

*Fully-Funded Research Opportunity:*

# PhD Fellowship: Unite! ULisboa

**Supervisors:** João F. Ferreira (IST, ULisboa, 🌐), Maurizio Morisio (Politecnico di Torino, 🌐), and Marco Torchiano (Politecnico di Torino, 🌐)

## Learning-Based Program Repair: Advancing Software Quality and Productivity

### Description

Learning-Based Program Repair refers to a type of Automated Program Repair technique that uses Artificial Intelligence / Machine Learning algorithms to automatically identify and fix software bugs.

Recent work on learning-based program repair has achieved impressive results [0,1]. However, it is still impossible to repair many bugs, especially those of a semantic nature or that are domain-specific [2,3,4].

The main goal of this project is to create new Learning-Based Program Repair techniques that advance the state of the art in this area. In particular, we will investigate how Large Language Models that have been trained on code (e.g., OpenAI Codex [5], Github Copilot [6]) can be used for program repair, potentially in combination with other techniques such as Self-Supervised Learning [1,2].

Automated program repair will revolutionize the software development industry, by dramatically improving software quality, reducing costs associated with software maintenance, and freeing up valuable developer time.

We will explore high-impact / high-value application domains. Examples include smart contracts for blockchain systems [7] and infrastructure code [8], but selected candidates will have the opportunity to apply the developed research in any domain of their interest.

---

## Requisites

The candidates should have an MSc in Computer Science or related areas.

Preference is given to candidates who satisfy one or more of the following criteria:

1. Interest in the areas of Software Engineering and Artificial Intelligence
2. Practical experience with modern Artificial Intelligence techniques, such as Deep Neural Networks and Large Language Models
3. Experience with Automated Program Repair methods and techniques
4. Experience at publishing at reputable CS events related to the topics addressed in this project

## Further Information

This project will be developed in collaboration with researchers from Universidade de Lisboa (ULisboa) and Politecnico di Torino (POLITO). The selected candidate will have the opportunity to spend time in Lisbon (Portugal) and Turin (Italy).

For more information, please contact João F. Ferreira ([joao@joaoff.com](mailto:joao@joaoff.com))

## References

[0] Goues, C.L., Pradel, M. and Roychoudhury, A., 2019. Automated program repair. Communications of the ACM, 62(12), pp.56-65.

[1] Yasunaga, M. and Liang, P., 2021. Break-it-fix-it: Unsupervised learning for program repair. In International Conference on Machine Learning.

[2] Ye, H., Martinez, M., Luo, X., Zhang, T. and Monperrus, M., 2022. SelfAPR: Self-supervised Program Repair with Test Execution Diagnostics. In 37th IEEE/ACM International Conference on Automated Software Engineering.

[3] Ye, H., Martinez, M. and Monperrus, M., 2022. Neural program repair with execution-based backpropagation. In 44th International Conference on Software Engineering.

[4] Silva, André, 2022. A Comparative Study of Self-Supervised Learning Techniques for Functional Program Repair, MSc Thesis (Supervised by Ferreira), Instituto Superior Técnico, ULisboa.

[5] OpenAI Codex. Available at: <https://openai.com/blog/openai-codex> (Accessed: 22 May 2022)

[6] GitHub Copilot: <https://github.com/features/copilot> (Accessed: 22 May 2022)

---

[7] Saavedra, N., and J. F. Ferreira. "GLITCH: Automated Polyglot Security Smell Detection in Infrastructure as Code." In 37th IEEE/ACM International Conference on Automated Software Engineering (CORE A\*). 2022.

[8] Durieux, T., Ferreira, J. F., Abreu, R., and Cruz, P. Empirical review of automated analysis tools on 47,587 Ethereum smart contracts. In ACM/IEEE 42nd International conference on software engineering (CORE A\*). 2020.