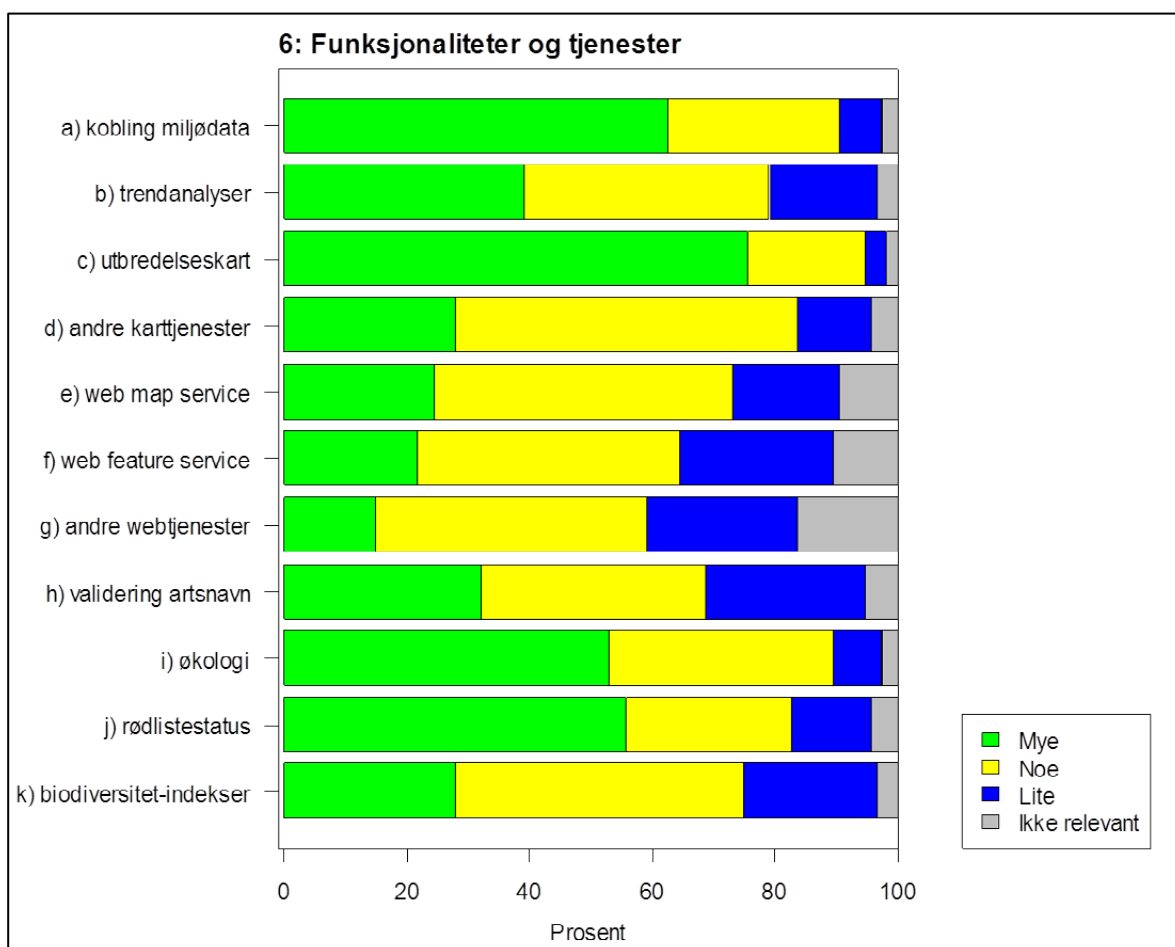
De aller fleste respondentene uttrykker behov for gode metadata. Flest uttrykte behov for informasjon om geografisk stedfesting på punktnivå (55 %) mens noen færre uttrykte behov for informasjon om mer detaljert romlig utbredelse på flatenivå (45 %).

Behovet for informasjon om prøvetaking og analyser ble ansett som noe mindre (41 %) enn vi forventet. Dette kan ha sammenheng med at de fleste var opptatt av forekomst av arter, men ikke nødvendigvis mengdeangivelse. Dermed kan også detaljer om metodikk for kvantifisering av forekomst være noe mindre viktig. Behovet for informasjon om tidsintervaller ble også ansett som noe mindre (44 %).

Når det gjelder informasjon om datakvalitet ble uttrykte 59 % av respondentene behov for presisjon på stedfesting av data. 51 % av respondentene uttrykte behov for høy presisjon på tidsangivelse mens kun 38 % uttrykte behov for presis mengdeangivelse. Dette stemmer overens med andre svar om at stedfesting blir ansett som noe viktigere enn tidfesting og at utbredelseskart er viktigere enn tidsseriedata.

Det ble påpekt at det ligger svært mye nyttig informasjon i gamle rapporter (både fra offentlig forvaltning og forskningsinstitusjoner) som ikke er å få tak i lenger, og at det er ønskelig at mest mulig av denne informasjonen digitaliseres og legges ut på Internett. En respondent mente imidlertid erfaringsmessig at kvalitetsforskjellen på vitenskapelige artikler og enklere fagrapporter er for stor til at data fra fagrapporter kan inkluderes som publiserte data.

3.1.6 Om behov for verktøy og tjenester



Figur 7: Om behov for verktøy og tjenester

I spørreundersøkelsen ga vi noen eksempler på mulige verktøy og tjenester for å laste ned data i brukervennlige formater og for å utføre enkle analyser. 77 % av respondentene ga uttrykk for at verktøy for å lage og laste ned utbredelseskart var mest nyttig for dem. 64 % av respondentene vektla nytten av verktøy som gjør det mulig å koble biologiske data med relevante miljødata (som f.eks. klimadata, vegetasjonsdata, geologiske data o.l.). 58 % vektla nytten av verktøy for å søke og laste ned rødlistedata og annen økologisk informasjon om enkeltarter. Om lag en tredjedel av respondentene var interessert i verktøy for validering av artsnavn.

Kun et mindretall av respondentene så nytte i analyseverktøy for beregning av biodiversitetsindeks (29 %) og for å søke og laste ned tidsseriedata (40 %). Slike dataanalyser utføres nok

mest av forskere, men det kan tenkes at mange av disse foretrekker å beregne indekser og utføre analyser i egne programpakker for å ha best mulig oversikt og kontroll med resultatene. Noen respondenter uttalte bekymring for at resultater fra slike online verktøy ville bli usikre og kunne bli misbrukt dersom kvaliteten på data og algoritmer ikke er god nok.

Respondentene uttrykte selv behov for andre online tjenester som f.eks. Artskontroll (muligheten for å få artsbestemt materiale på grunnlag av innsendt materiale/ foto) og en nasjonal oversikt over personer med kompetanse til å bestemme ulike organismegrupper. Generelt ble det uttrykt et stort behov for en større samordning av eksisterende/fremtidige nettjenester med Norge Digitalt og offentlige kartinnsynsløsninger.

3.1.7 Oppsummering av resultater

Generelt avdekker kartleggingen et stort behov for bedre oversikt og enklere tilgang til miljø- og biodiversitetsdata. Omlag 70 % av respondentene sier de er interessert i å ta i bruk nye kilder og verktøy for bruk av slike data. Forskere uttrykker størst behov for enklere tilgang til observasjonsdata, tidsseriedata og miljødata. Naturforvaltere er mest interessert i utbredelseskart for utvalgte arter. Respondentene uttrykte størst interesse for data fra terrestre økosystemer og ferskvannøkosystemer inkludert våtmark. Den noe lavere interessen for marinbiologiske data kan skyldes at disse allerede er bedre organisert og lettere tilgjengelig enn data fra terrestre og ferskvannøkosystemer. De ulike taksonomiske artsgruppene var omtrent like mye etterspurt, men undersøkelsen viser imidlertid at det er mest interesse for data om utvalgte arter og naturtyper. Videre var det spesielt stort behov for data om truede/sårbare og introduserte/invaderende arter. Det var generelt noe mindre behov for data om høstbare arter. De aller fleste respondentene vektla behovet for tilgang til gode metadata og brukervennlige verktøy for å kunne søke etter og laste ned data, lage utbredelseskart og koble biologiske data med miljødata. Om lag en tredjedel av respondentene var også interessert i verktøy for validering av artsnavn. Et mindretall av respondentene uttrykte behov for analyseverktøy for nedlasting av tidsseriedata og beregning av biodiversitetsindekser. Slike dataanalyser utføres nok mest av forskere, men det kan tenkes at mange av disse foretrekker å beregne indekser og utføre analyser i egne programpakker for å ha best mulig oversikt og kontroll med resultatene. Generelt ble det uttrykt et stort behov for en bedre samordning med Norge Digitalt og offentlige kartinnsynsløsninger.

3.2 Kartlegging av metadata (arbeidspakke 2)

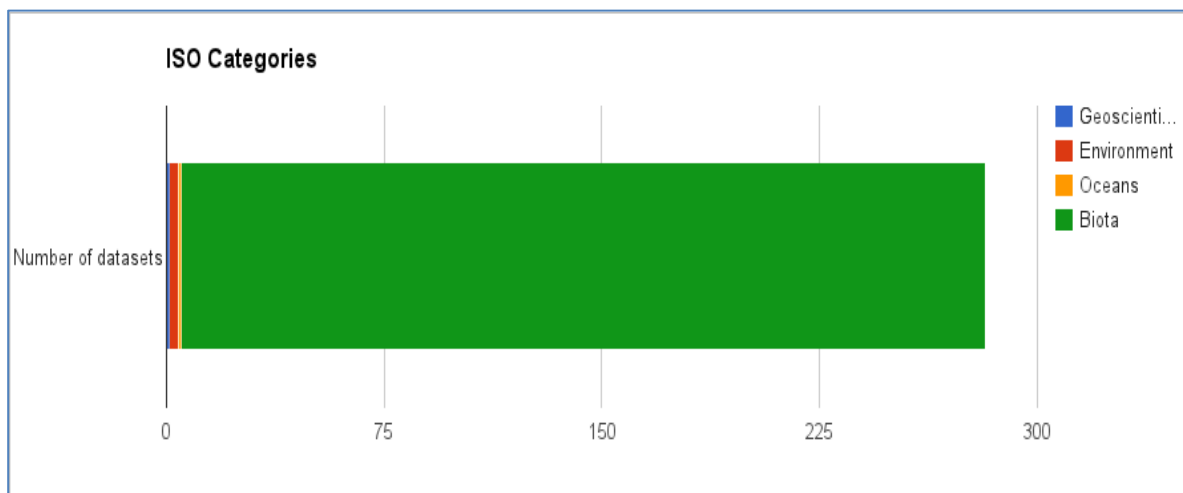
LifeWatch har gjennom prosjektperioden gjennomført en foreløpig kartlegging av biodiversitetsdatasett som er tilgjengelig i Norge. Kartleggingen brukte teknologi utviklet for NFR prosjektet Norsk Marint Datasenter (NMDC) for å samle og gjøre tilgjengelig resultatene. Metadata som ble samlet inn er tilgjengelig i portalen [Biodiversity datasets](#).

