Crowd Requirements Engineering

Instructions for corpus annotation

To enhance the Clean CaDET Tutor intelligent tutoring system, we analyze student feedback in Serbian. The feedback consists of students' textual responses to the question, "How would you improve the Clean CaDET platform as a software?". Our task is to develop an AI-based solution for automatically generating valid user development requirements (Crowd Requirements Engineering - CrowdRE). Your task is to annotate texts with their intent (bug report or new development requests) and the topic covered by the text using these guidelines.

Annotation is done at the sentence level. One sentence can contain one topic, each tied to one intent. Sentences are derived from the text. Along with the sentences, the corresponding texts (responses/comments) are provided. The guidelines include instructions for identifying the intent and topic defined by the annotation model. Instructions are given in the form of explanations and examples. The annotation process is shown in Figure 1.

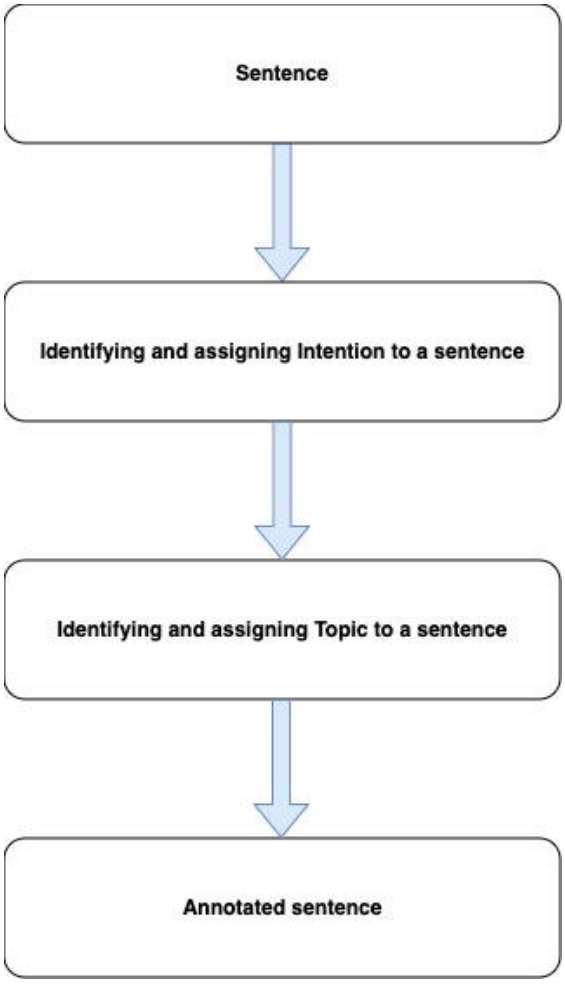

Figure 1. Annotation process

1. Annotation steps and general tips

1. Read the sentence carefully, **multiple times** if necessary.
2. Analyze the sentence to determine the writer's intention and assign it **one** of the predefined values for the **Intention** attribute described in section **2. Intention**.
3. Analyze the sentence to determine the topic described and assign it **one** of the predefined values for the **Topic** attribute described in section **3. Topic**.
4. Intention and topic can be expressed **explicitly or implicitly**. **Explicit** writing conveys information directly and precisely, making it easy to understand. **Implicit** writing is indirect and suggestive, allowing for multiple interpretations. Implicitness requires the annotator to infer and fill gaps to understand the author's intended meaning.
5. If intention or topic cannot be assigned to a sentence without considering the **context**, annotators are advised to look at other sentences that make up the response and then assign appropriate attributes to the annotated sentence as described in section **4. Context**.

2. Intention

The **Intention** attribute represents the intention or goal of the comment. Intention categories include *Bug report*, *Feature request*, and *Rating*. For sentences where the intention cannot be determined from the content, the value *Other intention* is defined. The hierarchy of intentions is shown in Figure 2. Individual intentions are explained and illustrated below.

