

# HEUROBOX

## Heuristic Robots Experience

**Expert Evaluation for Human Robot Interaction**

**Mondragon University – Faculty of Engineering**

**Design Innovation Center (DBZ)**

## Introduction

We conducted a systematic literature review to identify methodologies that evaluate the human–robot interaction (HRI) from a human-centred approach. The review provided insight into how evaluations are conducted in HRI. The results showed the most evaluated factors and how they are measured considering different types of measurements: qualitative and quantitative, objective and subjective. No validated model to assess UX in HRI was identified.

In a study on “User eXperience Evaluation Methods” by Väänänen-Vainio-Mattila et al. (2008), they proposed a set of requirements for good UX evaluation in industrial environments. Although they stated that it is not possible to have a single method that meets all the requirements because some of them may be contradictory or even unrealistic, heuristic evaluation meets most of them. It is a common technique based on evaluators’ criticism for quickly identifying design problems (Molich & Nielsen, 1990) because of its simplicity, low cost, and wide applicability. The iterative use of heuristics is key to improve HRI workspaces (Clarkson & Arkin, 2007). Although there are several lists of heuristics to evaluate robotics or HMIs from different points of view, no list of heuristics has been identified in the literature that is specific to industrial HRI and considers i) aspects of safety, trust and perceived safety, ii) physical ergonomics, iii) cognitive ergonomics and emotions, iv) inclusivity, v) robot types, and vi) robot type and functionalities. In HRI, the robot's behavior influences the safety, comfort, and acceptance of the person in the robotic system

Therefore, we make a contribution in the field by developing a new set of heuristics to evaluate the UX in HRI.

HEUROBOX will be complete and acquire more scientific contribution by taking your expertise into account through your evaluation of HEUROBOX. This evaluation can be contributed into two folds:

1. Your evaluation of the general usability of HEUROBOX in terms of readership and practice
2. Your possible comments on HEUROBOX

The expected time to complete this survey is as below:

- Read and understand the explanation of HEUROBOX (section 2): 30 minutes
- Complete evaluation (section 3): 30 minutes
- **Total expected time: 60 minutes**

Your time and expertise are valuable to our research in particular and the research community in general. Lastly, we strongly appreciate your acceptance of this contribution.

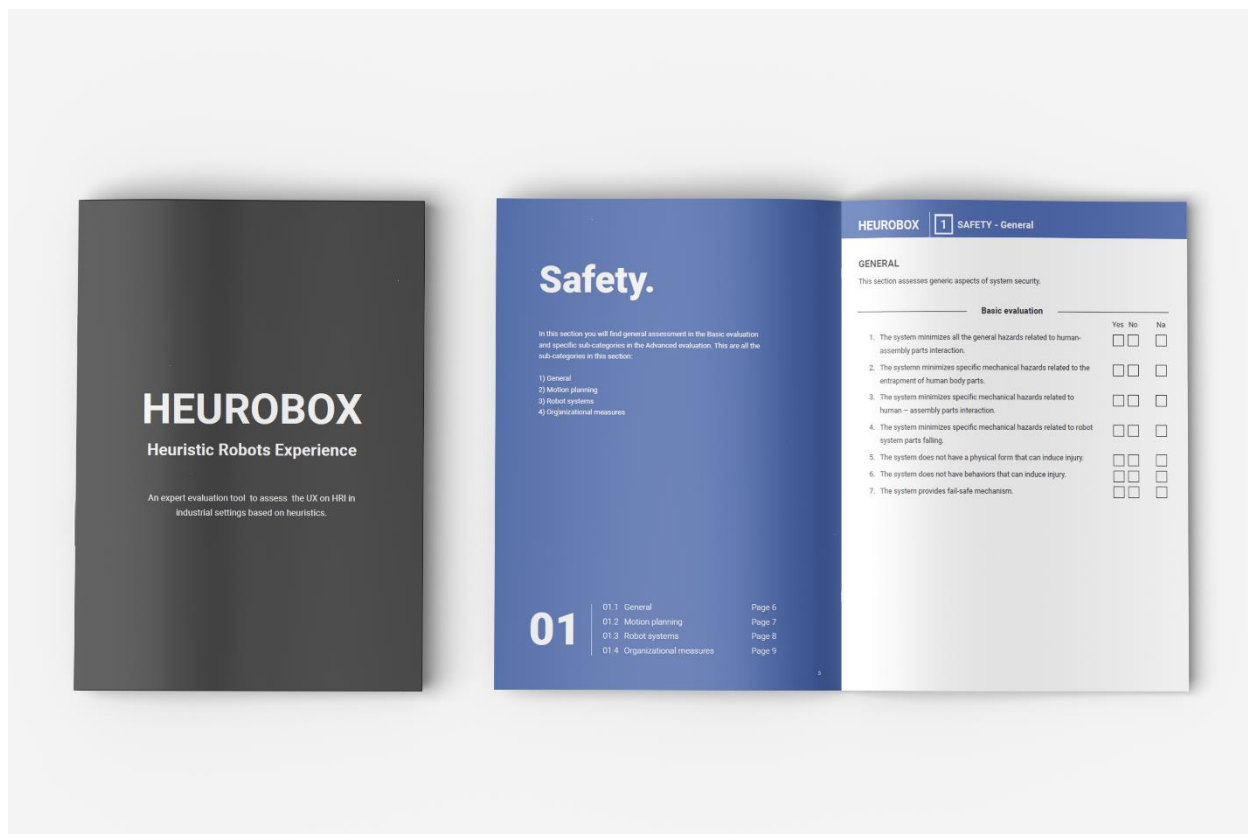
The following section explains HEUROBOX and its working mechanism. The last section is a full part of the expert evaluation on (i) HEUROBOX and (ii) the usability of the heuristic evaluation on HRI.

