

# Stand HERMES und Software CaRD

Zur Bereitstellung von Metadaten für die Software-Indikatorik

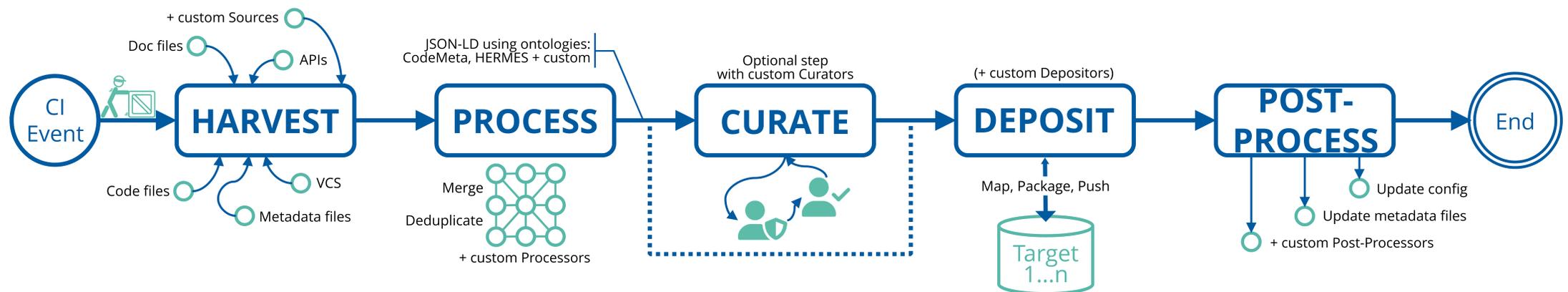
# Übersicht HERMES

## Pipeline zur Veröffentlichung von Research-Software



<https://hermes.software-metadata.pub>

- HERMES sammelt Metadaten über ein Software-Projekt zur Veröffentlichung
  - Quellen: Git-History, CITATION.cff, codemeta.json, ...
  - Ziele: Zenodo, Rodare, Jülich Data, ... (Invenio, Dataverse)
- HERMES läuft lokal oder in CI/CD, ist kein Online-Service
- Programmablauf in mehreren Stufen, konfigurierbar und durch Plugins anpassbar
- Datenmodell: RDF, hauptsächlich basierend auf CodeMeta



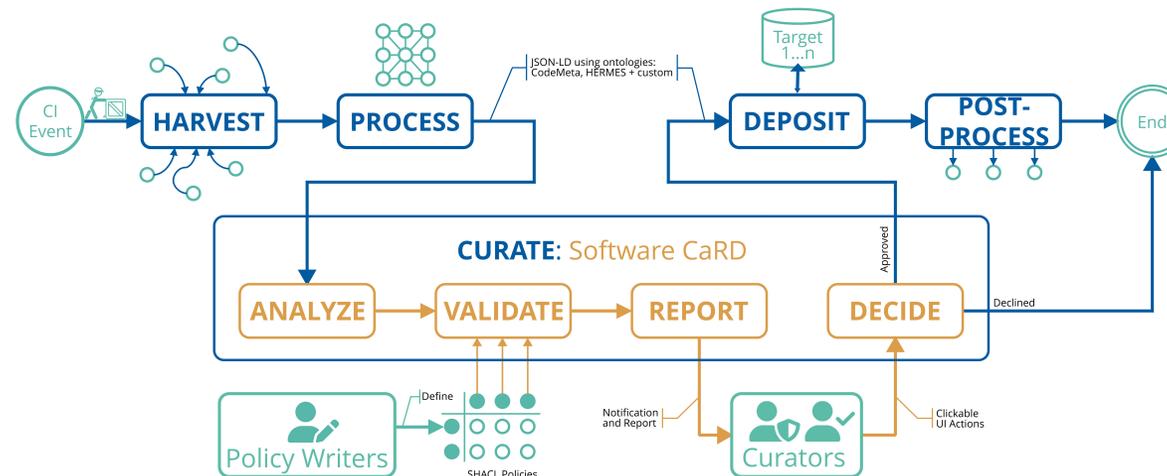
# Projekt Software CaRD

Poster: [10.5281/zenodo.15341915](https://zenodo.org/record/15341915)

