



European  
Commission



# CROSSMINER

**CROSSMINER enables in-depth analysis, evidence-based selection and monitoring of open source components, and facilitates knowledge extraction from large open source software repositories to support development of systems for a wide range of applications by reusing the best available open source software. Selected open source components are monitored to raise alerts related to quality, and to give suggestions to reduce development effort and increase the quality of new products and services.**

## AT A GLANCE

### Project title

Developer-Centric Knowledge Mining from Large Open-Source Software Repositories

### Project coordinator

The Open Group (UK)

### Partners

Athens University of Economics & Business (GR)  
Bitergia (ES)  
Centrum Wiskunde & Informatica (NL)  
Eclipse Foundation Europe (DE)  
Edge Hill University (UK)  
FrontEndART (HU)  
OW2 Consortium (FR)  
Softteam (FR)  
University of L'Aquila (IT)  
University of York (UK)  
Unparallel Innovation (PT)

### Duration

01.2017 – 12.2019

### Total cost

4.519.007 €

### EC Contribution

4.519.007 €

### Programme

H2020-ICT-2016-1

### Further information

[www.crossminer.org](http://www.crossminer.org)

## Context and motivation

Open source software (OSS) is computer software provided under a license that permits developers to study, change, and improve the software for free. A report by Standish Group states that adoption of open source software models has resulted in savings of about €58 billion per year to consumers.

Unlike commercial software, which is developed with a business motivated commitment to provide support and updates, OSS technologies are most often developed in a public, collaborative, and loosely-coordinated manner. This has important implications for the quality of different OSS technologies, as well as the level of support that different OSS communities provide to software developers.

## Challenge

While there are several high quality and mature OSS projects, there are also many OSS projects that are dysfunctional due to one or more of the following:

- The development team behind the OSS project invests little time in further development and support
- Development of the OSS project has been discontinued due to lack of commitment or motivation

- Documentation is limited or of poor quality making the software difficult to understand and to update
- The community around the OSS project is small, questions receive late/no response, and defects get repaired slowly or ignored

Consequently, developing new software systems by reusing existing OSS components raises challenges in:

- Searching for OSS candidate components
- Evaluating and selecting the most suitable OSS components from amongst identified candidates
- Adapting selected OSS components to fit specific requirements of new software products and services

## Solution

CROSSMINER will deliver an integrated platform for development of complex software systems by (1) enabling monitoring, in-depth analysis and evidence-based selection of OSS components, and (2) facilitating knowledge extraction from large open source software repositories. The following key scientific and technology innovations will be developed:

- OSS analysis tools to extract and store actionable knowledge from a collection of OSS projects
- Natural language analysis tools to extract quality metrics related to community communication channels and bug tracking systems of OSS projects
- System configuration analysis tools to provide an integrated development and operational view of OSS projects
- Workflow-based extractors that simplify creation of customised analysis and knowledge extraction from OSS projects
- Cross-project analysis tools for understanding a wide range of OSS project relationships (e.g. dependencies and conflicts) based on developer defined similarity measures
- Advanced integrated development environments allowing developers to



easily adopt the CROSSMINER knowledge base and analysis tools while receiving OSS project alerts, recommendations, and feedback to improve developer productivity

The outcomes of the CROSSMINER innovations will contribute to the definition of a knowledge base addressing complex classifications of OSS projects and will automate identification of complementary and competing OSS projects, project incompatibilities, and prediction of the future of given OSS projects based on the evolution of other projects having similar historical characteristics.

## Expected impact

The key industrial impacts targeted by the project technologies and innovations are:

- Reduction in time-to-market of new generations of software enabled products and services as OSS continues to speed innovation and improve productivity
- Substantial productivity increase in all aspects of the software lifecycle especially for distributed systems
- Ability to meet quality levels required by a fast growing number of software-enabled products and services
- Increased reuse of code, design and functional requirements in the development of new software

CROSSMINER innovations are driven by six industrial user partners in the domains of IoT, multi-sector IT services, API co-evolution, software analytics, software quality assurance, and OSS forges, who will also validate the expected impacts from the project technologies are fulfilled.