

CSIP - Report

BELMONT -Recommendations

Fill critical global e-infrastructure gaps; improve data management and exploitation; coordinate and integrate disparate organizational and technical elements; share best practices; and foster new data literacy to enable actionable and societally beneficial science. The five recommendations are:

1. Adopt Data Principles that establish a global, interoperable e-infrastructure data layer with cost-effective solutions to widen access to data, and ensure its proper management and long-term preservation. Researchers should be aware of, and plan for, the costs of data intensive research.
2. Foster communication, collaboration and coordination with the wider research community and across Belmont Forum projects through a Data and e-Infrastructure Coordination Office established within a Belmont Forum Secretariat.

Belmont - Recommendations

3. Promote active data planning and stewardship in all Belmont Forum agency-funded research to enable harmonization of the e-infrastructure data layer through enhanced project data planning, monitoring, review and sharing.
4. Determine international and community best practice to inform Belmont Forum research e-infrastructure policy, in harmony with the evolving research practices and digital ecosystem, through identification and analysis of cross-disciplinary research case studies.
5. Support the development of a cross-disciplinary training curriculum to expand human capacity in technology and data-intensive analysis methods for global change research and increase the number of scientists with cross-cutting skills and experience in best practice.

Reproducibility in Science

Elements of reproducibility underpin all science, including global change research.

They include: reuse of data and code; need for data repositories and sharing platforms; standards required for sharing code and data effectively and accurately; citation and incentive mechanisms; capture and sharing of workflows; and ensuring statistical reproducibility in the computational and data science software stack.

Accurate capture and free exchange of data and information is inherent in this. Reproducibility is thus not drawn out separately in this report but is interwoven into its conclusions and recommendations. The term “reliability” of data is emerging as a possible alternative descriptor of the issues involved.