



## *AnaEE : data ownership, data policies and licenses*

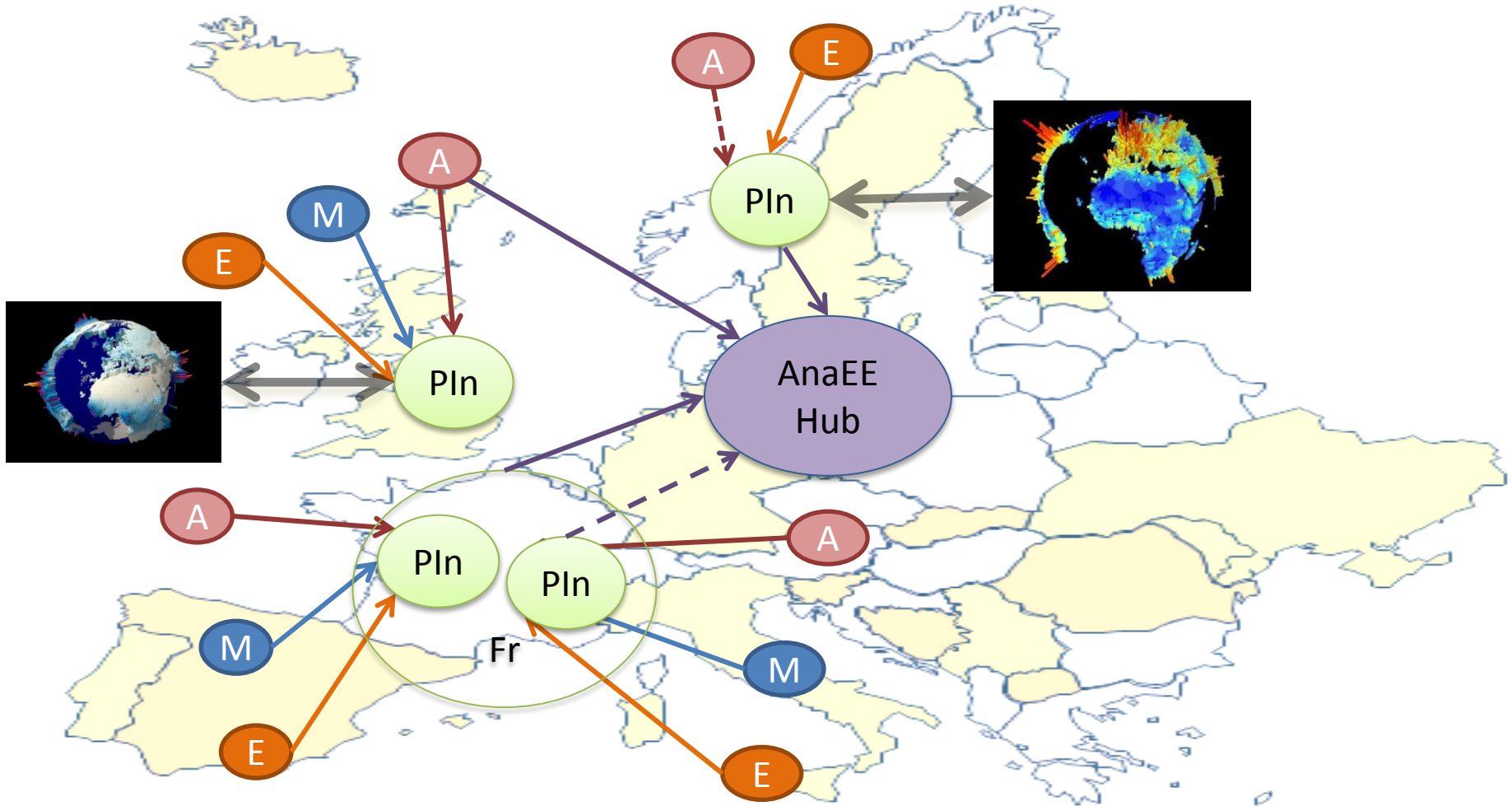
---

Presented by Daniel Rasse for the AnaEE  
WP4 Team (Chris Rawlings, WP leader).

