

# Scientific report

## 1. Project information

### Project title:

Romanian: Abordări noi bazate pe soft computing pentru evoluția software-ului: modificări și gestionarea defectelor

English: nOvel appRoaCHEs based on Soft compuTing for softwaRe evolution: change and defect management

**Acronym:** ORCHESTRA

**<https://www.cs.ubbcluj.ro/~avescan/orchestra>**

## 2. Project objectives

### 2.1. Objective 3. Quality attributes and bug prediction

Activities

A 3.1. Synthetization of the most actual knowledge in the domain of bug prediction and quality attributes; ((partially in 2022, partially in 2023))

A 3.2. Development of a code smell detector for sql queries;

A 3.3. Validate the proposed models using open source software systems.

With the aim of advancing error prediction methodologies and augmenting software quality attributes, our principal research objective is to systematically synthesize the latest knowledge within the domain. As a crucial step toward achieving this goal, we are actively engaged in the development of soft computing approaches for predicting defects related to quality attributes. This multifaceted initiative spans both 2022 and 2023, emphasizing our commitment to an extensive and thorough exploration of innovative methodologies.

Our focus centers on investigating the potential of soft computing and software metrics in bug prediction and quality attribute analysis, ultimately aiming to elevate the overall quality of software systems. The research investigation started with a meticulous replication study of a pivotal paper in the domain of software defect prediction, originating from the Springer journal, *Frontiers of Computer Science*.

The research investigations related to this objective were finalized in several research manuscripts and submitted to ISI journals and international conferences:

- A. Vescan, R. Găceanu, C. Serban, *Exploring the Impact of Data Preprocessing Techniques on Composite Classifier Algorithms in Cross-Project Defect Prediction*, *Frontiers of Computer Science*, 2023 (under review)
- A. Vescan and R. Găceanu, "Cross-Project Defect Prediction using Supervised and Unsupervised Learning: a Replication Study," *2023 27th International Conference on*

*System Theory, Control and Computing (ICSTCC)*, Timisoara, Romania, 2023, pp. 440-447, doi: 10.1109/ICSTCC59206.2023.10308464.

- A. Moldovan, A. Vescan, *Title (not shown due to double anonymous review)*, manuscript submitted to an International Conference (under review)

*Frontiers of Computer Science Journal (under review)*

*Exploring the Impact of Data Preprocessing Techniques on Composite Classifier Algorithms in Cross-Project Defect Prediction*

The initial phase involved a Systematic Literature Review (SLR), where three databases (ACM, IEEE, Springer) were used to extract the latest papers in the domain of defect prediction. Each member meticulously assessed papers based on predefined inclusion/exclusion criteria, scrutinizing proposed approaches, datasets (whether public or private), and the resulting findings.

Having identified 30 papers through this rigorous process, team members collectively made informed decisions to progress to the next stage. Subsequently, we embarked on the replication of results from the selected paper, and datasets as those from the replicated paper to ensure accuracy and consistency. Our approach extends beyond mere replication; we systematically checked each results table, seeking not only to duplicate the findings but also to identify potential enhancements. Encouragingly, our preliminary findings indicate successful replication, coupled with substantial improvements in certain cases.

To leverage our comprehensive research on software defect prediction, we submitted our manuscript to the *Journal of Systems and Software*, after leveraging the review received from the JSS journal and enhancing the article in a second stage. This marks a significant milestone in our ongoing exploration of innovative approaches to software quality enhancement.

Our article introduces three composite algorithms—Average Voting (AvgVoting), MaxVoting, and Bagging—each integrating multiple machine classifiers to enhance cross-project defect prediction. The experiments employ pre-processing techniques such as normalization and standardization, along with feature selection.

Replicating the original findings using raw data, the study affirms the efficacy of the proposed algorithms. Notably, when normalization is applied, improved results are achieved compared to the original paper. The article underscores that even more promising outcomes emerge with the incorporation of feature selection. MaxVoting stands out for its superior F-measure performance, while BaggingJ48 excels in terms of cost-effectiveness. These findings not only validate the initial study using raw data but also demonstrate enhancements through thoughtful preprocessing and feature selection.

Our research contributes substantively to the ongoing discourse in defect prediction methodologies, promising valuable insights for the software engineering community.

