

- Definitions like ORL crucial to identify reliable operational systems, especially for emerging AI applications
 - Invite Google to calculate ORL for their weather forecasts?
- Go through ORL and ETOOFS Guide to assess if they are relevant for AI applications, and update if not
- Class4 intercomparison allows centres to better verify proposed developments, not just current systems
 - Further this by leveraging OceanBench and other tools better
- Can be a bottleneck getting developments from academia to operational centres due to different expectations about robustness of coding and testing – common toolboxes and clear standards/best practices could help