

AnaEE DK FAIR data projekt 2019

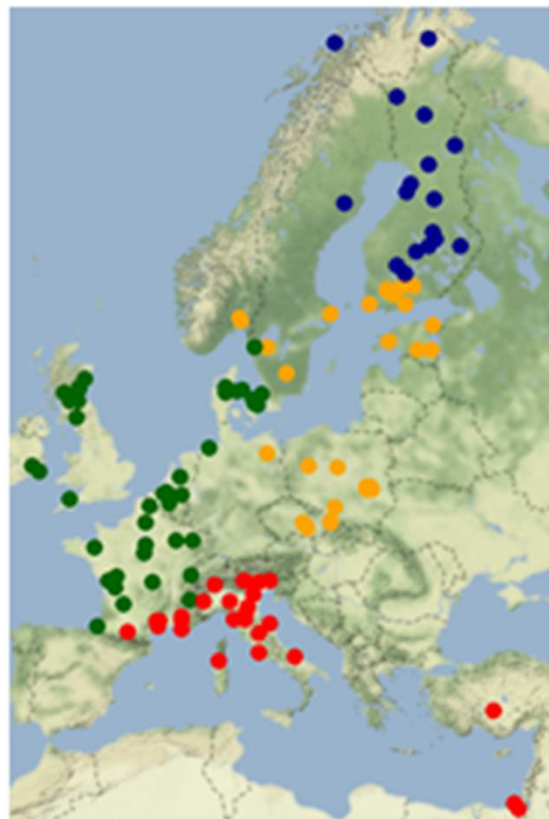
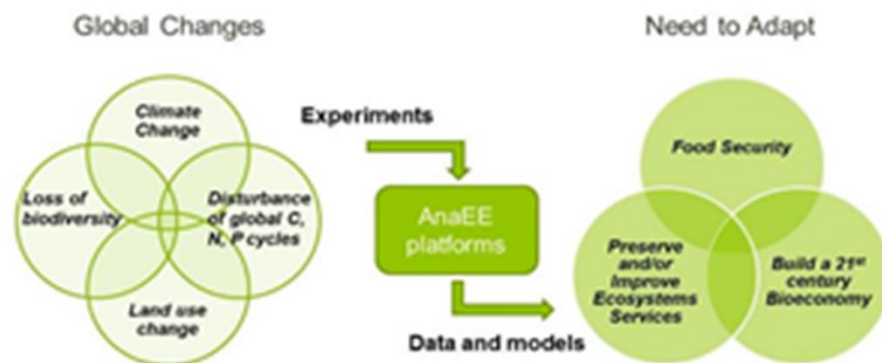
Katrine Vendelboe