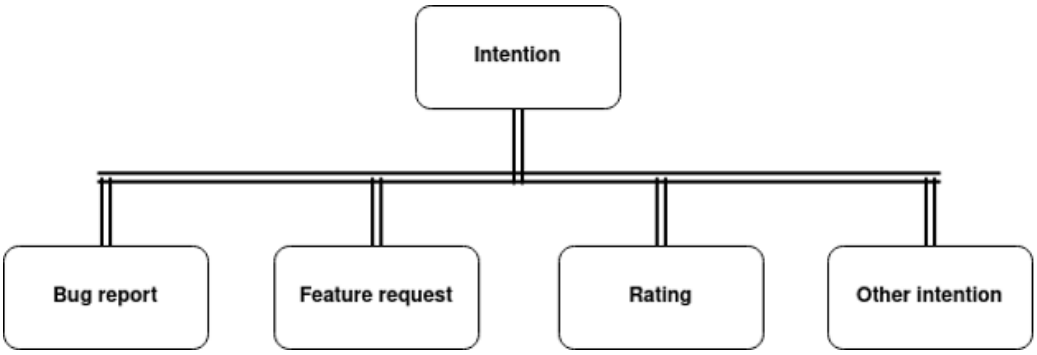

Figure 2. **Intention** attribute values

Bug report aims to inform the developer about a problem or defect in the entire software or any of its functionalities that the user/writer has encountered. Examples of sentences with the intention of *Bug report*:

Sentence	Intention	Topic
I can't proceed to the next task.	Bug report	Reliability
The notes button is overlaid while the night mode button is somehow above the picture (in a white box??).	Bug report	User Interface

A **Feature request** represents the user's/writer's request to add new functionalities or content to the software or to remove, change, or improve existing ones. Examples of sentences with the intention of *Feature request*:

Sentence	Intention	Topic
Increase the font size of certain text parts.	Feature request	User Interface
The answers to questions should be a bit less abstract	Feature request	Assessment item

Rating aims for the user/writer to express general satisfaction or dissatisfaction with the application. It focuses on general judgment and includes praises, derogations, and criticisms. Examples of sentences with the intention of *Rating*:

Sentence	Intention	Topic
The dark theme is great.	Rating	User Interface
I would delete it	Rating	Whole App

3. Topic

The **Topic** attribute assesses the software product, its aspects, and the context as specific topics on which the user/writer shares their opinion. Basic topics include the *Whole App*, its *Content*, *Usability*, *Reliability*, *Compatibility*, and *User Interface*. The *Content* topic includes two categories: *Instructional item* and *Assessment item*. The hierarchy of topics is shown in Figure 3. The annotator first assigns a lower hierarchical level topic to the sentence. If there is no basis for such annotation, a higher hierarchical-level topic is assigned to the sentence. Individual topics are explained and illustrated below.