Selve kartleggingen foregikk todelt med en samling av de biodiversitetsmetadataene som Artsdatabanken hadde tilgjengelig, og gjennom direkte kontakt med 103 institusjoner og personer med spørsmål om registrering av datasett. Personene og institusjonene som ble kontaktet fikk tilsendt forespørselen på mail med et følgebrev som forklarte bakgrunnen for metadataregistreringen og hvordan den kunne gjennomføres. Gjennom denne prosessen ble 281 unike datasett registrert fra 31 Norske institusjoner og universiteter. Kartlagte metadata fra NFR-rapporten "Lange tidsserier for miljøovervåking og forskning", Rapport 1, 2 og 3 (NFR 2004) er ikke inkludert i denne metadatakartleggingen, men bør absolutt legges til grunn ved en eventuell realisering av en fullskala LifeWatch Infrastruktur i Norge.

Metadataverktøyet som ble brukt, er som tidligere nevnt utviklet for bruk i NMDC og gir mulighet til å skille registrerte metadata i ulike portaler. Verktøyet er [ISO 19115](#) kompatibel og er derfor kompatibel med gjeldende internasjonale initiativ, slik som Europeiske [INSPIRE](#) og [Amerikanske GCMD](#). Verktøyet er bygd opp for å kunne ta inn data på flere forskjellige måter,

både manuell registrering, importering av filer som inneholder metadata og høsting av data over [OAI-PMH protokollen](#). For registreringen av biodiversitetsmetadata for LifeWatch ble de to første metodene brukt.

Når man fordeler datasettene på ISO kategoriene som er registrert på datasettene får man en fordeling som viser at mesteparten av de registrerte datasettene som har blitt registrert har en biologisk tilknytning.



Figur 8: Datasett fordelt per ISO- kategori

Den påfølgende tabellen viser fordelingen av datasett på nøkkelord fra Global Change Master Directory (GCMD). Et datasett kan kobles mot flere nøkkelord så antall nøkkelord registrert er høyere enn totalt antall datasett.

GCMD KEYWORDS	NUMBER OF DATASET
EARTH SCIENCE > AGRICULTURE > AGRICULTURAL AQUATIC SCIENCES > FISHERIES	1
UNKNOWN > UNKNOWN > INVERTEBRATES	3
UNKNOWN > UNKNOWN > ANIMALS	4
UNKNOWN > UNKNOWN > PLANTS	37
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES	1
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > ARTHROPODS > CHELICERATES > ARACHNIDS	4
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > ARTHROPODS > CRUSTACEANS > OSTRACODS	12
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > ARTHROPODS > HEXAPODS > ENTOGNATHA	2
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > ARTHROPODS > HEXAPODS > INSECTS	42
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > ARTHROPODS > MYRIAPODS > MILLIPEDES	3
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > BRYOZOANS/MOSS ANIMALS	2
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > CNIDARIANS > JELLYFISHES	7
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > ECHINODERMS > SEA URCHINS	3
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES >	1

FLATWORMS/FLUKES/TAPEWORMS	
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > LAMP SHELLS	3
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > MOLLUSKS > GASTROPODS	13
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > SEGMENTED WORMS (ANNELIDS) > EARTHWORMS	1
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > SEGMENTED WORMS (ANNELIDS) > LEECHES	7
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > SPONGES	7
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/INVERTEBRATES > TUNICATES > SEA SQUIRTS	3
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/VERTEBRATES > AMPHIBIANS > SALAMANDERS	7
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/VERTEBRATES > BIRDS	1
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/VERTEBRATES > BIRDS > WADERS/GULLS/AUKS AND ALLIES	15
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/VERTEBRATES > FISH > SHARKS/RAYS/CHIMAERAS	6
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/VERTEBRATES > MAMMALS > CARNIVORES > SEALS/SEA LIONS/WALRUSES	1
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/VERTEBRATES > MAMMALS > RODENTS	9
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > ANIMALS/VERTEBRATES > REPTILES > TURTLES	7
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > FUNGI > LICHENS	22
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > FUNGI > MUSHROOMS	25
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > PLANTS > MACROALGAE (SEAWEEDS)	1
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > PLANTS > MICROALGAE > HAPTOPHYTES > COCCOLITHOPHORES	7
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > PLANTS > MOSESSES/HORNWORTS/LIVERWORTS	15
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > PROTISTS > MACROALGAE (SEAWEEDS) > GREEN ALGAE	1
EARTH SCIENCE > BIOLOGICAL CLASSIFICATION > PROTISTS > MACROALGAE (SEAWEEDS) > RED ALGAE	7
EARTH SCIENCE > BIOSPHERE > AQUATIC ECOSYSTEMS > COASTAL HABITAT	1
EARTH SCIENCE > BIOSPHERE > AQUATIC ECOSYSTEMS > PLANKTON > PHYTOPLANKTON	1
EARTH SCIENCE > BIOSPHERE > AQUATIC ECOSYSTEMS > PLANKTON > ZOOPLANKTON	1
EARTH SCIENCE > BIOSPHERE > ECOLOGICAL DYNAMICS > COMMUNITY DYNAMICS > BIODIVERSITY FUNCTIONS	1
EARTH SCIENCE > OCEANS > COASTAL PROCESSES > SEDIMENTATION	1
EARTH SCIENCE > OCEANS > MARINE ENVIRONMENT MONITORING	1
UNKNOWN > UNKNOWN > VERTEBRATES	1
UNKNOWN > UNKNOWN > PLANTS	1
UNKNOWN > UNKNOWN > MICROALGAE, MACROALGAE (SEAWEEDS),	1

Tabell 1: Fordelingen av datasett på nøkkelord fra Global Change Master Directory (GCMD)

3.3 Prinsipper, lover og retningslinjer for deling av miljø- og biodiversitetsdata (arbeidspakke 3)

I det følgende vil vi forsøke å beskrive dagens utfordringer knyttet til datadeling i relasjon til gjeldende nasjonale og internasjonale prinsipper, lover og retningslinjer for deling av miljø- og biodiversitetsdata.

Store mengder offentlig finansierte miljø- og biodiversitetsdata er i dag enten utilgjengelig eller vanskelig tilgjengelig for samfunnet. Dersom disse datamengdene gjøres fritt tilgjengelig på standardiserte formater vil nytteverdien av dataene langt overgå de bakenforliggende investeringer. I tillegg til dette oppnås en åpnere forvaltning og økt samhandling i offentlig sektor.

For å kunne realisere en slik datadeling er det nødvendig å se på gjeldende prinsipper, lover og retningslinjer som regulerer deling av data. Med noen få unntak er hovedregelen at offentlige finansierte data alltid skal gjøres fritt tilgjengelig for samfunnet. Ulike tekniske, organisatoriske, finansielle og juridiske forhold kan likevel virke begrensende på de ulike aktørenes evne og vilje til å dele data. Disse barrierene kommer spesielt til uttrykk i forholdet mellom offentlig og privat instituttsektor hvor de økonomiske rammevilkårene er svært ulike. Det er også viktig at arbeidet med å gjøre mer data åpent tilgjengelig skjer med utgangspunkt i samfunnets generelle og forskningens spesielle behov for kvalitetssikrede og relevante datasett.

Prinsipper, lover og retningslinjer som i særlig grad regulerer forhold rundt deling av data i Norge er [OECD- prinsippene om tilgang til offentlig finansierte data](#) (OECD Principles and Guidelines for Access to Research Data from Public Funding), [EU Viderebruksdirektiv](#) (Directive 2003/98/EF), [Offentlighetsloven](#) (Lov 19. mai 2006 nr. 16 om rett til innsyn i dokument i offentlig virksomhet), [INSPIRE- direktivet](#) (Directive 2007/2/EC of the European Parliament and of the Council of 15 May 2007 establishing an Infrastructure for Spatial Information in the European Community), [Geodataloven](#) (LOV 2010-09-03 nr. 56: Lov om infrastruktur for geografisk informasjon), [Geodataforskriften](#) (FOR 2012-08-08 nr. 797: Forskrift om infrastruktur for geografisk informasjon), [Miljøinformasjonsloven](#) (LOV 2003-05-09 nr. 31: Lov om rett til miljøinformasjon og deltakelse i offentlige beslutningsprosesser av betydning for miljøet) og [Naturopsynsloven](#) (LOV 1996-06-21 nr. 38: Lov om statlig naturopsyn).

I 2004 undertegnet kunnskapsministrene fra alle nasjoner av OECD (Organisasjonen for økonomisk samarbeid og utvikling), som inkluderer de fleste utviklede land i verden, en erklæring som i hovedsak sier at alle offentlig finansierte data bør gjøres offentlig tilgjengelig. Etter en forespørsel og en intens diskusjon med dataproduserende institusjoner i medlemslandene, publiserte OECD i 2007 OECDs prinsipper og retningslinjer for tilgang til forskningsdata fra offentlig støtte som en soft-Law anbefaling. Det norske Kunnskapsdepartementet har anerkjent OECD prinsippene og anbefaler norske institusjoner [å gjennomføre disse](#).

3.3.1 OECD- prinsippene

OECD- prinsippene er retningsgivende og har som hovedformål å sikre en åpen, bærekraftig og kostnadsoptimal tilgang til offentlig finansierte data. Prinsippene setter krav til infrastruktur for deling av data med tanke på interoperabilitet, fleksibilitet, kvalitet, dokumentasjon (meta-data), samt opplærings- og belønningsmekanismer. OECD påpeker behovet for datadelingsregler og anbefaler at disse blir samordnet med retningslinjer for nasjonale forskningsprogram og tilpasset hensynet til blant annet nasjonal sikkerhet, opphavsrett, konkurranseopplysninger, navngivelse, eiendomsrett, videre deling, bruksbegrensninger, finansielle avtaler, etiske regler, lisensavtaler, juridiske avtaler og bærekraftig arkivføring. OECD presiserer at opphavsrett knyttet til privat dataprodusent i offentlige finansierte prosjekt ikke nødvendigvis skal være begrensende for åpen deling av data. I samarbeidsprosjekter anbefaler OECD at ansvarsforhold og regelverk for datadeling formaliseres så tidlig som mulig i oppstarten. OECD påpeker behovet for at datadeling må utformes i samsvar med gjeldende etiske standarder og verdier i forsk-

ningsmiljøene. Oppdragstaker kan for en kortere periode gis eksklusiv tilgang på innsamlede data for egen vitenskapelig publisering, men dette skal avtales eksplisitt mellom finansieringspart og forskningsmiljø i oppstart av et prosjekt. OECD vektlegger at det offentlige bør sikre tilstrekkelige finansielle rammevilkår for oppdragstakers dataforvaltning. Forskningsmidler skal ikke finansiere dataforvaltning. Eventuelle uttakskostnader bør ikke overskride faktiske påløpte kostnader ved uttak av data.

