Finanziato dall'Unione europea
NextGenerationEU

Ministero dell'Università e della Ricerca

Segretariato Generale

Direzione Generale della Ricerca

PRIN: PROGETTI DI RICERCA DI RILEVANTE INTERESSE NAZIONALE – Bando 2022 PNRR
Prot. P2022553SL

PART A

1. Line of intervention

South line/Linea Sud

2. Research project title

FRINGE: context-aware FaiRness engineerING in complex software systEms

3. Duration of the project (months)

24 months

4. Strategic emerging Topics - 5. Related Cluster

Strategic emerging topic: CIRCULAR ECONOMY
Cluster: Digital, Industry and Space
Sub Cluster: 6. Human-centred and ethical development of digital and industrial technologies ensured through a two-way engagement in the development of technologies, empowering end-users and workers, and supporting social innovation.

6. Main ERC field

PE - Physical Sciences and Engineering

7. Other ERC field

8. ERC subfields

1.	PE6_3 Software engineering, programming languages and systems
2.	
3.	

9. Keywords

nº	Testo inglese
1.	Software Engineering
2.	Software Engineering for Artificial Intelligence
3.	Mining Software Repositories
4.	Software Fairness Engineering
5.	Ethical Artificial Intelligence
6.	Machine Learning

10. Principal Investigator

PALOMBA (Surname)	**FABIO** (Name)
Ricercatore a t.d. - t.pieno (art. 24 c.3-b L. 240/10) (Qualification)	
03/08/1989 (Date of birth)	**PLMFBA89M03F839I** (Personal identification code)
Università degli Studi di SALERNO (Organization)	
(Phone number)	**fpalomba@unisa.it** (E-mail address)

Declarations

☑ I declare that I have not participated as PI in PRIN 2022 call (n. 104 02/02/2022)

☑ I declare that I have participated as associated PI in PRIN 2022 call (n. 104 02/02/2022)

Current funding and applications submitted

Age limits derogation

The principal investigator and the substitute are both under 40 at the time of the publication of the call. They intend to benefit from the derogations to the age limits for the amount allocated to under 40 PI;

11. List of research units (RU)

nº	Associated Investigator	Qualification	University/ Research Institution	Registered office (address)	Does operating office coincide with registered office?	Operating office (address) *	Contract of availability of operating office	Declaration of availability of the operating office and copy of contract	e-mail ac
1.	PALOMBA Fabio	Ricercatore a t.d. - t.pieno (art. 24 c.3-b L. 240/10)	Università degli Studi di SALERNO	Via Giovanni Paolo II, 132 - FISCIANO - SALERNO (SA)	yes	Address Via Giovanni Paolo II, 132, 84084 Fisciano SA			fpalomba@
2.	DI SORBO Andrea	Ricercatore a t.d. - t.pieno (art. 24 c.3-b L. 240/10)	Università degli Studi del SANNIO di BENEVENTO	Piazza Guerrazzi, 1 - BENEVENTO (BN)	yes	Address Via Traiano, 3, 82100 Benevento BN			andrea.disorb
3.	DI RUSCIO Davide	Professore Associato (L. 240/10)	Università degli Studi dell'AQUILA	Palazzo Camponeschi, Piazza Santa Margherita 2 - L'AQUILA (AQ)	yes	Address Palazzo Camponeschi, Piazza Santa Margherita 2 - L'Aquila AQ			davide.diruscio

12 - Substitute Principal Investigator (PI) (To be identified among one of the associated PIs participating in the project).*

DI SORBO (Surname)	**ANDREA** (Name)
Ricercatore a t.d. - t.pieno (art. 24 c.3-b L. 240/10) (Qualification)	
18/07/1984 (Date of birth)	**DSRNDR84L18G596E** (Personal identification code)
Università degli Studi del SANNIO di BENEVENTO	

(Organization)

(Phone number)

andrea.disorbo@yahoo.it
(E-mail address)

13. Brief description of the proposal

Machine learning (ML) pervades our daily lives, being embedded in any kind of device, service, and/or system, and enabling complex data-driven decision-making solutions. By 2025, over 463 exabytes of data will be generated daily and the digital transformation will further increase the pervasiveness of ML solutions and the subsequent risk that those solutions may unethically exploit sensitive pieces of data, potentially leading to forms of discrimination. This is why the European Union, through the Artificial Intelligence Act (https://artificialintelligenceact.eu/), raised the concern of having novel solutions to ensure trustworthy and ethical artificial intelligence able to protect ethical principles.

Project FRINGE (context-aware FaiRness engineerING in complex software systEms) studies solutions for engineering software fairness in the context of ML-Intensive systems. FRINGE foresees systems where data are continuously collected through heterogeneous interconnected smart devices and real-time monitoring devices. Collected data are filtered, aggregated, and consumed by ML models. The large amount of data available, the interconnection among them, and the continuous data acquisition/learning activities raise serious ethical concerns, as it is hard to verify that ML models act according to ethical principles and make predictions that do not perpetuate discrimination toward sensitive groups.

The overarching goal of the project is to provide software engineers, data scientists, and ML experts with a comprehensive set of methodologies, approaches, and software engineering (SE) solutions to improve the development, monitoring, and design of fairness-related properties of ML-intensive systems. FRINGE will help to sustain the digital transformation by empowering advanced artificial intelligence applications through instruments that can lead to the trustworthy and ethical exploitation of large amounts of data. More specifically, FRINGE aims to conceive: (a) a metamodel for ML fairness; (b) a fairness-aware monitoring infrastructure for ML-intensive systems, (c) solutions able to exploit contextual information to recommend and spot issues with fairness definitions, requirements and specifications for the specific task at hand; (d) design pattern recommenders and antipattern detection approaches.

The FRINGE activities will be carried out through 6 work packages: project management, four technical threads, and dissemination and exploitation. The activities will be based on concrete problems and case studies, i.e., one main testbed, plus a combination of exemplars and open-source systems. The project leverages the solid research records of its participants. While the partners have already collaborated, either in previous/ongoing projects or in joint publications, their combined expertise will help to reach the foreseen goals.

14. Total cost of the research project identified by items

Associated Investigator	item A.1	item A.2	item B	item C	item D	item E	item F	Total
PALOMBA Fabio	9.125	77.500	0	0	12.994	0	0	**99.619**
DI SORBO Andrea	9.000	77.500	0	0	12.975	0	0	**99.475**
DI RUSCIO Davide	9.875	77.500	0	0	13.106	0	0	**100.481**
Total	**28.000**	**232.500**	**0**	**0**	**39.075**	**0**	**0**	**299.575**

N.B. The Item D and TOTAL columns will be filled in automatically

- item A.1: enhancement of months/person of permanent and temporary employees
- item A.2: cost of contracts of non-employees, specifically to recruit
- item B: cost of equipment and tools
- item C: cost of consulting and other services
- item D: overhead
- item E: materials cost
- item F: other costs

PART B

B.1

---

1. State of the art

We report the state of the art for the key aspects treated by FRINGE.

Assembling fair ML-Intensive Systems

Studies have been conducted to treat software fairness throughout the lifecycle of ML-intensive systems, yet literature still lacks comprehensive and systematic/engineered solutions like those conceived by FRINGE. Biswas and Rajan (2021) empirically compared 37 different configurations of ML pipelines, showing that a number of data preparation strategies may induce ethical bias. A recent systematic literature review (Pessach and Shmueli, 2022) classified strategies for pipeline-oriented fairness enhancement. They reported on (1) *pre-process strategies* - which consist of data manipulation steps conducted by taking fairness into account before training an ML model; (2) *in-process strategies* - that modify training strategies based on algorithmic fairness; and (3) *post-process strategies* - which work on the basis of an output scoring, namely they aim to make the decisions made by a learner more fair a posteriori.

Some commercial instruments have also been developed. Fairlearn by Microsoft (2022) allows practitioners to assess fairness based on group similarity, i.e., by studying how the predictions made by a ML model vary when considering different groups of training instances. In contrast, AI Fairness 360 (IBM, 2022), enables the application of in-process strategies, making available 10 fairness algorithms.

Fairness-Aware Data-Preparation

Data-preparation steps are extremely impactful to deal with fairness in an ML-Pipeline workflow. Biswas and Rajan (2021) analyzed multiple data transformation patterns (e.g., feature generation, missing values removal) to reduce the sources of unfairness in training data. Balayn et al. (2021) introduced a data-centered approach to enforce fairness requirements and constraints on datasets characteristics. Zelaya (2019) used data rebalancing to measure volatility in two public fairness datasets. Chakraborty et al. (2021) introduced re-sampling and generation of synthetic instances, allowing the training data to reach equal distributions among different labels, and different protected attributes.

Fairness Design and Verification

Finkelstein et al. (2008, 2009) proposed an approach to balance fairness and non-functional requirements during requirements elicitation and analysis. Albarghouthi et al. (2017) showed that various formal definitions of fairness can be modeled as probabilistic program properties. FairSquare was conceived to verify probabilistic properties, accepting a wide range of decision-making programs. Recently, Chakraborty et al. (2021) built Fair-SMOTE, creating synthetic data that could reduce the risk of discrimination for minorities contained in training data. Peng et al. (2022) developed xFAIR as an extension of Fair-SMOTE. Differently from Fair-SMOTE, once built the model does not require from users access to protected attributes (those that identify the social groups that the data instances under consideration belong to). In this respect, the approach only collects and uses protected attributes in its building phase. Galhotra et al. (2017) devised Thermis, a fair-testing approach that exploits automated test suite generation and data rebalancing approaches to detect causes of fairness. Similarly, Perera et al. (2022) developed a tool to control the fairness of regression-based ML systems within a continuous integration scenario. Other fair-testing approaches are based on black-box testing. For instance, Aggarwal et al. (2019) combined symbolic execution and local explainability to verify the ethical level of an ML-intensive system.

With respect to the current literature, FRINGE aims to innovate under various perspectives. First, it will propose novel design practices and automated mechanisms to detect fairness issues–-this represents a premiere of our project. Second, it will embed the concept of “context” (fairness of ML models depend on the specific application contexts where they are applied) within testing mechanisms, providing approaches that could control fairness at a finer-level, being therefore more practical and useful.

Fairness Analytics

Calmon et al. (2017) created a framework that employs probabilistic measurements to help practitioners remove ethical bias. Similarly, Loukina et al. (2019) defined an analytical instrument to measure fairness in NLP instruments for educational purposes. Kusner et al. (2017) proposed a measurement model based on counterfactual fairness: this model applies a-posteriori analysis of the confusion matrix, where group fairness metrics are applied to analyze ethical bias.

Riazy et al. (2020) highlighted the lack of metrics and software analytics instruments for fairness. This paper motivates FRINGE: the project aims to propose novel metrics and analytics instruments, thus notably advancing the state of the art.

2. Detailed description of the project: methodologies, objectives, and results that the project aims to achieve; indicate deliverables and milestones outlining the project coherence as to the strategic themes, indicating clear and innovative objectives, setting out the project sector relevance and its positioning with reference to the state of art, describing the role and contribution of each research unit

MOTIVATIONS AND CHALLENGES

FRINGE (context-aware FaiRness engineerING in complex software systEms) studies solutions to engineer software fairness in Machine Learning (ML)-intensive systems. ML is increasingly being used to create data-driven decision-making systems for which an algorithmic solution is not feasible due to the complexity of the problem. By 2025 over 463 exabytes of data per day will be produced (McKinsey, 2022). Such data are increasingly retrieved by heterogeneous interconnected smart objects and real-time monitoring devices. The volume of data, the relations among the pieces of information collected, and the continuous data acquisition and learning activities conducted to improve the accuracy of ML-intensive systems, make it unfeasible for developers to properly verify that such systems act according to the ethical principles raised by the European Artificial Intelligence Act and make predictions that do not perpetuate discrimination against sensitive groups (Zhang, 2021). Recent reports showed the implications of fairness concerns for companies and society: 71% of customers would stop dealing with companies that unethically treat sensitive data (McKinsey, 2022) and over 40% of the world population has been affected by unfair decisions of ML-intensive systems in 2020 (HBR, 2020). The relevance of software fairness in ML-intensive systems has been also made popular by infamous incidents happened to the recruitment instrument employed by Amazon (Reuters, 2018) and the criminal recidivism predictions made by the commercial risk assessment software COMPAS (AI Incident Base, 2016).

To explain the motivations and challenges faced by FRINGE, let consider the following example, **which has been designed with the partner software company that will support the research team during this project, i.e., Kiranet**. This is an Italian provider of AI-based software solutions for medical IoT. According to the experience of the partner company, a key issue in medical AI is the data labeling during clinical assessment. Samulowitz et al. (2018) showed that due to gender stereotypes women are over-diagnosed for diseases like depression and under-diagnosed for other diseases like cancer, and Seyyed-Kalantari et al. (2021) found that women are diagnosed later than men for most diseases. As such, if the data labels in health registries are affected by disparities, then the ML models built on top of these data are likely to keep perpetuating such inequalities (Rajkomar, 2018). Therefore, novel instruments able to assist developers in dealing with fairness-related properties would reduce the risk of biased predictions.

While initial scientific contributions to fairness engineering have been made, recent experience demonstrates that developing fair ML-intensive systems is to a large extent an *open challenge*, as discussed in recent surveys (Köchling, 2020; Starke, 2022). Indeed, methods and techniques that support the analysis of fairness requirements, the design of fair ML algorithms, and the monitoring of fairness properties over time are still under-explored and significant improvements are needed.

FRINGE aims to exploit the long-lasting experience of the proposers in the diverse areas of SE to develop innovative solutions to tackle most of the aforementioned problems and deliver solutions for the creation of *fair ML-intensive systems*.

FRINGE focuses on the following challenges:

- **C1. Context-dependent definition of software fairness**: Defining software fairness is intrinsically challenging because it might depend on the specific use case implemented by the ML-intensive system, thus requiring context-aware methods to diagnose fairness-related properties.
- **C2. Multi-disciplinary context**: Engineering fairness requires a broad variety of skills that are not always available inside teams, including knowledge about SE practices, ML algorithms, and ethical and societal design and principles.
- **C3. Data and feature heterogeneity**: ML-intensive systems are often trained with data collected from heterogeneous sources and engineered through the definition of features having different nature (e.g., audio, video, text, etc.). The lack of methods and instruments to ensure fairness-aware data and feature extraction and cleaning makes it hard for practitioners to deal with fairness as part of the data and feature engineering process.
- **C4. Limited knowledge on fairness design and verification**. Little is known about how to design ML-intensive systems for fairness, which factors can impact the level of fairness of those systems, and how fairness-related properties can be verified, especially in a continuous data acquisition and learning scenario.

These challenges must be addressed in a holistic manner, with an engineered approach that can address them by introducing proper methodologies, techniques, and tools. Defining such an approach is the main aim of FRINGE.

FRINGE OBJECTIVES

FRINGE aims to provide methodologies, techniques, and approaches to analyze fairness requirements and assist the design and monitoring of fairness-related properties of ML-intensive systems. FRINGE will target interdisciplinary teams consisting of software engineers, data scientists, and ML experts, reducing the distances among such specialists and facilitating their collaboration.