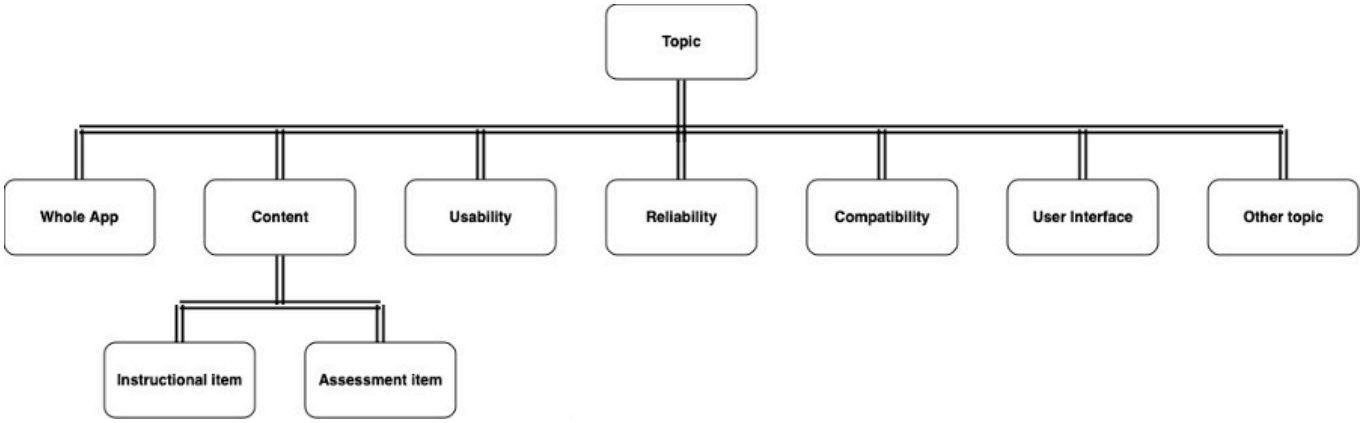

Figure 3. **Topic** attribute values

Whole App refers to the sentences that consider the software as a whole without its functionalities. Examples of sentences with the topic *Whole App*:

Sentence	Topic	Intention
I like the tutor as it is now, maybe I'll have some remarks later during the learning.	Whole App	Rating
It's a fun app to use	Whole App	Rating

Instructional item refers to the instructional materials available in the software - instructional videos, hints, texts, and illustrations. Examples of sentences with the topic *Instructional item*:

Sentence	Topic	Intention
The educational content is very clearly and plainly written and easy to follow.	Instructional item	Rating
There could be some hints on the WEB part as well	Instructional item	Feature request

Assessment item refers to knowledge assessment elements - multiple-choice questions, drag-and-drop tasks, and programming challenges. Examples of sentences with the topic *Assessment item*:

Sentence	Topic	Intention
I would add a quiz that covers all the material learned so far.	Assessment item	Feature request
All tasks are very good and instructive, but there are too many.	Assessment item	Rating

Content refers to all content available through the software that cannot be categorized into one of the previous two categories (*Instructional item* and *Assessment item*). Examples of sentences with the topic *Content*:

Sentence	Topic	Intention
Use of animations as examples of functionalities	Content	Feature request
A little less text	Content	Feature request

Usability refers to the ease with which users can learn and use the software. This includes intuitiveness, efficiency, ease of understanding, navigation, and completing tasks. Examples of sentences with the topic *Usability*:

Sentence	Topic	Intention
Automatic downloading of materials should be possible.	Usability	Feature request
Easier viewing and navigation through lessons should be enabled.	Usability	Feature request

Reliability refers to the software's ability to function correctly and consistently over time. This also includes the software's ability to recover from failures and continue operating without interruption. Examples of sentences with the topic *Reliability*:

Sentence	Topic	Intention
At one point, only 2 out of 4 questions appeared, and they kept repeating.	Reliability	Bug report
The only thing I found wrong is the 101% accuracy.	Reliability	Bug report

Compatibility refers to the software's ability to work with other software. This allows data and functionalities to be shared between different software applications and platforms, enabling efficient communication and information exchange. Examples of sentences with the topic *Compatibility*:

Sentence	Topic	Intention
It should be possible to handle the tutor with other development environments.	Compatibility	Feature request
VS Code and Tutor do not cooperate with each other	Compatibility	Bug report

User Interface is the point of interaction between the user and the software. It is how users interact with the software and control it, as well as how the software presents information and feedback to the user. The user interface includes graphical elements such as icons, buttons, and menus, as well as text-based commands that the user enters to interact with the software. Examples of sentences with the topic *User Interface*:

Sentence	Topic	Intention
In some places, Tutor is written with a capital letter, and in others with a lowercase letter.	User Interface	Bug report
I like the dark theme very much, but I don’t like the blue color of the side menu.	User Interface	Rating

4. Context

For sentences where intention or topic cannot be assigned without considering the context, it is necessary to look at the other sentences that make up the comment. Examples of sentences whose intention or topic cannot be accurately determined without considering the context:

Sentence	I would change that color to another one :)
The entire comment	I like the dark theme very much, but I don’t like the blue color of the side menu. I would change that color to another one :)
Intention	Feature request
Topic	User Interface

Sentence	Everything else is fine.
The entire comment	The only thing I found wrong is the 101% accuracy. Everything else is fine.
Intention	Rating
Topic	Whole App