*ICSTCC Conference paper*

*Cross-Project Defect Prediction using Supervised and Unsupervised Learning: a Replication Study*, DOI: <https://ieeexplore.ieee.org/document/10308464>

This paper investigates the effectiveness of various software defect prediction methods. This research is important because accurately predicting software defects can greatly enhance

the quality and reliability of software, leading to improved user satisfaction and reduced development costs.

The proposed study focuses on software defect prediction, a crucial aspect of software quality assurance. It aims to determine the most effective methods for predicting defects in software projects. This is important because identifying potential defects early in the development process can save time and resources, and improve the overall quality of the software.

The paper conducts a replication study, changing researchers and methods used in previous studies. It employs the Goal-Question-Metric (GQM) model to construct research questions, focusing on confirming the results of original studies and exploring the impact of various preprocessing techniques. It is noteworthy that a comprehensive review of related work in software defect prediction is also performed, covering supervised and unsupervised methods, within-project and cross-project defect prediction, and the use of different datasets. This review discusses previous studies that compared unsupervised and supervised learning methods for software defect prediction, highlighting the importance of feature selection techniques and the performance of various clustering algorithms.

The study conducts two sets of experiments using supervised (Random Forest, Logistic Model Tree, Naive Bayes) and unsupervised (EM, DBSCAN, KMeans) methods. The first set of experiments uses all available features, while the second set employs Principal Component Analysis (PCA) for feature selection. These experiments are performed in the context of cross-project defect prediction.

The results indicate that supervised methods perform better when considering all features, with Naive Bayes being the most effective. When fewer features are used, the results are similar or better. For unsupervised methods, the newly considered methods (EM, DBSCAN, KMeans) perform better than the original methods when fewer features are considered.

The study confirms the findings of the original papers when using all features for supervised approaches. It also demonstrates that unsupervised methods can achieve better results with fewer features, offering a new perspective compared to previous studies.

The paper concludes that both supervised and unsupervised methods are effective for software defect prediction, with the choice of features playing a significant role in the performance of these methods. The study's findings contribute to the ongoing research in software defect prediction, providing insights into the effectiveness of different methods and the impact of feature selection.

This research is valuable for software developers and project managers, as it provides a deeper understanding of how to effectively predict software defects, ultimately leading to higher quality software products.

*International Conference paper (under review)*

A. Moldovan, A. Vescan, *Title (not shown due to double anonymous review), manuscript submitted to an International Conference*

Outlier Detection through Connectivity-based Outlier Factor for Software Defect Prediction was investigated. The results are under review at a workshop within an International Conference.

## **2.2. Objective 4. Regression testing (TCP) considering bug prediction information**

### Activities

A 4.1. Synthetization of the most actual knowledge in the domain of regression testing (Test Case Prioritization)

A 4.2. Development of methods for optimally testing the software considering the changes performed on the source code and bug prediction information

A 4.3. Validate the proposed models using open source software systems.

In this objective our aim is to investigate an innovative approach, namely the use of bug prediction information for regression testing, especially for the test case prioritization problem.

Regression testing is an important quality assurance technique in the dynamic world of software development. It involves retesting a software system after changes, such as upgrades, bug patches, or interface changes, to ensure that the new code hasn't harmed the old functionality. This form of testing is critical for ensuring the long-term integrity of software, especially as new features and updates are constantly implemented.

The importance of regression testing lies in its ability to uncover faults that may have been introduced unintentionally into previously stable programs. Without it, a new feature may damage old functionality, resulting in a bad user experience or even severe program failures. Regression testing is especially important in complex and large software systems where changes are frequent and interdependencies within the codebase are complex.

While regression testing is essential, it frequently requires a huge number of test cases, which can be time-consuming and resource-intensive. TCP (Test Case Prioritization) comes into play here. TCP is the process of ordering test cases to maximize some desired outcome, such as early fault detection or reduced testing time. The idea is to execute the most significant or relevant tests as early in the testing cycle as possible.

TCP has numerous implications. For starters, it improves testing efficiency, allowing teams to find and fix issues more quickly. This is especially useful in agile and continuous integration contexts where speed and efficiency are essential. Second, TCP aids in the efficient use of limited testing resources by ensuring that the most crucial portions of the software are thoroughly tested. Finally, by prioritizing test cases based on their importance, TCP can improve overall testing efficacy, resulting in higher quality software and a better end-user experience.

