



Use cases and identifier schemes for persistent software source code identification

The Challenge:

Software, and in particular source code, plays an important role in both industrial and academic research. However, unlike research data and scientific articles, the overwhelming majority of software source code is developed and used outside the academic world, in industry and in developer communities. As a consequence, software in general is either not formally identified or referenced at all, or is identified and referenced through methods that are totally different from the ones typically used in scholarly publications.



Produced by: **RDA/FORCE11 Software Source Code Identification WG**

<https://www.rd-alliance.org/groups/rdaforce11-software-source-code-identification-wg>

What is the solution?

This document captures the current state-of-the-art of the practice of software identification, including use cases and identifier schemes from different academic domains and industry, mapping and clarifying the usage of different identifiers, including both intrinsic and extrinsic identifiers. This should provide solid ground on which to build recommendations for the academic community, and help academic and industrial stakeholders to adopt solutions compatible with each other and especially with the software development practice of tens of millions of developers worldwide.

What is the impact?

Provides an overview of the current *state-of-the-art of the practice of software identification*, including use cases and identifier schemes from different academic domains and in industry.

Find out more about the use cases and identifier schemes for persistent software source code identification



September 2020