Beispiel-Policies: [10.5281/zenodo.17043330](https://zenodo.org/record/17043330)

## Kurations- und Reporting-Dashboard als Curation-Plugin

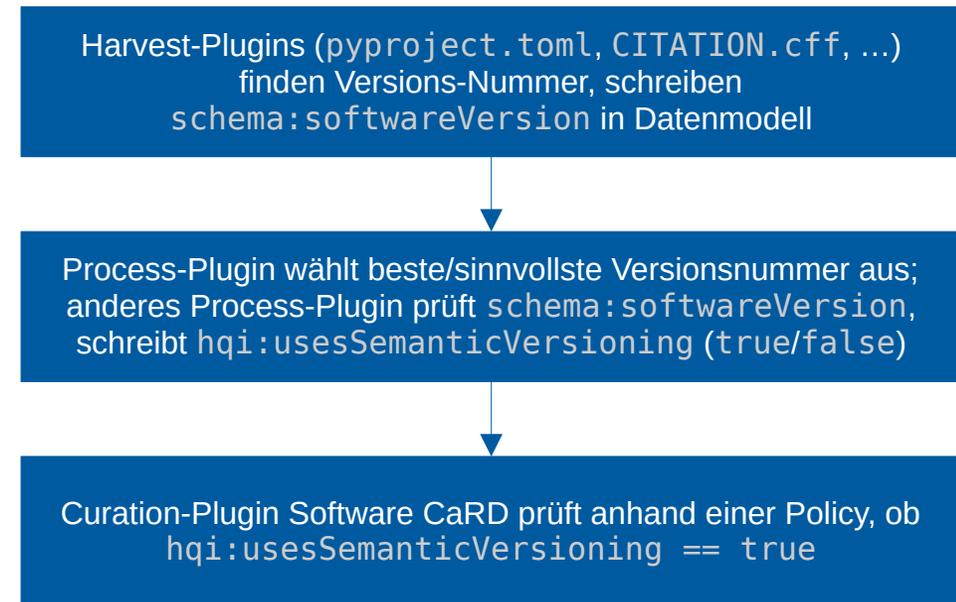
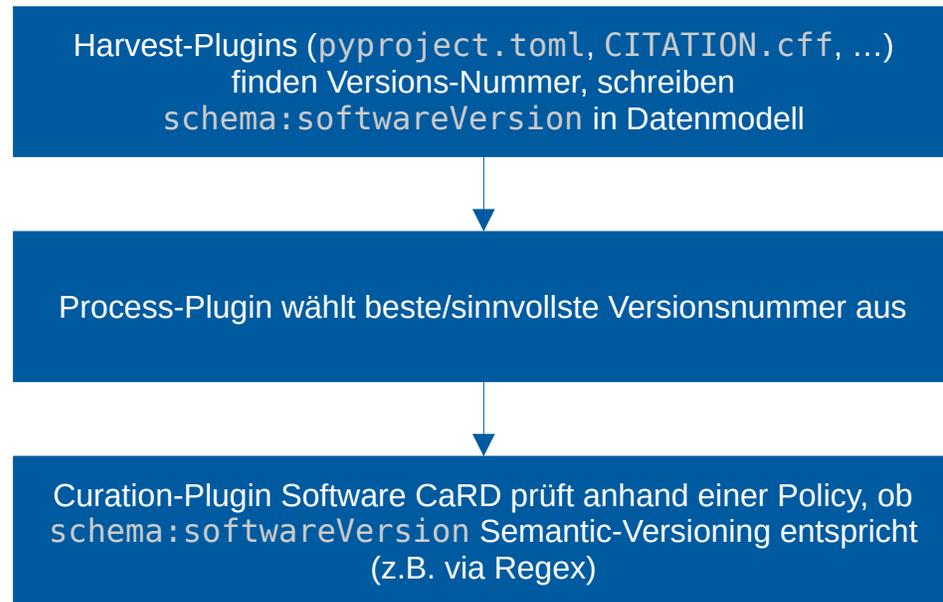
- Überprüfung gesammelter Metadaten anhand von Policies, grafische Darstellung der Ergebnisse
  - Z.B. erlaubte Lizenz genutzt, korrekte Email-Adressen angegeben, Beschreibung vorhanden, ...
  - Sicherstellung hochwertiger Metadaten **vor** der Publikation
- Erstellung von generischen, konfigurierbaren Policies in SHACL
- Checks laufen auf HERMES' Datenmodell
  - Wenn zusätzliche Metadaten benötigt → Harvest- und Process-Plugins
  - Veröffentlichung der Metadaten abhängig von Mapping in Deposit-Plugin