The incorporation of bug prediction into TCP can help to improve the testing process. Bug prediction models predict which areas of the software are more likely to have defects based on historical information and other criteria. Testing efforts can be more properly focused on the regions of the codebase that are most prone to errors by adding this information into TCP.

We have started this objective by investigating the most recent papers published in the field of regression testing, focusing on the test case prioritization approaches. We have continued by developing our own methods for TCP and software bug prediction. We have published and sent for review several papers in this context, as detailed in the following.

The research investigations related to this objective were finalized in several research manuscripts, submitted to international conferences (some of them being already published) and one manuscript is going to be finalized and submitted:

- A. Vescan, R. Găceanu, A. Szederjesi-Dragomir, "Neural Network-Based Test Case Prioritization in Continuous Integration," 2023 38th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW), Luxembourg, Luxembourg, 2023, pp. 68-77, doi: 10.1109/ASEW60602.2023.00014.
- R. Găceanu, A. Szederjesi-Dragomir, A. Vescan, *Title (not shown due to double anonymous review)*, manuscript submitted to an International Conference (under review)
- A. Szederjesi-Dragomir, R. Găceanu, A. Vescan, *Title (not shown due to double anonymous review)*, manuscript submitted to an International Conference (under review)
- A. Vescan, R. Găceanu, A. Szederjesi-Dragomir, *Title (not shown due to double anonymous review)*, manuscript submitted to an International Conference (under review)
- A. Vescan, R. Găceanu, C. Serban, *TCP and bug prediction*, ISI journal manuscript, (to be submitted)

*ASE-ASYDE Conference paper*

*Neural Network-Based Test Case Prioritization in Continuous Integration*

DOI: <https://ieeexplore.ieee.org/abstract/document/10298714>

This study makes an important contribution to the field of software testing, focusing on Test Case Prioritization (TCP) in continuous integration systems. This study tackles the important problem of guaranteeing that changes or additions to software code do not introduce new defects. Efficiently managing and prioritizing test cases is critical for enhancing the software development process by identifying possible issues early on and lowering testing time and resources.

An intensive literature review followed by a series of experiments were used to investigate this problem in a systematic manner. The examination of the literature established the current state of knowledge in TCP and identified areas for improvement. The experiments were designed to test the effectiveness of neural network-based classification models in prioritizing test cases.

The investigations made use of three industry datasets: IOFROL, Paint Control, and GSDTSR. These real-world datasets give practical significance to the research by illustrating the applicability and efficacy of the suggested TCP technique in real-world industry contexts. The tests focused on various parts of TCP, employing various test case features and taking into account fault information at varying intervals.

The results from these experiments were very good. The NEUTRON approach, proposed for TCP in this study, showed better prioritization in terms of the Normalized Average Percentage of Faults Detected (NAPFD) metric compared to random permutation. NEUTRON's results were also comparable to other sort-based solutions and competitive with existing state-of-the-art approaches, especially when considering larger testing budgets.

Furthermore, one of the paper's significant contributions is the transformation and reduction of the original datasets without data loss. One dataset, for example, was reduced from 32,261 lines to 1,962 lines, greatly simplifying data handling and processing for research purposes. This decrease in dataset size is critical for computational efficiency and accessibility, particularly for academics and practitioners with limited computing resources.

The paper also made these transformed datasets available to the public, increasing the study's accessibility and utility to the broader scientific community. The files are available [https://figshare.com/articles/dataset/Neural\\_Network-based\\_Test\\_Case\\_Prioritization\\_in\\_Continuous\\_Integration\\_NEUTRON-Dataset/23727300/1](https://figshare.com/articles/dataset/Neural_Network-based_Test_Case_Prioritization_in_Continuous_Integration_NEUTRON-Dataset/23727300/1). This makes it possible for other researchers to duplicate the study, use the datasets for relevant research, or compare new approaches to these datasets.

In conclusion, the proposed approach validated through thorough experiments on industrial datasets, and providing the transformed, publicly available datasets constitute a robust foundation for future research and practical applications in software development and testing.

*International Conference paper (under review)*

R. Gaceanu, A. Szederjesi-Dragomir, A. Vescan, *Title (not shown due to double anonymous review), manuscript submitted to an International Conference*

The paper investigates the effectiveness of soft computing approaches applied to the problem of test case prioritization in resource constrained environments. A thorough literature review as well as extensive experiments have been conducted in this sense. The literature review showcased the current state of the research in this matter together with open issues that may be addressed. The results of the experiments are comparable with state of the art approaches and reveal promising directions for further investigation. The paper is currently under review at an international conference.