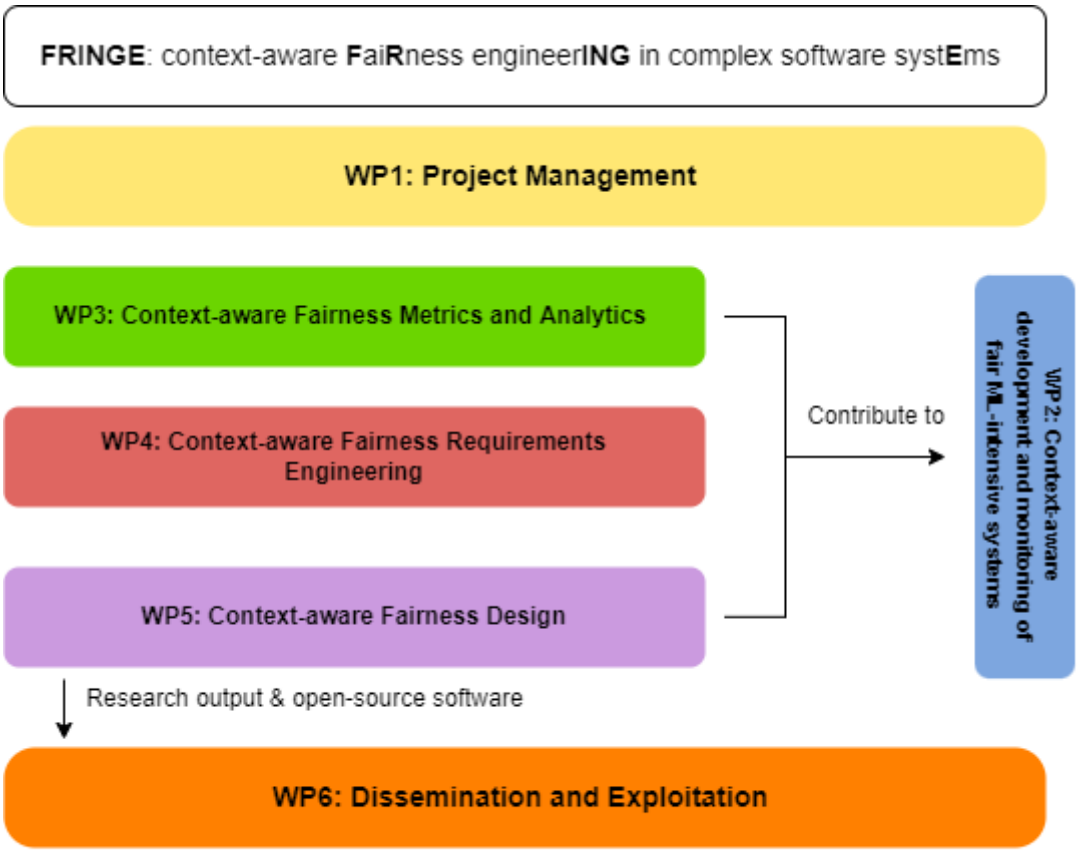

Figure 1: Overview of the Work Package of the FRINGE Project.

The project objectives are:

- **O1. Support to context-aware fairness definition analysis**: Definition of automated tools that can extract relevant information to help practitioners analyze the context the ML-intensive system is being developed for and the fairness definition to use.

- **O2. Support to context-aware fairness metrics and analytics**: Definition of solutions to monitor software fairness properties over time, e.g., by measuring the level of fairness at a given point of the evolution or reasoning on the compromise between fairness and other ML quality requirements.

- **O3. Support to context-aware fairness requirements engineering**: Definition of (semi-)automated methods and instruments to allow practitioners to deal with fairness-related requirements while engineering features and gathering data for ML training.

- **O4. Context-aware fairness design**: Provision of methodologies and tools to support the design of ML-intensive systems with respect to fairness in a given context.

- **O5. Consolidation of developed solutions** : Definition of a fairness metamodel and set of prototypes to support practitioners in the engineering processes to handle fairness throughout the lifecycle of ML-intensive systems.

WORK PACKAGES, METHOD, AND DELIVERABLES

The objectives will be organized in the 6 Work Packages (WPs) depicted in Figure 1. One of the work packages (WP1) concerns project management, four (WP2-WP5) map onto O1-O5, and one (WP6) regards dissemination and exploitation.

For the sake of understandability, the name and goal of each WP is summarized in Table 1.

FRINGE - context-aware FaiRness engineerING in complex software systEms	
WP1: Project Management **Leader:** UNISA; **Contributors:** UNISANNIO, UNIVAQ.	
Goal: This WP coordinates the project development, ensuring timely progress, and productive coordination and collaboration among the units.	**Key output:** Shared infrastructure to contribute to the project and annual reports on the status of the project.
WP2: Context-aware development of fair ML-intensive systems **Leader:** UNIVAQ; **Contributors:** UNISA, UNISANNIO.	
Goal: The WP fulfills O1 and O5 by conceiving a *catalog of fairness definitions* and *guidelines* for when each definition should be used. We will develop a *fairness definition recommender system*, leveraging the systematic knowledge formalized in the catalog to recommend the most relevant fairness definition. A monitoring infrastructure to continuously measure the fairness of ML-intensive systems will be also devised.	**Key output:** A fairness metamodel, a recommender system to suggest the best fairness definition to use depending on the context, and a monitoring infrastructure to inform the fairness specification.
WP3: Context-aware fairness metrics and analytics **Leader:** UNISANNIO; **Contributors:** UNISA, UNIVAQ.	
Goal: This WP fulfills O2 and aims at defining a *taxonomy of fairness metrics* and implementing a *fairness metrics recommender system*.	**Key output:** A taxonomy of fairness metrics and a fairness analytical framework.
WP4: Context-aware fairness requirements engineering **Leader:** UNISANNIO; **Contributors:** UNISA, UNIVAQ.	
Goal: This WP fulfills O3 and defines a *fairness requirements processing framework*, supporting the elicitation, specification, and maintenance of context-aware fairness requirements.	**Key output:** A fairness requirements recommender.
WP5: Context-aware fairness requirements engineering **Leader:** UNISA; **Contributors:** UNISANNIO, UNIVAQ.	
Goal: This WP fulfills O4, defining *novel taxonomies of design patterns and antipatterns* and *(semi-)automated approaches to support developers in adopting or avoiding them.*	**Key output:** A fairness design recommender.
WP6: Dissemination and Exploitation **Leader:** UNIVAQ; **Contributors:** UNISA, UNISANNIO.	
Goal: This WP will coordinate dissemination activities and drive the open-source release of the artifacts produced.	**Key output:** A publication strategy and guidelines to enable open-science and open-data.

Table 1: Overview of the Work Packages of the FRINGE Project.

WP1: Project Management (Leader: UNISA)

WP1 includes the following tasks.

T1.1 - *Project development infrastructure setup [M1-M6]*
Contributions: All

The task sets up the collaboration and communication instruments. We will set up a shared infrastructure to upload the material produced. Moreover, (a)synchronous communication channels will be defined to share news and messages about the project.

Deliverables. A *shared collaboration and communication infrastructure*.

T1.2 - *Technical coordination [M1-M24]*
Contributions: All

The task conducts continuous coordination and verification of the activities. We will establish two annual in-person meetings where the partners will report on the current achievements and more frequent remote status updates to coordinate collaborations.

Deliverables. An *annual report* describing the achievements made.

WP2. Support to context-aware development of fair ML-intensive systems (Leader: UNIVAQ)
WP2 includes the following tasks:

T2.1 - *Development of a metamodel for software fairness [M1-M12]*
Contributions: All

We will capture the proper fairness definitions that should be used in different contexts in a catalog, to organize and systematize the definitions, notions, principles, and theories related to the different contexts, enabling us to see the bigger picture. The catalog will be defined in terms of a metamodel encoding all the identified concepts and relationships concerning software fairness.

Deliverables. A systematically defined *metamodel of software fairness definitions*.

T2.2 - *Development of recommender systems supporting the adoption of fairness definitions [M7-M18]*
Contributions: UNIVAQ

We will conceive a novel recommender system, providing users with fairness definitions relevant for the system under development. The envisioned recommender system will ask users to capture the potentially desired fairness properties, will ease consulting the catalog, and will enable the extraction of the information needed for specific tasks.

Deliverables. A *recommender system* to automatically provide users with relevant fairness definitions for the specific task at hand.

T2.3 - *Fairness-aware monitoring for ML-intensive systems [M13-M24]*
Contributions: UNIVAQ, UNISA

By relying and enhancing existing technologies, we will design a monitoring infrastructure to continuously measure the fairness of ML-intensive systems. The intended monitoring system will leverage the patterns and antipatterns defined in WP5 to identify potential fairness issues to be further investigated.

Deliverables. A *monitoring system* to automatically raise alerts about fairness issues, identified according to the patterns and antipatterns defined in WP5.

WP3. Context-aware fairness metrics and analytics (Leader: UNISANNIO)
The WP contains the following tasks.

T3.1 - *Definition of a taxonomy of fairness metrics for software systems [M1-M6]*
Contributions: UNISANNIO, UNIVAQ

We will first organize the defined metrics for measuring fairness in a *novel taxonomy of fairness metrics*. Such a taxonomy will particularly focus on the specific fairness definition and data formats (tabular, text, image, etc.) each metric applies to, providing guidelines on when and how to use it. The taxonomy will be built by surveying the literature and exploring the collective knowledge shared in discussions on Q & A services (e.g., StackOverflow).

Deliverables. A *taxonomy of fairness metrics*.

T3.2 - *Development of a knowledge base (KB) for synthesizing information on fairness metrics [M7-M12]*
Contributions: UNISANNIO, UNIVAQ

We will develop *a novel KB of fairness metrics* to promote knowledge sharing across software communities. The KB will be initially filled with the information collected in the taxonomy developed in T3.1, but it will also provide ready-to-use user-friendly interfaces for easily extending the initial set of metrics.

Deliverables. A *knowledge base for synthesizing information on fairness metrics*.

T3.3 - *Development of a fairness metrics recommender system [M7-M24]*
Contributions: UNIVAQ, UNISA

On top of the KB conceived in T3.2, we will develop a *recommender system* to help software engineers select the best set of metrics to adopt for monitoring the fulfillment of specific fairness requirements. The recommender system will leverage information retrieval techniques to select the metrics collected in the KB that best fit with the given fairness requirements, recommending the metrics to adopt, the right measurement strategies, and code examples on how to compute them.

Deliverables. A *fairness metrics recommender system*.

T3.4 - *Integration within the conceptual metamodel [M19-M24]*
Contributions: All.

We will develop an extendible model linking (1) fairness metrics to the KB through *gathered by* relations, (2) fairness metrics to fairness definitions through *measured by* relations, (3) fairness metrics to examples of fairness requirements through *measurable by* relations, and (4) fairness metrics to the recommender system through *recommended by* relations. We will then integrate the model within the metamodel in WP2.

Deliverables. A *conceptual model* representing the outcomes of the WP and relations among them.

WP4. Context-aware fairness requirements engineering (Leader: UNISANNIO)
The WP is composed of the following tasks.

T4.1 - *Development of a software documentation analyzer [M7-M24]*
Contributions: UNISANNIO.

We will devise a *tool to analyze software documentation and specify fairness requirements* that inspects formal and informal software documents and outputs fairness requirement specifications. We will collect a corpus of different kinds of software documents from large ML-intensive systems and experiment with approaches merging various natural language processing (NLP) techniques, e.g., dependency parsing, textual analysis, and named entity recognition, to mine information of interest from such documents. Based on the information extracted, the tool will leverage template-based techniques to specify the expected fairness properties.

Deliverables. A *software documentation analyzer* to inspect formal and informal documentation and formally specify each fairness requirement.

T4.2 - *Definition of a taxonomy of fairness issue categories [M1-M6]*
Contributions: UNISANNIO, UNIVAQ.

We will define a *taxonomy of fairness issue categories*, enumerating the likely causes, characteristics, and consequences for the different types of fairness-related issues. Fairness constraint violations may emerge during the operation phase of a complex software system, signaled by users and developers using issue reports and pull requests. We will collect issues from large ML-based systems, including open-source software, to inspect them and build an initial set of fairness-based issue categories. This initial set will be then extended and iteratively refined according to feedback collected through interviews and surveys targeting data scientists and SE researchers and practitioners.

Deliverables. A *taxonomy of fairness issue categories*.

T4.3 - *Development of a fairness issue detector [M7-M24]*
Contributions: UNISANNIO, UNISA.

To help developers maintain software fairness, we will leverage textual analysis and NLP techniques to develop a *novel tool for detecting fairness issues* that (1) automatically identifies the fairness-related issues reported by users and developers and (2) the specific categories they fall into. Specifically, the fairness issues collected in T4.2 will be used to experiment with ML-based solutions, automatically identifying and classifying the various types of fairness reported issues.

Deliverables. A *fairness issue detector*.

T4.4 - *Integration within the conceptual metamodel [M19-M24]*
Contributions: All.

We will devise a model linking (1) the information sources to the software documentation analyzer through *analyzed by* relations, (2) the inferrable information to the software documentation analyzer through *inferred by* relations, (3) templates of fairness requirement specifications to the inferrable information through *based on* relations, (4) templates of fairness requirement specifications to the software documentation analyzer through *specified by* relations, and (5) the types of fairness issues to the fairness issue detector through *identified by* relations. We will then integrate the model within the metamodel designed in WP2.

Deliverables. A *conceptual model* representing the outcomes of the WP and relations among them.

WP5. Context-aware fairness design (Leader: UNISA)
This WP is split into the following tasks.

T5.1 - *Definition of a catalog of fairness design patterns and antipatterns [M1-M6]*
Contributions: UNISA.

We will define *novel taxonomies of context-aware fairness design patterns and antipatterns*. These will describe the methods that practitioners should use or avoid when designing a fair ML-intensive system that is continuously retrained and the practices that may help assess how fairness should be treated in different contexts. We will exploit data mined from large ML-systems, surveys, interviews, and focus groups with data scientists and software engineers. We will then apply grounded theory research to define the taxonomies and iteratively refine them.

Deliverables. A *taxonomy of fairness design patterns and antipatterns*.

T5.2 - *Definition of automated techniques [M7-M12]*
Contributions: UNISA, UNIVAQ.

We will devise *automated solutions for detecting antipatterns* and *recommending design patterns*, relying on static analysis to define automated linters that support the automated improvement process. We will consider additional sources of information, like the textual analysis of the training data coming from **O2**, and techniques, like ML/DL mechanisms to combine static and textual analysis. We will also define *semi-automated refactoring approaches* that recommend how to refactor training data and features, optimizing fairness while preserving accuracy. Thus, we will use meta-heuristic methods, e.g., genetic algorithms and neuroevolution.

Deliverables. (1) *Design pattern recommenders and antipattern detection approaches*; (2) *refactoring techniques*.

T5.3 - *Definition of automated test case generators [M7-M18]*
Contributions: UNISA, UNISANNIO.

We hypothesize that the design patterns and antipatterns may be verified through the definition of testing-like approaches, e.g., feature engineering issues may be subject to verification procedures. We will define *automated test case generation approaches* that dynamically generate adversarial data samples to challenge the ML-intensive system and verify whether the expected level of fairness is met in presence of perturbations of data that might be observable when the system operates.

Deliverables. *Automated test case generators*.

T5.4 - *Integration within the conceptual metamodel [M19-M24]*
Contributions: All.

We will finally develop a model connecting (1) design patterns to the recommender systems developed through *recommended by* relations, (2) antipatterns and detection and refactoring approaches through *detectable by*, and *refactorable with* relations, and (3) both design patterns and antipatterns to the corresponding verification approaches released through *verifiable with* relations. The model will be then integrated within the metamodel in WP2.

Deliverables. A *conceptual model* representing the outcomes of the WP.

WP6. Dissemination and Exploitation (Leader: UNIVAQ)
This WP will coordinate dissemination activities and drive the open-source release of the artifacts produced.

T6.1 - *Dissemination*
Contributions: All

The task will coordinate the dissemination strategy, by identifying academic and practitioner venues to disseminate the project's achievements and high schools interested in seminars and organization of role-playing games to increase the awareness of the students on themes concerned with software fairness. The task will also curate the social media activities (e.g., project's website and social media accounts).

Deliverables. (1) A *publication and dissemination strategy* of the partners; (2) *website and social media accounts* of the project.

T6.2 - *Open source release of the FRINGE artifacts*
Contributions: All

The task will coordinate with the project management—which set up the shared infrastructure—to define guidelines to make the outcomes of the project publicly available. Secondly, the task will curate the advertisement of the products released.

Deliverables. (1) A *set of guidelines to make the outcomes of the project open source*; (2) The *creation of content to advertise the open-source releases*.

RELEVANCE AND COHERENCE OF THE RESULTS
We are not aware of any comprehensive approach providing software engineers with automated support to elicit, specify, design, and verify context-aware fairness requirements. The results of FRINGE will be relevant under two perspectives.

On the foundational side, the state of the art still provides limited knowledge. There are no conceptual models able to map the engineering methods and approaches to deal with software fairness in ML-intensive systems. At the same time, while literature has proposed some metrics and strategies to detect/mitigate bias, it did not consider that relevant information on when and how to use each metric can be disseminated over different sources (e.g., scientific publications, collaborative platforms, Q&A websites). Also, there is no catalog of best and bad practices to deal with software fairness. The meta-model proposed in WP2 and *the set of taxonomies proposed within WP3, WP4, and WP5 will provide a brand new theoretical ground for software fairness engineering*. These taxonomies will advance the state of the art by proposing a comprehensive view of how fairness-related requirements, metrics, issues, patterns, and antipatterns, can be elicited, acquired, and managed over the development of ML-intensive systems. The shared infrastructure set up in WP1 will enable open-science and allow researchers to build on top of the findings achieved by FRINGE.

