# HiPEAC info 68

JANUARY 2023

**HiPEAC conference 2023**

## HiPEAC Vision 2023: A race against time

## Subhasish Mitra's NanoSystems and Dream Chips

## Drive through of adaptive computing, with Ivo Bolsens

## Liliana Cucu-Grosjean on probability and safety-critical systems



Fully autonomous vehicles will only be possible thanks to artificial intelligence (AI), yet the approach to AI is radically different to the approach to functional safety, a prerequisite for mobility solutions. The European Union (EU) project SAFEXPLAIN seeks to change that, delivering explainable – and therefore certifiable – AI fit for use in safety-critical contexts. HiPEAC caught up with SAFEXPLAIN coordinator Jaume Abella (Barcelona Supercomputing Center) to find out more.

# 'You can't transfer control to the driver if the car doesn't have a steering wheel'

***Why is it essential to include AI in safety-critical systems such as cars, trains and satellites?***

Autonomous systems are the next big revolution in transportation and related domains. So far, only AI-based systems have been proven capable of providing accurate-enough perception and navigation solutions for transportation.

AI therefore has to work in these systems, but, as of today, there is no way to build safe AI-based systems in general. Until now, there have been workarounds for if the AI system fails, for example by transferring back the control to the driver. That's not an option if your car doesn't even have a steering wheel.

***What are the current roadblocks to applying AI in safety-critical systems?***

The problem is that the AI software development process is fundamentally at odds with software development processes for safety-critical systems. The latter take a top-down approach, starting from specific safety goals and requirements, and ensuring that the software is correct by design.

In the case of AI, however, the implementation is based on experimentation and intuition. It uses a bottom-up approach using representative data, which is then tuned empirically. This makes it very difficult to trace the reasons behind the decisions it makes, which means that the AI can't be certified for use in safety-critical systems without the aforementioned workarounds.

***What does SAFEXPLAIN set out to do?***

SAFEXPLAIN will devise deep learning (DL) solutions – a key subset of all AI solutions – starting from functional safety requirements and providing explainability and traceability. The project will also provide recommendations to adapt functional safety standards so that they can certify aspects of software that haven't been foreseen until now. This will allow the certification of DL software, as long as it can demonstrate that it has followed a particular methodology.

This implies a shift in thinking on the part of certification authorities. Traditionally, certification has allowed failure rates for hardware to allow for external factors such as radiation, but software has been assumed correct (infallible) by design. This new approach would allow certification of software that has a possibility of failing, but with multiple levels of diverse redundancy implemented to ensure that safety is maintained.

SAFEXPLAIN will also deliver DL software implementations that both meet the above requirements and run efficiently and with time predictability on relevant high-performance computing (HPC) platforms. In addition, it will provide multiple safety patterns with their corresponding DL realizations to meet varying safety requirements, along with industrial tools and case-study integrations proving the feasibility of the SAFEXPLAIN approach.

To achieve this, the project is bringing the functional safety and AI communities together and helping them to understand one another. In effect, my main role is to act as a translator, facilitating the definition of common terms that everyone can understand. I'm confident that the baseline solutions already exist within the AI community; it's just a case of identifying the necessary properties for safety-critical systems and bridging the gap between the two worlds.

*SAFEXPLAIN project consortium*