*International Conference paper (under review)*

A. Szederjesi-Dragomir, R. Gaceanu, A. Vescan, *Title (not shown due to double anonymous review), manuscript submitted to an International Conference*

The NEUTRON model was incorporated into a tool that allow developers to prioritize test cases based on various information from executed cycles with different test cases being available.

*International Conference paper (under review)*

A. Vescan, R. Gaceanu, A. Szederjesi-Dragomir, *Title (not shown due to double anonymous review), manuscript submitted to an International Conference*

The investigation regarding Test Case Prioritization and defect prediction allowed us to envision a moder view on TCP based on the current trend in testing and on software development. This vision was incorporated into a paper that is currently under review.

### **2.3. Objective 5. Regression testing (TCP) considering requirements-tests-faults traceability**

#### Activities

A 5.1. Synthetization of the most actual knowledge in the domain of regression testing (Test Case Prioritization) (2023)

A 5.2. Development of methods for optimally testing the software considering the changes and requirements-tests-faults traceability matrix (in 2024)

A 5.3. Validate the proposed models using open source software systems. (in 2024)

In ensuring the reliability of software system regression testing plays an important role. Nowadays, maintaining test suite in the continuous integration environments is challenging considering various aspects related to available resource, time and best selected test cases. In this respect, Test Case Prioritization is a solution that reorders the test suite for better and sooner fault detection. Yet, current methods use manual detection of artifact dependencies (requirement, code, test cases, faults, cycles).

The research investigations related to this objective were finalized in several research manuscripts, submitted to an international conference, being already published.

- I. -C. Rotaru and A. Vescan, "Test Case Prioritization Based on Neural Network Classification with Artifacts Traceability," *2023 38th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW)*, Luxembourg, Luxembourg, 2023, pp. 78-87, doi: 10.1109/ASEW60602.2023.00015.

*ASE-ASYDE Conference paper*

*Test Case Prioritization Based on Neural Network Classification with Artifacts Traceability*

DOI: <https://ieeexplore.ieee.org/document/10298744>

We have investigated this aspect regarding artifact traceability, by providing an automatic detection of dependency from various artifacts in the context of Behavior-Driven Development (BDD).

- The first contribution of the investigation is related to the design and implementation of an automatic traceability component to retrieve dependencies based on BDD artifacts (requirements, source code, test cases, and faults).
- The second contribution refers to the integration of the discovered traces as features in a neural network classification model for test cases for further prioritization. Various architectures were used for the neural network classification model.

A dataset consisting of two real-world BDD projects were used for the validation of the models.

## **3. Research visits and collaborations**

### **3.1. Research visit to Austria**

Between October 22-29, 2023, 2 members of the team (Andreea Vescan and Radu Gacean) visited the collaborator prof. Alexander Egyed at the Johannes Kepler University

(JKU), Linz, Austria. During this visit to Johannes Kepler University, there were meetings and discussions both with members of the ORCHESTRA group and PhD students and professors from the Software-Intensive Systems department of the Johannes Kepler University.

Prof. Alexander Egyed introduced us to the PhD. students he coordinates, but also to his colleagues from the Software-Intensive Systems department. There were meetings in which JKU members presented to grant members their research topics and the results obtained during the studies, but also meetings in which the grant members presented the objectives and how they reached their fulfillment.

### **3.1.1. Presentations of the researchers from Austria**

The first presentation was from prof. Paul Grünbacher, who presented ECCO, an alternate version control system. Other presentations included quality constraints checks in safety critical systems from Christoph Mayr-Dorn, applications of large language models in software in the software development life-cycle from Anamaria Roberta Preda, transitioning from monolithic systems to microservices from Saad Shafiq and Design Space which is a shared environment that allows several platforms to integrate and collaborate.

### **3.1.2. Presentations of the researchers from Romania**

On October 23, 2023 Andreea Vescan and Radu Gaceanu prepared and presented *Test Case Prioritization approaches in Regression Testing* detailing several contributions as well as research directions that are currently being pursued in this field.