# Zusammenspiel von HERMES und Software CaRD

HERMES sammelt Informationen, Software CaRD überprüft

- Policies in SHACL haben Einschränkungen, z.B. bei I/O, komplexen Berechnungen, ...
  - Abhilfe durch mehr Harvest-/Process-Plugins
- Oft prinzipiell zwei Möglichkeiten zur Implementierung
- Beispiel: **F2 Level 2**, *The software uses structured (e.g. semantic) versioning.*



# Nutzung von HERMES für die Indikatorik

## helmholtz.software als Indikator-Sammelstelle

- Zentral bereitgestellte HERMES- und CI/CD-Konfiguration, auch Einbindung in `hermes init` möglich
- Für Projekte, die schon HERMES / CI/CD nutzen:
  - Übernehmen relevanter Teile in das eigene Setup
  - Updates z.B. bei Publikation neuer Version
  - Senden von Metadaten an RSD via API (mit User-Authentifizierung) im Deposit-Schritt
- Für andere Projekte:
  - Prozess wird unabhängig von Software-Entwicklungs-Workflow angestoßen (z.B. durch Reporting-Team oder automatisiert)
  - Senden von Metadaten an RSD via API (mit System-Authentifizierung) oder Einlesen von Datei

# Nutzung von HERMES für die Indikatorik

## Abdeckung des minimalen Polygons

- HERMES deckt minimales Polygon größtenteils ab
- Fehlende Metadaten (siehe auch [HERMES-Issues](#)) / unklare Bestimmbarkeit:
  - Existenz bestimmter Files (README, LICENSE, ...) wäre hilfreich zu wissen → [Issue #271](#)
  - **T2 Level 1**, *All files are provided in some structured/unstructured way inside the publication repository.*
  - **T3 Level 1**, *Every developer is free to use his/her own style of coding.*
  - **T4 Level 1**, *The code follows a modular structure allowing for component reusability.*
- Nur gewisse hinreichende Bedingungen für alle Kriterien prüfbar, erschöpfende Suche nicht möglich
  - Beispiel: **A3 Level 1**, *“How to install” information is provided.* unter Nutzung von CodeMeta-Plugin
  - Kriterium ist nur dann erfüllt, wenn Datei `codemeta.json` existiert und ihr Feld `buildInstructions` eine URL enthält
  - Folge: Researcher müssen CodeMeta-Datei bereitstellen
- „HQI“-Plugin(s) nötig, die gesammelte Metadaten in erfüllte Kriterien „übersetzen“

# Aktuelle Entwicklungen

- Dataverse-Plugin in der Finalisierungsphase
- DLR refaktoriert Datenmodell – Stand?
- Software CaRD Hackathon, 6. – 8. Oktober

## Projektantrag SQuaRE (*Software Quality Reporting Ecosystem*)

- Eingereicht bei Science Serve
- Automatisierte Berechnung des Indikators mit HERMES
  - Entwicklung zusätzlicher Plugins, z.B. für Somesy → weitere Programmiersprachen-Ökosysteme
- Visualisierungen in helmholtz.software
  - RSD als zentrale Sammelstelle für Indikatorik-Daten
  - Entwicklung eines Austauschformats für Indikatorik-Daten
- Einfache Nutzbarkeit für Researcher, Coder, Administration