On the operational side, the automated support provided to practitioners is still limited to some preliminary techniques that may not support the engineering activities throughout the development of ML-intensive systems. FRINGE aims to provide *a set of automated analyzers and recommender systems that largely advance the current knowledge and practice on fairness engineering*. With these instruments, FRINGE will support practitioners to treat fairness throughout the engineering process.

The results brought by FRINGE will contribute to the human-centered and ethical development of technologies based on ML algorithms, hence *supporting circular economy and social innovation and addressing key goals established by the PNRR in Cluster 4, Point 6*.

MILESTONES

Table 2 reports the key milestones (fixed at months 6, 12, 18, and 24, respectively) and the tasks we expect to complete within the project.

MILESTONE 1: Project setup and knowledge acquisition (M6)			
Task	**Description**	**Version**	**Months**
T2.1	Development of a metamodel for software fairness	intermediate	M1-M12
T3.1	Definition of a taxonomy of fairness metrics for software systems	consolidated	M1-M6
T4.2	Definition of a taxonomy of fairness issue categories	consolidated	M1-M6
T5.1	Definition of a catalog of fairness design patterns and antipatterns	consolidated	M1-M6
T6.1	Website and social media accounts creation	consolidated	M1-M6

MILESTONE 2: Solutions for processing information sources (M12)			
Task	**Description**	**Version**	**Months**
T2.1	Development of a metamodel for software fairness	consolidated	M1-M12
T2.2	Development of recommender systems supporting the adoption of fairness definitions	intermediate	M7-M18
T2.3	Fairness-aware monitoring for ML-intensive systems	intermediate	M13-M24
T3.2	Development of a knowledge base for synthesizing information on fairness metrics	consolidated	M7-M12
T3.3	Development of a fairness metrics recommender system	intermediate	M7-M24
T4.1	Development of a software documentation analyzer	intermediate	M7-M24
T4.3	Development of a fairness issue detector	intermediate	M7-M24
T5.2	Definition of automated techniques to support fairness design	consolidated	M7-M12
T5.3	Definition of automated test case generators	intermediate	M7-M18
T6.2	Open source release of the FRINGE artifacts	intermediate	M7-M24

MILESTONE 3: Solutions for generating outputs (M18)			
Task	**Description**	**Version**	**Months**
T2.2	Development of recommender systems supporting the adoption of fairness definitions	consolidated	M7-M18
T2.3	Fairness-aware monitoring for ML-intensive systems	intermediate	M13-M24
T3.3	Development of a fairness metrics recommender system	intermediate	M7-M24
T4.1	Development of a software documentation analyzer	intermediate	M7-M24
T4.3	Development of a fairness issue detector	intermediate	M7-M24
T5.3	Definition of automated test case generators	consolidated	M7-M18
T6.2	Open source release of the FRINGE artifacts	intermediate	M7-M24

MILESTONE 4: Refinement and finalization (M24)			
Task	**Description**	**Version**	**Months**
T2.3	Fairness-aware monitoring for ML-intensive systems	consolidated	M13-M24
T3.3	Development of a fairness metrics recommender system	consolidated	M7-M24
T4.1	Development of a software documentation analyzer	consolidated	M7-M24
T4.3	Development of a fairness issue detector	consolidated	M7-M24
T3.4	Integration of WP3 outcomes within the conceptual metamodel	consolidated	M19-M24
T4.4	Integration of WP4 outcomes within the conceptual metamodel	consolidated	M19-M24
T5.4	Integration of WP5 outcomes within the conceptual metamodel	consolidated	M19-M24
T6.2	Open source release of the FRINGE artifacts	consolidated	M7-M24

Table 2: Milestones expected in the context of FRINGE.

CASE STUDY AND EMPIRICAL ASSESSMENT

FRINGE will identify a set of case studies used to guide the development and assessment activities. We plan to use a combination of open-source and industrial applications.

Relevant open-source opportunities are already available through the general-purpose open-source repositories, e.g., GitHub, which store plenty of fairness-critical ML-intensive systems that may serve as case studies. For instance, we will consider the case of news recommenders, e.g., the Microsoft Recommenders project (https://github.com/microsoft/recommenders). These systems typically learn from users' news click behavior to recommend relevant news; as such, suspicious activities made by a user might be even encouraged by news recommenders that provide the user with additional, relevant insights on how to pursue the suspicious activity.

We will also consider other recommender systems, like bots and conversational agents (e.g., https://github.com/topics/bot-list). Most of them are used as decision support systems, e.g., to support hiring decisions in recruiting. The sensitivity of the tasks performed represents a valuable case study to further analyze.

Finally, we will investigate the possibility to include industrial cases by exploiting the numerous contacts of the partners with companies interested in crafting fairness-aware ML-intensive systems. Among them, Kiranet s.r.l.—which supports the proposal, as

mentioned in the attached support letter—will allow us to experiment with medical IoT, providing access to ML-intensive systems that collect sensitive data from medical environments and recommend treatments to doctors and patients.

The assessment of the proposed solutions will be conducted under multiple perspectives. We will run empirical studies and validate the effectiveness of the proposed techniques. Moreover, we will design observational studies where practitioners will be monitored for specific time periods to understand the way the proposed instruments are used and how practitioners react to the information provided. Whenever possible, we will complement the validations with survey studies and semi-structured interviews where practitioners will be involved to provide their feedback about the effectiveness and practical usefulness of the outcomes of the project.

---

3. Detailed description of the project team and planning; indicating the research team components – PI and associated PIs - and their relative expertise/track record, gender equality of the composition, the interrelation and coherence of the team components. RUs- and the feasibility of the project, thus outlining the congruity between objectives, timing and costs

RESEARCH UNITS

UNISA The Unisa unit leader and project's PI works with the Department of Computer Science (DI) at the University of Salerno. The DI has been instituted as an evolution of the previous Information Science Institute, which offered one of the first four computer science degrees in Italy. The Department has a strong interdisciplinary connotation and covers the vast majority of the research areas in Computer Science, e.g., from software security to networks, from artificial intelligence to software engineering and human-computer interaction. The DI currently features 53 faculty members (22 full, 17 associate, and 14 assistant professors). 16 faculty members (30%) are female.
The Department has obtained recognition as a department of excellence for 2018/2022 in implementation with the provisions of art. 1, paragraphs 319 and 320, of the Budget Law 2017 (Law 232/2016).
For what concerns the Evaluation of Research Quality (Valutazione Qualità della Ricerca - VQR) for the years 2015-2019 the Department has obtained the overall first position for the area 01 (Computer Science).

The specific research unit involved in the FRINGE project is composed of two faculty members (Fabio Palomba and Filomena Ferrucci) and three Ph.D. students (Giulia Sellitto, Stefano Lambiase, and Gilberto Recupito). All of them have expertise in software engineering for artificial intelligence, software quality, and social software engineering, which are the main themes of the FRINGE project. They are all currently working on software fairness, management of non-functional requirements of artificial intelligence systems, and the evolution of machine learning models embedded in IoT systems. The unit has been assembled to optimize the expertise on the project and the gender balance of the members.

UNISANNIO The Unisannio unit leader and project's Deputy PI works with the Department of Engineering (DING) at the University of Sannio.
The DING has been instituted in 2013 based on law 240/2010 and in continuity with the previous Engineering Department. The Department has a strong interdisciplinary nature, combining a strong tradition in information engineering with leading groups from other areas, including in particular energy engineering and civil engineering. The Department currently features 74 faculty members (17 full, 36 associate, and 21 assistant professors). 16 faculty members (21%) are female.
The Department has obtained recognition as a department of excellence for 2018/2022 in implementation with the provisions of art. 1, paragraphs 319 and 320, of the Budget Law 2017 (Law 232/2016).
For what concerns the Evaluation of Research Quality (Valutazione Qualità della Ricerca - VQR) for the years 2015-2019 the Department has obtained the overall third position for the areas 09 (Industrial and Information Engineering) and 08 (Civil engineering), and the University has obtained for the same areas, the first position for Universities of the same size. The Department researchers are involved in 9 academic spin-offs.

Two faculty members (Andrea Di Sorbo and Maria Tortorella) and one Ph.D. student (Arnaldo Sgueglia) compose the specific research unit involved in the FRINGE project. The expertise in natural language processing, software metrics, software analytics, and machine learning for software engineering brought by this research unit is fundamental to achieving specific objectives of the project. Currently, all the components of the research unit are working on software maintenance and evolution, the management of non-functional requirements in software systems, and the robustness of machine learning models. The unit has been assembled by both considering the expertise required for the development of the project and gender diversity.

UNIVAQ The University of L'Aquila (UNIVAQ) unit leader and project's associated PI works with the Department of Engineering Computer Science and Mathematics (DISIM) at the University of L'Aquila. The University of L'Aquila is one of the medium-sized universities in Italy with around 20,000 students and seven Departments that promote, coordinate and organize both scientific research and teaching and training activities. The University of L'Aquila appears this year for the second time among the top 100, precisely in 87th position in the world and 2nd in Italy, after the Alma Mater of Bologna, in the "Times Higher Education'' ranking – Impact." The University of Aquila played a crucial role during the various phases of the reconstruction following the earthquake of April 2009 through the implementation and promotion of strategic research and development projects. Thanks to research and

development activities implemented by the University, the Ministry of Economic Development has included the city of L'Aquila among the 5 sites in Italy and in Europe where it is possible to experiment with 5G technology. The University of L'Aquila has also promoted collaborations with major companies (Leonardo, Telespazio, Ansaldo-STS and FCA - https://www.fcagroup.com) to test solutions for connected vehicles that rely on 5G location and connectivity. The University of L'Aquila is also active in various areas of Artificial Intelligence and Computational Logic with research and technology transfer activities to which it contributes with 8 patents (two international), one of which is funded by the INFM and one co-financed from the University of L'Aquila. The University of L'Aquila is a member of the CINI AIIS "Artificial Intelligence and Intelligence Systems" National Laboratory.

Two faculty members (Davide Di Ruscio and Antinisca Di Marco), one assistant professor (Phuong Nguyen) and one PhD student (Giordano D'Aloisio) compose the specific research unit involved in the FRINGE project. They are working on different software engineering, data mining, and model-driven engineering aspects. In particular, their recent research focuses on modeling and metamodeling fairness of ML-intensive systems, further than conceiving recommender systems to support complex software engineering tasks.

INTERRELATION AND COHERENCE OF THE TEAM COMPONENTS

For the successful development of each technical work package (WP2-WP5), different and specific expertise is required. In particular, WP2 requires specific competencies in modeling and model-driven engineering, software engineering for AI (including fairness and calibration), and recommender systems. The development of WP3 demands special skills in software metrics and prediction models, software engineering for AI, and recommender systems. Proficiency in natural language processing, software analytics, and machine learning for software engineering is required to achieve WP4 objectives. Finally, expertise in software analytics, crowd-based software engineering, and software refactoring is fundamental for developing WP5. Table 3 summarizes the key expertise required for the development of each technical work package.

Expertise	WP2	WP3	WP4	WP5
NLP, including text mining	-	-	X	-
Software analytics	-	-	X	X
Modeling and model-driven engineering	X	-	-	-
Crowd-based software engineering	-	-	-	X
Software metrics and prediction models	-	X	-	-
Software engineering for AI	X	X	-	-
Machine learning for Software Engineering	-	-	X	X
Software refactoring	-	-	-	X
Recommender systems	X	X	-	-

Table 3: Expertise required for the development of the work packages.

The project development will leverage the heterogeneous expertise of the project participants and the collaborations they will establish. In particular, the Univaq research unit will lead WP2 based on its solid track record in modeling and model-driven engineering, software engineering for AI, and recommender systems. The Unisannio research unit will lead WP3 and WP4, relying on its specific proficiency in software metrics and software analytics, beyond the strong expertise in leveraging natural language processing, text mining, and machine learning techniques to extract requirements from software informal documents and support software engineering activities, as from the publication list. Finally, the Unisa research unit, with its competencies shown in paramount publications about software analytics, crowd-based software engineering, and software refactoring, will lead WP5.

Therefore, as shown in Table 4, each unit will provide unique expertise, main role, and will collaborate on specific objectives:

Partner	Expertise	Main Role	Collaborations
UniSa	Software analytics; Crowd-based software engineering: Software engineering for AI, including fairness and calibration; Machine learning for software engineering; Software refactoring	Coordinate the project development and provide solutions for the identification of anti-patterns and recommendation of design patterns to deal with fairness constraints in different contexts	Collaborations with Unisannio to develop analytical instruments and design approaches for software fairness engineering. Collaborations with Univaq to devise a conceptual meta-model for ML fairness. Support to the dissemination and exploitation activities, maximizing the awareness and impact of the project.
UniSannio	NLP, including text mining; Software analytics; Crowd-based software engineering; Software metrics and prediction models; Machine learning for software engineering	Provide solutions for the development of a framework supporting the elicitation, specification, monitoring, and maintenance of context-aware fairness requirements	Collaborations with Univaq to develop a fairness monitoring approach and devise a conceptual meta-model. Collaborations with Unisa to define analytical and design instruments for software fairness. Support to the dissemination and exploitation activities, maximizing the awareness and impact of the project.
UnivAq	Modeling and model-driven engineering; Software Engineering for AI, including fairness and calibration; Machine Learning for Software Engineering; Recommender Systems	Manage integration and provide solutions for the development of a metamodel for context-aware software fairness. UnivAq will also conceive recommender systems to help software engineers adopt the proper fairness definitions for the system under development.	Collaborations with Unisa and Unisannio to devise the conceptual meta-model of software fairness. Collaborations with Unisannio to develop a fairness monitoring approach. Support to the dissemination and exploitation activities, maximizing the awareness and impact of the project.

Table 4: Expertise, Main Role, and Collaborations among the FRINGE Partners.

To further elaborate on the interrelation of the team components, Figure 2 overviews the dependencies among the work packages of the project. In particular, the figure reports the individual, conceptual outcomes of the WPs and how the other WPs will contribute to their development. As shown, WP2 will first develop the fairness meta-model and monitoring approach, which will be later integrated through the contributions provided by WP3, WP4, and WP5. The monitoring approach will inform the development of a fairness specification to which a ML-intensive system should adhere. Furthermore, WP3 has dependencies with WP4 and WP5. Specifically, the fairness analytical framework developed within WP3 will be used to measure fairness-related properties that will be elaborated by the recommenders developed within WP4 and WP5. In addition, the fairness requirements recommender output by WP4 will inform the decisions taken while designing the recommenders in WP5.