The first presented research direction focused on TCP based on neural networks with artifacts traceability. A paper that explores this topic was already published at ASE conference [https://ieeexplore.ieee.org/abstract/document/10298744?casa\\_token=cBKkRIPGdAsAAAAA:mHFZFMzMxd7odLXZH-QiVN58Uhrm4NWshl2PsAfg569EBgpIkdp7JVr6-u0b1txKIq5Ji3anVbueeQ](https://ieeexplore.ieee.org/abstract/document/10298744?casa_token=cBKkRIPGdAsAAAAA:mHFZFMzMxd7odLXZH-QiVN58Uhrm4NWshl2PsAfg569EBgpIkdp7JVr6-u0b1txKIq5Ji3anVbueeQ). The paper proposes a novel approach that combines automatic traceability of software artifacts with neural network classification for TCP. This method uses Behavior-Driven Development (BDD) practices to automatically retrieve dependencies between requirements, source code, test cases, and faults. A systematic literature review was also conducted to explore existing TCP techniques in the context of BDD and Continuous Integration (CI), focusing on approaches for traceability from tests to source code. This review helped in identifying gaps in current methodologies and informed the development of the proposed solution. The proposed approach demonstrated promising fault detection rates, indicating the efficiency of automatic traceability in improving TCP. The neural network models showed better reliability in fault detection compared to the baseline technique. The research contributes to the field by offering a novel integration of artifact traceability and neural network classification for TCP. This approach differs from existing methods that often rely on manual input for artifact dependencies, proposing a more automated and efficient solution.

The second research direction that was presented approaches the TCP problem in continuous integration contexts. A paper that explores this topic was already published at the ASE conference (<https://ieeexplore.ieee.org/abstract/document/10298714> ). TCP is crucial in software development and it is particularly important in continuous integration, where code



changes are frequently integrated and tested. The experiments from this study were conducted on three industrial datasets: IOFROL, Paint Control, and GSDTSR. The models were trained on the IOFROL dataset and tested on the other two. The effectiveness of the models was measured using the Normalized Average Percentage of Faults Detected (NAPFD) metric. The obtained results are comparable or superior to other state-of-the-art approaches, particularly at higher budget levels (75% and 100%). The study also proposes transformed versions of the industrial datasets IOFROL, Paint Control, GSDTSR. The size of the transformed versions is significantly reduced without any information loss thus making the machine learning processes faster. The datasets are publicly available so other researchers could consider using the transformed version of these datasets in the future.

*Meeting Tuesday, 23 October 2023, 12:30 – 13:30*

Topic: “Regression testing using soft computing approaches”  
Institute of Software Systems Engineering (ISSE), Johannes Kepler University (JKU), Austria

The seminar "Regression testing using soft computing approaches" is at its second session, the first one being held in Romania on the 20th of June. Researchers from both universities discussed open issues in regression testing and the advantages of incorporating machine learning methods as well as soft computing approaches in order to enhance the TCP process.

*Meeting Friday, 26 October 2023, 11:00 – 12:00*

Topic: “Quality attributes and bug prediction”  
Institute of Software Systems Engineering (ISSE), Johannes Kepler University (JKU), Austria

The seminar "Quality attributes and bug prediction" is at its second edition, the first one taking place in Romania on the 19th of June, 2023. Researchers from both universities approached the topic of quality attributes and bug prediction and discussed possible applications in other contexts like test case prioritization.

### **3.2. Research visit to Romania**

Prof. Dr. Alexander Egyed, Full Professor and Chair for Software-Intensive Systems at Johannes Kepler University, Austria (JKU), was invited to deliver a workshop as part of the ORCHESTRA grant. The workshop, titled "Evolving Product Lines – One Product at a Time," took place on June 21, 2023, at the FSEGA Building in Cluj-Napoca, Romania. The event provided both in-person and online participation options, highlighting the significance of the subject matter in the realm of software systems engineering.

During Prof. Dr. Alexander Egyed's research visit to our university, research meetings took place within the research team of the ORCHESTRA project. These meetings provided a platform for fruitful discussions on potential collaborations and the proposal of future research endeavors. The exchange of ideas and insights during this visit further enriched the ongoing research efforts within the ORCHESTRA project, fostering a collaborative environment for advancing knowledge and innovation in software systems engineering.

