



Data Papers in Civil Engineering: Requirements, Preparation, and Best Practices

Yong Kwon Cho, Ph. D., Professor
School of Civil & Environmental Engineering, Georgia Institute of Technology Co-
Editor-in-Chief, ASCE Journal of Computing in Civil Engineering



1

OVERVIEW

Introducing Data Papers

A new article type in the ASCE Journal of Computing in Civil Engineering

“Computing in civil engineering has become increasingly reliant on access to large, well-documented data sets — from sensor networks, simulation pipelines, robotic field systems, and computer vision models.”

— Cho & El-Gohary, JCCE Editorial (2026)

EDITORIAL • FREE ACCESS

Introducing Data Papers: A New Article Type in the Journal of Computing in Civil Engineering

Yong K. Cho, Ph.D., M.ASCE
Nora El-Gohary, Ph.D., A.M.ASCE

Journal: J. Comput. Civ. Eng., Vol. 40, Issue 3, Feb 2026
Type: Editor's Note
DOI: [10.1061/JCCE5.CPENG-7616](https://doi.org/10.1061/JCCE5.CPENG-7616)

[READ ARTICLE →](#)



ASCE JCCE / Data Paper: Concept and Application



Yong K. Cho, Ph.D. — Georgia Tech

2

MOTIVATION

Why Data Papers?

Elevating data sets to first-class research outputs

Data ≠ Supplement anymore

Data sets are now treated as primary research outputs — peer-reviewed, citable, and built for reuse rather than tucked into supplementary appendices.

PROMOTES

Reproducibility · Transparency · Reuse

✗ Traditional papers fall short

Focus stays on methods + results. Underlying data is partial or preprocessed, with limited reusability.



Funders push transparency

NSF and similar agencies now require formal data management plans and preservation. E.g. NSF FAIRO



Publishers reinforce the shift

ASCE and other publishers require data availability statements to accelerate science.

ASCE JCCE / Data Paper: Concept and Application



Yong K. Cho, Ph.D. — Georgia Tech

3

DEFINITION

What is a Data Paper?

A peer-reviewed paper about the dataset itself

Data Paper

=

A paper about the dataset itself

✗ NOT about

- model performance
- algorithm novelty
- experimental benchmarks

✓ Focuses on

- Dataset structure
- Dataset collection
- Dataset validation
- Dataset usage

ASCE JCCE / Data Paper: Concept and Application



Yong K. Cho, Ph.D. — Georgia Tech

4

SCOPE

Where it Fits in Civil Engineering Computing

Domains and dataset types welcomed by JCCE



EXAMPLE DATASET TYPES

- 01 LiDAR datasets**
Point clouds, 3D scans, semantic segmentation labels
- 02 Simulation datasets**
Synthetic environments, generated training samples
- 03 Sensor time-series**
IoT, structural health monitoring, environmental sensors
- 04 Annotated vision datasets**
Image / video annotations for construction, infrastructure

ASCE JCCE / Data Paper: Concept and Application

GT Georgia Tech.

Yong K. Cho, Ph.D. — Georgia Tech

5

and legal/IRB compliance

REQUIREMENTS

Key Requirements

Three pillars every JCCE data paper must satisfy

01 Open & Accessible

- Public repository (DOI required)
- No restricted access
- Follows FAIR principles
- Ethics/IRB compliance

Findable · Accessible · Interoperable · Reusable

02 Documentation

- README + comprehensive metadata
- Format explanation
- Annotation details and labels

Standards-aligned, machine-readable

03 Reusability

- Usage code & scripts provided
- Clear directory structure
- Examples for downstream use

Plug-and-play for other researchers

ASCE JCCE / Data Paper: Concept and Application

GT Georgia Tech.

Yong K. Cho, Ph.D. — Georgia Tech

6

OWNERSHIP & LICENSE

Data Ownership and License

Authors retain dataset copyright; ASCE holds copyright of the published article



Authors

Retain copyright of the dataset



ASCE

Holds copyright of the published article

Common License Options

CC BY 4.0	CC BY-NC 4.0	Custom / Restricted
<p>REUSE Academic + commercial reuse with attribution</p> <p>BEST FOR Maximizing openness and reuse</p>	<p>REUSE Academic, educational, non-profit reuse</p> <p>BEST FOR Data with commercial-use constraints</p>	<p>REUSE Depends on stated terms</p> <p>BEST FOR Special cases — may reduce reusability</p>

Note · If commercial reuse is restricted, the reason should be clearly stated.

ASCE JCCE / Data Paper: Concept and Application



Yong K. Cho, Ph.D. — Georgia Tech

7

SUBMISSION

General Submission Information

Standard JCCE journal format with dataset usage code



MANUSCRIPT FORMAT

30

max double-spaced pages

- Follows the standard JCCE (ASCE) journal paper format
- Structured as a full journal manuscript — not a supplementary document
- Includes figures, tables, and references



DATASET USAGE CODE

Provide code, scripts, or utilities used to process and use the dataset.

ENSURES

- Reproducibility of preprocessing & annotations
- Ease of integration into research workflows
- Practical reusability of the dataset



Code can be hosted in GitHub or another public code repo

ASCE JCCE / Data Paper: Concept and Application



Yong K. Cho, Ph.D. — Georgia Tech

8

STRUCTURE

Required Paper Structure

Standard sections every data paper should include

1 Introduction
Purpose, target users, motivation, and FAIR commitments

2 Dataset Description
Contents, scope, scale, types, formats, and class taxonomy
Folder structure and file organization

3 Methodology / Data Collection
How data was collected and created. **Provide enough detail to reproduce the dataset**

4 Data Validation
Annotation and data quality checks; consistency/integrity
Focus on **Data correctness & reliability, not model performance**

5 Usage Notes
Access (DOI, license), use cases, applications, and known limitations, tips for effective use

6 References
Cite related datasets, tools, standards, and prior work

★ **Key takeaway:** Methodology shows **how data is created** · Validation answers **why it's reliable** · Usage Notes show **how data is used**

ASCE JCCE / Data Paper: Concept and Application

GT Georgia Tech.

Yong K. Cho, Ph.D. — Georgia Tech

9

PEER REVIEW

Review Criteria

What reviewers evaluate — and what they explicitly do not

WHAT REVIEWERS EVALUATE

Dataset relevance
Fits JCCE scope and computing-in-CE community

Documentation quality
README, metadata, and clear schema

Accessibility
Open availability with persistent DOI

Validation rigor
Annotation QA, integrity, and consistency

Reuse potential
Generalizability and downstream value

IMPORTANT

Reviewers do NOT need to re-run experiments.

They judge:

Credibility & Usability

Verification focuses on whether the dataset is well-documented, openly accessible, and actually useful — not on benchmarking model performance.

ASCE JCCE / Data Paper: Concept and Application

GT Georgia Tech.

10

SUMMARY

Key Takeaways

What to remember when preparing your data paper



Data papers

elevate datasets to first-class research outputs.

Position the dataset as your main contribution.



Accessibility

Public DOI, no restricted access



Documentation

Metadata, schema, usage notes



Validation

Quality checks and known limits



Reusability

Code, scripts, and clear structure