



Reproducibility Guidelines – IPIN 2026

1. WRITING THE DATA AND SOFTWARE AVAILABILITY SECTION

Include a Data and Software Availability section describing:

- What is available: datasets, software, scripts, computational environments, and documentation.
- Where it is available: opt for locations offering persistent links, i.e., DOIs: Zenodo, Figshare, OSF (with registration).
- Conditions of access: licenses, copyright, or other potential restrictions.

If some resources cannot be shared, you are asked to disclose in this section the nature of the limitation (e.g., sensitive data, proprietary code) and to provide alternative solutions that allow readers to understand and reproduce the work, such as:

- Synthetic or anonymized datasets with the same structure.
- Sample subsets of datasets or templates illustrating the data structure.
- Providing results (statistics, figures), based on the demonstration data (synthetic/anonymized/sample datasets that are shared), to allow comparison and verification of reproduction.
- Partial code sharing where feasible, e.g., sharing the main workflow while clearly indicating any modules that cannot be disclosed, as black box functions.

When some or all resources cannot be shared, authors are only asked to disclose in this section the nature of the limitation (e.g., sensitive data, proprietary code) without revealing confidential or commercial details.

2. USING OPEN DATA IN YOUR PAPER

- Use/share Open Research Data whenever possible. If restrictions prevent sharing (e.g., due to privacy or licensing limitations), declare the nature of the restriction and provide alternative approaches to support reproducibility, such as synthetic datasets or data templates.
- Provide thorough documentation and metadata for all datasets, including content, provenance, assumptions in the data collection, detailed parameters for collection (e.g., geographic extent, time interval, or timestamps), all data cleaning steps, and data types. Clearly define potentially used training/validation/test subsets to facilitate coherent reuse.
- Use citable repositories that provide persistent identifiers, i.e., a DOI or equivalent, and ensure long-term accessibility. Include dataset citations in the paper, specifying the identifier, version, and date.
- When sharing data, opt for open, non-proprietary formats to maximize long-term accessibility and interoperability.
- Specify the license of the datasets and, when possible, prefer open and permissive licenses to encourage reuse.

Following these recommendations helps ensure that your work is verifiable, reproducible, interpretable, and usable by the broader research community.

3. SHARING CODE AND COMPUTATIONAL WORKFLOWS

- Share code covering as many steps as possible: Data preprocessing, implementation of the methods and baselines, experimental setup and evaluation, and generation of results (tables, statistics, figures) presented in the paper.
- Clearly connect and relate the code to the paper's content, including tables, figures, maps, statistical values, and results.
- Ensure code quality and documentation: share code that is readable, well-structured, and accompanied by sufficient explanations to help others understand and use it effectively.
- Provide a complete machine-readable specification of the computational environment, such as a Dockerfile, an installation script, or a dependency manifest (e.g., a requirements.txt, a Conda YAML file, or a renv snapshot), listing all required dependencies and their exact versions ("version pinning").
- Provide (in the code repo) detailed reproduction instructions targeting human users in a README file, flowchart, or script that guides readers through running the code. Include indicative execution times, hardware requirements, and any assumptions necessary to execute the workflow.
- Host self-developed code/software in a public repository with a permissive license and persistent identifiers (e.g., DOI, SWHID, etc.) to ensure long-term accessibility and archiving of the precise version used.
- Declare licenses for all self-developed code components explicitly, to inform users of usage rights. In case of doubt, you may choose MIT or an equivalent permissive license.
- Request peer validation: Have a colleague follow the provided instructions to ensure that the code runs correctly and reproduces consistent results.

Following these recommendations will maximize the reproducibility, clarity, and usability of your computational workflows.

**WE ARE LOOKING FORWARD TO
WELCOME YOU AT IPIN 2026!**





4. FULLY REPORTING METHODS AND EXPERIMENTAL SETTING

Authors are asked to ensure that all proposed methods, computational workflows, and experimental procedures are clearly and fully described in the paper to support reproducibility. When preparing the paper, especially when the code is not openly shared, reflect on the following questions:

- Does the reader have sufficient information to accurately recreate the proposed method and the workflow and fairly represent the paper's contributions as a baseline in future work?
- Are all steps (preprocessing, methods and baselines description, experimental setup, evaluation, analysis, and visualisation) sufficiently described in the text?
- Are inputs, parameters, and procedures explained clearly enough to allow a representative replication?

The goal is to enable readers to understand, reproduce, replicate, and build upon the work, whether the full code or data are publicly available or not. Authors should provide as much detail as necessary to achieve this, including diagrams, pseudocode, or textual descriptions of data retrieval and computational steps.