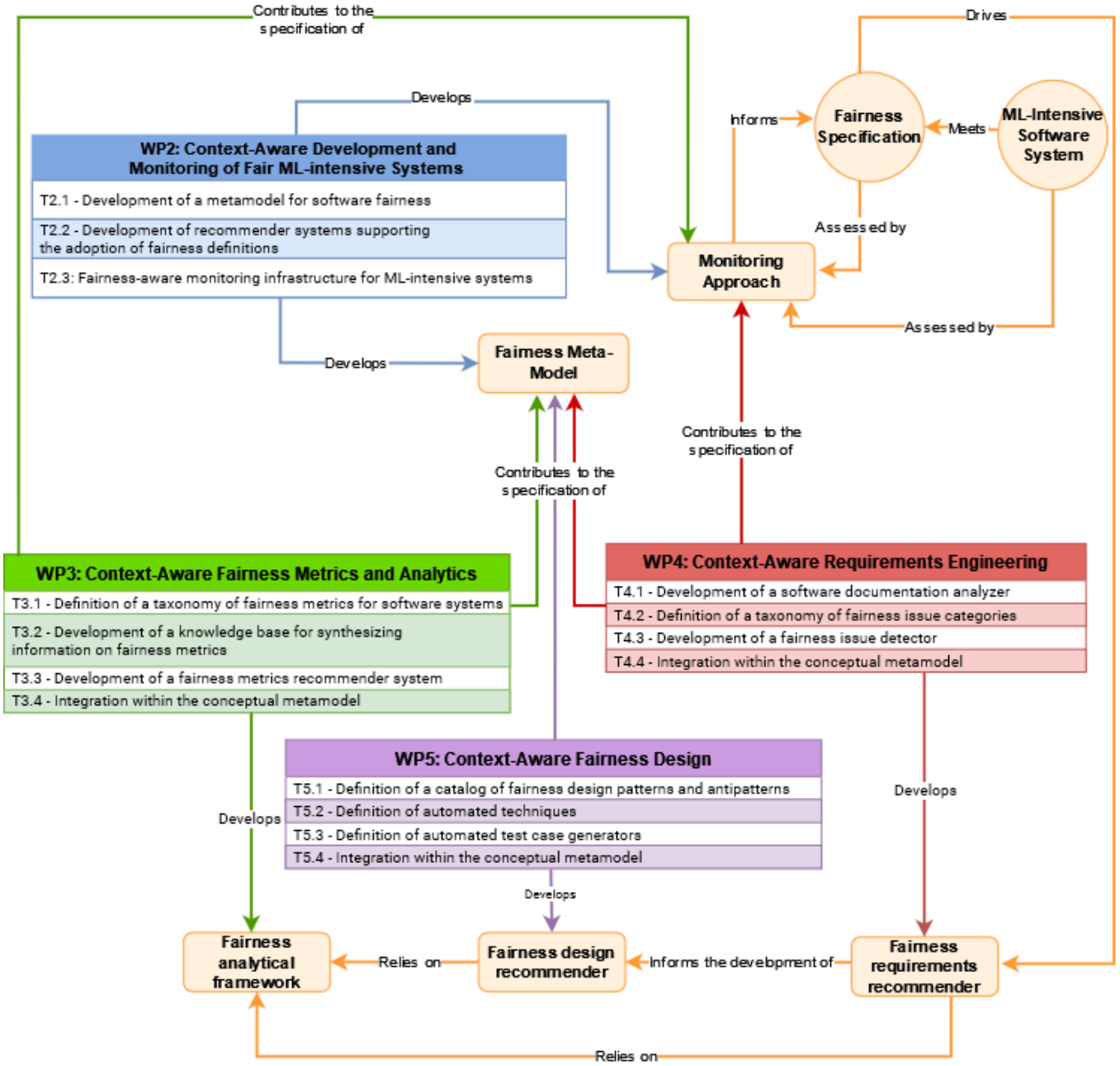

Figure 2: Dependencies among the work packages of the project.

It is worth considering that all the identified dependencies are “weak”, i.e., non-blocking ones. As such, the tasks of the individual WPs will be carried out independently and will be later refined according to the inputs provided by the other WPs. For instance, WP4 and WP5 will first develop catalogs and taxonomies and later, informed by the outcome of WP3, will work toward the definition of recommender systems.

FEASIBILITY OF THE PROJECT

Figure 3 overviews the Work Breakdown Structure (WBS) of the FRINGE project. We estimate the duration and cost of each task planned within the WPs.

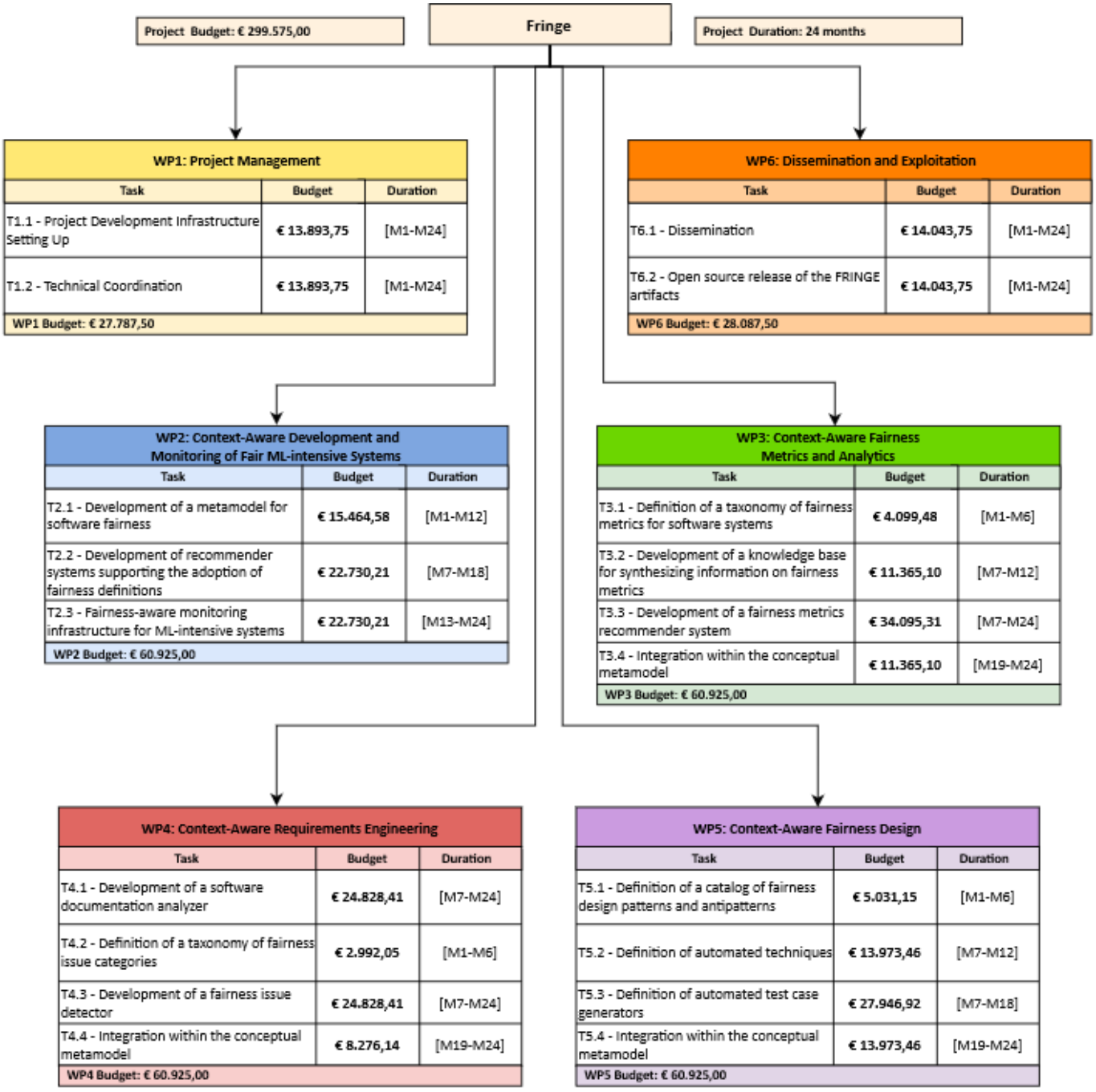

Figure 3: Work Breakdown Structure of the FRINGE Project.

As specified when presenting the work packages of the project (Part 2), all the tasks will be conducted within 24 months. While the cost of the tasks is proportional to their duration, it has been also determined by taking into account the amount of resources that each partner has committed on FRINGE and the collaborations among the research units.

The project plans to hire three researchers through research contracts. These contracts have a duration of 24 months, yet the researchers will be involved for 20 months on FRINGE. Each partner has committed enough person/months of permanent employees to cover the activities planned for the four months where the researchers will not work on FRINGE. At the same time, each partner commits to co-fund the researchers with external funding to cover the cost of the research contract not covered by the PRIN 2022 PNRR funding. In particular, the research units will rely on the budget of the following projects:

- UNISA will rely on the EMELIOT project funded by the MUR under the PRIN 2020 Program - Contract No. 2020W3A5FY. Research Unit Coordinator: Prof. Andrea De Lucia.

- UNISANNIO will rely on the EMELIOT project funded by the MUR under the PRIN 2020 Program - Contract No. 2020W3A5FY.

Research Unit Coordinator: Prof. Massimiliano Di Penta.

- UNIVAQ will rely on the EMELIOT project funded by the MUR under the PRIN 2020 Program - Contract No. 2020W3A5FY. Research Unit Coordinator: Prof. Davide Di Ruscio.

The cost for the technical WPs (i.e., WP2, WP3, WP4, WP5) will be greater than WP1 and WP6, as the activities of the technical WPs will require more person/hours. Indeed, the recruited researchers through three research contracts will be mainly involved to carry out the tasks of WP2, WP3, WP4, WP5, while the 60% of the time of the permanent employees will be devoted to the tasks of WP1 and WP6 and the remaining 40% to the tasks of the technical WPs.

The technical coordination of the project (Task 1.2 of WP1) will ensure constant monitoring of both duration and costs of the tasks and will make the development of the project as smooth as possible. Part of the two annual meetings will be devoted to the financial and time verification of the tasks. At the same time, asynchronous updates will be provided by each research unit on a monthly basis and will be shared with the partners through the mailing list of the project.

---

4. Detailed description of the Project impact, as such; indicating knowledge improvements, technological innovation and/or industrial applications, scientific community reinforcement, level of research internationalization, dissemination and exploitation of the results

KNOWLEDGE IMPROVEMENTS

The knowledge available on the engineering of fairness properties within ML-intensive systems is still limited, as highlighted in the state of the art. The FRINGE project has the ambition to bridge this gap by means of an increased awareness of how fairness-related properties can affect and can be treated during the engineering processes employed while developing, evolving, and maintaining ML-intensive applications. The conceptual meta-model will indeed provide an overall view of how software fairness can be managed during the lifecycle of ML-intensive software systems, hence contributing to an increase of the knowledge currently available in literature and practice.

Furthermore, the project will contribute to the definition of novel automated instruments that support the activities of practitioners, contributing to an increased knowledge on how to build human-centered and ethical technologies that positively affect society as a whole. In this sense, the instruments developed in the project address the objectives established by the PNRR with respect to the development of ethical technologies that support social innovation and empower (1) end-users through an improve reliability and trustworthiness of those technologies, and (2) workers who could take informed, ethical decisions on the processes to follow during their own activities.

To sum up, the knowledge improvement brought by the FRINGE project will be based on top-level publications and demonstrations, and it will be articulated in many directions:

- An improved awareness of how software fairness can be managed and improved over the lifecycle of ML-intensive systems, which will be achieved through the development of a conceptual meta-model;

- An improved understanding of how fairness can be affected within specific contexts, which will be obtained through a newfangled recommender system specifically conceived to provide users with relevant fairness definitions for the specific task at hand;

- An improved awareness of how fairness varies over time and how it can be measured, which will be achieved by means of new taxonomies of fairness metrics, a knowledge base for synthesizing information on fairness metrics, and a recommender system conceived to suggest practitioners how to measure fairness in specific contexts;

- Improved mechanisms and practices to elicit and diagnose fairness requirements, that will be ensured by novel documentation analyzers and automated generators of fairness specifications, other than automated approaches to classify fairness issue categories and detect them.

- An improved knowledge on how to handle fairness throughout the lifecycle of ML-intensive systems, which will be allowed through the innovative approaches and technologies conceived to support the design of fairness-related properties.

TECHNOLOGICAL INNOVATIONS AND INDUSTRIAL APPLICATIONS

The adoption of machine learning and artificial intelligence is significantly changing the way industries work. For instance, manufacturing processes involve artificial intelligence systems that collect data from heterogeneous interconnected smart devices, structured and unstructured sources, and real-time monitoring devices to elaborate them so that they can collaboratively provide insights into the core production operations to optimize them and disclose further revenue possibilities. Digital transformation can be fully achieved only if those advanced artificial intelligence applications are exploited to take full advantage of such a large amount of data while acting ethically toward the end users. This represents a critical challenge to face in the context of the digital transformation program established by the Italian government and European Commission. One of the main points raised by the European Artificial Intelligence Act (https://artificialintelligenceact.eu/) is indeed the definition of technologies enabling the development of trustworthy and ethical artificial intelligence able to preserve and protect ethical principles.

FRINGE is expected to provide relevant contributions in this respect by devising novel approaches and techniques underpinning software fairness in ML-intensive systems. They will permit the optimization of the development of machine learning models by explicitly considering contextual fairness requirements throughout the development, hence enabling the appropriate preparation and refinement of both data and ML training activities.

Overall, FRINGE will foster software solutions that can enhance industries under different perspectives. In particular, ML-Intensive systems analyzed, developed, and designed with FRINGE technologies will improve industrial applications by optimizing ethical requirements, avoiding congestion and hazards due to legal and ethical risks, and improving long-term sustainability.

Although the timeline of the project would not allow a full integration of the developed solutions within a unique, usable software framework, FRINGE solutions have the potential to reach already some achievements in terms of technological transfer. In particular, the industrial case studies of the project will enable a fruitful communication between the research team and the involved companies. Indeed, the application of the developed solutions will allow companies to receive feedback on their own project, hence allowing them to improve and optimize the development and design processes they have in place. At the same time, the fully-integrated conceptual meta-model will be transferred so that companies may optimize the requirements elicitation and analysis processes, other than the design principles applied to meet fairness requirements.

SCIENTIFIC COMMUNITY REINFORCEMENT

FRINGE aims at reinforcing the scientific community by advancing the state of the art with respect to the management and design of fairness requirements in ML-intensive systems. Furthermore, the project has the ambition to contribute to the open-science and open-data initiatives in an effort of further reinforcing the knowledge base on fairness engineering.

Several different communities will be highly impacted by the project results, including general communities of practice in software product development, machine learning, software maintenance and evolution, software product life-cycle management, and social software development. The relevant research communities maintain a large number of international journals and conferences. Only very few examples of relevant scientific journals and conferences that will be targeted by the project can be listed such as: IEEE Trans. on Softw. Eng. (TSE), ACM Trans. Softw. Eng. and Methodology (TOSEM), Software and Systems Modeling (Springer), Empirical Software Engineering (Springer), Int. Conf. on Soft. Eng. (ICSE), ACM Europ. Sw. Eng. Conf./Symp. on the Found. Sw. Eng. (ESEC/FSE), Int. Conf. on Automated Soft. Eng. (ASE), Int. Conf. on Model-Driven Eng. Languages & Systems (MODELS), ACM Conf. Series on Recommender Systems (RecSys), Int. Conf. on Machine Learning (ICML), Int. Conf. on Human Factors in Computing Systems (CHI).

As an additional effort to reinforce the scientific community, the project will feature special themes and special issues in top-tier software engineering journals, e.g., ACM Trans. Softw. Eng. and Methodology, and magazines, e.g., IEEE Software. Through these initiatives, the project aims at facilitating the establishment of a communities of practice that might let both researchers and practitioners be exposed to sensitive themes concerned with fairness and ethics of ML-intensive systems and possibly let them approach together to build on top of the engineering approaches developed within the FRINGE project.

LEVEL OF RESEARCH INTERNATIONALIZATION

The FRINGE activities will be conducted by the involved partners in collaboration with top-level research scientists and well-known institutions. Over the past years, the associated PIs have indeed established networks of collaborations that led them to interact with top researchers in the field of Software Engineering for and with Artificial Intelligence, Model-Driven Software Engineering, Software Analytics, Crowd-based Software Engineering, and Recommendation Systems, other than to apply for financial funding on themes concerned with the definition of engineering approaches for the development of complex software systems. It is therefore the intention of the involved partners to keep maintaining the ongoing collaborations and establishing further international collaborations that will contribute to advance the state of the art on software fairness modeling and engineering.

In addition, the activities of the project will interact with those of ongoing projects supported by European institutions, like the COSMOS project (DevOps for Complex Cyber-Physical Systems - https://www.cosmos-devops.org, financed by the European Commission through the Horizon 2020 program, Project No. 957254-COSMOS) and the TED project (Test Code Quality: A New Dimension of Test Code Effectiveness, financed by the Swiss National Science Foundation through the Ambizione program, Project No. PZ00P2_186090). On the one hand, the COSMOS project aims at developing DevOps methodologies and techniques to support the continuous development and evolution of cyber-physical systems, which are often enabled by machine learning and artificial intelligence approaches - and that, therefore, might suffer from ethical and fairness concerns. On the other hand, the TED project aims at establishing novel verification and validation techniques to support the development of reliable and dependable complex systems, including ML-intensive software - as such, some of those techniques may be adapted to the specific context of fairness testing. In other terms, the research activities proposed by FRINGE will complement and be complemented by the ones conducted in ongoing European projects, hence further increasing the impact of FRINGE and its visibility, feasibility, and sustainability. Moreover, FRINGE will interact with SoBigData RI that entered in the ESFRI roadmap as a high priority EU research infrastructure. SoBigData RI has been funded by the European Commission (through the Horizon 2020 program, Projects No. 101079043- SoBigData PPP and No. 871042-SoBigData ++) and by Italian PNRR (proposal No. IR0000013-SoBigData.it) and it will focus on social mining as a toolbox aimed at understanding and modeling complex social phenomena and their evolution. Its main aims are democratizing the benefits of data science and Big Data within an ethical framework that harmonizes individual rights and collective interest; to focus on Social Mining to understand and model complex social phenomena and to build an Open Science ecosystem composed by adequate means for accessing big social data together with algorithms for extracting knowledge from them. SoBigData RI will push the FAIR (Findable, Accessible, Interoperable) and FACT (Fair, Accountable, Confidential and Transparent) principles. It will also orient resources from multiple perspectives: e-infrastructures and online services developers; big data analytics and AI; complex systems focussed on modeling social phenomena; ELSEC (Ethical, Legal, SocioEconomic and Cultural) aspects of data protection (as defined by the HLEG-AI); privacy preserving techniques.