3.3.2 Viderebruksdirektivet

EU vedtok i 2003 Viderebruksdirektivet i den hensikt å utnytte det store potensialet til viderebruk av data fra offentlig sektor. Direktivet forsterker Norges juridiske forpliktelser til å dele data fra offentlig sektor gjennom minimumsregler for tilgang til offentlig finansierte data for innsyn og viderebruk. Direktivet ble ratifisert av Norge og nedfelt i Offentlighetsloven fra januar 2009 og Fellesføringsene fra Fornyings-, Administrasjons- og Kirkedepartementet (2012).

3.3.3 Offentlighetsloven

Offentlighetsloven er ment å sikre medborgernes innsynsrett gjennom en åpen og transparent offentlig virksomhet. Offentlighetsloven implementerer regler om innsyn og prising koblet til dette, og går noe lenger ved å fastsette at betaling kun kan kreves i særlige tilfeller, og ikke skal overstige de reelle kostnadene knyttet til leveringen.

3.3.4 Geodataloven, Geodataforskriften og INSPIRE

Norge har gjennom Geodataloven og tilhørende Geodataforskrift implementert EU-direktivet INSPIRE (Infrastructure for Spatial Information in the European Community) som skal realisere en felles internasjonal infrastruktur for geografisk informasjon. Direktivet pålegger offentlige myndigheter å dele geodata elektronisk, og at de skal kunne søkes i, vises og lastes ned av allmenheten. Statens kartverk er som Nasjonal geodatakoordinator ansvarlig for gjennomføring av direktivet. Geodataloven skal bidra til god og effektiv tilgang til offentlig geografisk informasjon (geodata). Spesifiserte geodata skal deles mellom offentlige myndigheter, som i tillegg har et ansvar for metadata, samvirkingsevne, koordinering og kontroll. Rettssubjekter (dvs. offentlige organ eller andre som utfører lovpålagte forvaltningsoppgaver) er videre forpliktet til å opprette og drive et felles nett av offentlige tjenester til søk, visning, nedlastning, omforming og aktivering av geodata. Søkjetjenestene skal være gratis og også selve dataene skal som hovedregel være gratis tilgjengelige. Avslag på innsyn i et bestemt datasett vil gjerne være fattet med hjemmel i annet lovverk som f.eks. Offentlighetsloven, Miljøinformasjonsloven, m.fl.

3.3.5 Miljøinformasjonsloven

Miljøinformasjonsloven skal sikre allmenheten tilgang til miljøinformasjon, og dermed gjøre det enklere for allmenheten å verne miljøet, verne publikum mot helse- og miljøskade og å gi innsyn i beslutningsprosesser knyttet til miljøet. Forvaltningsorganer plikter å ha relevant miljøinformasjon, og gjøre den tilgjengelig for allmenheten. Unntak kan gis i h.h.t. Offentlighetsloven dersom begrenset tilgang veier tyngre enn hensynet til miljø- og samfunnsinteresser.

3.3.6 Naturoppsynsloven

Naturoppsynsloven pålegger enhver med oppsynstjeneste taushetsplikt om sårbart miljø med miljøverdier med hjemmel i Forvaltningsloven. Ved viderebruk av data overføres taushetsplikt til mottaker av data. Sensitive opplysninger som unntas offentlighet kan skjermes med begrenset tilgang eller aggregert stedfesting. Krav om tilgang på taushetsbelagt informasjon hjemles

normalt i Miljøinformasjonsloven og Offentlighetsloven. Aktører med naturforvaltningsansvar bør alltid få tilgang, og dette regnes ikke som offentliggjøring.

4 Drøfting

Generelt vil åpen deling av data umiddelbart synliggjøre både data, metoder og dataproducent for omverdenen. Forvaltningen og forskningssektoren samler til sammen et svært bredt spekter av data i sin oppgaveløsning. Ved å gjøre disse dataene åpent tilgjengelig kan en bidra til økt næringsutvikling, verdiskapning, innovasjon, forskning, effektivisering, gjennomsiktighet og demokratisering av samfunnet. Deling av data sikrer også at data vil kunne kvalitetssikres av flere eksterne aktører.

Forprosjektet LifeWatch Norge har (i likhet med tidligere utredninger) avdekket store brukerbehov for lettere tilgang til offentlig finansierte miljø- og biodiversitetsdata på tvers av institusjonsgrensene og landegrensene. Erfaringene med institusjonenes samarbeid om deling av data gjennom Artsdatabanken og GBIF har vært gode, og tilfredsstillende i dag delvis brukernes behov. Vår brukerundersøkelse viser imidlertid at det er et stort udekket behov for tilgang på flere typer observasjonsdata, samlingsdata og tidsseriedata. Det er også stor etterspørsel etter miljø- og klimadata da dette bare delvis er tilgjengelig fra dataeierne og Norge Digitalt.

De senere år har gitt samfunnet stadig bedre tilgang på offentlig sektorbasert kartinformasjon gjennom et mangfold av nettbaserte karttjenester og distribusjonskanaler for nedlasting av data. Flere av våre respondenter ga uttrykk for at det er behov for en bedre samordning mellom de ulike distribusjonskanalene. Bedre tilgang til data i seg selv dekker ikke forskningens og forvaltningens behov i sin helhet. Mange respondenter ga også uttrykk for at de har behov for differensiert funksjonalitet til f.eks. aggregering av data, utforming av utbredelseskart, statistiske analyser og muligheter for å koble observasjonsdata med miljødata og andre kartlag for videre GIS-analyser, GIS-modellering og numerisk modellering.

4.1 Begrepene Åpne data og Åpen tilgang

“Åpne data” er ideen om at data bør være fritt tilgjengelig for alle å bruke og publisere som de ønsker, uten restriksjoner fra opphavsrett, patenter eller andre mekanismer for kontroll. Målene for den åpne data bevegelsen er lik de andre “Åpne” bevegelser som åpen kildekode, åpent innhold, og åpen tilgang. Filosofien bak åpne data har lenge vært etablert, men begrepet “åpne data” i seg selv er nylig blitt mer populært med fremveksten av Internett og World Wide Web og spesielt med lanseringen av “Åpne data” statlige tiltak som <http://data.norge.no>.

Begrepet “Åpen tilgang” til vitenskapelige data ble institusjonelt etablert med dannelsen av [World Data Center System \(WDS\)](#), som forberedelse til det internasjonale geofysiske år i 1957-1958. International Council of Scientific Unions (nå [International Council for Science](#)) etablerte flere verdensdatasentre for å minimere risikoen for tap av data og for å maksimere data tilgjengelighet. Det kom senere på 1950-tallet ytterligere anbefalinger om å gjøre data tilgjengelig i maskinlesbar form. Den “Åpne science-data» bevegelsen var langt forut for Internett. I dag har tilgjengeligheten av raske, allestedsnærværende nettverk vesentlig endret innholdet i åpne vitenskapelige data. [ICSU \(International Council for Science\) World Data System Science Steering Committee](#) har nylig blitt enig om at WDS skal være en ledende global organisasjon i å utvikle og administrere et akkreditert rammeverk med tilhørende tjenester for vitenskapelige datasentre og datatjenesteleverandører, og at det skal være et koordinerende forum for utvikling, spesifisering og kunngjøring av beste praksis standarder og metoder for datapublisering og datainteroperabilitet.

4.2 Begrepet Maskinlesbart format

Begrepet “maskinlesbart format” har gjennom flere år vært brukt som et slags anbefalt standard format i tidligere kontrakter mellom oppdragsgiver og oppdragstaker om deling av data. Dette begrepet har så langt vi kjenner til aldri hatt noen entydig definisjon i Norge før Direktoratet for Forvaltning og IKT (Difi) definerte det som: “*Dataformater som kan leses av en datama-*

skin. Eksempler på maskinlesbare dataformat er XML, JSON og CSV. Slike formater kan lett deles på tvers av ulike IT-systemer (<http://veileder.data.norge.no/Vedlegg/begreper-og-definisjoner.html>).

Dataprodusentene har gjennom årene etablert sine data på maskinlesbart format med ulike filformater, databaseplattformer og nettbaserte innsynsløsninger etter hvert som datateknologi og internasjonalt standardiseringsarbeid har utviklet seg.

4.3 Deling av nyetablerte og historiske datasett

Etter at Artsdatabanken ble operativ med sin Artskartløsning (2007) er det blitt mer vanlig at mange oppdragsgivere i kontrakts form krever at oppdragstaker skal publisere nyetablerte data gjennom Artskart. Dette gjør det lettere for oppdragstaker å budsjettere for dataforvaltningskostnader i sine innledende tilbud.

Tilrettelegging for deling av historiske data (både analoge og diverse maskinlesbare formater) via standardiserte dataprotokoller vil ofte utgjøre et stort løft for mange institusjoner. Så lenge oppdragstaker har overholdt sine tidligere kontraktforpliktelser (med krav om leveranser av data til oppdragsgiver på maskinlesbart format) bør kostnader knyttet til tilgjengeliggjøring via standardiserte dataprotokoller være myndighetenes overordnede ansvar og ikke den enkelte institusjon. Dersom den opprinnelige prosjekt- og oppdragsfinansieringen ikke er tilstrekkelig for å innløse og realisere kravet om offentliggjøring av data, bør oppdragsinstitusjonene ta dette opp med oppdragsgiver for å klargjøre tilleggsfinansiering.

Dersom relevant offentlig finansiering blir tilgjengelig til dette formål, bør den enkelte dataprodusent ta ansvar for å prioritere dette arbeidet i sin virksomhetsplanlegging. Det fins allerede flere offentlige støtteordninger med relevans som Artsdatabankens "Tilskudd til kvalitetssikring og tilrettelegging av stedfestede artsdata" og Miljøverndepartementets "Infrastrukturmidler til instituttsektoren". Sistnevnte støtteordning er forbeholdt miljøinstituttene. For at dataprodusentene skal bli i stand til å møte samfunnets økende databehov bør imidlertid rammen for disse støtteordningene utvides tilstrekkelig.

4.4 utfordringer og barrierer knyttet til deling og tilgjengeliggjøring av data

Det eksisterer flere utfordringer og barrierer knyttet til økt deling og tilgjengeliggjøring av offentlig finansierte miljø- og biodiversitetsdata. Disse er i hovedsak relatert til institusjonenes manglende harmonisering av teknologi og standarder, ulike økonomiske rammevilkår, ulike bedriftskulturer, ulike oppdragsgivere og finansieringskilder, forskernes holdninger til deling av data og metadata (informasjon om data), forskernes behov for akademisk merittering ved deling av data, manglende strategier for datadeling samt mangel på adekvate kontraktmal og tilstrekkelig oppfølging av disse fra oppdragsgiver. Vilkår og strategier for deling og tilgjengeliggjøring av private data og utenlandske offentlige data (f.eks. fra våre naboland) er ikke vurdert i denne sammenheng.

