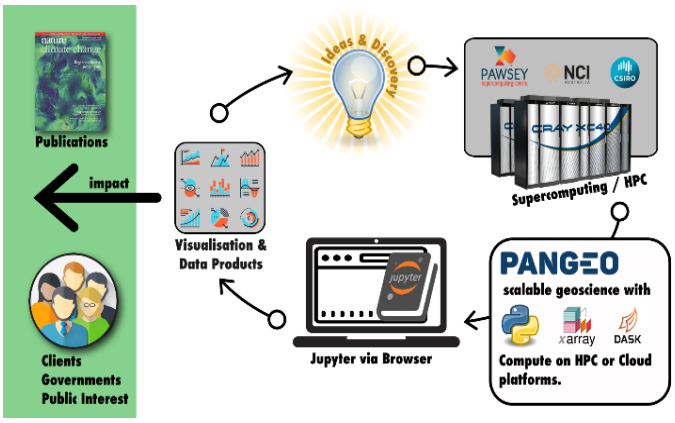
CSIRO endorsed as a Pangeo institutional partner

CSIRO, through the initiative and efforts of the Decadal Forecasting Project (DCFP) and the Coasts Program, have been endorsed as an institutional partner of the community platform for big data geoscience, Pangeo.

What is Pangeo?

The geoscientific community is facing a "Big Data" crisis as our rapidly growing datasets become too large for the legacy software tools to handle, which creates a major obstacle to scientific progress. Pangeo is a community effort for big data with the mission of cultivating an ecosystem for developing, distributing, and sustaining the next generation of open-source analysis tools required to meet our big data challenges.



The Significance of Institutional Partnership

The DCFP is energised and excited by the potential benefit of the Pangeo community effort and instigated CSIRO becoming a Pangeo institutional partner, a development recently endorsed by the Pangeo Steering Council.

Institutional partnership with Pangeo formalises our relationship and connects us with the global Pangeo effort. Pangeo brings together forward-looking researchers, software engineers, and other technical experts into a community that are committed to developing the new solutions for tackling the Big Data science questions.

Pangeo Steering Council Member, Associate Professor Ryan Abernathey made the comment that “We‘re extremely happy to see the level of effort and enthusiasm with which CSIRO has embraced Pangeo. I truly believe our work has the potential to transform science, make us all more productive, and enable new discoveries. Having a formal commitment from CSIRO is a major milestone for us”.

Our formalised membership of Pangeo recognises the significant effort that CSIRO staff have made in advancing the scalable geoscience workflows needed to handle the big data science challenges facing DCFP researchers. This is a challenge that other researchers will face as their scientific fields embrace big data.

Overcoming Big Data Challenges

In climate forecasting, the forecasts datasets are now often measured in terabytes (1000 gigabytes) and require new ways to extract useful information. Both the UK’s Met Office and USA’s National Centre for Atmospheric Research (NCAR) are working with the Pangeo effort, as they also require new solutions to handle the data they’re producing.

In May 2019, CSIRO staff held a dedicated Pangeo-focused session at the recent Collaborative Conference on Computational and Data Intensive Science 2019 conference in Canberra, bringing together researchers from the Met Office and NCAR. The session raised awareness and facilitated a discussion of our roadmap for adopting Pangeo geoscience principles.

CSIRO scientists have also held several Pangeo training sessions across the Australian geoscience community and a Pangeo hackathon at the CSIRO Oceans and Atmosphere site in Hobart. For the future success of the DCFP, Pangeo will become indispensable, by enabling the scientific inquiry of our big climate datasets. Importantly, Pangeo will also help other CSIRO researchers who are using big data in their work.