Institution	Expert	Main Research Area
Maastricht University The Netherlands	**Prof. Michel Dumontier** Role: Distinguished Professor & Director of the Institute of Data Science Department: Institute of Data Science Email: michel.dumontier@maastrichtuniversity.nl	Software Fairness Data Science Artificial Intelligence Biomedical Informatics
Motivation and Contribution. He is the co-founder of the FAIR (Findable, Accessible, Interoperable and Reusable) data principles. His feedback and advice will inform the development of fairness-aware methodologies, techniques, and tools, other than the impact on society of the solutions developed within the project.		
The College of William and Mary USA	**Prof. Denys Poshyvanyk** Role: Chancellor Professor & Director of the Software Engineering, Maintenance, and Evolution Research Unit Department: Computer Science Email: denys@cs.wm.edu	Software Engineering Machine Learning Engineering Artificial Intelligence Natural Language Processing
Motivation and Contribution. He is an expert of the development of engineered artificial intelligence and natural language processing solutions. His feedback and advice will inform the selection and design of engineer automated solutions to treat software fairness.		
Università della Svizzera Italiana Switzerland	**Prof. Gabriele Bavota** Role: Associate Professor of Software Engineering & Director of the Software Analytics Research Team Department: Faculty of Informatics Email: gabriele.bavota@usi.ch	Software Engineering Artificial Intelligence Mining Software Repositories Empirical Software Engineering
Motivation and Contribution. He is an expert of empirical methods to assess artificial intelligence-enabled software engineering solutions. His feedback and advice will inform the definition and empirical validation of the solutions developed within the project.		

Table 5: Members of the Strategic Advisory Board.

Last but not least, FRINGE will establish a Strategic Advisory Board which will provide advice and recommendations to let the project be significantly more efficient, effective and strategically profitable. The Strategic Advisory Board will be composed by well-known experts on the core themes of the project, i.e., software fairness, machine learning and natural language processing, and software

analytics. Specifically, it will be composed of the three members reported in Table 5—the table also reports the specific motivation and expected contribution that led us to account for their feedback.

The members of the Strategic Advisory Board already supported the vision of the FRINGE project (the support letters are attached to the proposal). They will be involved every six months as part of an online meeting with all the members of the research units involved in the project. In addition, they will be invited to join the periodic in-person project meetings.

DISSEMINATION AND EXPLOITATION

The FRINGE partners have identified multiple dissemination and exploitation opportunities. The activities of the project will be first disseminated and advertised through keynotes and presentations to international conferences and workshops in the field of software engineering and machine learning. In addition, the partners aim at interacting with practitioners in order to make the results achieved actionable and exploitable. In this sense, two main actions will be conducted. On the one hand, the partners will actively participate in established international practitioner's forums attended by thousands of developers, data scientists and companies, like DevTernity (https://devternity.com) and AIDev (https://aidevworld.com). On the other hand, the key achievements of the project will be made accessible to practitioners through posts and discussions on Medium (https://medium.com), a well-known community accounting for over 60 million users. These activities will increase the visibility of the results and the overall practitioners' awareness with respect to themes connected to fairness engineering.

The FRINGE partners will also conduct educational activities aiming at increasing the awareness of students on the themes of software fairness. The Associated PIs currently teach university courses that well fit the objectives of the project (e.g., the PI teaches a course on Software Engineering for Artificial Intelligence) and plan to integrate the knowledge accumulated during the project development into these academic courses. Furthermore, they will also engage with high-school students through seminars and organization of role-playing games that would help students understand the impact of software fairness on society and the relevance of the activities conducted within the FRINGE project.

Lettera d'intenti

Bando PRIN 2022 - Progetti di Ricerca di Rilevante Interesse Nazionale - Ministero dell'Università e della Ricerca - D.D. n. 1409 del 14-09-2022 Manifestazioni di interesse

Oggetto: Bando PRIN 2022 - Progetti di Ricerca di Rilevante Interesse Nazionale - Ministero dell'Università e della Ricerca - D.D. n. 1409 del 14-09-2022

Lettera di intenti per la manifestazione di interesse a collaborare al progetto di ricerca FRINGE: context-aware FaiRness engineerING in complex software systEms.

La sottoscritta **ROSA COSCIONE**, in qualità di **LEGALE RAPPRESENTANTE** dell'impresa **KIRANET SRL**, con sede legale in **VIA ALDO MORO, 56 - CITTA' AVERSA**, CF e P.IVA **03121520617**, conferma la volontà di collaborare con l'Università degli Studi di Salerno all'attività di ricerca dal titolo "*FRINGE: context-aware FaiRness engineerING in complex software systEms*", relativo al bando in oggetto. In particolare, l'impresa, qualora tale progetto venisse ammesso al finanziamento, si impegna a:

- supportare le attività di ricerca previste dal progetto permettendo ai ricercatori coinvolti di condurre survey e/o di accedere ai gold standard dataset utilizzati per lo sviluppo e la valutazione di AI-based systems, allo scopo di acquisire conoscenza sul processo di sviluppo ed ingegnerizzazione di moduli software basati su machine learning che rispettino vincoli di fairness nonchè sulle pratiche di valutazione e miglioramento dell'etica e della fairness di tali sistemi messe in atto dall'azienda;
- accogliere i componenti dell'unità di ricerca dell'Università degli Studi del Sannio/di Salerno/de L'Aquila presso la propria sede operativa in **Viale della Libertà n. 33 – Aversa (CE)** per pianificare/svolgere attività di ricerca previste dal progetto e oggetto del presente accordo.

Il referente aziendale è individuato nella persona dell'Ing. **RAFFAELE CHIANESE** (raffaele.chianese@kiranet.it). Per l'intera durata del periodo di svolgimento del progetto di ricerca nessun onere sarà a carico dell'impresa, e nessun compenso sarà richiesto all'Università.

Luogo, data
Aversa, 08/11/2022

Firma e timbro

KIRANET S.R.L.
P. IVA 03121520617

Maastricht University

November 24, 2022

To: Dr. Fabio Palomba
Department of Computer Science
University of Salerno
Via Giovanni Paolo II, 132, 84084 Fisciano SA
E-mail fpalomba@unisa.it

Subject: PRIN 2022 PNRR - Directorial Decree n. 1409 (14-9-2022). Letter of intent to be part of the Strategic Advisory Board to be constituted under the research project "*FRINGE: context-aware FaiRness engineerING in complex software systEms*".

Dear Dr. Fabio Palomba,

Should your project be accepted for funding, I would be pleased to be part of the Strategic Advisory Board that will be constituted to support the research activities under the project "*FRINGE: context-aware FaiRness engineerING in complex software systEms*" for the funding call of "PRIN 2022 PNRR - Directorial Decree n. 1409 (14-9-2022)".

Sincerely,

Michel Dumontier

Distinguished Professor of Data Science
Director, Institute of Data Science
Maastricht University

Visiting address
Paul-Henri Spaaklaan 1 (PHS1)
6229 GT, Maastricht
The Netherlands
T +31 (0)43 388 49 96
info-ids@maastrichtuniversity.nl

Postal address
Institute of Data Science
Maastricht University
Post box 616
6200 MD, Maastricht
The Netherlands

michel.dumontier@maastrichtuniversity.nl
https://maastrichtuniversity.nl/ids

WILLIAM & MARY
CHARTERED 1693

DEPARTMENT OF COMPUTER SCIENCE

Letter of intent
PRIN 2022 PNRR - Directorial Decree n. 1409 (14-9-2022)

To: Dr. Fabio Palomba
Department of Computer Science
University of Salerno
Via Giovanni Paolo II, 132, 84084 Fisciano SA
E-mail fpalomba@unisa.it

Subject: PRIN 2022 PNRR - Directorial Decree n. 1409 (14-9-2022). Letter of intent to be part of the Strategic Advisory Board to be constituted under the research project *"FRINGE: context-aware FaiRness engineerING in complex software systEms"*.

Dear Dr. Fabio Palomba,

I am writing to let you known that, should your project be accepted for funding, I gladly accept to be part of the Strategic Advisory Board that will be constituted to support the research activities under the project *"FRINGE: context-aware FaiRness engineerING in complex software systEms"*, submitted to apply for the funding "PRIN 2022 PNRR - Directorial Decree n. 1409 (14-9-2022)".

11/24/22

Denys Poshyvanyk, Ph.D.
Chancellor Professor of Computer Science
Computer Science Department
William and Mary
McGlothlin-Street Hall 105
Williamsburg, VA 23185
http://www.cs.wm.edu/~denys

Letter of intent
PRIN 2022 PNRR - Directorial Decree n. 1409 (14-9-2022)

To: Dr. Fabio Palomba
Department of Computer Science
University of Salerno
Via Giovanni Paolo II, 132, 84084 Fisciano SA
E-mail fpalomba@unisa.it

Subject: PRIN 2022 PNRR - Directorial Decree n. 1409 (14-9-2022). Letter of intent to be part of the Strategic Advisory Board to be constituted under the research project *"FRINGE: context-aware FaiRness engineerING in complex software systEms"*.

Dear Dr. Fabio Palomba,

I am writing to let you known that, should your project be accepted for funding, I gladly accept to be part of the Strategic Advisory Board that will be constituted to support the research activities under the project *"FRINGE: context-aware FaiRness engineerING in complex software systEms"*, submitted to apply for the funding "PRIN 2022 PNRR - Directorial Decree n. 1409 (14-9-2022)".

Lugano, November 25 2022

Prof. Gabriele Bavota
Software Institute - Università della Svizzera Italiana
(Lugano, Switzerland)

5. Financial aspects: costs of each research unit

nº		Funds of the Ministry of University and Research (euro)
1.	PALOMBA Fabio	99.619
2.	DI SORBO Andrea	99.475
3.	DI RUSCIO Davide	100.481

		299.575

6. Bibliography

McKinsey report:
https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/data-ethics-what-it-means-and-what-it-takes, 2022.

HBR - Harvard Business Review: https://hbr.org/2020/10/ai-fairness-isnt-just-an-ethical-issue, 2020.

Reuters:
https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G, 2018.

AI Incident Base: https://incidentdatabase.ai/cite/40, 2016.

Zhang, Jie M., and Mark Harman. "" Ignorance and Prejudice" in Software Fairness." 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE). 2021.

Köchling, Alina, and Marius Claus Wehner. "Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development." Business Research 13.3 (2020): 795-848.

Starke, Christopher, et al. "Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature." Big Data & Society 9.2 (2022): 20539517221115189.

Samulowitz A, Gremyr I, Eriksson E, Hensing G. ''Brave Men' and 'Emotional Women': A Theory-Guided Literature Review on Gender Bias in Health Care and Gendered Norms towards Patients with Chronic Pain', Pain Res Manag. 2018;2018:6358624, 2018.

Seyyed-Kalantari, L., Liu, G., McDermott, M., Chen, I.Y. and Ghassemi, M. 'CheXclusion: Fairness gaps in deep chest X-ray classifiers' BIOCOMPUTING 2021: Proceedings of the Pacific Symposium p. 232-243.

Rajkomar, A., Hardt, M., Howell, M.D., Corrado, G. and Chin, M.H., 2018. 'Ensuring fairness in machine learning to advance health equity', Annals of Internal Medicine, 169(12), p.866-872.

Albarghouthi, A., D'Antoni, L., Drews, S., & Nori, A.V. (2017), FairSquare: probabilistic verification of program fairness. In Proceedings of the ACM on Programming Languages, OOPSLA 2017.

Biswas, S., & Rajan, H. (2021). Fair preprocessing: towards understanding compositional fairness of data transformers in machine learning pipeline. 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, p. 981-993.

Pessach, D., & Shmueli, E. (2022). A Review on Fairness in Machine Learning. ACM Computing Surveys (CSUR), 55(3), 1-44.

Microsoft (2022). The Fairlearn tool: https://fairlearn.org (Last access: November 2022).

IBM (2022). The AI Fairness 360 tool: https://aif360.mybluemix.net (Last access: November 2022).

C. V. Gonzalez Zelaya, "Towards Explaining the Effects of Data Preprocessing on Machine Learning," 2019. 35th International Conference on Data Engineering (ICDE), 2019, pp. 2086-2090.

Finkelstein, A., Harman, M., Mansouri, S. A., Ren, J., & Zhang, Y. (2008). “Fairness analysis” in requirements assignments. 16th IEEE International Requirements Engineering Conference, pp. 115-124.

Finkelstein, A., Harman, M., Mansouri, S. A., Ren, J., & Zhang, Y. (2009). A search based approach to fairness analysis in requirement assignments to aid negotiation, mediation and decision making. Requirements engineering, 14(4), 231-245.

Chakraborty, J., Majumder, S., & Menzies, T. (2021). Bias in machine learning software: why? how? what to do?. 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, p. 429-440.

Galhotra, S., Brun, Y., & Meliou, A. (2017). Fairness testing: testing software for discrimination. 11th Joint meeting on foundations of software engineering, p. 498-510.

Aggarwal, A., Lohia, P., Nagar, S., Dey, K., & Saha, D. (2019). Black box fairness testing of machine learning models. 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, p. 625-635.

Perera, A., Aleti, A., Tantithamthavorn, C., Jiarpakdee, J., Turhan, B., Kuhn, L., & Walker, K. (2022). Search-based fairness testing for regression-based machine learning systems. Empirical Software Engineering, 27(3), 1-36.

Riazy, S., Simbeck, K., & Schreck, V. (2020). Systematic literature review of fairness in learning analytics and application of insights in a case study. International conference on computer supported education, p. 430-449.

Calmon, F., Wei, D., Vinzamuri, B., Natesan Ramamurthy, K., & Varshney, K. R. (2017). Optimized pre-processing for discrimination prevention. Advances in neural information processing systems, 30.

Loukina, A., Madnani, N., & Zechner, K. (2019). The many dimensions of algorithmic fairness in educational applications. 14th Workshop on Innovative Use of NLP for Building Educational Applications, p. 1-10.

Kusner, M. J., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual fairness. Advances in neural information processing systems, 30.

Balayn, A., Lofi, C. & Houben, GJ. Managing bias and unfairness in data for decision support: a survey of machine learning and data engineering approaches to identify and mitigate bias and unfairness within data management and analytics systems. The VLDB Journal 30, 739–768 (2021).

Peng, K., Chakraborty, J., Menzies, T. (2022) xFAIR: Better Fairness via Model-based Rebalancing of Protected Attributes, to appear.

VERMA, Sahil; RUBIN, Julia. Fairness definitions explained. 2018 IEEE/ACM International Workshop on Software Fairness (Fairware). IEEE, 2018. p. 1-7.

BRUN, Yuriy; MELIOU, Alexandra. Software fairness. 26th ACM joint meeting on european software engineering conference and symposium on the foundations of software engineering. 2018. p. 754-759.