4.4.1 Teknologi, standarder og økonomiske rammevilkår

Harmonisering av teknologi og standarder på tvers av institusjonsgrenser kan oppnås ved implementering av standardiserte databaser og interoperable dataprotokoller. Dette er i bunn og grunn et praktisk, teknisk forhold som i stor grad vil styres av den enkelte institusjons prioriteringer, økonomiske rammevilkår og interne IKT- kompetanse.

Mange institusjoner vil gjerne dele alle sine data, men har ikke økonomisk evne til å prioritere dette og iverksette nødvendige tiltak. Dette gjelder spesielt de ikke- statlige forskningsinstitut-

tene hvor kostnader til forvaltning av nyere data utelukkende dekkes inn gjennom dedikerte midler i forskningsprosjekter eller som tilskuddsordninger gjennom Miljøverndepartementet og Kunnskapsdepartementet via Artsdatabanken. Tilrettelegging av historiske datasett for online datautveksling vil medføre ekstra kostnader som privat forskningssektor ikke har dekning for og som derfor kan gå på bekostning av forskningsarbeidet. Situasjonen er uten tvil annerledes i offentlig forvaltning og forskningssektor, hvor budsjetttrammene gir større handlingsrom for en mer helhetlig dataforvaltning.

4.4.2 Bedriftskultur og holdninger

Ulike bedriftskulturer kan være en utfordring for nasjonal fremdrift i arbeidet med åpen deling, tilgjengeliggjøring og dokumentasjon (metadata). Dersom bedrifter eller enkeltpersoner i bedriftene har avvikende agendaer, kan dette utgjøre en direkte barriere. Dette er et organisatorisk ansvar som påligger den enkelte institusjon. Det er ikke akseptabelt for samfunnet at individuelle avgjørelser på personnivå utgjør en barriere for åpen deling og tilgang til offentlig finansierede data. Den enkelte institusjon bør utvikle strategier og datapolitikk for hvordan data og metadata skal håndteres internt og eksternt.

Gjennom forprosjektets kartlegging av metadata i arbeidspakke 2 opplevde vi en noe lavere respons fra fagmiljøene enn forventet. Delvis kan dette skyldes at grensesnittet for vår metadatakartlegging er basert på den internasjonale metadatastandard ISO 19115/139 (kompatibel med Europeiske INSPIRE og Amerikanske GCMD). Flere dataprodusenter har i ettertid gitt uttrykk for at standardens krav til registrering av metadata er altfor omfattende, komplisert og tidkrevende. Det er derfor betydelig grunn til å anta at den lave responsen delvis kan forklares ut fra dette. Men lav respons på forespørsler om metadataregistrering er også blitt rapportert fra lignende prosjekter (Schmidt-Kloiber *et al.* 2012). Motviljen mot å rapportere metadata blant dataprodusenter kan ha mange forklaringer, inkludert de følgende:

1. Å oppgi detaljert og tilstrekkelig dokumentasjon for at data skal kunne gjenbrukes krever arbeidstid, og derfor økonomiske ressurser, som ofte er begrenset hos forskere. Mange forskere anser derfor at slik tidsbruk vil gå på bekostning av "ekte" forskning. En mulig løsning kan være at oppdragsgivere stiller betingelser om at metadata (og eventuelt data) fra nye prosjekter skal lagres og gjøres offentlig tilgjengelig, og i tillegg bevilger ekstra ressurser til slike aktiviteter. Beniston *et al.* (2012) har foreslått implementering av retningslinjer, eller til og med et EU-direktiv om god styring av datadeling, inkludert generelle regler for dataformater og standarder, datalagring etter fullføringen av et prosjekt, og tilgang til dataene. En annen mulighet kan være andre slags incentiver enn økonomiske som kan oppveie for arbeidsbelastningen ved å rapportere metadata. I sosiale nettverk beregnet på forskere (f.eks. www.mendeley.com og www.researchgate.net) lagrer enkeltpersoner informasjon om egne publikasjoner, ferdigheter m.m. Det er mulig forskere kan motiveres til også å lagre informasjon om tilgjengelige datasett gjennom slike kanaler.
2. Noen forskere frykter at å informere om metadata innebærer å gi avkall på den intellektuelle eiendomsretten til selve dataene. Det er derfor viktig å understreke at å publisere metadata er ikke nødvendigvis er det samme som å gjøre dataene automatisk tilgjengelig for andre. En metadata-base bør videre ha grundig informasjon om rettigheter, kriterier for bruk av dataene, kontaktinformasjon for dataeiere, osv.
3. Dataeiere kan frykte at deres data vil bli brukt til upassende formål. Det forekommer en bekymring for at andre forskere bruker dataene på egen hånd, med risiko for feiltolkning og påfølgende gale slutninger, og/eller at de selv ikke vil motta noen anerkjennelse for bruken av dataene. En metadata-base skal imidlertid først og fremst informere potensielle brukere om datasettenes eksistens og opphav. Slik informasjon kan resultere i nye samarbeidsprosjekter mellom dataeiere og nye brukere, som også kan komme dataeierne til gode.

I en del tilfeller fører usikkerhet og lavt kunnskapsnivå om konsekvenser av åpen deling til at forskere og dataeiere blir skeptiske. Direktoratet for Forvaltning og IKT (Difi) har i sitt arbeid med "[Veileder for tilgjengeliggjøring av offentlige data](#)" hatt mange henvendelser om dette: "Vi vet ikke hva dataene våre kommer til å bli brukt til"; "Kvaliteten på dataene våre er for dårlig"; "Pris og kostnader ved tilgjengeliggjøring blir for høye"; "Tilgjengeliggjøring av data er utenfor kjerneaktivitetene våre"; "Vi har ingen kontroll over hvilke datasett som settes sammen med våre egne"; "Dataene våre inneholder personopplysninger"; "Tilgjengeliggjøring av data er avslørende" og "Det er for komplisert!". Difi-veilederen er et tiltak for å støtte offentlige virksomheter i arbeidet med å gjøre data tilgjengelig for gjenbruk. Veilederen gir en innføring i hvordan offentlige data kan gjøres tilgjengelig for viderebruk, og inneholder blant annet eksempler på hva åpne data kan brukes til, argumenter for hvorfor offentlige virksomheter bør dele sine data og praktiske råd til hvordan det kan gjøres. Veilederen er spesielt rettet mot offentlig sektor, men kan selvfølgelig bli brukt av alle. Vår vurdering er at den ikke er tilstrekkelig relevant for forskningsdata.

4.4.3 Behovet for akademisk merittering ved deling av data

Åpen deling av forskningsdata har endel andre utfordringer enn deling av rene forvaltningsdata. Akademisk merittering er viktig for forskernes karriere og anerkjennelse. Dette er tradisjonelt oppnådd gjennom formidling av resultater i vitenskapelige forum og anerkjente publikasjoner. Per i dag er ikke åpen deling av data direkte akademisk meritterende for en forskers karriere. Indirekte er det ingen tvil om at åpenhet vil gavne den enkelte forsker både med hensyn til ekstern kvalitetssikring av åpne data, synliggjøring av eget arbeid og ikke minst fri tilgang til andres data.

Fremveksten av elektronisk publisering av litteratur har identifisert behovet for mekanismer for sitering av "online" referanser på måter som kan sikre gjenkjenning og gjenfinning i mange år inn i fremtiden. Elektronisk publisering av datasett gir relaterte, men mer komplekse utfordringer. Det avhenger av evnen til pålitelig å kunne identifisere, lokalisere, få tilgang til, tolke og verifisere versjon, integritet, og eierskap til digitale datasett.

Datasiteringsstandarder og god praksis kan danne grunnlag for økt interesse, anerkjennelse og belønning for vitenskapelige data. Den raskt voksende mengden av tilgjengelige digitale data gir mulighet til å tillate undersøkelse og vurdering av konklusjoner eller analyser basert på data (både eksperimentelle og innsamlede). Tilgjengeliggjøring av data gjennom nye former for vitenskapelig publisering gir nye og uforutsette brukere muligheten til å kombinere datasett eller utføre nye analyser på datasettene. Problemet med å sitere online data kompliseres av mangel på etablert praksis for å henvise til deler eller delsett med data. Det er derfor viktig at nødvendige internasjonale standarder, insentiver og konvensjoner for å støtte datasitat, bevaring og tilgjengelighet blir etablert. En rekke tiltak i ulike organisasjoner, land og disipliner er allerede under utvikling med dette som formål.

Et viktig sett av tekniske og politiske tilnærminger på dette området har allerede blitt lansert av den amerikanske [National Information Standards Organization \(NISO\)](#). I tillegg har en del andre standardiseringsorganer uttalt seg om permanente identifikatorer og tilgjengeliggjøring av disse. Permanente identifikatorer (PI) er lett vedlikeholdsvennlige identifikatorer som tillater å referere til et digitalt objekt - en fil eller et sett med filer, for eksempel en e-print (artikkel, papir eller en rapport), et bilde eller en installasjonsfil for en programvare. De eneste interessante permanente identifikatorene er de som er permanent aktive (som betyr, du kan "klikke" på dem).

I motsetning til en enkel hyperkobling skal permanente identifikatorer fortsette å gi tilgang til ressursen, selv når den flyttes til en annen tjener eller til en annen organisasjon. Det er flere "standarder" som er i en moden fase av utviklingen:

- Uniform Resource Name (URN)

- Persistent URL (PURL)
- Digital Object Identifier (DOI)
- National Bibliography Numbers (NBNS)
- Archival Resource Key (ARK)
- Open URL
- Universally Unique Identifier (UUID)

[Digital Object Identifier \(DOI\)](#) ble en ISO-standard i mai 2012. DOI-systemet ble initiert av Den internasjonale DOI- stiftelsen i 1998 (non-profit organisasjon initiert av flere forlag og organisasjoner). DOI- Stiftelsen er autorisasjonskontor for ISO 26324. Hittil har cirka 60 millioner DOI-navn blitt tildelt gjennom et voksende forbund av registreringsbyråer over hele verden. ISO 26324:2012 gir syntaks, beskrivelse og oppløsning for funksjonelle komponenter av Digital Object Identifier- systemet. Standarden gir også de generelle prinsippene for etablering, registrering og administrasjon av DOI navn.

Flere [CODATA- komiteer](#) (Committee on Data for Science and Technology), [ICSTI](#) (International Council for Scientific and Technical Information) og representanter fra ulike organisasjoner har i fellesskap satt sammen gruppen "[The Task Group on Data Citation Standards and Practices](#)" for å undersøke en rekke sentrale problemstillinger knyttet til sitering av data, bidra til å koordinere aktiviteter i dette området og fremme felles praksis og standarder i det vitenskapelige miljøet.