## 4. Events

### 4.1. Research Seminars

#### *Seminar 1: Quality attributes and bug prediction,*

The first seminar had as topic quality attributes and bug prediction. Quality attributes in software, such as reliability, maintainability, and usability, are crucial indicators of a software product's overall quality and user satisfaction. By using historical data and various metrics, machine learning models may predict potential defects in software, allowing developers to proactively address issues before they impact the end user. This proactive approach to identifying and fixing bugs contributes to improved reliability, as the software is less likely to fail unexpectedly. It also enhances maintainability, as the codebase is kept cleaner and less cluttered with defects, making it easier to manage and update.

The first session was organized in Cluj-Napoca, during the visit of professor Alexander Egyed when each participant shared research findings in their teams about various approaches regarding bug prediction.

- First session:
  - Monday, 19 June 2023, 11:00 – 12:00  
Babeş-Bolyai University, FSEGA Building, Room C404

The second session was organized in Linz, during the visit of the team members from Cluj-Napoca, when the participants discussed ongoing research related to bug prediction.

- Second session:
  - Friday, 26 October 2023, 11:00 – 12:00  
Institute of Software Systems Engineering (ISSE), Johannes Kepler University (JKU), Austria

#### *Seminar 2: Regression testing using soft computing approaches*

The second seminar had as topic regression testing, discussing existing approach on various strategies, in particular Test Case Prioritization. Regression testing is a crucial phase in software development, ensuring that new code changes do not negatively affect the existing functionality of the software. The integration of soft computing approaches into regression testing represents a significant advancement in this field. Soft computing, a collection of methodologies that include fuzzy logic, neural networks, genetic algorithms, and machine learning, offers a more adaptable and tolerant approach to problem-solving compared to traditional hard computing methods. By employing these techniques, regression testing can be made more efficient and effective. For instance, neural networks can be used to predict the impact of code changes, identifying the most critical areas that need testing. Genetic algorithms can optimize the selection and prioritization of test cases, ensuring that the most significant tests are run first, saving time and resources. Fuzzy logic can handle the uncertainty and

imprecision inherent in software development. Machine learning algorithms in general can learn from past testing cycles, continuously improving the regression testing process over time.

The first session was organized in Cluj-Napoca, during the visit of professor Alexander Egyed when we discussed various solutions on what are the best information to be included in the prioritization of test cases.

- First session:
  - Tuesday, 20 June 2023, 9:00 – 10:00  
Babeş-Bolyai University, FSEGA Building, Room C335

The second session took place in Linz, during the visit at professor Egyed's research group. We had several discussion about traceability between various artifacts and how to incorporate them in the prioritization of test cases.

- Second session:
  - Tuesday, 23 October 2023, 12:30 – 13:30  
Institute of Software Systems Engineering (ISSE), Johannes Kepler University (JKU), Austria

## **4.2. Research Workshop**

### **4.2.1. Focus on mind. Mindfulness science and practice. From mind overload to presence, stillness & focus in research teams**

by **Nicoleta PFEFFER-BARBELA**,  
MSc., founder citySTILLE Mindfulness Center Vienna

The workshop organized as part of the ORCHESTRA grant took place at the Faculty of Mathematics and Computer Science. The event, detailed in the announcement, represented an invitation to participants.

In attendance, Mrs. Nicoleta PFEFFER-BARBELA, the Founder of citySTILLE Mindfulness Center Vienna, presented on the compelling topic: "Focus on Mind: Mindfulness Science and Practice - From Mind Overload to Presence, Stillness & Focus in Research Teams." The workshop aimed to delve into the intricate relationship between mindfulness, scientific understanding, and its practical applications in the context of research teams.

The event occurred on Wednesday, June 21, 2023, from 11:00 AM to 12:00 PM, within the FSEGA Building, specifically in Room C404, located at Teodor Mihali Street, No. 58-60, 400591, Cluj-Napoca, Romania.

Abstract: In a today's world awash in distraction, stress, and information overload—all of which can affect focus, creativity and wellbeing – mindfulness becomes a valuable tool for supporting researchers and learners.

Research demonstrates that mindfulness improves a person's ability to concentrate (Sedlmeier et al., 2012),

decreases the fear of being judged, and enhances open-minded thinking while reducing aversive self-conscious thinking (Brown, Ryan, & Creswell, 2007). These points map directly onto key characteristics of creative habits of working, thinking, and being in the world, including: relaxation or flow states (improved concentration), risk-taking (requiring a lack of fear about judgment), and curiosity or open-mindedness/openness to experience (reducing self-conscious experience) (Prabhu, Sutton, & Sauser, 2008). These effects suggest that mindfulness supports the skills associated with focus and creativity.