7. Time schedule of the research activities (GANTT CHART)

Milestone 1 Project setup and knowledge acquisition (M6)

ACTIVITY	ASSIGNED TO	I year						II year					
		BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6	BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6
Development of a metamodel for software fairness	PALOMBA F DI SORBO A DI RUSCIO D	X	X	X									
Definition of a taxonomy of fairness metrics for software systems	DI SORBO A DI RUSCIO D	X	X	X									
Definition of a taxonomy of fairness issue	DI SORBO A	X	X	X									

categories	DI RUSCIO D												
Definition of a catalog of fairness design patterns and antipatterns	PALOMBA F	X	X	X									
Website and social media accounts creation	PALOMBA F DI SORBO A DI RUSCIO D	X	X	X									

Milestone 2 Solutions for processing information sources (M12)

ACTIVITY	ASSIGNED TO	I year						II year					
		BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6	BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6
Development of a metamodel for software fairness	PALOMBA F DI SORBO A DI RUSCIO D				X	X	X						
Development of recommender systems supporting the adoption of fairness definitions	DI RUSCIO D				X	X	X						
Fairness-aware monitoring for ML-intensive systems	PALOMBA F DI RUSCIO D				X	X	X						
Development of a knowledge base for synthesizing information on fairness metrics	DI SORBO A DI RUSCIO D				X	X	X						
Development of a fairness metrics	PALOMBA F				X	X	X						

recommender system	DI RUSCIO D												
Development of a software documentation analyzer	DI SORBO A				X	X	X						
Development of a fairness issue detector	DI SORBO A DI RUSCIO D				X	X	X						
Definition of automated techniques to support fairness design	PALOMBA F DI RUSCIO D				X	X	X						
Definition of automated test case generators	PALOMBA F DI SORBO A				X	X	X						
Open source release of the FRINGE artifacts	PALOMBA F DI SORBO A DI RUSCIO D				X	X	X						

Milestone 3 Solutions for generating outputs (M18)

ACTIVITY	ASSIGNED TO	I year						II year					
		BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6	BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6
Development of recommender systems supporting the adoption of fairness definitions	DI RUSCIO D							X	X	X			
Fairness-aware monitoring for ML-intensive systems	PALOMBA F DI RUSCIO D							X	X	X			

Development of a fairness metrics recommender system	PALOMBA F DI RUSCIO D							X	X	X			
Development of a software documentation analyzer	DI SORBO A							X	X	X			
Development of a fairness issue detector	PALOMBA F DI SORBO A							X	X	X			
Definition of automated test case generators	PALOMBA F DI SORBO A							X	X	X			
Open source release of the FRINGE artifacts	PALOMBA F DI SORBO A DI RUSCIO D							X	X	X			

Milestone 4 Refinement and finalization (M24)

ACTIVITY	ASSIGNED TO	I year						II year					
		BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6	BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6
Fairness-aware monitoring for ML-intensive systems	PALOMBA F DI RUSCIO D										X	X	X
Development of a fairness metrics recommender system	PALOMBA F DI RUSCIO D										X	X	X
Development of a software documentation analyzer	DI SORBO A										X	X	X

Development of a fairness issue detector	PALOMBA F DI SORBO A										X	X	X
Integration of WP3 outcomes within the conceptual metamodel	PALOMBA F DI SORBO A DI RUSCIO D										X	X	X
Integration of WP4 outcomes within the conceptual metamodel	PALOMBA F DI SORBO A DI RUSCIO D										X	X	X
Integration of WP5 outcomes within the conceptual metamodel	PALOMBA F DI SORBO A DI RUSCIO D										X	X	X
Open source release of the FRINGE artifacts	PALOMBA F DI SORBO A DI RUSCIO D										X	X	X

8. Time schedule of the expenses

nº	Research Units	Expenses	I year						II year					
			BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6	BIM. 1	BIM. 2	BIM. 3	BIM. 4	BIM. 5	BIM. 6
1.	PALOMBA Fabio	item A1	X	X	X	X	X	X	X	X	X	X	X	X
2.	PALOMBA Fabio	item A2			X	X	X	X	X	X	X	X	X	X
3.	PALOMBA Fabio	item B												
4.	PALOMBA Fabio	item C												

5.	PALOMBA Fabio	item D	X	X	X	X	X	X	X	X	X	X	X	X
6.	PALOMBA Fabio	item E												
7.	PALOMBA Fabio	item F												
8.	DI RUSCIO Davide	item A1	X	X	X	X	X	X	X	X	X	X	X	X
9.	DI RUSCIO Davide	item A2			X	X	X	X	X	X	X	X	X	X
10.	DI RUSCIO Davide	item B												
11.	DI RUSCIO Davide	item C												
12.	DI RUSCIO Davide	item D	X	X	X	X	X	X	X	X	X	X	X	X
13.	DI RUSCIO Davide	item E												
14.	DI RUSCIO Davide	item F												
15.	DI SORBO Andrea	item A1	X	X	X	X	X	X	X	X	X	X	X	X
16.	DI SORBO Andrea	item A2			X	X	X	X	X	X	X	X	X	X
17.	DI SORBO Andrea	item B												
18.	DI SORBO Andrea	item C												
19.	DI SORBO Andrea	item D	X	X	X	X	X	X	X	X	X	X	X	X
20.	DI SORBO Andrea	item E												
21.	DI SORBO Andrea	item F												

B.2

1. Scientific Curriculum of the Principal Investigator

- Researcher unique identifier: ORCID Id	0000-0001-9337-5116
- URL for web site:	https://docenti.unisa.it/027888/home
Academic age (years from the beginning of scientific activity, i.e. years from first publication or from the beginning of PhD or	10

Medical Specialisation School)	
Previous positions	- 01/01/2018 - 31/07/2019. Senior Research Associate at the Department of Informatics, University of Zurich (Switzerland); - 01/01/2017 - 31/12/2017. Research Fellow at the Department of Software Technology (ST), Technical University Delft (The Netherlands).
Prizes and awards	- 2017 IEEE Computer Society Best Ph.D. Thesis (Italy Chapter); - 2 ACM/SIGSOFT Distinguished Paper Awards (IEEE/ACM International Conference on Automated Software Engineering 2013, IEEE/ACM International Conference on Software Engineering 2015); - 1 IEEE/TCSE Distinguished Paper Awards (IEEE International Conference on Software Analysis, Engineering, and Reengineering 2022); - 1 Best Paper Honorable Mention (ACM Computer Supported Cooperative Work 2018); - 2 Best Paper Awards (IEEE International Conference on Software Analysis, Engineering, and Reengineering 2018, Euromicro Conference Series on Software Engineering and Advanced Applications 2022); - 13 Distinguished Reviewer Awards from major software engineering journals (e.g., Springer's Empirical Software Engineering Journal) and conferences (e.g., IEEE/ACM International Conference on Automated Software Engineering); full list available on the Dr. Palomba's website.
Visiting academic positions	- 05/01/2021 - 05/03/2021. Visiting Professor at the Jheronimus Academy of Data Science (The Netherlands); - 01/06/2021 - 30/09/2021. Visiting Professor at the Jheronimus Academy of Data Science (The Netherlands); - 01/09/2015 - 31/12/2015. Visiting Ph.D. Student at the Department of Software Technology (ST), Technical University Delft (The Netherlands).
Teaching activities and PhD supervision	At the Computer Science BSc, MSc, and Ph.D. degree of the University of Salerno (UNISA), Dr. Palomba teaches the courses of Fundamentals of Artificial Intelligence (BSc - 6 ECTS) and Software Engineering for AI (MSc - 6 ECTS). He is also co-lecturer of the courses of Software Engineering (BSc - 9 ECTS) and Empirical Research Methods (Ph.D. - 3 ECTS). From 2019 to 2022, he also taught the course of Software Dependability (MSc degree in Computer Science - 9 ECTS). He advised or currently advises a total of 101 BSc and MSc thesis. At UNISA, Dr. Palomba is the advisor of four Ph.D. students (Emanuele Iannone, Giammaria Giordano, Gilberto Recupito, Vincenzo De Martino). He is the co-advisor of five more Ph.D. students (Valeria Pontillo, Manuel De Stefano, Stefano Lambiase, Giulia Sellitto, Dario Di Dario) and one research assistant (Carmine Ferrara). Formerly, he co-advised one Ph.D. student at UNISA (Fabiano Pecorelli) and two at University of Zurich (Larissa Braz and Giovanni Grano).
Other work experience (e.g. consultancy if any)	None.
- Administrative role and position responsibility	- Member of the "Commissione per i Piani di Studio", Department of Computer Science, Università degli Studi di Salerno from 01-09-2021 to today; - Responsabile di Sede of the Working Group of the CINI laboratory for "System and Service Quality" from 01-02-2022 to today.
- Scientific organisations/Coordination of academic activities	Dr. Palomba is a Steering Committee Member of the IEEE/ACM International Conference on Program Comprehension (ICPC, CORE: A). He has served as Program Co-Chair of the IEEE/ACM International Conference on Program Comprehension (ICPC 2021, CORE: A) and will serve as Program Co-Chair of the IEEE International Conference on Software Analysis, Engineering, and Reengineering (SANER 2024, CORE: A). He has also co-chaired conferences tracks (e.g., Industrial and Novel Ideas track) of multiple international conferences such as the IEEE/ACM International Conference on Automated Software Engineering (ASE 2022, CORE: A*), the 2022 International Conference on Source Code Analysis and Manipulation (SCAM 2022, CORE: C), the International Conference on Software Analysis, Evolution, and Reengineering (SANER 2022, CORE: A), the

	International Conference on Mining Software Repositories (MSR 2022), and the IEEE/ACM International Conference on Mobile Software Engineering and Systems (MobileSoft 2022). He has served as Associate Chair of the ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2021 and 2023). He has been the co-founder of the International Workshop of Machine Learning Techniques for Software Quality Evaluation and of the International Workshop on Software Quality Assurance for Artificial Intelligence. He also served as Scientific Director of the 15th International Summer School on Software Engineering (held at the University of Salerno). Dr. Palomba has served or is currently serving as Program Committee Member of all the major software engineering conferences, e.g., ICSE, ESEC/FSE, ASE. In addition, he has served as Program Committee Member of the Genetic and Evolutionary Computation Conference (GECCO).
Editorial activity	Dr. Palomba has been or is currently part of the Editorial Board of all major software engineering journals. Since 2022, he is Editorial Board Member of the Elsevier's Information and Software Technology Journal (IST). Since 2021, he is Editorial Board Member of the Springer's Empirical Software Engineering Journal (EMSE), where he already served as Review Board Member since 2016, and the e-Informatica Software Engineering Journal (EISEJ). Since 2020 he is a Review Board Member of the IEEE Transactions on Software Engineering. Since 2019, he is Editorial Board Member of Elsevier's Journal of Systems and Software (JSS), and Elsevier's Science of Computer Programming (SCICO). For his reviewing activities, he was the recipient of 13 Distinguished/Outstanding Reviewer Awards. From 2019 to 2022, he was also Social Media Director of ACM Transactions on Software Engineering and Methodology (TOSEM).
Membership of scientific societies	- IEEE Computer Society; - ACM.
Funding (current and past)	

Anno	Project title	Person months	Funding organisation
2022	Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care	3	Ministero della Salute (PNRR Complementare Salute)
2019	Test Code Quality: A New Dimension of Test Code Effectiveness	2	Swiss National Science Foundation (SNSF)

Significant career breaks	None.
- H-Index (in Scopus):	39
- Total number of publications in peer-reviewed journals	51
- Total IF	241,624
- n. and total IF of publications where the candidate is first author or equivalent (for the disciplines where the position in the list of authors correspond to the role in the work presented)	Dr. Palomba is the first author of 11 articles published in journals, with a total IF of 62,219.
- N. and total IF of the publications where the candidate is last or corresponding author (for the disciplines where the	Dr. Palomba is the last author of 4 articles published in journals, with a total IF of 13,002.

position in the list of authors correspond to the role in the work presented)	

2. Scientific Curriculum of the associated PIs

1. DI SORBO Andrea

- Researcher unique identifier: ORCID Id	0000-0002-3192-739X
- URL for web site:	https://www.unisannio.it/user/9355/contatti
Academic age (years from the beginning of scientific activity, i.e. years from first publication or from the beginning of PhD or Medical Specialisation School)	7
Previous positions	- 16/07/2019 - 15/07/2022. Research Fellow at the Department of Engineering, University of Sannio (Italy); - 01/11/2017 - 15/07/2019. Postdoctoral researcher at the Department of Engineering, University of Sannio (Italy).
Prizes and awards	- 2013. Best graduate award in Computer Engineering at the University of Sannio
Visiting academic positions	- 11/11/2019 - 11/02/2020. Visiting Researcher at the University of Zurich (Switzerland).
Teaching activities and PhD supervision	At the University of Sannio, Dr. Di Sorbo currently teaches Algorithms and Data Structures (BSc degree in Computer Engineering) and Machine Learning and Empirical Methods (Ph.D. degree in Computer Engineering); From 2019 to 2020, he also taught Programming Techniques (MSc degree in Computer Engineering), while in 2018 he was co-lecturer of the course of Statistical Analysis of Biological Data and Advanced Programming (MSc degree in Genetic and Molecular Biotechnologies). He advised or co-advised a total of 10 BSc and MSc theses. Dr. Di Sorbo is currently the co-advisor of one Ph.D. student at the University of Sannio (Arnaldo Sgueglia). Formerly, he co-advised two Ph.D. students at the University of Sannio (Sara Forootani and Anna Vacca).
Other work experience (e.g. consultancy if any)	None
- Administrative role and position responsibility	None
- Scientific organisations/Coordination of academic activities	Dr. Di Sorbo serves and has served as a program committee member in flagship international conferences in the software engineering domain, such as ICSE 2023, ASE 2022, SEAA 2019, ARES 2019/2020/2021, and MobileSOFT 2020/2021/2022/2023. He is also co-chair and organizer of the International Workshop on Natural Language-based Software Engineering (NLBSE). The workshop is at its second edition and is co-located with ICSE, the most important international conference in the software engineering field.
Editorial activity	Dr. Di Sorbo currently serves as guest editor for the “NLBSE’22: Natural Language-based Software to Support Software Engineering Processes” special issue of the Science of Computer Programming (SCICO) journal and guest associate editor for the Frontiers in Big Data (e-ISSN 2624-909X) journal. Formerly, he has served as a guest editor for the “User Feedback and Software Quality in the Mobile Domain” special issue of the Information and Software Technology (IST) journal edited by Elsevier, and as a guest editor for the “Data and privacy protection in surveillance economy” topic of the Frontiers in Big Data journal. Also, dr. Di Sorbo is frequently

	invited to review research papers submitted to top-tier journals, like IEEE Transactions on Software Engineering (TSE), IEEE Transactions on Mobile Computing (TMC), Expert Systems with Applications (ESWA), ACM Transactions on Software Engineering and Methodology (TOSEM), Empirical Software Engineering (EMSE), etc.
Membership of scientific societies	None

Funding (current and past)	**Anno**	**Project title**	**Person months**	**Funding organisation**

Significant career breaks	None
- H-Index (in Scopus):	13
- Total number of publications in peer-reviewed journals	14
- Total IF	55.905
- n. and total IF of publications where the candidate is first author or equivalent (for the disciplines where the position in the list of authors correspond to the role in the work presented)	Dr. Di Sorbo is first author of 4 articles published in journals with a total IF of 18.462
- N. and total IF of the publications where the candidate is last or corresponding author (for the disciplines where the position in the list of authors correspond to the role in the work presented)	Dr. Di Sorbo is last or corresponding author of 2 articles in journals with a total IF of 8.967

2. DI RUSCIO Davide

- Researcher unique identifier: ORCID Id	0000-0002-5077-6793
- URL for web site:	http://people.disim.univaq.it/diruscio/
Academic age (years from the beginning of scientific activity, i.e. years from first publication or from the beginning of PhD or Medical Specialisation School)	18
Previous positions	--
	2021 ESEC/FSE 2021 Distinguished Reviewer Award 2020 Best Paper Award: PT. “CrossRec: Supporting software developers by recommending third-party libraries”. J. Syst. Softw. 2020 Diamond Best Paper Award: “CrossRec: Supporting software developers by