Projektleder

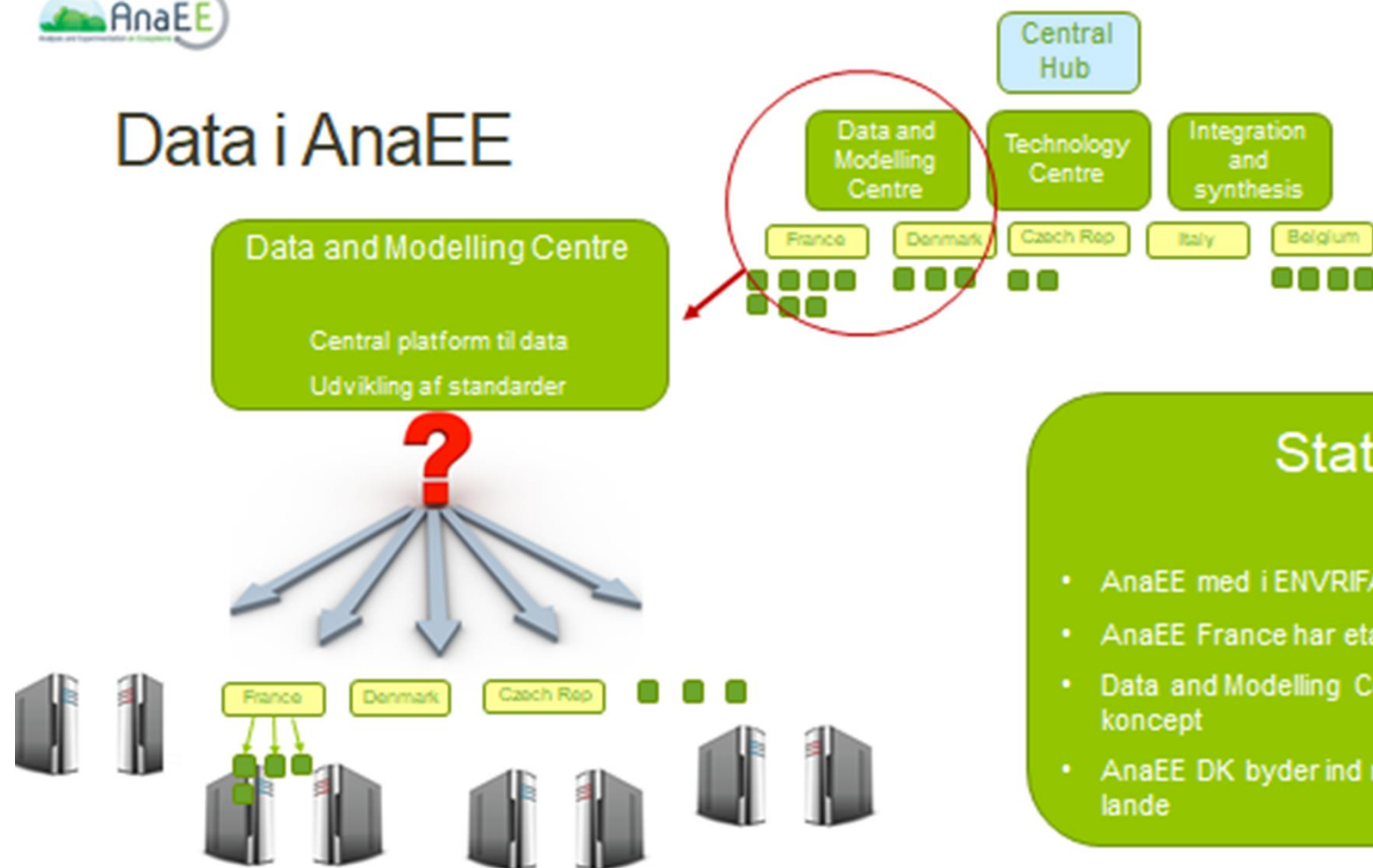
Outline

- AnaEE infrastruktur
- Data i AnaEE
- Implementering af FAIR data i AnaEE DK - projektet
- Konklusioner

AnaEE infrastruktur



Data i AnaEE



AnaEE DK FAIR data projekt

- 6 måneder: juli-december 2019
- Ledet af KU
- Rådgivning fra DelC
- Deltagere fra 4 danske AnaEE platforme

Primære mål

- Læring for deltagerne
- 3 delvist FAIR datasæt med DOI og landing page
- Feedback til DelC
- Forberedelse til AnaEE FAIR data process

Processen

- 1 dags workshop– 1 Juli
 - 1 data manager fra hver af de 4 platforms (KU PLEN og IGN, DTU og AU)
 - DeiC/ DM Forum rådgivere

- Identificering af passende datasæt
 - Meteorologiske data – mest simple og bruges af alle

Formål

At forstå FAIR og komme til enighed om projektplan

Udfordring

Overvejelse om brug af variable/levende datasæt (nye data hver time)

Beslutning: Konstante, afsluttede, kvalitetssikrede datasæt med opdatering halvårligt

Processen fortsat

- Identificering af passende repositories
 - Mulighed for DOI
 - Landing page muligheder
- Beslutning om fælles nomenklatur
 - Kombination af ICOS, WMO og FAO – accepterede standarder indenfor forskningsfeltet
- IPR check
 - AUs data er fra DMI – men tilladelse er på plads

Udfordring

Manglende interne tilbud fra universiteterne

KU: ERDA, som er mangelfuld

AU: Repository kan ikke udmønte DOI

Processen fortsat

- Metadata
- Enighed om Quality Assurance
- Udmøntning af DOI via ERDA (KU)

Udfordring

Forståelse for maskinlæsbare metadata – hvad er det?

Udfordring

Kvalitetssikring på meteorologiske data er håndholdt – ingen standarder – ingen mulighed for at gøre maskinlæsbart

Processen fortsat

Opdagelse af ERDAs begrænsninger

Landing page format mangelfuldt

Visning af metadata – maskinlæsbart – ikke særligt menneskelæsbart

Manglende versioningsmuligheder grundet sikkerhedsforanstaltninger

Ingen mulighed for at lave udkast til senere publicering

... flere tekniske udfordringer

- Møde med KUs folk bag ERDA – feedback og forklaringer
- Landing page for AnaEE DK data – manuelt sat op

https://erda.ku.dk/vgrid/AnaEE_DK/

Målene og resultaterne

Fra projektbeskrivelsen

A minimum output (project requirements) is thought to be:

- ✓ *Agreed initial quality control of datasets*
- ✓ *Making the dataset Findable, via persistent identifiers (PIDs)*
- ✓ *Making the dataset Accessible via a Landing Page*
- ✓ *Preferably some degree of Interoperable*

Yderligere resultater

- ✓ Deltagerne har fattet FAIR og ved hvordan det KAN gribes an
- ✓ Der er givet direkte feedback til institutionerne
- ✓ Opdagelse af nu- og fremtidens udfordringer