# AnaEE – developing a distributed RI

- AnaEE will be a distributed RI comprising experimental, analytical and modelling platforms in different locations
- Every platform will have scientific interests beyond those of being participants in AnaEE
- Every platform will have pre-existing relationships with other national and probably International collaborations

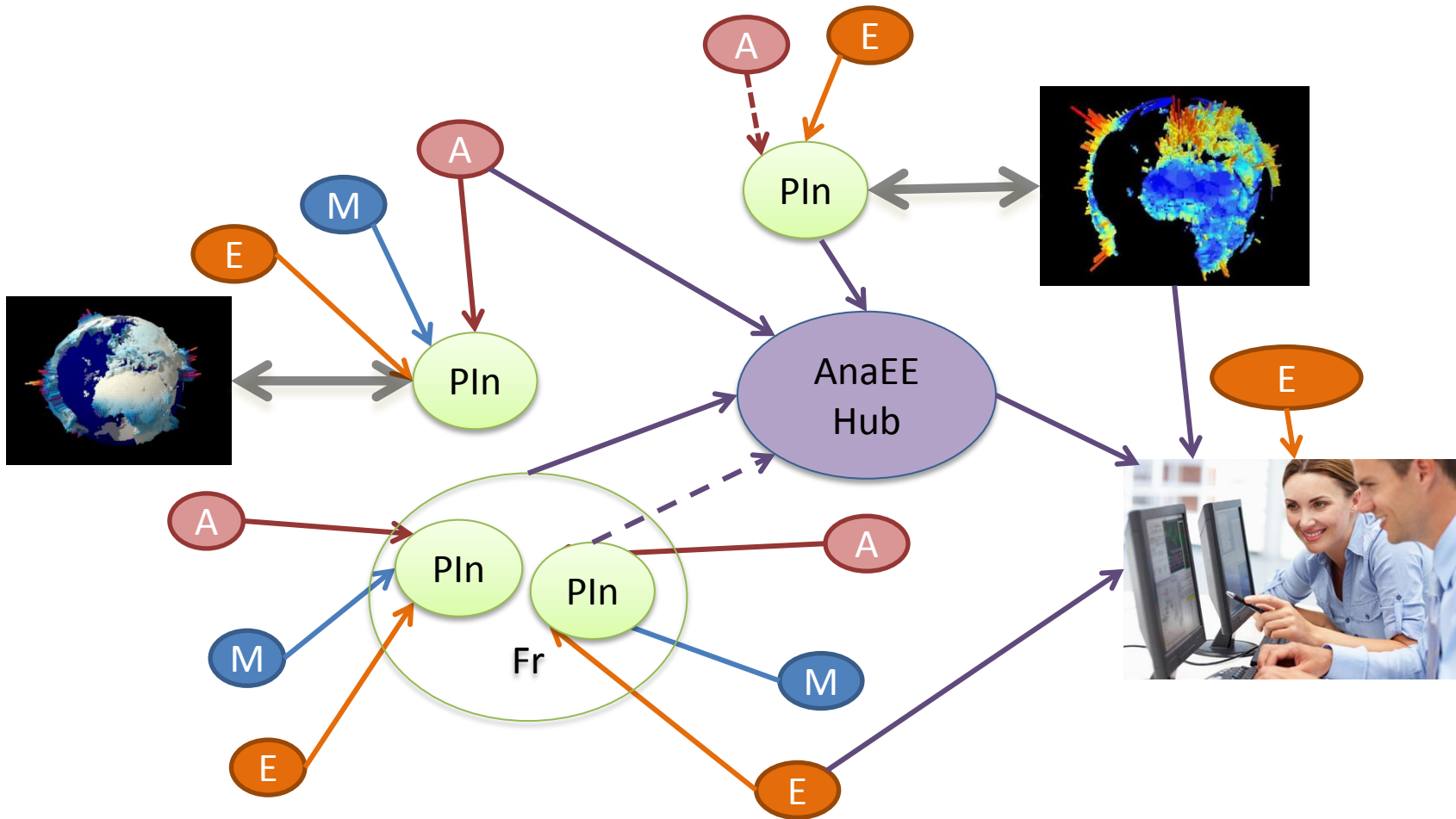
# Complex AnaEE Organisation



Platforms: Analytical, Modelling and Experimental



# How users will see AnaEE



Platforms: Analytical, Modelling and Experimental



# General considerations

- As a Research Infrastructure supported by European national public research funding agencies, AnaEE will comply with the recently published Guidelines on Open Access to Scientific Publications and Research Data
- This means ensuring that data collected on AnaEE platforms is made accessible as quickly as possible to potential users following any necessary steps to ensure research quality and to protect the rights of investigators to interpret project data and secure any potential intellectual policy