Prizes and awards	recommending third-party libraries". J. Syst. Softw. 2020 Most Influential Paper Award (MIP) received at the 13th ACM SIGPLAN Int. Conf. on Software Language Engineering (SLE) for the paper: R. Eramo, A. Pierantonio, D. Di Ruscio and A. Cicchetti: JTL: a bidirectional and change propagating transformation language. SLE 2010 2019 Best paper award: P. T. Nguyen, J. Di Rocco, D. Di Ruscio, A. Pierantonio, L. Iovino "Automated Classification of Metamodel Repositories: A Machine Learning Approach" (MODELS 2019) 2017 Best reviewer award at ACM/IEEE 20th Int. Conf. on Model Driven Engineering Languages and Systems 2013 Winner of the context "20 talenti per l'Italia" Working Capital Telecom Italia
Visiting academic positions	15/04/2009 - 15/10/2009. Visiting Researched at ISTI-CNR, Pisa, Laboratorio di Ingegneria del Software, dell'Ing. Antonia Bertolino 1/02/2006 - 1/05/2006. Visiting PhD student at ATLAS group Université de Nantes (Francia), under the supervision of Prof. Jean Bézivin;
Teaching activities and PhD supervision	MSc DEGREE COURSES Software Engineering for Autonomous Systems Software Engineering for the Internet of Things BSc DEGREE COURSES Multimedialità per le scienze sociali e dell'educazione Multimedialità e informatica per le scienze sociali CURRENT PHD STUDENTS I. Felicien, working on low-code techniques and tools supporting the development of IoT systems C. Di Sipio, working on the low-code development of recommendation systems A. Indamutsa, working on reuse and integration of low-code artifacts A. Sahay, working on mechanisms to compose model transformations in low-code development environments CURRENT POST-DOCS Dr. J. Di Rocco (co-advised with Prof. A. Pierantonio) working on traceability mechanisms to model management in MDE Dr. P. Nguyen working on the development of recommendation systems to support the development of complex software systems R. Rubei, working on the definition of recommendation systems supporting the development of complex software systems
Other work experience (e.g. consultancy if any)	PROJECT EVALUATIONS 2018-2019 Reviewer for National Centre of Science and Technology, Ministry of Education and Science, Republic of Kazakhstan 2016 Reviewer for EU COST Association -Open Call OC-2016-1 in the roles of EXTERNAL EXPERT and RAPPORTEUR 2015 Reviewer for Austrian Science Fund MAIN COORDINATION ACTIVITIES IN RESEARCH PROJECTS 2017-2020 Technical Leader for Uni L'Aquila of the EU H2020 MegaM@Rt2 project since 2019 co-PI of the EU H2020-MSCA-ITN-2018 Lowcomote project 2017-2019 PI, Scientific and Technical Leader, WP leader for EU H2020 CROSSMINER project since 2018 PI Uni L'Aquila and Workpackage Leader of the EU H2020 TYPHON project since 2018 Steering committee member of the project POR FESR ABRUZZO ConnectPA 2014-2017 Tech Leader of the MiSE project "Smart digital ecosystem for the Customer Experience Enhancement" 2012-2015 co-PI Uni L'Aquila and WP Leader of the EU FP7 STREP OSSMETER project 2008-2011 WP Leader EU FP7 MANCOOSI project
	- APPOINTED BY THE COORDINATOR of the Doctoral Program in Information and Communication Technology (University of L'Aquila) as SUPPORT for the MANAGEMENT AND COORDINATION of the Doctoral Program in Information and Communication Technology (University of L'Aquila) from 01-12-2018 to today

- Administrative role and position responsibility	- REFERENTE DI ATENEO per il Placement, Università degli Studi dell'Aquila from 1-2-2022 to today MEMBER of the “Commissione Orientamento di Ateneo”, Università degli Studi dell'Aquila from 14-10-2018 to today - MEMBER of the COMMISSION “Orientamento e Tutorato” - Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Università degli Studi dell'Aquila from 18-07-2018 to today - MEMBER of the CINI laboratory for SmartCities & Communities from 01-01-2017 to today - MEMBER of the CINI laboratory for “Informatica e Scuola” from 20-12-2020 to today
- Scientific organisations/Coordination of academic activities	PROGRAM COMMITTEE MEMBER for more than 130 international events including: - ICSE 2023 - 45th IEEE/ACM International Conference on Software Engineering - ESEC/FSE 2021 - ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering - Artifact Evaluation@ICSE 2021 - Artifact Evaluation Track at the 43rd International Conference on Software Engineering - OSS 2021 - 17th International Conference on Open-Source Systems - ICSE 2013 - Formal Demo Selection Committee, 35th International Conference on Software Engineering - SLE 2013/2014 - International Conference on Software Language Engineering - MODELS 2017/2019/2020 - ACM/IEEE International Conference on Model Driven Engineering Languages and Systems - ECMFA 2019/2020/2021 - European Conference on Modelling Foundations and Applications - SEAA 2018/2019/2020/2021 - Euromicro Conference on Software Engineering and Advanced Applications PROGRAM CHAIR - 11th International Conference on the Quality of Information and Communications Technology - thematic track on Quality Aspects in Model-Driven Engineering (QUATIC 2018) - 8 th Software Language Engineering Conference (SLE 2015) - 7 th International Conference on Model Transformation 2014 (ICMT2014) and co-chairs of several co-located events at MODELS, and ICSE BOARD MEMBER of MODELS 2018, 2021, and 2022 GENERAL CHAIR of SATToSE 2014, 7th Seminar Series on Advanced Techniques & Tools for Software Evolution ORGANIZING COMMITTEE / WORKSHOP CHAIR of STAF 2015, 4th edition of Software Technologies: Applications and Foundations Conference STEERING COMMITTEES 2013-2018: Int. Conf. on Model Transformation (ICMT) since 2016: Workshop on Modelling in Software Engineering at ICSE (MiSE) since 2015: Software Language Engineering (SLE) Conf. since 2014: Seminar Series on Advanced Techniques & Tools for Software Evolution (SATTOSE)
Editorial activity	since 2020: Editorial Board Member IEEE Software since 2019: Editorial Board Member J. of Object Technology (JOT) since 2019: Editorial Board Member Springer Int. Journal on Software and Systems Modeling (SoSyM) since 2017: Editorial Board Member IET Software J. 2015-2016: Guest Editor of the special issue “Flexible Model-Driven Engineering” of the Elsevier Computer Languages, Systems & Structures j. 2014: Guest Editor of the special issue “Success Stories in Model-Driven Engineering” of the Elsevier Science of Computer Programming j. 2013: Guest Editor of the special issue “Experimental Software and Toolkits” of the Elsevier Science of Computer Programming j. 2012-2013: Guest Editor of the special issue “Extreme Modeling” of the J. of Object Technology

	2011-2012: Guest Editor of the special issue “Model Comparison In Practice” of the J. of Object Technology
Membership of scientific societies	ACM SigSoft

Funding (current and past)

Anno	Project title	Person months	Funding organisation
2018	Polyglot and Hybrid Persistence Architectures for Big Data Analytics	50.8	EU H2020
2017	Developer-Centric Knowledge Mining from Large Open-Source Software Repositories	55.5	EU H2020
2012	Automated Measurement and Analysis of Open Source Software	58.5	EU FP7 STREP

Significant career breaks	None
- H-Index (in Scopus):	26
- Total number of publications in peer-reviewed journals	45
- Total IF	147.51
- n. and total IF of publications where the candidate is first author or equivalent (for the disciplines where the position in the list of authors correspond to the role in the work presented)	86.303
- N. and total IF of the publications where the candidate is last or corresponding author (for the disciplines where the position in the list of authors correspond to the role in the work presented)	51.247

3. Main Principal Investigator's scientific publications (Max. 20)

1. Borovits N., Kumara I., Di Nucci D., Krishnan P., Dalla Palma S., Palomba F., Tamburri D. A., van den Heuvel W. -J. (2022). FindICI: Using machine learning to detect linguistic inconsistencies between code and natural language descriptions in infrastructure-as-code. EMPIRICAL SOFTWARE ENGINEERING, vol. 27, p. 1-30, ISSN: 1382-3256, doi: 10.1007/s10664-022-10215-5 - **Articolo in rivista**
2. Fabio Palomba, Damian Tamburri (2021). Predicting the Emergence of Community Smells using Socio-Technical Metrics: A Machine-Learning Approach. THE JOURNAL OF SYSTEMS AND SOFTWARE, vol. 171, p. 1-16, ISSN: 0164-1212, doi: 10.1016/j.jss.2020.110847 - **Articolo in rivista**
3. Stefano Dalla Palma, Dario Di Nucci, Fabio Palomba, Damian Andrew Tamburri (2021). WITHIN-PROJECT

DEFECT PREDICTION OF INFRASTRUCTURE-AS-CODE USING PRODUCT AND PROCESS METRICS. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, p. 1, ISSN: 0098-5589, doi: 10.1109/TSE.2021.3051492 - **Articolo in rivista**

4. Azeem M, Palomba F, Shi L, Wang Q (2019). Machine Learning Techniques for Code Smell Detection: A Systematic Literature Review and Meta-Analysis. INFORMATION AND SOFTWARE TECHNOLOGY, ISSN: 0950-5849, doi: 10.1016/j.infsof.2018.12.009 - **Articolo in rivista**

5. Carmine Vassallo, Sebastiano Panichella, Fabio Palomba, Sebastian Proksch, Harald C. Gall, Andy Zaidman (2019). How Developers Engage with Static Analysis Tools in Different Contexts. EMPIRICAL SOFTWARE ENGINEERING, p. 1-39, ISSN: 1382-3256, doi: 10.1007/s10664-019-09750-5 - **Articolo in rivista**

6. Damian A. Tamburri, Fabio Palomba, Rick Kazman (2019). Exploring Community Smells in Open-Source: An Automated Approach. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, p. 1, ISSN: 1939-3520, doi: 10.1109/TSE.2019.2901490 - **Articolo in rivista**

7. Fabio Palomba, Dario Di Nucci, Annibale Panichella, Andy Zaidman, Andrea De Lucia (2019). On the impact of code smells on the energy consumption of mobile applications. INFORMATION AND SOFTWARE TECHNOLOGY, vol. 105, p. 43-55, ISSN: 0950-5849, doi: 10.1016/j.infsof.2018.08.004 - **Articolo in rivista**

8. GRANO, GIOVANNI, Palomba, Fabio, Gall, Harald C. (2019). Lightweight Assessment of Test-Case Effectiveness using Source-Code-Quality Indicators. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, p. 1, ISSN: 0098-5589, doi: 10.1109/TSE.2019.2903057 - **Articolo in rivista**

9. Palomba F, Tamburri D, Arcelli Fontana F, Oliveto R, Zaidman A, Serebrenik A (2019). Beyond Technical Aspects: How Do Community Smells Influence the Intensity of Code Smells?. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, ISSN: 0098-5589, doi: 10.1109/TSE.2018.2883603 - **Articolo in rivista**

10. Palomba, Fabio, Zanoni, Marco, Arcelli Fontana, Francesca, De Lucia, Andrea, Oliveto, Rocco (2019). Toward a Smell-aware Bug Prediction Model. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol. 45(2), p. 194-218, ISSN: 0098-5589, doi: 10.1109/TSE.2017.2770122 - **Articolo in rivista**

11. Pascarella L, Palomba F, Bacchelli A (2019). Fine-Grained Just-In-Time Defect Prediction. THE JOURNAL OF SYSTEMS AND SOFTWARE, ISSN: 0164-1212, doi: 10.1016/j.jss.2018.12.001 - **Articolo in rivista**

12. Di Nucci, Dario, Palomba, Fabio, De Rosa, Giuseppe, Bavota, Gabriele, Oliveto, Rocco, De Lucia, Andrea (2018). A Developer Centered Bug Prediction Model. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol. 44, p. 5-24, ISSN: 0098-5589, doi: 10.1109/TSE.2017.2659747 - **Articolo in rivista**

13. Palomba F, Panichella A, Zaidman A, Oliveto R, De Lucia A (2018). The Scent of a Smell: An Extensive Comparison between Textual and Structural Smells. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, ISSN: 0098-5589 - **Articolo in rivista**

14. Palomba, Fabio, Bavota, Gabriele, Penta, Massimiliano Di, Fasano, Fausto, Oliveto, Rocco, Lucia, Andrea De (2018). On the diffuseness and the impact on maintainability of code smells: a large scale empirical investigation. EMPIRICAL SOFTWARE ENGINEERING, p. 1188-1221, ISSN: 1382-3256, doi: 10.1007/s10664-017-9535-z - **Articolo in rivista**

15. Palomba, Fabio, Panichella, Annibale, Zaidman, Andy, Oliveto, Rocco, De Lucia, Andrea (2018). The Scent of a Smell: An Extensive Comparison between Textual and Structural Smells. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol. 44(10), p. 977-1000, ISSN: 0098-5589, doi: 10.1109/TSE.2017.2752171 - **Articolo in rivista**

16. DI NUCCI, DARIO, Palomba, Fabio, Oliveto, Rocco, De Lucia, Andrea (2017). Dynamic Selection of Classifiers in Bug Prediction: An Adaptive Method. IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE, vol. 1, p. 202-212, ISSN: 2471-285X, doi: 10.1109/TETCI.2017.2699224 - **Articolo in rivista**

17. Tufano M, Palomba F, Bavota G, Oliveto R, Di Penta M, De Lucia A, Poshyvanyk D (2017). When and Why Your Code Starts to Smell Bad (and Whether the Smells Go Away). IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, ISSN: 0098-5589 - **Articolo in rivista**

18. PALOMBA, FABIO, G. Bavota, M. Di Penta, R. Oliveto, D. Poshyvanyk, DE LUCIA, Andrea (2015). Mining

Version Histories for Detecting Code Smells. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol. 41, p. 462-489, ISSN: 0098-5589, doi: 10.1109/TSE.2014.2372760 - **Articolo in rivista**

19. Palomba, Fabio, SALZA, PASQUALE, Ciurumelea, Adelina, Panichella, Sebastiano, Gall, Harald, FERRUCCI, Filomena, DE LUCIA, Andrea (2017). Recommending and Localizing Change Requests for Mobile Apps Based on User Reviews. In: Proceedings - 2017 IEEE/ACM 39th International Conference on Software Engineering, ICSE 2017. p. 106-117, Association for Computing Machinery, Inc. (ACM), ISBN: 9781538638682, Buenos Aires, Argentina, 2017, doi: 10.1109/ICSE.2017.18 - **Contributo in Atti di convegno**

20. Moritz Eck, Fabio Palomba, Marco Castelluccio, Alberto Bacchelli (2019). Understanding Flaky Tests: The Developer's Perspective. In: ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. p. 830-840, ISBN: 978-1-4503-5572-8, doi: 10.1145/3338906.3338945 - **Contributo in volume (Capitolo o Saggio)**

4. Main scientific publications of the associated PIs (Max. 20, for each associated PI)

1. **DI SORBO Andrea**

1. Di Sorbo, Andrea, Panichella, Sebastiano, Visaggio, Corrado Aaron, Di Penta, Massimiliano, Canfora, Gerardo, Gall, Harald C. (2021). Exploiting Natural Language Structures in Software Informal Documentation. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol. 47, p. 1587-1604, ISSN: 0098-5589, doi: 10.1109/TSE.2019.2930519 - **Articolo in rivista**
2. Rani P., Panichella S., Leuenberger M., Di Sorbo A., Nierstrasz O. (2021). How to identify class comment types? A multi-language approach for class comment classification. THE JOURNAL OF SYSTEMS AND SOFTWARE, vol. 181, ISSN: 0164-1212, doi: 10.1016/j.jss.2021.111047 - **Articolo in rivista**
3. Panichella S., Canfora G., Di Sorbo A. (2021). “Won't We Fix this Issue?” Qualitative characterization and automated identification of wontfix issues on GitHub. INFORMATION AND SOFTWARE TECHNOLOGY, vol. 139, ISSN: 0950-5849, doi: 10.1016/j.infsof.2021.106665 - **Articolo in rivista**
4. Kallis R., Di Sorbo A., Canfora G., Panichella S. (2021). Predicting issue types on GitHub. SCIENCE OF COMPUTER PROGRAMMING, vol. 205, ISSN: 0167-6423, doi: 10.1016/j.scico.2020.102598 - **Articolo in rivista**
5. Andrea Di Sorbo, Sebastiano Panichella, Corrado Aaron Visaggio, Di Penta M, Gerardo Canfora, Harald C. Gall (2015). Development Emails Content Analyzer: Intention Mining in Developer Discussions (T). In: Proceedings - 2015 30th IEEE/ACM International Conference on Automated Software Engineering, ASE 2015. p. 12-23, Los Alamitos, CA:IEEE Computer Society, ISBN: 978-1-5090-0025-8, Lincoln, NE, USA, November 9-13, 2015 - **Contributo in Atti di convegno**
6. Andrea Di Sorbo, Corrado Aaron Visaggio, Gerardo Canfora, Sebastiano Panichella, Carol V. Alexandru, Junji Shimagaki, Harald C. Gall (2016). What would users change in my app? summarizing app reviews for recommending
software changes. In: (a cura di): Thomas Zimmermann Jane Cleland-Huang Zhendong Su, Proceedings of the 24th ACM SIGSOFT International Symposium onFoundations of Software Engineering, FSE 2016, Seattle, WA, USA,November 13-18, 2016. p. 499-510, ACM, ISBN: 978-1-4503-4218-6, Seattle, WA, USA, November 13-18, 2016, doi: 10.1145/2950290.2950299 - **Contributo in Atti di convegno**
7. Sebastiano Panichella, Andrea Di Sorbo, Emitza Guzman, Visaggio C, Gerardo Canfora, Harald C. Gall (2015). How can i improve my app? Classifying user reviews for software maintenance and evolution. In: 2015 IEEE 31st International Conference on Software Maintenance and Evolution, ICSME 2015 - Proceedings. p. 281-290, Institute of Electrical and Electronics Engineers Inc., ISBN: 978-1-4673-7532-0, Bremen, Germany, September 29 - October 1, 2015, doi: 10.1109/ICSM.2015.7332474 - **Contributo in Atti di convegno**
8. Russo E. R., Di Sorbo A., Visaggio C. A., Canfora G. (2019). Summarizing vulnerabilities’ descriptions to support experts during vulnerability assessment activities. THE JOURNAL OF SYSTEMS AND SOFTWARE, vol. 156, p. 84-99, ISSN: 0164-1212, doi: 10.1016/j.jss.2019.06.001 - **Articolo in rivista**
9. Di Sorbo A., Laudanna S., Vacca A., Visaggio C. A., Canfora G. (2022). Profiling gas consumption in solidity smart contracts. THE JOURNAL OF SYSTEMS AND SOFTWARE, vol. 186, ISSN: 0164-1212, doi: 10.1016/j.jss.2021.111193 - **Articolo in rivista**
10. Di Sorbo A., Panichella S. (2021). Exposed! A case study on the vulnerability-proneness of Google Play Apps. EMPIRICAL SOFTWARE ENGINEERING, vol. 26, ISSN: 1382-3256, doi: 10.1007/s10664-021-09978-0 - **Articolo in rivista**
11. Di Sorbo A., Grano G., Corrado Aaron Visaggio., Panichella S. (2020). Investigating the criticality of user-reported issues through their relations with app rating. JOURNAL OF SOFTWARE, ISSN: 2047-7481, doi: 10.1002/smr.2316 -

