The importance of open science principles for expanding and diversifying the user base of marine satellite data

Hayley Evers-King and Ben Loveday
PML Applications Ltd with EUMETSAT
@HayleyEversKing, hek@pml.ac.uk

+ many people who we’ve worked with over the last few years!
Open Science in EO and it’s benefits

• Copernicus (open data policy)
• But data being ‘open’ as in ‘free/not limited in usage’ is not enough.
• Other principles (https://opendatacharter.net):
  • Timely/comprehensive
  • Accessible
  • Comparable
  • Improved governance and engagement
  • Inclusive

= DEMOCRATISATION OF DATA

Why? Ethically commendable but also means greatest impact and scope for continuity of work.
What has this looked like for us in practice?

• Improvement in metadata and documentation
• MOOCs
• Hackathons
• Training (400 in 2 years)
  • Participant led (we teach them what they decide they need)
  • Exclusively in open source software: SNAP, Python, Git
• Hosted processing – Linux VMs
• Vital skills for next generation (requirements), inherent shareability for training.
• Sharing experience of open science formally
  • Credit, value to others.
A code base for training (but possibly more)

- Aims of code:
  - Teach about the data and its uses
  - Don’t terrify people (but don’t limit them)!
  - Provide common workflows
  - Use transferable packages
  - Offer skills growth applicable to (increasingly) diverse user community

- [https://gitlab.com/ben.loveday/EUMETSAT_training_jupyter_notebooks](https://gitlab.com/ben.loveday/EUMETSAT_training_jupyter_notebooks)
  - Will soon be ported to EUMETSAT GitHub.
  - Continuous expansion based on requests
What could we do to be more ‘principled’ in our work?

• Thoughts for discussion:
  • Continue and expand great work of deriving community standards for protocols
  • Raising use of community repositories
    • Great history of doing this with data
    • Share useful snippets
    • Standards to be discussed! Licencing and attribution.
    • What are ‘common workflows’ for Ocean Colour (and beyond?)
  • Diversify those we work with (data science communities)
    • Lots for us to learn
    • Potential for greater impact
    • More discussion of this in R2O workshop!