Andre viktige grupper er [DataCite initiativet](#) samt den sammensatte gruppen fra [SCOR](#) (Scientific Committee on Oceanic Research), [IODE](#) (International Oceanographic Data and Information Exchange), [MBLWHOI](#) (Marine Biological Laboratory/Woods Hole Oceanographic Institution Library) som møtes årlig for å diskutere implementeringen av datalager i ulike datasentre og samarbeid med tilhørende nasjonal og internasjonal innsats, og høre om hvordan datapublikasjon blir håndtert i andre disipliner og samhandling med utgivere av vitenskapelige tidsskrifter. Denne gruppen gjennomførte i 2012 en undersøkelse om informasjonskilder som er direkte eller perifert fokusert på siteringspraksis og henvisninger for digitale data. En [bibliografi](#) er gjort tilgjengelig online. Flere eksperter i Storbritannia, Canada, USA, Tyskland og Norge er intervjuet om deres praksis i forhold til sitering av data. Resultatet viste varierende kunnskap og praksis om temaet.

[GBIF](#) og [Pensoft Publishers](#) lanserte i 2012 et nytt [verktøy for publisering av biodiversitetsdata](#). Dette verktøyet genererer faktaark (i journals form) om hvert enkelt datasett basert på registrerte metadata som publiseres til GBIF-portalene. Dette faktaarket kan så sendes til vurdering for publisering i en vitenskapelig journal. Per i dag kan man publisere slike faktaark i Pensoft journalene [PhytoKeys](#), [ZooKeys](#), [BioRisk](#) og [NeoBiota](#). To nye journaler [Biodiversity Data Journal](#) og [Nature Conservation](#) vil om kort tid også akseptere publisering av slike faktaark om data.

4.4.4 Dataforvaltning, datastrategier og kontraktsforhold

Dataforvaltning og tilgjengeliggjøring av offentlig finansierte data er et institusjonelt ansvar som bør være basert på bevisste strategivalg for fremtiden. Datastrategi(er) må forankres i organisasjonen og brukes som et styringsverktøy for videre planlegging og fremtidige prioriteringer.

Internasjonalt har ulike finansieringskilder for vitenskapelig forskning nå begynt å kreve dataforvaltningsplaner som en del av deres valg og godkjenningprosesser. En dataforvaltningsplan vil hjelpe forskeren og hans institusjon med å synliggjøre alle relevante aspekter innen dataforvaltning. Dette er en sjekklister over hva en må ta stilling til for å sikre at data forvaltes i henhold til pålagte krav fra lovverk, standarder og oppdragsgivere. Det internasjonale [Data Curation Center \(DCC\)](#) i Storbritannia har laget en [mal](#) for hva som bør inkluderes i en dataforvaltningsplan. DCC har også laget en [online løsning](#) for å lage og vedlikeholde en dataforvaltningsplan. Eksempler på dataforvaltningsplaner er også tilgjengelig fra [Australian](#)

[National Data Service \(ANDS\)](#). I Norden har Danmark opprettet et [nasjonalt dataarkiv for forskningsdata](#).

Med hensyn til status for dataforvaltningsplaner og strategiutvikling i Norge er vårt inntrykk at det er svært stor variasjon blant dataeierne på dette området. Så vidt vi kjenner til fins det heller ingen nasjonale maler for dataforvaltningsplaner i Norge. Selv om de ulike dataprodusentene er svært ulike kunne en felles mal hjulpet flere på vei. GBIF har utviklet anbefalte retningslinjer for strategiutvikling og handlingsplaner for publisering av biodiversitetsdata ("[Best practice guide for Data Discovery and Publishing Strategy and Action Plans](#)"). Hensikten med en slik strategi må være at det er institusjonen som avgjør hvordan data skal forvaltes og deles, og ikke den enkelte ansatte. Dette er en utfordrende problemstilling spesielt innen forskningsmiljøene. Strategien må utformes på en slik måte at den synliggjør ansvarsforhold, rettighetsproblematikk, nytteverdi og synergier ved datadeling.

For å unngå misforståelser i ettertid er det veldig viktig at ansvarsforhold og regelverk for datadeling formaliseres så tidlig som mulig i oppstarten av et kartleggings- eller forskningsprosjekt. Dette pålegger oppdragsgiver et betydelig ansvar for utarbeiding av konsise kontraktmaler og oppfølging av disse i forhold til oppdragsgiver. Oppdragstaker forpliktes gjennom dette til å samle inn, tilrettelegge og dokumentere data på en standardisert måte.

4.5 Juridiske og administrative rammevilkår for deling av offentlig finansierte data

I henhold til lovverket gjelder pålegget om åpen deling av offentlige finansierte data kun offentlig forvaltning eller andre som utfører lovpålagte forvaltningsoppgaver. Dette er hjemlet i Geodataloven som gir hjemmel for å kreve at offentlige organ eller andre som utfører lovpålagte forvaltningsoppgaver skal opprette og drive et felles nett av offentlige tjenester til søk, visning, nedlastning, omforming og aktivering av geodata. Slike tjenester og tilhørende data skal som hovedregel være gratis og offentlig tilgjengelig. Videre sikrer Offentlighetsloven samfunnet innsynsrett så lenge annet forbehold om dette ikke er hjemlet i annet lovverk. Miljøinformasjonsloven kan også gi hjemmel om innsyn.

Forskningsinstitusjoner som ikke utfører lovpålagte forvaltningsoppgaver er ikke rettssubjekter for Geodataloven. Med unntak av eksterne krav hjemlet i Offentlighetsloven og Miljøinformasjonsloven er ikke slike institusjoner lovpålagt å dele offentlig finansierte data dersom det ikke foreligger spesifikke krav om dette i kontrakts form.

OECD- prinsippene bør legges til grunn for hvordan forskningsinstitusjonene skal forholde seg til deling av offentlig finansierte data. Behovet for datadelingsregler understrekes av OECD, og at disse må samordnes med retningslinjer for nasjonale forskningsprogram. OECD presiserer at opphavsrett knyttet til privat dataprodusent i offentlige finansierte prosjekt ikke nødvendigvis skal være begrensende for åpen deling av data. Det anbefales at ansvarsforhold og regelverk for datadeling formaliseres så tidlig som mulig i oppstarten av et samarbeidsprosjekt. Dette må imidlertid tilpasses gjeldende etiske standarder og verdier i forskningsmiljøene. En oppdragstaker kan gis tidsbegrenset eksklusiv tilgang på data for vitenskapelig publisering. Dette skal imidlertid avtales i oppstarten av et prosjekt. For å øke forskningsinstitusjonenes evne og kapasitet til å dele data, vektlegger OECD at det offentlige bør sikre tilstrekkelige finansielle rammevilkår slik at kostnader til dataforvaltning ikke går ut over prioriterte forskningsoppgaver.

Taushetsplikt om sårbare miljøverdier er hjemlet i Forvaltningsloven. Ved viderebruk av data overføres taushetsplikt til mottaker av data. Sensitive opplysninger som unntas offentlighet kan skjermes med begrenset tilgang eller aggregert stedfesting. Håndtering av sensitive natur- og miljødata er strengt regulert gjennom Miljøinformasjonsloven og Naturoppsynsloven. Forvaltningsorganer plikter å ha relevant miljøinformasjon, og gjøre denne tilgjengelig for allmennheten. Unntak kan gis i henhold til Offentlighetsloven dersom begrenset tilgang veier tyngre enn

hensynet til miljø- og samfunnsinteresser. Sensitive opplysninger som unntas offentlighet kan skjermes med begrenset tilgang eller aggregert stedfesting. Aktører med naturforvaltningsansvar bør alltid få tilgang, da dette ikke regnes som offentliggjøring. Tilgang på taushetsbelagt informasjon hjemles normalt i Miljøinformasjonsloven og Offentlighetsloven. I [paragraf 24 i Offentlighetsloven](#) er det hjemlet adgang til å unnta sensitiv informasjon fra full offentlighet. Motivet for dette er å skjermes informasjon som kan misbrukes i miljøkriminalitet.

Direktoratet for naturforvaltning har utformet [retningslinjer for håndtering av sensitiv artsinformasjon](#). Det dreier seg i all hovedsak om 22 pattedyr- og fuglearter der det er hi-, hekke- og mytelokaliteter som skal skjermes. Universitetsmuseene har i tillegg ei liste på 7 arter (de fleste er planter) som de som dataleverandører ikke vil eksponere med full geografisk lokalisering. Retningslinjene har også bestemmelser for hvordan slik informasjon skal synliggjøres i saksdokumenter som skal vises offentlighet. Der slik sensitiv informasjon brukes offentlig skal detaljerte funnstedsopplysninger skjermes ved at opplysninger som lokalitetsnavn, nøyaktige koordinater utelates eller at funnsted georefereres med rutenett 5x5km eller grovere.

Gjennom årets tildelingsbrev til offentlig forvaltning krever Fornyings-, administrasjons- og kirkedepartementet at "etatene skal gjøre egnede og eksisterende rådata tilgjengelige i maskinlesbare formater" så lenge disse ikke er taushetsbelagte. Det presiseres at dette kun gjelder data der kostnadene antas å være beskjedne.

4.6 Norsk Lisens for Offentlige Data (NLOD)

[Direktoratet for forvaltning og IKT](#) (Difi) har utarbeidet en lisens til deling og viderebruk av åpne data; [Norsk lisens for offentlige data](#) (NLOD). NLOD tillater lisenstakeren å bruke, bearbeide og dele datasett, også kommersielt. Det kreves at lisensgiveren blir navngitt og dataene brukes på eget ansvar og på en forsvarlig måte. I sin veileder i tilgjengeliggjøring av offentlige data oppsummerer Difi NLOD som en liste over rettigheter, vilkår og bemerkninger:

Du har lov til:

*å kopiere og tilgjengeliggjøre
å endre og/eller sette sammen med andre datasett
å kopiere og tilgjengeliggjøre en endret eller sammensatt versjon
å benytte datasettet kommersielt*

På følgende vilkår:

*at du navngir lisensgiver slik lisensgiver ber om, men ikke på en måte som indikerer at disse har godkjent eller anbefaler deg eller din bruk av datasettet
at du ikke bruker dataene på en måte som fremstår som villedende, og heller ikke fordreier eller uriktig fremstiller dataene*

Med den forståelse:

*at data som inneholder personopplysninger og er taushetsbelagt ikke er omfattet av denne lisensen og ikke kan viderebrukes
at lisensgiver fraskriver seg ethvert ansvar for informasjonens kvalitet og hva informasjonen brukes til*

Det finnes en engelsk versjon av NLOD, og bruk i internasjonal sammenheng er videre forenklet ved at lisensen anses å være kompatibel med [Creative Commons Navngivelse 3.0 Unported \(CC BY 3.0\)](#) dersom hele datasettet består av data som ikke er berørt av åndsverkloven, noe som data fra offentlige organ i utgangspunktet ikke skal være.