# AnaEE Platforms

- Not all data generated by a platform will be “part of” AnaEE
- Each platform will need capacity for local data management and quality control
  - Local data managers will need to understand data flows
- Local institutions and national funding agencies will want an important say concerning policies on:
  - Access Management, Intellectual property, Data sharing
- All these need to be consistent to simplify picture for users
- Significant challenge
- Outline current proposals for discussion here
  - Key issues only for this discussion
- More engagement desired and further debate needed

# Data Access Model – Responsibilities and Timelines

| Data Category   | Responsibility and timing of activities          |                     |                        |                                     |
|---|--|---------------------|------------------------|-------------------------------------|
|   | Collection and curation                          | Access by DMC       | Management and Archive | Public Access                       |
| Metadata describing site and facilities   | Site.<br>At start of operation as AnaEE platform | Immediate           | DMC.                   | DMC<br><1mth                        |
| Metadata describing all routine measurements made on platform.  | Site.<br>At start of operation as AnaEE platform | Immediate           | DMC                    | DMC<br><1mth                        |
| Meteorological data.<br>Summary data  | Site. Daily.<br>DMC                              | <1mth<br>By mo, yr. | Site and DMC<br>DMC    | DMC. Site optional.<br><3Mth<br>DMC |
| Routine environmental measurements from platform from automatic data loggers (e.g. Soil temperature).<br>Summary Data | Site<br>DMC                                      | <1Mth<br>By mo, yr. | Site and DMC<br>DMC    | DMC. Site optional.<br><3Mth<br>DMC |
| Optional additional measurement from sites, to be agreed in collaboration agreement.                                  |  |                     |                        |                                     |



# Intellectual Property Framework

---



# IP Generation and Use

- All policies and procedures will be open and published
  - to ensure transparency and equality of access.
- IP Rights will be maintained by generating parties.
- The AnaEE Central hub will be given widest possible rights to manage and license to third parties.
  - Assignment of IP is not generally considered practical or desirable.
- Oversight of IP management by the AnaEE hub should be the responsibility of an appropriate executive-level committee.
- Owners and generators of IP should have priority access for use within the AnaEE partnership
  - but would require licenses to use data for purposes involving external partners.

# IP Usage

- Use by third parties would be under clear terms set out in licenses (to be agreed).
  - Standard licenses for not-for-profit use
  - e.g. based on Creative Commons models would be appropriate.
- All publications by third-parties must acknowledge AnaEE and the party(s) generating the IP supplied to the third-party
- Use by for-profit organisations (industry) will be negotiated on a case-by-case basis by the Central Hub.
  - No exclusive licenses would be granted.
- Revenue generated by third-party use will be held by the Central Hub and used for the benefit of the Project.

# Other Considerations

- There is a need to differentiate IP rights relating to
  - Baseline monitoring data from the RI e.g. the routine environmental measurements from a platform
  - Project data will be available after an agreed embargo period designed to protect the intellectual contribution of the research team generating the data
- Need to ensure that IP framework is consistent with platform and data access policies

# Licensing

## Creative Commons Licenses:

- No Rights Reserved
- Attribution (CC BY)
- Attribution no derivatives (CC BY-ND)
- Attribution-NonCommercial-ShareAlike (CC BY-NC-SA)
- Attribution-ShareAlike (CC BY-SA)
- Attribution-NonCommercial (CC BY-NC)
- Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

# Licensing

- All of the Creative Common licenses can possibly used by AnaEE.
- CC BY-NC-SA would be the most appropriate license for the majority of data products. (*This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms*)
- This does not, however, prevent other licenses to be used when the data are being used for non-research purposes. It is recommended that a subset of CC licenses are selected that are most likely to be relevant to AnaEE to help simplify discussions for all stakeholders. This is the approach that TERN has taken

# *Data Licensing between DMC and Platforms*

Objective: provide clarity and establish the freedom to operate for the AnaEE centres while avoiding any future misunderstandings about re-use and the development of derivative data products

No decision made yet (the CC4 license **CC BY-SA** might be the most appropriate)

# Data Publishing

- AnaEE will support the deposition of partial or complete datasets from AnaEE platforms and projects in a publicly accessible repository.
- Data depositions will be accompanied by suitable metadata for automated data discovery and will use DOIs to provide permanently citable references to the datasets.
- Where appropriate, data publications in data journals such as Nature Data will be used to provide independent peer reviewed and citable works.