Plans for Data & Software in INSPIRE

Software Citation and Recognition in HEP Workshop, November 2022

Micha Moskovic (CERN)
INSPIRE Product Manager

Introduction

- INSPIRE is the main information platform for High-Energy Physics
- Main collection is Literature: 1.5M records
- Includes papers about Software

GEANT4--a simulation toolkit #2
GEANT4 Collaboration • S. Agostinelli (Genoa U.) et al. (Aug, 2002)
Published in: *Nucl.Instrum.Meth.A* 506 (2003) 250-303
pdf   links   DOI   cite   claim   reference search   16,213 citations

- Data & Software not directly supported currently (was partially implemented on old platform)

A special case: HEPData integration

Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV   #1

ATLAS and CMS Collaborations • Georges Aad (Marseille, CPPM) et al. (Jun 7, 2016)

Published in: *JHEP* 08 (2016) 045 • e-Print: 1606.02266 [hep-ex]

Note: 70 pages plus author lists + cover page (104 pages total), 32 figures, 22 tables, submitted to JHEP. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2015-07/ and at http://cms-results.web.cern.ch/cms-results/public-results/publications/HIG-15-002/

📄 pdf   🔗 DOI   ⤷ cite   ☰ datasets   ✎ edit   🔖 claim   🔍 reference search   ⟲ 1,705 citations

https://www.hepdata.net/record/ins1468068

Only for HEPData currently!

iNSPIRE HEP

Slogan: "First class Data & Software"

High-level objectives

- Help & incentivize researchers to make their Data & Software FAIR
- Collect & index all Data & Software products relevant for HEP
- Give credit to authors of Data & Software
- Avoid manual work for INSPIRE team

On roadmap for 2023

Content selection

- In principle: any Data & Software products relevant for HEP
- In practice: harvest metadata from relevant trusted repositories
    - Must adhere to FAIR principles
    - Must use DOIs
    - Must have APIs for automated retrieval
- Examples:
    - Zenodo (INSPIRE HEP Community)
    - HEPData
    - CERN OpenData
- No direct submissions

Organization

- New Data & Software collection on same footing as Literature
- Search in metadata
- Linking to other collections
  - Literature → related Data / Software records
  - Data / Software record → Author profiles
  - Data & Software tab on Author profiles
  - Data / Software → Literature citations

Citations

- Count citations like for papers
- Use DOI only for identification
- Aggregate different versions (similar to arXiv handling)
- Distinguish from Literature citations
- Standard citation formats like for Literature (some open questions)

Summary

- INSPIRE currently handles Software papers
- Will add Data & Software records in medium term
- Automated ingestion from a limited number of repositories
- Track & count citations to Data & Software by DOI

Further questions, comments: micha.moshe.moskovic@cern.ch

iNSPIRE HEP