I motsetning til Creative Commons kan NLOD også omfatte hele databaser som enheter, og er i slike tilfeller kompatibel med [Open Data Commons Attribution License \(versjon 1.0\)](#). NLOD er en nyutviklet, anbefalt nasjonal standard innen deling av offentlige data, kompatibel med relevante, velkjente internasjonale lisensavtaler. NLOD foreslåes som standardlisens for deling av data gjennom LifeWatch Norge.

4.7 Nasjonal Infrastruktur for deling av offentlige miljø og biodiversitetsdata

Det er gjennom forprosjektet påvist et stort behov innen forskning, forvaltning og samfunnet for enklere tilgang til offentlige data gjennom en nasjonal infrastruktur. Realisering av en slik infrastruktur må tilpasses brukernes behov, internasjonale standarder og harmoniseres med parallelle initiativ. Norge deltar aktivt i det nordiske LW- forprosjektet og har i tillegg tett kontakt med LifeWatch Europa. Dette gir oss tilgang til et rikt erfaringsgrunnlag fra andre institusjoner og operative støttefunksjoner. I tillegg har vi den fordelen av å kunne bygge videre på de erfaringer og den teknologi som er utviklet gjennom GBIF, Artsdatabanken og Norsk Marint Data-senter.

5 Anbefalinger og tiltak

Vi har gjennom forprosjektet identifisert rådende behov og utfordringer, og vil på bakgrunn av dette anbefale relevante tiltak for realisering av en nasjonal infrastruktur for deling av offentlig finansierte miljø- og biodiversitetsdata. Utgangspunktet er at det i dag eksisterer store mengder relevante biodiversitetsdata som av ulike årsaker er lite, eller overhodet ikke tilgjengelig for samfunnet.

5.1 Utvikling av kontraktmaler

For å forbedre tilgangen på nyetablerte data anbefaler vi at det legges ned en større innsats fra oppdragsgiver (sektormyndighet, forskningsrådet eller andre) på utforming av standardiserte kontrakter hvor det i tillegg til rene oppdragsformaliteter også tydelig fremgår hvilket dataformat leveransen skal være på.

Den gjeldende bruk av det upresise begrepet “maskinlesbart format” bør utgå og erstattes med referanser til mer spesifikke, interoperable fagstandarder. For publisering av observasjonsdata til Artsdatabanken og GBIF er [Darwin Core](#) en påkrevd standard. For multimediedata er utvidelsen [Audubon Core](#) å anbefale. For mer komplekse romlige datamodeller bør [INSPIRE-direktivets dataspesifikasjoner](#) være styrende.

Ansvar for å etablere standardiserte kontraktmaler bør legges til den enkelte oppdragsgiver i tett dialog med oppdragstakere, Artsdatabanken, GBIF, Nasjonal Geodatakoordinator (Kartverket) og Direktoratet for forvaltning og IKT (Difi). Oppdragsgiver må sørge for god oppfølging underveis for å sikre at data blir levert og dokumentert i henhold til kontrakt.

5.2 Utvikling av strategier og handlingsplaner

Det bør utvikles helhetlige og spesifikke strategier for målstyring av datafangst, dataforvaltning og datadeling innen de ulike fagsektorene. Strategiutviklingen må forankres både i egen organisasjon, overordnet sektormyndighet og evt. Forskningsrådet der forskningssektoren er involvert. Hensikten med dette arbeidet er å avklare roller, ansvarsforhold og forventninger mellom oppdragsgiver og oppdragstaker. En slik strategi må implementeres på ulike nivå som nasjonalt myndighetsnivå (f.eks. Direktorat) og institusjonsnivå. Strategiarbeidet på institusjonsnivå vil kunne være mer handlingsrettet mot konkrete interne tiltak, mens strategiarbeidet på nasjonalt myndighetsnivå bør være mer fokusert på å utvikle retningslinjer og veiledere. GBIF har gjennom sin [“Best practice guide for Data Discovery and Publishing Strategy and Action Plans”](#) skissert en mulig måte å organisere dette arbeidet på. Vår anbefaling er at sektormyndighetene og forskningsmiljøene iverksetter felles tiltak på dette området.

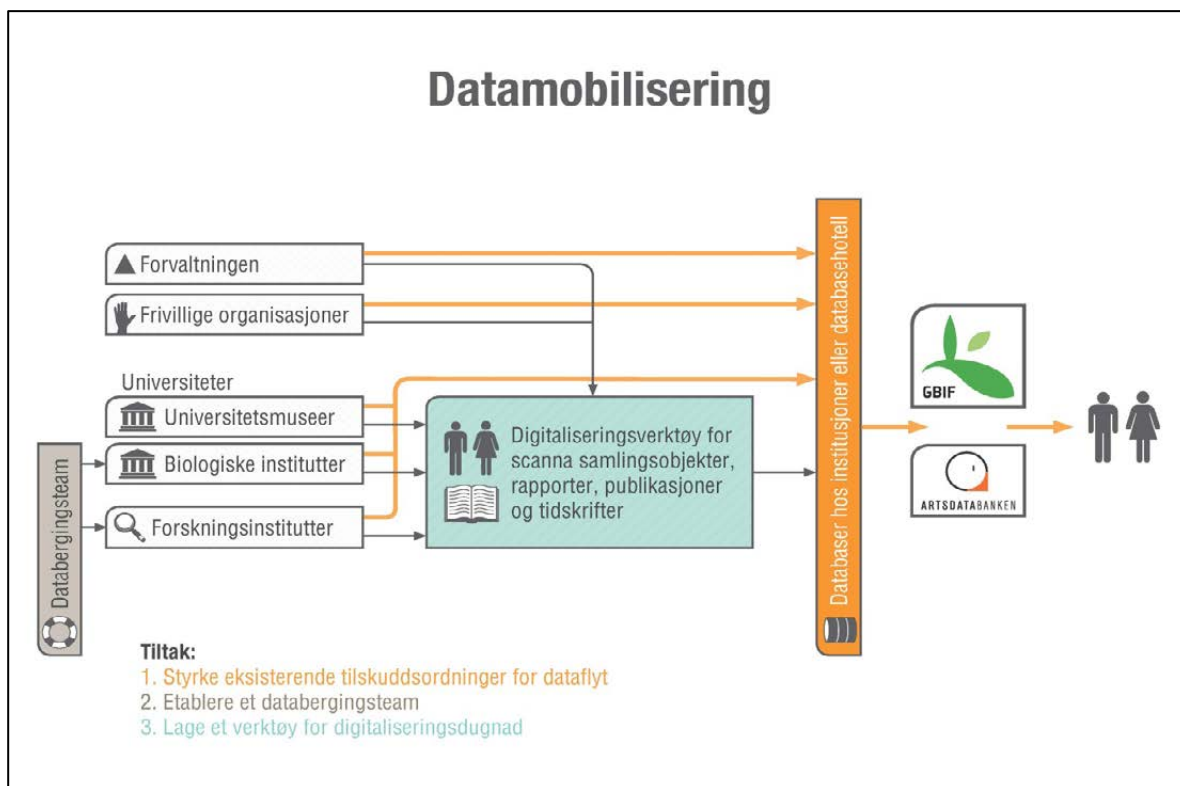
5.3 Kapasitetsbygging og utvikling av datamobiliseringsverktøy

I en del tilfeller mangler institusjonene nødvendig IKT- kompetanse for å kunne legge til rette for økt datadeling. I tillegg har flere institusjoner store mengder eldre og ustrukturerte data som ligger lagret både analogt og på diverse maskinlesbare formater. Som nevnt tidligere kan slike forhold danne praktiske barrierer for å dele data selv om institusjonen har som intensjon å dele data.

Vår anbefaling er at det foretas målrettede kapasitetsbyggingstiltak for å assistere de enkelte institusjoner i dette arbeidet. Det bør foretas behovsanalyser for å sikre at nytteverdien av tilgjengeliggjøring står i forhold til den totale kostnadsramme for arbeidet.

Med ekstra finansiering vurderer Artsdatabanken å opprette et operativt Databergingsteam som besitter nødvendig kompetanse og teknologi, og som kan utføre handlingsrettet innsats

for datamobilisering i samarbeid med den enkelte institusjon. En slik innsatsgruppe vil kunne være et egnet tiltak i dette arbeidet. I tillegg er det behov for å tilrettelegge digitaliseringsverktøy for analog eller skanna informasjon og en dataflytmodell som kan omfatte “crowdsourcing” med publikumsinnsats og etterkontroll i den enkelte institusjon.



Figur 9: Prinsippkisse for datamobilisering

GBIF har utviklet datamobiliseringsverktøyet [Integrated Publishing Toolkit \(IPT\)](#) for lettere tilgjengeliggjøring av observasjonsdata og metadata fra ulike datakilder som tekstfiler, regneark og databaser publisert som DarwinCore-arkiv ut på Internett. IPT er en Java-basert web-applikasjon som kan installeres og kjøres på standard web- serverplattformer som Windows, Linux, Unix og Apple OS X. Vi vil anbefale IPT som et egnet datamobiliseringsverktøy for LifeWatch Norge.

5.4 Evaluering av rammevilkår og virkemidler rettet mot instituttsektoren

For å gjøre privat instituttsektor bedre i stand til å kvalitetssikre og dele offentlig finansierte data anbefaler vi departementene og forskningsrådet å ta en gjennomgang av eksisterende virkemidler og instituttsektorens økonomiske rammevilkår for å se om disse er tilstrekkelige i forhold til dagens ambisjonsnivå.

6 Utvidelse av det nasjonale LW- konsortiet

Forprosjektets brukerundersøkelse har påvist et stort behov for en nasjonal infrastruktur for deling av miljø- og biodiversitetsdata. For å kunne realisere en operativ norsk LifeWatch Infrastruktur må hele konseptet aller først sikres en tverrsektoriell forankring og organisering.

For å komme videre med dette arbeidet anbefaler vi at nåværende LW- konsortium i samarbeid med Norges Forskningsråd arrangerer en interdepartemental LifeWatch konferanse innen Juni 2013. Som et minimum bør Miljøverndepartementet, Fiskeri- og kystdepartementet, Kommunal- og regionaldepartementet, Kunnskapsdepartementet og Landbruks- og matdepartementet være representert på denne konferansen. Dette bør være første fase i utvidelsen av det nasjonale LW-konsortiet, hvor hovedformålet bør være å klarlegge potensialet for merverdi og effektivisering for forskning og forvaltning LifeWatch konseptet innebærer.

