Smart Software Development Assistant

Call for partners for H2020 ICT-16-2018 *Software Engineering* call *

Armand PUCCETTI Basile STARYNKEVITCH Franck VEDRINE

CEA, LIST, Palaiseau, France†

December 2017

1 H2020 ICT-16-2018 call *Software Technologies*

Look on *H2020 Work Programme 2018-2020*
5.1. Information and Communication Technologies[1] pages 36 and 37:

> New advances in ICT technology influence the way software is developed. Software is increasingly becoming a pervasive and enabling technology and the impact of software defined infrastructures in the software development & management processes will span across multiple technology domains (e.g HPC, IoT, Big Data, Cloud, Artificial Intelligence). There is a need for novel and generic software engineering methods and tools that are applicable across different domains [...]
>
> Scope a) **Integrated programming models & techniques for exploiting the potential of virtualised and software defined infrastructures** [research & innovation actions] [...]

2 Project core ideas

Increase the software development productivity and quality by developing an **open source** toolset for software developers and associated methodologies. This applies mainly to the software developments with **large software components** (written in C or C++ or Java, etc...[2]) and their existing ecosystem, such as many open-source projects. Productivity gains will be achieved by the toolset that digests in batch mode[3], on a powerful workstation or server, the entire source code and other related artefacts; then it extracts and builds some high-level persistent data, model and knowledge from it, using a combination of static source code analysis, natural language processing and machine learning techniques. Later, these elements will enable the software developer to query and use this data/model/knowledge interactively for more efficient development activities (designing, coding, debugging, maintenance,...). Quality will be improved as the toolset explicits abstract and human-understandable models/knowledge and high-level properties of the pre-existing and newly developed software. Time-to-market of

*`https://ec.europa.eu/digital-single-market/events/cf/ict-proposers-day-2017/document.cfm?doc_id=37698`

†email: *firstname*`.`*familyname*`@cea.fr`; phone: +33 1 6908 {8304/6595/9132}

[1] on `http://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-leit-ict_en.pdf`

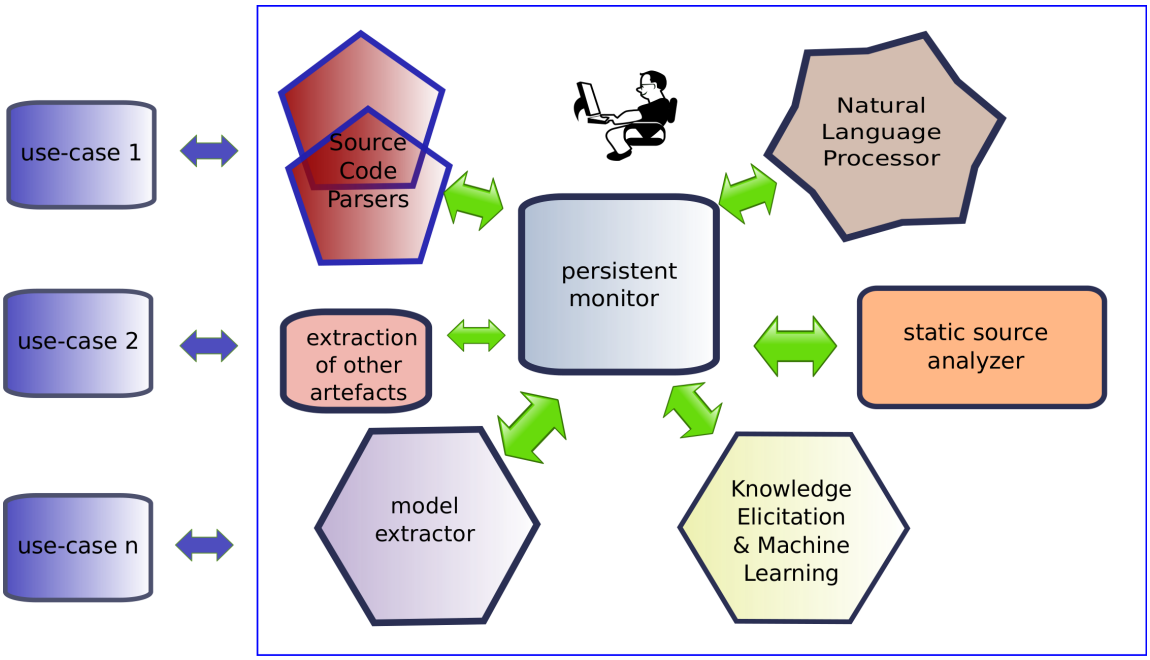
[2] If an industrial partner needs to handle some other programming language, we might consider that too!

[3] The processing by our toolset of the open source components will last much more than its compilation time, so we consider a digest time of many hours or several days of computer time for handling a free software component of several millions lines of code.

software enabled-products will be shortened as the resulting abstract models will help the developers to understand easily and quickly the various elements of the code.

We are aware that existing sound formal methods approaches (extensively used in Frama-C[4], with abstract interpretation and weakest precondition techniques) are powerful but don't scale well to large code bases, that model extraction approaches are useful but not enough, and that natural language processing can extract some knowledge from documentation (but don't use enough information inside the source code). So we propose to combine these techniques (also using machine learning, see figure 1) in a collaborative research & development project providing an open source toolset prototype.

Figure 1: toolset architecture



3 Potential partners and new ones

We already have academic partners expert in static analysis, in natural language processing and in machine learning; we also have some industrial partners bringing several ecosystems (e.g. GUI development in medical devices, using C++ and Qt; image processing for robotics, using C++ OpenCV).

We need more **partners, notably bringing a software ecosystem around Cloud or IoT, in Java or some other language** (other than C or C++).

You can download this sheet on `starynkevitch.net/Basile/partnersearch-ict16.pdf` and forward it to your colleagues.

git commit `648dd9f14be8706`

[4] `https://frama-c.com/`