We invite you to have a glimpse into the some neuroscience facts, experience stillness of mind that naturally leads to focus and inspiration!

#### **4.2.2. Evolving Product Lines – One Product at a Time, by prof. dr. Alexander Egyed**

The workshop organized as part of the ORCHESTRA grant took place at the Faculty of Mathematics and Computer Science. Professor Alexander Egyed, a distinguished Full Professor and Chair for Software-Intensive Systems at Johannes Kepler University, Austria (JKU), delivered a compelling presentation during the event. The presentation, titled "Evolving Product Lines – One Product at a Time," delved into insightful perspectives on the evolution of product lines within software-intensive systems.

The workshop, held on Wednesday, June 21, 2023, from 12:30 PM to 1:30 PM in Room C335 of the FSEGA Building on Teodor Mihali Street, No. 58-60, 400591, Cluj-Napoca, Romania, drew participants both in-person and online. The online participation option was facilitated through a Teams Meeting Link, along with the relevant Meeting ID and passcode.

**Abstract:** To keep pace with the increasing demand for custom-tailored software systems, companies often apply a practice called clone-and-own. Thereby, a company builds an initial version of a system and then copies and adapts it with every new customer. Instead of a single, configurable system the company ends up with a portfolio of multiple, similar variants. Clone-and-own has widespread industrial use because it requires no major upfront investments compared to infrastructures facilitating reuse. Yet, it lacks a methodology for systematic reuse and for addressing the resulting feature interaction problems. This talk proposes ECCO (Extraction and Composition for Clone-and-Own), a novel approach to actively support software engineers in applying clone-and-own. A software engineer selects the desired features and ECCO finds the proper software artifacts for the copying step and then guides the software engineer during the manual completion by hinting which software artifacts may be missing or may need

adaptation. We evaluated our approach on 4 case studies, covering 305 variants having up to 344KLOC, and found that precision and recall of composed products quickly reached a near optimum.

Bio Prof. Egyed's impressive professional background includes a Doctorate degree from the University of Southern California, USA, and extensive industry experience before joining academic institutions. He is widely recognized for his contributions to software and systems design, particularly in variability, consistency, and traceability. With over 200 refereed scientific publications and numerous presentations, including being named an IBM Research Faculty Fellow, Prof. Egyed's expertise significantly enriches the field of software engineering.

### **4.2.3. Training sessions**

Training sessions were organized with the aim to improve the scientific writing skills. Each participant received several consultation hours for feedback and review of the writing submitted. The course *Academic Writing Skills* was delivered by Rachel Appleby.

The course consisted of individual sessions for each participant and also shared sessions with all participants. The topics includes various academic writing elements, from *How to write an abstract* to *Academic phrases*, and *tools* to help improve academic writing.

## **5. Project progress summary and dissemination of results**

### **5.1. Project progress summary**

Objective 6 regarding the *Project management* activities was employed using various tools to help obtain optimal results of the team during the project implementation.

The activities of the project started with a kickoff meeting that had the aim to meet with the entire team, to discuss activities of the project for 2023, to outline the expectations of each member and to establish the norms in the team. Several other meetings took place, specific for each of the three objectives of 2023 of the project, with discussions on the identified problems that may interfere with the research (available dataset, applied methods and methodology, design of experiments). Alternative solutions were discussed and established with the team.

The results obtained within the project are:

- related to dissemination: project webpage, workshops and presentations during research visits in Austria and Romania, posters of the seminars and workshops,
- related to project management: project monitoring plan, training session on communications and time management, training session on team

- related to scientific activities: knowledge transfer sessions between group members and with external collaborators during the research visits, workshops and presentations during research visits in Austria and Romania.
- manuscripts sent to ISI journals: one for objective 3, one to be send in 2024 as a result of the investigation in 2023, and one for the objective 2 (reviewed and resubmitted to another ISI journal).
- conferences:
  - published: three papers at two international conferences
  - under-review: 2 paper for Objective 3, 4 papers for Obj4

