



LAUREA MAGISTRALE IN SOFTWARE ENGINEERING AND IT MANAGEMENT  
UNIVERSITÀ DI SALERNO  
CORSO DI INGEGNERIA GESTIONE EVOLUZIONE DEL SOFTWARE

PROF. ANDREA DE LUCIA  
DOTT. MANUEL DE STEFANO,  
EMANUELE IANNONE

## **Proposed Changes**

Repository GitHub:

<https://github.com/rebeccadimatteo/csDetector/tree/Explainability/explainability>

<https://github.com/rebeccadimatteo/CADOCS>

2023

Rebecca Di Matteo

Leonardo Monaco

<b>1</b>	<b>Contesto del Progetto</b>	<b>2</b>
<b>2</b>	<b>Obiettivi del progetto</b>	<b>3</b>
<b>3</b>	<b>CADOCS: Conversational Agent for the Detection of Community Smells</b>	<b>5</b>
<b>4</b>	<b>Proposed Change Request</b>	<b>6</b>
4.1	CR_1 Estensione di CSDETECTOR con modulo di ML basato sull'Explainability	7
4.2	CR_2 Estensione di CADOCS con un interfaccia grafica per rappresentare l'Explainability . . . . .	8
4.3	CR_3 Refactoring del codice . . . . .	9

---

## Contesto del Progetto

---

L'ingegneria del software è un'attività incentrata sull'uomo che coinvolge vari stakeholder con background diversi che devono comunicare e collaborare per raggiungere obiettivi condivisi. L'emergere di conflitti tra le parti interessate può portare a effetti indesiderati sulla manutenibilità del software, ma è spesso inevitabile nel lungo periodo.

I Community Smells, ovvero le pratiche di comunicazione e collaborazione non ottimali, sono stati definiti per individuare i conflitti ricorrenti tra gli sviluppatori.

Per facilitare un uso più ampio delle informazioni relative ai Community Smells da parte dei praticanti, prendiamo in considerazione CADOCS, un agente conversazionale client-server che si basa su un precedente strumento di rilevamento di Community Smells: CSDETECTOR. Quest'ultimo è uno strumento che rileva automaticamente i Community Smells con un approccio basato sull'apprendimento automatico; esso parte dall'analisi dei repository pubblici su GITHUB ed utilizza l'elaborazione del linguaggio naturale per comprendere l'intento dietro i messaggi di commit degli sviluppatori, le richieste pull, e problemi. Con questa conoscenza, lo strumento calcola le metriche socio-tecniche, che verranno utilizzate per addestrare un modello di apprendimento automatico in grado di prevedere i Community Smells su determinati repository.

L'idea alla base di questo progetto è di potenziare CADOCS aggiungendo nuovi moduli basati su ML, andando a creare dei modelli di Explainability all'interno di CSDETECTOR così da comprendere la spiegabilità del tool nel determinare le metriche socio-tecniche e la presenza o non dei Community Smells.

---

### Obiettivi del progetto

---

Per fornire agli utenti nuove conoscenze, si vuole raffinare il tool CADOCS in modo da dare ai praticanti uno strumento migliore in grado di esprimere anche l'Explainability dei modelli che ricavano, mediante le metriche socio-tecniche, la presenza dei Community Smells. Explainable Artificial Intelligence è un insieme di metodi e processi che consentono agli utenti di comprendere e considerare attendibili i risultati e l'output creati dagli algoritmi di Machine Learning. L'intelligenza artificiale spiegabile viene utilizzata per descrivere un modello di intelligenza artificiale, il relativo impatto previsto ed i potenziali errori.

Le funzionalità del tool verranno estese mediante l'analisi della spiegabilità dei modelli di predizione dei Community Smells per mezzo di due librerie: **Python LIME e SHAP**, al fine di comprendere quali metriche socio-tecniche siano più influenti nella predizione dei suddetti. L'Explainability sarà inserita all'interno di CADOCS mediante un'interfaccia grafica che rappresenterà l'output di una delle due librerie. La scelta dell'Explainability da rappresentare sarà affettuta sulla base di uno studio relativo ai risultati di un questionario che, sarà somministrato agli studenti del Dipartimento di Informatica dell'Università degli Studi di Salerno. Nel questionario verranno rappresentati gli output delle due librerie con delle domande a riguardo. I risultati ottenuti ci permetteranno di rispondere alle domande di ricerca che ci siamo posti per comprendere quale fra i due framework ha un'Explainability più chiara e una predizione più veritiera nel determinare le metriche socio-tecniche per la ricerca dei Community Smells, così da rappresentarla graficamente su CADOCS. Nello stato attuale CADOCS non contiene un approfondimento relativo all'Explainability. L'inserimento di tale

approfondimento può portare l'utente ad avere una maggiore fiducia sui risultati del processo decisionale che porta all'identificazione dei Community Smells. Inoltre, basandosi su questi risultati l'utente può comprendere quale metrica socio-tecnica è più influente rispetto ad un'altra nel determinare il Community Smells e dunque attuare delle strategie per risolverlo.