# Heuristic evaluation

## What does it for?

Heuristic evaluation is a method of usability evaluation based on the inspection of different sections. In short, it consists of checking the quality of a set of principles called heuristic principles, i.e., checklist items. It is an agile and low resources method, which is known to detect usability errors<sup>1</sup>.

Thus, the evaluation tool presented below is based on the usage of a heuristic evaluation, proposing a series of principles to be analysed, obtained from various studies and own experience.



<sup>1</sup> Relative percentage to 4-5 evaluators, How to Conduct a Heuristic Evaluation, Jakob Nielsen, 1995.

## How is it built?

The evaluation template consists of four categories:

Category	Subcategory	Definition
Safety	General	Assesses generic aspects of system security.
	Motion planning	Assesses Motion planning aspects of system security.
	Robot systems	Analyses the characteristics of the robot system and workspaces.
	Organizational measures	Analyses the organizational measures to ensure safe conditions.
Ergonomics	Physical ergonomics	Concerns with human anatomical, anthropometric, physiological, and biomechanical characteristics as they relate to physical activity (International Ergonomics Association, 2019).
	Cognitive ergonomics	Focuses on how well the use of a product matches the cognitive capabilities of users. It draws on knowledge of human perception, mental processing, and memory.
Functionality	System	Covers the usability of the system in terms of functions available.
	Information	Defines the assessment of how system information is presented to users.
	Task	Based on the efficiency of the task and how well it is implemented.
	Error handling	Assess the actions to capture, discover and recover from an error in the system.
	Assistance	Defines documentation or help (from to the system) methodologies.
Interface	General	Involves general evaluation of the interface.
	Visual	Analyses the interface based on information shown on displays or screens.
	Gesture	Accounts for gesture inputs to the system.
	Haptic	Assess the system output through touch interactions.
	Voice	Evaluates the information presented and transmitted through sounds and voice.

In addition, in each category, the basic level and the advanced level are differentiated. The basic level is intended to cover the essential aspects of an HRI assessment. In total, it includes 84 heuristic principles. The advanced level encompasses aspects of specific elements or functions, amounting to a total of 228 (including the ones of the basic level) heuristic principles.

Category	Subcategory	Total Nº of heuristic	Nº of heuristics in the Basic evaluation	Nº of heuristics in the Advanced evaluation
<b>Safety</b>		<b>30</b>	<b>7</b>	<b>23</b>
	General	7	7	0
	Motion planning	10	0	10
	Robot systems	9	0	9
	Organizational measures	4	0	4
<b>Ergonomics</b>		<b>49</b>	<b>13</b>	<b>36</b>
	Physical ergonomics	29	4	25
	Cognitive ergonomics	20	9	11
<b>Functionality</b>		<b>78</b>	<b>36</b>	<b>42</b>
	System	18	6	12
	Information	28	13	15
	Task	12	7	5
	Error handling	14	7	7
	Assistance	6	3	3
<b>Interfaces</b>		<b>71</b>	<b>28</b>	<b>43</b>
	General	28	28	0
	Visual	22	0	22
	Voice	11	0	11
	Haptic	6	0	6
	Gesture	4	0	4
<b>TOTAL</b>		<b>228</b>	<b>84</b>	<b>144</b>

### General structure

Before starting the evaluation, the user has to fill in the technical datasheet. Here the information about the robot itself and the robotic system is collected, in order to know how the human-robot relationship is (coexistence, cooperation, collaboration...). Finally, there are three tasks that the evaluator has to perform before starting the evaluation.

HEUROBOX 0 TECHNICAL DATASHEET

Date:

Robot brand:

Robot model:

Year of robot:

	Yes	No
The person and the robot have a shared workspace.	<input type="checkbox"/>	<input type="checkbox"/>
The person and the robot have direct contact.	<input type="checkbox"/>	<input type="checkbox"/>
The person and the robot have the same task.	<input type="checkbox"/>	<input type="checkbox"/>
The person and the robot have simultaneous process.	<input type="checkbox"/>	<input type="checkbox"/>
The person and the robot have sequential process.	<input type="checkbox"/>	<input type="checkbox"/>

Choose the types of interface the robot has:

☐ Visual ☐ Voice ☐ Haptic ☐ Gesture

Choose the context the robot has:

☐ Laboratory ☐ Real

Before starting the assessment, do the following tasks:

☐ Make an observation of the workspace.

☐ Execute all robot functionalities.

