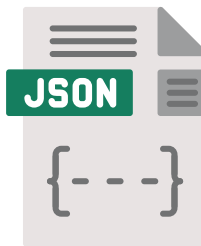
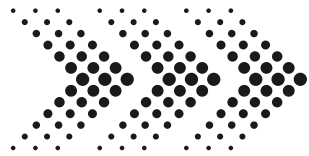


# AUTOMATED METADATA EXTRACTION COMPLIANT WITH MACHINE-ACTIONABLE SOFTWARE MANAGEMENT PLANS

DHWANI SOLANKI, SUHASINI VENKATESH, DIETRICH REBHOLZ-SCHUHMANN, AND LEYLA JAEI CASTRO  
ZB MED INFORMATION CENTRE FOR LIFE SCIENCES



TRANSFORMING GITHUB REPOSITORIES INTO FAIR AND MACHINE ACTIONABLE METADATA

## Machine-Actionable Software Management Plan (maSMP)



A structured document that organizes all the important metadata about a software project in a machine-readable format.

1



Manually tracking is time-consuming

2



Information gets lost or outdated

3

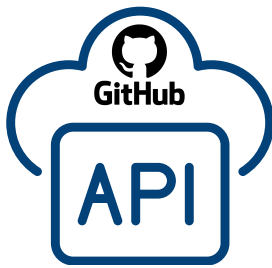


Hard to find, cite, and reuse

## Automated Metadata Extraction



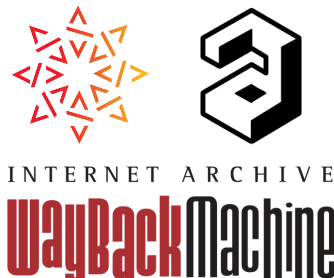
GitHub URL



Extract from GitHub API



Parse Repository Files



Enrich with External Sources



JSON-LD



**F**indable

Extracts metadata from GitHub and third-party sources, boosting discoverability and ensuring reliable citations with persistent identifiers.



**A**ccessible

Metadata is stored in JSON-LD, an open standard for easy sharing, ensuring that key information is always retrievable.



**I**nteroperable

Uses standardized metadata schemas to integrate with research repositories and linked open data sources.



**R**eusable

Provides structured, machine-readable metadata for software preservation, enriched with reliable external platforms for long-term usability.