---

### CADOCS: Conversational Agent for the Detection of Community Smells

---

Per raggiungere il nostro obiettivo, abbiamo in programma di estendere le funzionalità di CADOCS, il quale è un agente conversazionale che viene utilizzato per rilevare automaticamente i Community Smells. CADOCS fa uso di CSDETECTOR, ovvero uno strumento che rileva automaticamente i Community Smells con un approccio basato sull'apprendimento automatico. CSDETECTOR partendo dall'analisi dei repository pubblici su GITHUB, utilizza l'elaborazione del linguaggio naturale per comprendere l'intento dietro i messaggi di commit degli sviluppatori, le richieste pull, e problemi. Con questa conoscenza, lo strumento calcola le metriche socio-tecniche, che verranno utilizzate per addestrare un modello di apprendimento automatico in grado di prevedere i Community Smells su determinati repository.

Dal punto di vista dei requisiti svilupperemo:

- un potenziamento di CADOCS aggiungendo nuovi moduli basati su ML.  
Inserendo all'interno di CSDETECTOR un modello basato sull'Explainability che vada ad analizzare l'output restituito da quest'ultimo tool
- una rappresentazione grafica dell'Explainability su CADOCS
- un refactoring del codice per migliorare la chiarezza e l'organizzazione del codice.

---

### Proposed Change Request

---

In questa sezione, si spiegherà come si intende modificare e migliorare lo strumento CSDETECTOR e CADOCS. Le change request saranno elencate ed eseguite in ordine di priorità.

Abbiamo tre gradi di priorità :

- alto
- medio
- basso

Le change request con priorità alta saranno sviluppate per prime poi successivamente quelle a priorità media e infine quelle a priorità bassa, il grado di priorità va a definire la necessità di implementare alcune modifiche prima di altre. Quando andiamo ad effettuare delle modifiche dobbiamo considerare l'effort di quest'ultime.

Abbiamo tre gradi di effort :

- alto
- medio
- basso

Le change request con un effort alto hanno una quantità di lavoro maggiore rispetto ad altre con effort medio o basso.

## 4.1 CR\_1 Estensione di CSDETECTOR con modulo di ML basato sull'Explainability

Descrizione
<p>Creazione del modello basato sull'Explainability che prende in input le metriche socio-tecniche e i Community Smells forniti in output da CSDETECTOR.</p> <p>Il modello sarà eseguito su due librerie Python LIME e SHAP che forniranno in output il grado di Explainability.</p>
Motivazione
<p>Nello stato attuale CADOCS non contiene un approfondimento relativo all'Explainability. L'inserimento di tale approfondimento può portare l'utente ad avere una maggiore fiducia sui risultati del processo decisionale che porta all'identificazione dei Community Smells, inoltre basandosi su questi risultati quest'ultimo può comprendere quale metrica socio-tecnica è più influente rispetto ad un'altra nel determinare il Community Smells e dunque attuare delle strategie per risolverlo.</p>
Priority
Alta [X] Media [] Bassa[]
Effort
Alta [X] Media [] Bassa[]
Conseguenze della non accettazione
<p>Il tool non contenendo un approfondimento sull'Explainability, non faciliterà l'utente nella comprensione dei Community Smells estratti e nella risoluzione di quest'ultimi.</p>



## 4.2 CR\_2 Estensione di CADOCS con un interfaccia grafica per rappresentare l'Explainability

Descrizione
Creazione di un interfaccia grafica per rappresentare l'Explainability all'interno di CADOCS.
Motivazione
<p>Nello stato attuale CADOCS non contiene una rappresentazione grafica dell'Explainability.</p> <p>Tale inserimento può portare facilità e immediatezza per l'utente nel comprendere come si arrivi ad avere determinati Community Smells e sfruttare tali risultati per applicare delle strategie per risolverli.</p>
Priority
Alta [X] Media [] Bassa[]
Effort
Alta [X] Media [] Bassa[]
Conseguenze della non accettazione
Il tool non contenendo un interfaccia grafica rappresentante l'Explainability non faciliterà l'utente nella comprensione dei Community Smells estratti.

### 4.3 CR\_3 Refactoring del codice

Descrizione
Riorganizzare, ristrutturare e rendere più chiaro il codice esistente di CADOCS garantendo che il comportamento complessivo del codice non cambi.
Motivazione
Il Refactoring del codice potrebbe rendere CADOCS maggiormente organizzato, più chiaro e leggibile.
Priority
Alta [] Media [X] Bassa[]
Effort
Alta [] Media [X] Bassa[]
Conseguenze della non accettazione
Non attuando il Refactoring del codice il tool potrebbe risultare poco chiaro e poco organizzato.