Bakgrunnen for det brede fokuset vårt er at disse saksområdene omfatter et vidt spekter av produsenter, dataeiere og brukere av miljø- og biodiversitetsdata. Etter første fase, hvor de mest sentrale departementer og aktører (se ovenfor) har tatt en gjennomgang, er det naturlig å utvide spekteret, og nedenfor følger en liste over aktuelle organisasjoner som kan inviteres til videre deltagelse:

- Universitetene (inkludert alle naturhistoriske museer)
- Relevante Høgskoler
- Norsk regnesentral
- Norsk institutt for luftforskning (NILU)
- Norsk institutt for by- og regionforskning (NIBR)
- Norsk institutt for kulturminneforskning (NIKU)
- Norsk institutt for naturforskning (NINA)
- Norsk institutt for vannforskning (NIVA)
- Bioforsk (Bioforsk Jord og miljø)
- CICERO Senter for klimaforskning.
- Norsk institutt for Skog og landskap
- Norges Geologiske Undersøkelser
- Norsk Polarinstitut
- Bjerknessenteret for klimaforskning
- Nansensenteret
- Havforskningsinstituttet
- Bioforsk
- Internasjonalt forskningsinstitutt i Stavanger
- Christian Michelsen forskningssenter
- Konsulentbransjen
- Direktoratet for Forvaltning og IKT
- Direktoratet for mineralforvaltning
- Direktoratet for naturforvaltning/ Klima- og forurensningsdirektoratet
- Direktoratet for samfunnssikkerhet og beredskap
- Fiskeridirektoratet
- Jernbaneverket
- Kystverket
- Mattilsynet
- Nasjonal Sikkerhetsmyndighet
- Norges vassdrag og energidirektorat
- Oljedirektoratet
- Sjøfartsdirektoratet
- Vegdirektoratet

- Statens kartverk
- Fylkesmennene
- Fylkeskommunene
- Kommunenes Sentralforbund

Ved siden av tverrsektoriell forankring er hensikten med en slik konferanse å konsolidere et utvidet LW- konsortium som skal jobbe videre med å utrede mulighetene for finansiering av forarbeid og realisering av et storskala LifeWatch- prosjekt.

7 Realisering av et storskala LifeWatch- prosjekt

Forskning er grunnlaget for løsninger på mange av de kunnskapsutfordringer samfunnet står overfor. Med effektiv tilgang til data og verktøy gjennom en avansert forskningsinfrastruktur frigjøres mer tid til forskning og det brukes mindre tid på tilrettelegging av data fra ulike kilder. Forskningsrådet ønsker å styrke sin innsats for innovasjon i offentlig sektor og har derfor startet arbeidet med strategiutvikling hvor grunntanken er å styrke kunnskapsflyten mellom forskning, næringsliv/industri og utdanningssektoren. Vi mener at en nasjonal LifeWatch forskningsinfrastruktur for miljø- og biodiversitetsdata generelt kan bidra til økt kvalitet, innovasjon og effektivitet innen forskningsmiljøenes, næringslivets og den offentlige forvaltningens oppgaveløsning.

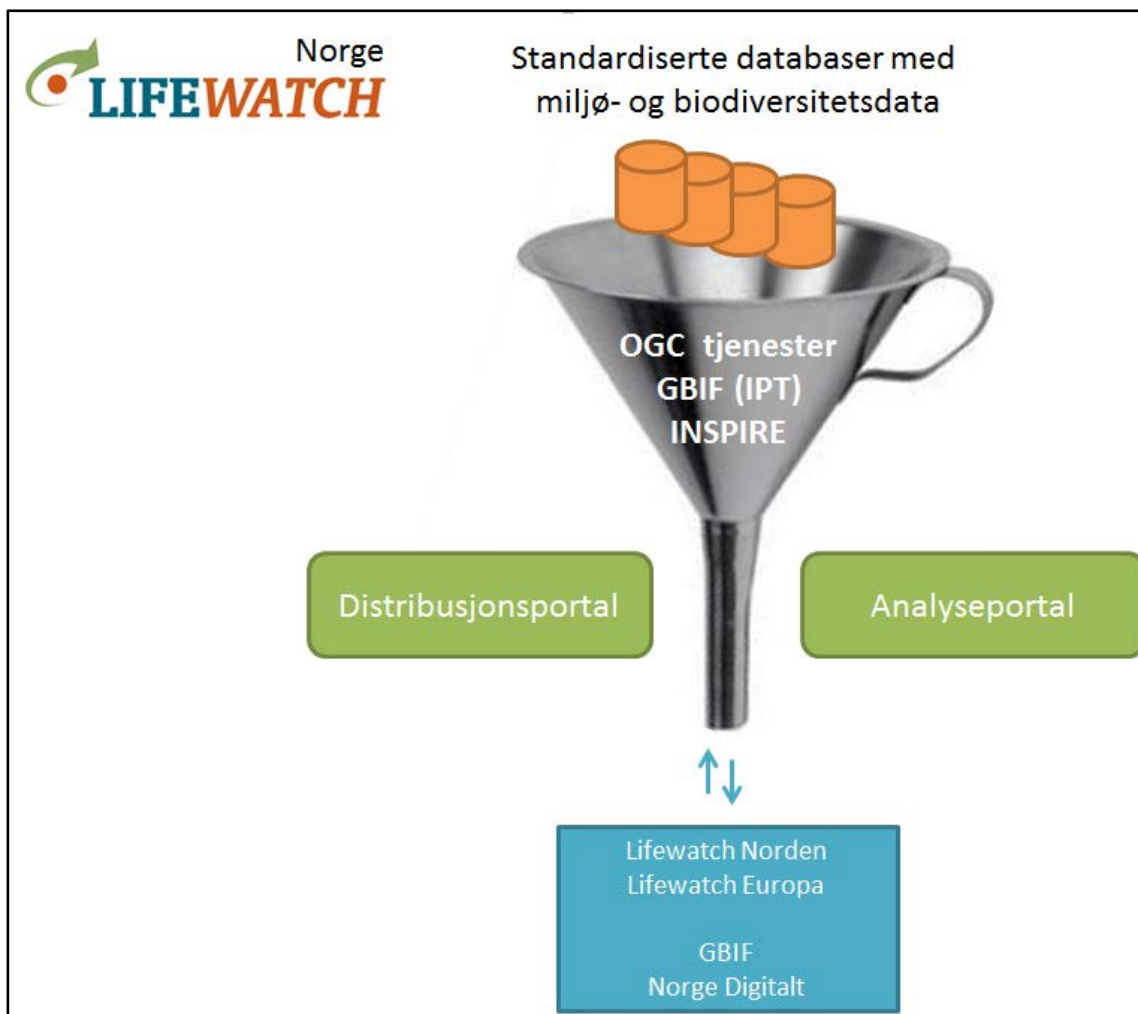
LifeWatch kom inn på det norske ESFRI- veikartet i 2010. Etter dette har det europeiske hovedprosjektet etablert sin egen driftsorganisasjon (2012) hvor hovedformålet er å bistå nasjonale prosjekter i etableringsfasen. LifeWatch ERIC blir med all sannsynlighet formelt vedtatt på nyåret i 2013. Arbeidet med å identifisere felles satsningsområder og å organisere kommunikasjonskanaler mellom europeisk og nasjonalt nivå er satt i gang.

Parallelt med dette har de nordiske landene med støtte fra NordForsk gått sammen og etablert et nordisk forprosjekt for å identifisere mulighetene for en nordisk storskala infrastruktur. Dette forprosjektet skal avsluttes høsten 2013, og vil komme med konkrete tilrådinge for veien videre. Fra europeisk nivå blir det nordiske initiativet tatt vel i mot ut fra tanken om at dette kan katalysere fremdriften i det europeiske prosjektet. De nordiske landene drar nytte av det arbeidet som er lagt ned i etableringen av Artsdatabankene i Norge og Sverige. Finland har nylig vedtatt å etablere sin egen Artsdatabank.

De nordiske LifeWatch- prosjektene har noe ulik status. [Det svenske LW konsortiet](#) er kommet lengst og har sikret seg langsiktig økonomisk støtte til utviklingen av en nasjonal infrastruktur fra det svenske Vetenskapsrådet. Det finske LW- konsortiet søkte i november 2012 om tilsvarende finansieringsrammer, og sjansene skal være gode for å få dette realisert i 2013. Finnene sier at det nordiske samarbeidet har vært viktig i denne sammenhengen. Med bakgrunn i nordisk samarbeid, homogenitet og finansielle situasjon er potensialet for å etablere et nordisk tyngdepunkt i det europeiske LW samarbeidet meget stort. Det europeiske LifeWatch er tenkt som en distribuert organisasjon med noen få sentrale servicefunksjoner og et nettverk av distribuerte, regionale fagsentre. Et nordisk samarbeid vil også kunne danne grunnlaget for etableringen av et nordisk fagsenter.

Utforming og organisering av en fremtidig nasjonal LW infrastruktur må ta utgangspunkt i de behov som er avdekket gjennom forprosjektet. Videre er det meget viktig at denne infrastrukturen samordnes godt med andre relevante initiativ som f.eks. [Norsk Marint Datasenter](#), [Nor-Store](#), [Artsdatabanken](#), [GBIF](#), [Statens kartverk](#) og [Norge Digitalt](#).

Hovedkomponentene i den nasjonale LW- infrastrukturen tenkes bygd opp som en distribusjonsportal og en analyseportal (se prinsippsskisse nedenfor).



Figur 10: Prinsippskisse for den nasjonale LW-infrastrukturen

Distribusjonsportalen er beregnet for brukere som vil søke opp grunnlagsdata og koble disse med miljødata for videre analyser i egne verktøy. Løsningen skal inneholde:

- Standard funksjonalitet for søk, innsyn og nedlasting
- Funksjonalitet for kobling med miljødata (f.eks. klima, geologi, hydrologi osv.)

Analyseportalen er beregnet for brukere som ønsker tilrettelagt funksjonalitet for:

- Søk på datakilde (artsdata, naturtypedata, miljødata og metadata)
- Filtrering av data (taksonomi, geografi, tidsrom og datakvalitet),
- Beregninger (statistikk, aggregering og trendanalyser)
- Presentasjon (kart, diagram og tabell)
- Skjermvisning og nedlasting

Alle typer data skal leses online fra dataeierens originaldatabaser via standardiserte WEB-protokoller og tjenester basert på [Open Geospatial Consortium \(OGC\)](#), INSPIRE sine datasettspesifikasjoner, GBIF DarwinCore og maskinlesbare formater som støttes av GBIF sitt data-mobiliseringsverktøy Integrated Publishing Toolkit (IPT).

Det er på det nåværende tidspunkt også vanskelig å si noe om hvordan et fullskala LifeWatch-prosjekt bør organiseres. De sentrale støttefunksjonene i LifeWatch Europa kan her bistå med veiledning og rådgivning. I tillegg vil det nordiske samarbeidet og deling av erfaringer derfra være nyttige. Det vil være essensielt at forvaltningsinstitusjonene deltar aktivt i et framtidig LW-konsortium, for på denne måten å sikre koblingen mellom forskning og forvaltning når det gjelder miljødata og tidsserier.