Articolo in rivista

12. Canfora G., Di Sorbo A., Forootani S., Pirozzi A., Visaggio C. A. (2020). Investigating the vulnerability fixing process in OSS projects: Peculiarities and challenges. COMPUTERS & SECURITY, vol. 99, ISSN: 0167-4048, doi: 10.1016/j.cose.2020.102067 - **Articolo in rivista**
13. Zampetti F., Di Sorbo A., Visaggio C. A., Canfora G., Di Penta M. (2020). Demystifying the adoption of behavior-driven development in open source projects. INFORMATION AND SOFTWARE TECHNOLOGY, vol. 123, ISSN: 0950-5849, doi: 10.1016/j.infsof.2020.106311 - **Articolo in rivista**
14. Vacca A., Di Sorbo A., Visaggio C. A., Canfora G. (2021). A systematic literature review of blockchain and smart contract development: Techniques, tools, and open challenges. THE JOURNAL OF SYSTEMS AND SOFTWARE, vol. 174, ISSN: 0164-1212, doi: 10.1016/j.jss.2020.110891 - **Articolo in rivista**
15. Canfora G., Di Sorbo A., Forootani S., Martinez M., Visaggio C. A. (2022). Patchworking: Exploring the code changes induced by vulnerability fixing activities. INFORMATION AND SOFTWARE TECHNOLOGY, vol. 142, ISSN: 0950-5849, doi: 10.1016/j.infsof.2021.106745 - **Articolo in rivista**
16. Forootani S., Di Sorbo A., Visaggio C. A. (2022). An Exploratory Study on Self-Fixed Software Vulnerabilities in OSS Projects. In: 2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER). p. 90-100, ISBN: 978-1-6654-3786-8, doi: 10.1109/SANER53432.2022.00023 - **Contributo in Atti di convegno**
17. Vacca A., Fredella M., Di Sorbo A., Visaggio C. A., Canfora G. (2022). An Empirical Investigation on the Trade-off between Smart Contract Readability and Gas Consumption. In: 2022 IEEE/ACM 30th International Conference on Program Comprehension (ICPC). vol. 2022, p. 214-224, ISBN: 978-1-4503-9298-3, Pittsburgh, PA, USA, 16-17 May 2022, doi: 10.1145/3524610.3529157 - **Contributo in Atti di convegno**
18. Sgueglia A., Di Sorbo A., Visaggio C. A., Canfora G. (2022). A systematic literature review of IoT time series anomaly detection solutions. FUTURE GENERATION COMPUTER SYSTEMS, vol. 134, p. 170-186, ISSN: 0167-739X, doi: 10.1016/j.future.2022.04.005 - **Recensione in rivista**
19. Azeem M. I., Panichella S., Di Sorbo A., Serebrenik A., Wang Q. (2020). Action-based recommendation in pull-request development. In: Proceedings - 2020 IEEE/ACM International Conference on Software and System Processes, ICSSP 2020. p. 115-124, Association for Computing Machinery, Inc, ISBN: 9781450375122, kor, 2020, doi: 10.1145/3379177.3388904 - **Contributo in Atti di convegno**
20. Mercaldo F., DI SORBO, Andrea, Visaggio C. A., Cimitile A., Martinelli F. (2018). An exploratory study on the evolution of Android malware quality. JOURNAL OF SOFTWARE, vol. 30, ISSN: 2047-7481, doi: 10.1002/smr.1978 - **Articolo in rivista**

---

2. DI RUSCIO Davide

1. Bettini, L, Di Ruscio, D, Iovino, L, Pierantonio, A (2022). Supporting safe metamodel evolution with edelta. INTERNATIONAL JOURNAL ON SOFTWARE TOOLS FOR TECHNOLOGY TRANSFER, ISSN: 1433-2779, doi: 10.1007/s10009-022-00646-2 - **Articolo in rivista**
2. Bettini, Lorenzo, Di Ruscio, D., Iovino, L., Pierantonio, A. (2022). An executable metamodel refactoring catalog. SOFTWARE AND SYSTEMS MODELING, vol. 21, p. 1689-1709, ISSN: 1619-1366, doi: 10.1007/s10270-022-01034-9 - **Articolo in rivista**
3. Di Rocco, Juri, Di Ruscio, Davide, Di Sipio, Claudio, Nguyen, Phuong Thanh, Pierantonio, Alfonso (2022). MemoRec: a recommender system for assisting modelers in specifying metamodels. SOFTWARE AND SYSTEMS MODELING, vol. 1, ISSN: 1619-1366, doi: 10.1007/s10270-022-00994-2 - **Articolo in rivista**
4. Di Ruscio D., Kolovos D., de Lara J., Pierantonio A., Tisi M., Wimmer M. (2022). Low-code development and model-driven engineering: Two sides of the same coin?. SOFTWARE AND SYSTEMS MODELING, ISSN: 1619-1366, doi: 10.1007/s10270-021-00970-2 - **Articolo in rivista**
5. Juri Di Rocco, Davide Di Ruscio, Claudio Di Sipio, Phuong Nguyen, Riccardo Rubei (2022). HybridRec: A recommender system for tagging GitHub repositories. APPLIED INTELLIGENCE, ISSN: 0924-669X, doi: 10.1007/s10489-022-03864-y - **Articolo in rivista**
6. Nguyen, Phuong Thanh, Di Rocco, Juri, Rubei, Riccardo, Di Sipio, Claudio, Di Ruscio, Davide (2022). DeepLib: Machine translation techniques to recommend upgrades for third-party libraries. EXPERT SYSTEMS WITH APPLICATIONS, ISSN: 0957-4174, doi: 10.1016/j.eswa.2022.117267 - **Articolo in rivista**
7. Rubei R., Di Ruscio D., Di Sipio C., Di Rocco J., Nguyen Phuong (2022). Providing upgrade plans for third-party libraries: a recommender system using migration graphs. APPLIED INTELLIGENCE, ISSN: 0924-669X, doi: 10.1007/s10489-021-02911-4 - **Articolo in rivista**
8. Di Rocco J., Di Ruscio D., Di Sipio C., Nguyen Phuong, Rubei R. (2021). Development of recommendation systems for software engineering: the CROSSMINER experience. EMPIRICAL SOFTWARE ENGINEERING, vol. 26, ISSN: 1382-3256, doi: 10.1007/s10664-021-09963-7 - **Articolo in rivista**
9. NGUYEN THANH PHUONG, Di Ruscio D., Pierantonio A., Di Rocco J., Iovino L. (2021). Convolutional neural networks for enhanced classification mechanisms of metamodels. THE JOURNAL OF SYSTEMS AND SOFTWARE, vol. 172, ISSN: 0164-1212, doi: 10.1016/j.jss.2020.110860 - **Articolo in rivista**
10. Nguyen, Phuong, Di Rocco, Juri, Di Sipio, Claudio, Di Ruscio, Davide, Di Penta, Massimiliano (2021). Recommending API Function Calls and Code Snippets to Support Software Development. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, p. 1, ISSN: 0098-5589, doi: 10.1109/TSE.2021.3059907 - **Articolo in rivista**

11. Phuong Nguyen, Juri Di Rocco, Iovino L., Di Ruscio Davide, Pierantonio A. (2021). Evaluation of a machine learning classifier for metamodels. SOFTWARE AND SYSTEMS MODELING, vol. 20, p. 1797-1821, ISSN: 1619-1366, doi: 10.1007/s10270-021-00913-x - **Articolo in rivista**
12. Basciani, Francesco, DEmidio, Mattia, Di Ruscio, Davide, Frigioni, Daniele, Iovino, Ludovico, Pierantonio, Alfonso (2020). Automated Selection of Optimal Model Transformation Chains via Shortest-Path Algorithms. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol. 46, p. 251-279, ISSN: 0098-5589, doi: 10.1109/TSE.2018.2846223 - **Articolo in rivista**
13. Capiluppi A., Di Ruscio D., Di Rocco J., Nguyen Phuong Thanh, Ajienka N. (2020). Detecting Java software similarities by using different clustering techniques. INFORMATION AND SOFTWARE TECHNOLOGY, vol. 122, ISSN: 0950-5849, doi: 10.1016/j.infsof.2020.106279 - **Articolo in rivista**
14. Nguyen P. T., Di Rocco J., Di Ruscio D., Di Penta M. (2020). CrossRec: Supporting software developers by recommending third-party libraries. THE JOURNAL OF SYSTEMS AND SOFTWARE, vol. 161, ISSN: 0164-1212, doi: 10.1016/j.jss.2019.110460 - **Articolo in rivista**
15. Nguyen Phuong, Di Rocco J., Rubei R., Di Ruscio D. (2020). An automated approach to assess the similarity of GitHub repositories. SOFTWARE QUALITY JOURNAL, vol. 28, p. 595-631, ISSN: 0963-9314, doi: 10.1007/s11219-019-09483-0 - **Articolo in rivista**
16. Rubei R., Di Sipio C., Nguyen Phuong, Di Rocco J., Di Ruscio D. (2020). PostFinder: Mining Stack Overflow posts to support software developers. INFORMATION AND SOFTWARE TECHNOLOGY, vol. 127, ISSN: 0950-5849, doi: 10.1016/j.infsof.2020.106367 - **Articolo in rivista**
17. Juan De Lara, Esther Guerra, Davide Di Ruscio, Juri Di Rocco, Jesus Sanchez Cuadrado, Ludovico Iovino, Alfonso Pierantonio (2019). Automated Reuse of Model Transformations through Typing Requirements Models. ACM TRANSACTIONS ON SOFTWARE ENGINEERING AND METHODOLOGY, ISSN: 1049-331X, doi: 10.1145/3340108 - **Articolo in rivista**
18. Juri Di Rocco, Davide Di Ruscio, Johannes Härtel, Ludovico Iovino, Ralf Lämmel, Alfonso Pierantonio (2019). Understanding MDE projects: megamodels to the rescue for architecture recovery. SOFTWARE AND SYSTEMS MODELING, vol. 19, p. 1-23, ISSN: 1619-1366, doi: 10.1007/s10270-019-00748-7 - **Articolo in rivista**
19. Franzago, Mirco, Ruscio, Davide Di, Malavolta, Ivano, Muccini, Henry (2018). Collaborative Model-Driven Software Engineering: a Classification Framework and a Research Map. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, p. 1, ISSN: 0098-5589, doi: 10.1109/TSE.2017.2755039 - **Articolo in rivista**
20. AUTILI, Marco, Bertolino, Antonia, DE ANGELIS, GUGLIELMO, DI RUSCIO, DAVIDE, Di Sandro, Alessio (2016). A Tool-Supported Methodology for Validation and Refinement of Early-Stage Domain Models. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol. 42, p. 2-25, ISSN: 0098-5589, doi: 10.1109/TSE.2015.2449319 - **Articolo in rivista**

5. Main staff involved (max 10 professors/researchers for each research unit, in addition to the PI or associated PIs), highlighting the expected time commitment

List of the Research Units

Unit 1 - PALOMBA Fabio

Personnel of the research unit

nº	Surname Name	Qualification	University/ Research Institution	e-mail address	Months/person expected
1.	PALOMBA Fabio	Ricercatore a t.d. - t.pieno (art. 24 c.3-b L. 240/10)	Università degli Studi di SALERNO	fpalomba@unisa.it	3,0
2.	FERRUCCI Filomena	Professore Ordinario	Università degli Studi di SALERNO	fferrucci@unisa.it	1,0
3.	SELLITTO Giulia	Dottorando	Università degli Studi di SALERNO	gisellitto@unisa.it	1,0
4.	LAMBIASE Stefano	Dottorando	Università degli Studi di SALERNO	S.LAMBIASE7@STUDENTI.UNISA.IT	1,0
5.	RECUPITO	Dottorando	Università degli	grecupito@unisa.it	1,0

	Gilberto	Studi di SALERNO			

Unit 2 - DI SORBO Andrea

Personnel of the research unit

nº	Surname Name	Qualification	University/ Research Institution	e-mail address	Months/person expected
1.	DI SORBO Andrea	Ricercatore a t.d. - t.pieno (art. 24 c.3-b L. 240/10)	Università degli Studi del SANNIO di BENEVENTO	andrea.disorbo@yahoo.it	3,0
2.	TORTORELLA Maria	Professore Associato confermato	Università degli Studi del SANNIO di BENEVENTO	tortorella@unisannio.it	1,5
3.	SGUEGLIA Arnaldo	Dottorando	Università degli Studi del SANNIO di BENEVENTO	arnaldo.sgueglia@gmail.com	1,0

Unit 3 - DI RUSCIO Davide

Personnel of the research unit

nº	Surname Name	Qualification	University/ Research Institution	e-mail address	Months/person expected
1.	DI RUSCIO Davide	Professore Associato (L. 240/10)	Università degli Studi dell'AQUILA	davide.diruscio@univaq.it	0,5
2.	DI MARCO Antinisca	Professore Associato (L. 240/10)	Università degli Studi dell'AQUILA	antinisca.dimarco@univaq.it	0,5
3.	NGUYEN Thanh Phuong	Ricercatore a t.d. - t.pieno (art. 24 c.3-b L. 240/10)	Università degli Studi dell'AQUILA	phuong.nguyen@univaq.it	1,0
4.	D'ALOISIO Giordano	Dottorando	Università degli Studi dell'AQUILA	giordano.daloisio@graduate.univaq.it	2,0

---

6. Information on the new contracts for personnel to be specifically recruited

nº	Associated or principal investigator	Number of expected research contracts	Number of expected PhD scholarships	Overall expected time commitment (months)
1.	PALOMBA Fabio	1	0	20
2.	DI SORBO Andrea	1	0	20
3.	DI RUSCIO Davide	1	0	20

	Total	**3**	**0**	**60**

7. PI "Do No Significant Harm (DNSH)" declaration, in compliance with article n. 17, EU Regulation 852/2020. (upload PDF)

Upload:

"The data contained in the application for funding are processed exclusively for carrying out the institutional functions of MUR. The CINECA, Department of Services for MUR, is data controller. The consultation is also reserved to universities, research institutes and institutions (each for its respective competence), MUR - Directorate-General Research- Office III, CNVR, CdV, and the reviewers in charge of the evaluation peer review.
MUR also has the right to the dissemination of the main economic and scientific data related to the funded projects.".

Date 29/11/2022 ore 22:57