## 5.2. Publications. Conferences and Journals.

- Ioana-Claudia Rotaru, Andreea Vescan, *Test Case Prioritization based on Neural Network Classification with Artifacts Traceability*, the 5th International Workshop on Automated and verifiable Software sYstem DEvelopment (ASYDE) at the 38th IEEE/ACM International Conference on Automated Software Engineering (ASE 2023), **Conference CORE Rank A\***, <https://conf.researchr.org/track/ase-2023/ase-2023-workshop-asyde?#event-overview>
- Andreea Vescan, Radu Găceanu, Arnold Szederjesi-Dragomir, *Neural Network-based Test Case Prioritization in Continuous Integration*, the 5th International Workshop on Automated and verifiable Software sYstem DEvelopment (ASYDE) at the 38th IEEE/ACM International Conference on Automated Software Engineering (ASE 2023), **Conference CORE Rank A\***, <https://conf.researchr.org/track/ase-2023/ase-2023-workshop-asyde?#event-overview>
- Andreea Vescan, Radu Găceanu, *Cross-Project Defect Prediction Using Supervised and Unsupervised Learning: A Replication Study*, 27th International Conference on System Theory, Control and Computing, October 11-13, 2023, Timisoara, Romania, [https://controls.papercept.net/conferences/conferences/STCC23/program/STCC23\\_ContentListWeb\\_2.html#thb3](https://controls.papercept.net/conferences/conferences/STCC23/program/STCC23_ContentListWeb_2.html#thb3)

### Conference with ISI proceedings

The research investigations for the objectives in 2023 were integrated into a manuscript that was submitted to an ISI journal.

## 5.3. Dissemination of results

The Objective 7 of the project regarding *Dissemination and exploitation* was operationalized by several activities that are outlined in what follows.

Creating and maintaining the web page of the project with information regarding various meetings, visits and events that took place under the ORCHESTRA grant umbrella.

The webpage contains the summary of the project, the objectives, the team members, along with contacts. The webpages were updated with the presentations and workshops that took place during the visits of the researchers to Austria and to Romania.

The workshop meetings in Romania were disseminated on the webpage of the project, on the Faculty of Mathematics and Computer Science webpage, also by emails to various

research collaborators both from universities (prof. dr. Andrian Marcus from University of Texas at Dallas, prof. dr. Alexander Serebrenik from Eindhoven University of Technology, prof. dr. Alina Campan and prof. dr. Marius Truta from Northern Kentucky University) and from IT firms (EVOZON, Nagarro, Altom, Synopsys, Porsche engineering, Yonder).

As a result, the provided performance criteria were met: for the third scientific objective that is split into 2022 and 2023 years, the submission to an ISI/WoS journal was done in 2023 and one published paper to an international conference, for the fourth scientific objective two papers were published at an international conference, four more papers are under review (submitted to international conference); research visits, knowledge transfer sessions between group members and with external collaborators. Furthermore, the project objectives for 2023 have been met, and all associated activities have been completed and carried out in accordance with the project implementation plan.

### **Executive Summary of Activities**

Software maintenance and evolution play an important role in obtaining a qualitative software system, either during the development phase or after release. The changes that occur during software development are one of the important factors that threaten the quality of the software system, therefore methods and tools to achieve a better state of the system is a must.

The main objective of the project is to explore and develop an innovative set of soft computing techniques for software maintenance and evolution-related activities regarding software changes. In 2023 two directions were approached entirely (test case prioritization in continuous integration and using information from requirements-test-faults) and another one (test case prioritization using traceability matrix) was started with the final activity in 2024.

Two research visits were employed during 2023, professor Alexander Egyed from Austria (Institute of Software Systems Engineering (ISSE), Johannes Kepler University (JKU)) visited our research group in Romania in June, and two researchers from the group visited in October the research team of professor Alexander Egyed. Workshops were organized during the visit in Romania, with a presentation on *Evolving Product Lines – One Product at a Time*. Another presentation was provided by Nicoleta Pfeffer-Barbela, MSc., founder citySTILLE Mindfulness Center Vienna with the title *Focus on mind. Mindfulness science and practice. From mind overload to presence, stillness & focus in research teams*. During the visits, several other presentations and seminars took place to disseminate the project work, the respective expertises of the two teams, and help foster potential new collaborations.

Project Director,

Assoc. prof. dr. Vescan Andreea