I prosjektfasen må det etableres en egen styringsgruppe som har ansvar for daglig oppfølging av prosjektet. Denne styringsgruppen bør bestå av prosjektleder, sekretariat, deltagere fra konsortiet og Forskningsrådet.

Arbeidsoppgaver bør deles inn i spesifikke arbeidspakker. Ansvarlig leder for hver arbeidspakke rapporterer jevnlig til prosjektleder. Det bør i tillegg etableres en referansegruppe som består av uavhengige eksperter.

Prosjektet bør videre lage kommunikasjonsstrategier for samspillet internt i konsortiet, og i forhold til de nordiske og europeiske prosjektene. Dersom prosjektet på sikt går over i en driftsfase og organiseres som en ERIC setter dette helt andre krav til driftsorganisasjonen.

Konsortiet har som ambisjon å søke finansiering av et storskala LW-prosjekt fra Forskningsrådets nasjonale [INFRASTRUKTUR-program](#). Neste utlysning vil sannsynligvis bli i oktober 2014. EU sitt nye rammeprogram for forskning og infrastrukturbygging [HORIZON 2020](#) (oppstart 2014) kan også være en aktuell finansieringskilde dersom det norske konsortiet søker sammen med en eller flere andre europeiske partnere som f. eks de nordiske landene.

8 Referanser

Beniston, M., M. Stoffel, R. Harding, M. Kernan, R. Ludwig, E. Moors, P. Samuels & K. Tockner, 2012. Obstacles to data access for research related to climate and water: Implications for science and EU policy-making. *Environmental Science & Policy* 17: 41-48.

[Chavan, V. S., Stood, R. K., and A. H. Arino. 2010. Best practice guide for Data Discovery and Publishing Strategy and Action Plans. Global Biodiversity Information Facility \(GBIF\) Secretariat.](#)

[Direktoratet for forvaltning og IKT \(DIFI\) 2011. Norsk lisens for offentlige data.](#)

[Direktoratet for forvaltning og IKT \(DIFI\) 2012. Veileder i tilgjengeliggjøring av offentlige data.](#)

[Direktoratet for Naturforvaltning. Retningslinjer for håndtering av sensitiv informasjon om arter.](#)

[European Commission 2003. Viderebruksdirektivet \(Directive 2003/98/EF\).](#)

[European Commission 2007. INSPIRE- direktivet \(Directive 2007/2/EC of the European Parliament and of the Council of 15 May 2007 establishing an Infrastructure for Spatial Information in the European Community\).](#)

[European Commission 2012. European Strategy Forum on Research Infrastructures \(ESFRI\).](#)

[European Commission 2012. European Research Infrastructure Consortium \(ERIC\).](#)

[Fornyings- og administrasjonsdepartementet \(FAD\). Følgerebrev ved tilgjengeliggjøring av offentlige data.](#)

[GBIF-TDWG Joint Multimedia Resources Metadata Task Group \(MRTG\), 2012. Audubon Core Multimedia Standard. Biodiversity Information Standards \(TDWG\).](#)

[International Council for Science 2012: Committee on Data for Science and Technology. Data Citation Standards and Practices Task Group.](#)

[International Organization for Standardization \(ISO\) 2003. ISO 19115/TC 211 Geographic Information/Geomatics Metadata Standard.](#)

[International Organization for Standardization \(ISO\) 2012. ISO 26324:2012 Information and documentation - Digital object identifier system.](#)

[Kunnskapsdepartementet 2009. St. Melding nr 30 \(2008/2009\), Klima for forskning.](#)

[Lovdata 2004. LOV 2003-05-09 nr 31: Lov om rett til miljøinformasjon og deltakelse i offentlige beslutningsprosesser av betydning for miljøet \(miljøinformasjonsloven\).](#)

[Lovdata 2009. LOV 2006-05-19 nr 16: Lov om rett til innsyn i dokument i offentlig verksemd \(offentleglova\).](#)

[Lovdata 2010. LOV 1996-06-21 nr 38: Lov om statlig naturoppsyn \(naturoppsynsloven\).](#)

[Lovdata 2012. LOV 2010-09-03 nr 56: Lov om infrastruktur for geografisk informasjon \(geodataloven\)](#)

[Lovdata 2012. Forskrift om infrastruktur for geografisk informasjon \(geodataforskriften\)](#)

[Miljøverndepartementet 2001. St. Melding nr. 42 \(2000/2001\), Biologisk mangfold. Sektoransvar og Samordning.](#)

[NASA 2012. Global Change Master Directory \(GCMD\).](#)

[NFR 2004. Lange tidsserier for miljøovervåking og forskning, Rapport 1, 2 og 3.](#)

[NFR 2012. Verktøy for forskning, Nasjonal strategi for forskningsinfrastruktur 2012- 2017.](#)

[NFR 2012. Nasjonal satsing på Forskningsinfrastruktur.](#)

[Open Data Commons. Open Data Commons Attribution License \(ODC-By\) v1.0.](#)

Schmidt-Kloiber, A., S. J. Moe, B. J. Dudley, J. Strackbein & R. Vogl, 2012. The WISER metadata-base: the key to more than 100 ecological datasets from European rivers, lakes and coastal waters. *Hydrobiologia* (in press). DOI 10.1007/s10750-012-1295-6

[SCOR/IODE/MBLWHOI Library Workshop on Data Publication, 5th Session, Woods Hole Oceanographic Institution, Woods Hole, USA, 9-10 October 2012 Paris, UNESCO, 10 October 2012 \(IOC Workshop Report No. 252\)](#)

[Darwin Core task group, 2009. The Darwin Core standard. Biodiversity Information Standards \(TDWG\).](#)

[The Organisation for Economic Co-operation and Development \(OECD\) 2007. Principles and Guidelines for Access to Research Data from Public Funding.](#)

[Uhlir, P. E. 2012. For Attribution - Developing Data Attribution and Citation Practices and Standards: Summary of an International Workshop; Board on Research Data and Information; Policy and Global Affairs; National Research Council ISBN 978-0-309-26728-1](#)

9 Appendiks 1: Spørsmål stilt i brukerundersøkelsen

Tema 1: Generelt om behov for biologiske data

Spørsmål:

- a) Hvor mye bruker du biologiske data i ditt arbeid eller i andre aktiviteter?
- b) Vil du ha nytte av bedre tilgang til biologiske data enn det du har i dag?
- c) Vil du ha nytte av bedre oversikt over biologiske data som er tilgjengelig fra ulike institusjoner i Norge?
- d) Er du interessert i å ta i bruk nye kilder og verktøy for bruk av biologiske data?
- e) I hvilken grad har du hittil benyttet deg av informasjon og andre tjenester fra Artsdatabanken? (www.artsdatabanken.no)

Tema 2: Dataformater

Spørsmål: I hvilken grad har du bruk for de følgende typene av data?

- a) Tid- og stedfestede artsobservasjoner, uten mengdeangivelse (forekomst)
- b) Tid- og stedfestede artsobservasjoner, med mengdeangivelse (f.eks. antall/km²; biomasse/liter)
- c) Lange tidsserier (f.eks. antall individer av en art på en gitt lokalitet over flere år)
- d) Utbredelseskart for utvalgte arter
- e) Miljødata som kan knyttes til de biologiske dataene (fysikk, kjemi, hydrologi, klima, forurensning, annen menneskelig påvirkning m.m.)

Tema 3: Økosystemer

Spørsmål: I hvilken grad har du bruk for de følgende typene av data?

- a) Terrestre økosystemer
- b) Ferskvannøkosystemer (elver, innsjøer, våtmark)
- c) Marine økosystemer (hav, kyst, brakkvann)

Tema 4: Grupper av arter

Spørsmål: I hvilken grad har du bruk for de følgende typene av data?

- a) Pattedyr
- b) Fugler
- c) Fisk
- d) Amfibier og reptiler
- e) Invertebrater
- f) Planter, terrestre
- g) Planter, akvatisk
- h) Sopp
- i) Introduserte eller invaderende arter
- j) Truede eller sårbare arter
- k) Høstede arter (jakt, fiske m.m.)

Tema 5: Metadata (informasjon om datasett)

Spørsmål: I hvilken grad tror du følgende informasjon om tilgjengelige datasett kan være nyttig for deg?

- a) Publisert informasjon om dataene (f.eks. artikler, rapporter)
- b) Geografisk informasjon: punkt-koordinater
- c) Geografisk informasjon: polygoner (flater)
- d) Økosystem-typer (f.eks. innsjøer)
- e) Naturtyper (f.eks. kalkfattig kysthei)
- f) Taksonomisk nivå (f.eks. artsnivå)
- g) Taksonomiske navn (f.eks. Crustacea)
- h) Tidsintervaller (f.eks. 1990-2010)
- i) Metodikk for prøvetaking og analyse
- j) Datakvalitet: presisjon for tidfesting
- k) Datakvalitet: presisjon for stedfesting
- l) Datakvalitet: presisjon for mengdeangivelse

Tema 6: Verktøy og tjenester

Spørsmål: I hvilken grad tror du følgende funksjonaliteter vil være nyttig for deg?

- a) Kobling av biologiske data med relevante miljødata
- b) Enkle trendanalyser for tidsserier
- c) Utbredelseskart
- d) Andre kart-tjenester
- e) Web map service, se http://en.wikipedia.org/wiki/Web_map_service
- f) Web feature service, se http://en.wikipedia.org/wiki/Web_Feature_Service
- g) Andre web-tjenester
- h) Artsnavntjenester: validering mot gyldige vitenskapelige navn
- i) Økologisk informasjon (habitat, levevis m.m.)
- j) Rødlistestatus
- k) Beregning av biodiversitet-indeks



Norsk institutt for naturforskning (NINA) er et nasjonalt og internasjonalt kompetansesenter innen naturforskning. Vår kompetanse utøves gjennom forskning, utredningsarbeid, overvåking og konsekvensutredninger.

NINAs primære aktivitet er å drive anvendt forskning. Stikkord for forskningen er kvalitet og relevans, samarbeid med andre institusjoner, tverrfaglighet og økosystemtilnærming. Offentlig forvaltning, næringsliv og industri samt Norges forskningsråd og EU er blant NINAs oppdragsgivere og finansieringskilder.

Virksomheten er hovedsakelig rettet mot forskning på natur og samfunn, og NINA leverer et bredt spekter av tjenester gjennom forskningsprosjekter, miljøovervåking, utredninger og rådgiving.

ISSN:1504-3312

ISBN: 978-82-426-2532-8

Norsk institutt for naturforskning

NINA Hovedkontor

Postadresse: Postboks 5685 Sluppen, NO-7485 Trondheim

Besøks/leveringsadresse: Tungasletta 2, NO-7047 Trondheim

Telefon: 73 80 14 00, Telefaks: 73 80 14 01

E-post: firmapost@nina.no

Organisasjonsnummer 9500 37 687

<http://www.nina.no>

Samarbeid og kunnskap for framtidens miljøløsninger