☐ Make an observation during task execution.

4

Then, the structure of each section follows a classification system of heuristics. Thus, there are general principles related to each section, items that correspond to the basic level and items that correspond to the advanced level. Finally, there is an area for scoring and comments. It should be noted that for each item there are three boxes: “yes”, “no” and “not applicable” (in case the item is not applicable or does not correspond to the robotic system in question). The items have been proposed in such a way that the appropriate response is positive. So, after all the checks have been carried out, the answers marked “yes” will correspond to the number of items that the robotic system adequately fulfills, while the “no” answers will correspond to aspects that need to be improved.

Heading to locate the user -  
**Category and subcategory**

Subcategory and brief description of what is to be assessed

Heuristics  
**Basic evaluation**

HEUROBOX 2 ERGONOMICS - Cognitive ergonomics

HEUROBOX 2 ERGONOMICS - Cognitive ergonomics

**COGNITIVE ERGONOMICS**  
Cognitive ergonomics focuses on how well the use of a product matches the cognitive capabilities of users. It draws on knowledge of human perception, mental processing, and memory.

**Basic evaluation**

	Yes	No	Na
1. The system makes components well identifiable and distinguishable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system makes components easy to orient, locate and find.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system makes the work intuitive (supports the formation of a mental model, reduces the choice reaction time, facilitates the learning transfer, promotes similarity).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system maximizes the user's trust.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The UI is designed with ergonomics and accessibility in mind, ensuring the robot UI is comfortable to work with for the necessary duration.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The system maximizes operator psychological wellbeing and satisfaction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The system avoids high speed motions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The robot is as slow as possible.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The system informs operators about the robot speed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Advanced evaluation**

	Yes	No	Na
10. The system makes menu items that need attention visually salient, do not attract attention unnecessarily.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The system provides fused sensor information to avoid making the user lose the data instantly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

14

12. The system supports the user in understanding the connection between user actions and system responses, for instance by providing feedback and using appropriate terminology.

13. The system ensures the robot performs in accordance with polite social etiquette.

14. The system provides fused sensor information to lower the cognitive load on user.

15. The system reduces the number of assembly sub-systems.

16. The system allows an immersive operator experience: sense of "being there".

17. The system implements smooth trajectories (which can be associated to natural human arm motions).

18. The system implements swing trajectories (not continuously straight).

19. The system involves operators into the definition of layout and work activities.

20. The system avoids misalignment in operator and robot use of production resources (avoid inefficiency).

Yes No Na

☐ ☐ ☐

☐ ☐ ☐

☐ ☐ ☐

☐ ☐ ☐

☐ ☐ ☐

☐ ☐ ☐

☐ ☐ ☐

☐ ☐ ☐

☐ ☐ ☐

15

Heuristics  
**Advanced evaluation**

At the end of each category there is a section to count how many heuristics the evaluated system does and does not meet.

HEUROBOX | 2 ERGONOMICS - Summary

02 ERGONOMICS

02.2 Physical Ergonomics: \_\_\_ / \_\_\_

02.2 Cognitive Ergonomics: \_\_\_ / \_\_\_

TOTAL SCORE: \_\_\_ / \_\_\_

Comments:

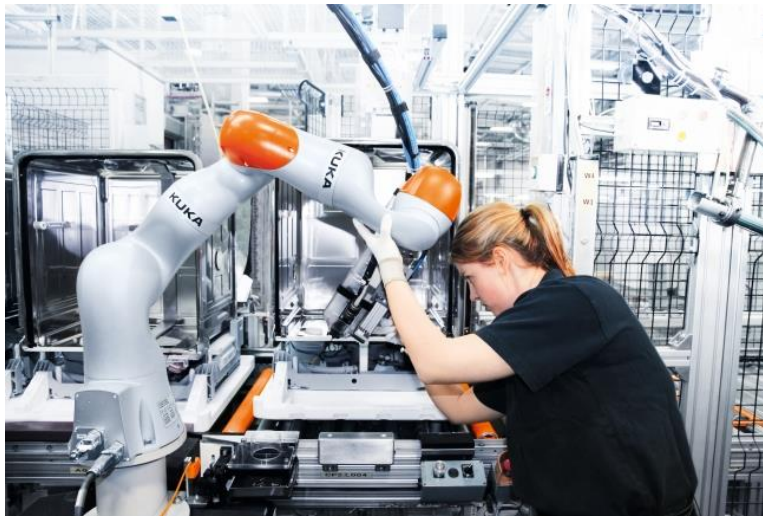
16

## How does it work?

1. Place the evaluator in context: The evaluator should inquire about the profile of the user and the function that the robot has, as this is the basis for a good evaluation.
2. Previous exercises: It is advisable to carry out various activities such as executing tasks.
3. Execution of the evaluation: Once you have chosen the level of evaluation, proceed with it.
4. Analysis of the results.

## An implementation example

For example, in the context shown in the figure below, we would fill HEUROBOX as follows.



To begin with, we fill in the technical datasheet.

HEUROBOX | 0 TECHNICAL DATASHEET

Date: 27/10/2022  
Robot brand: KUKA  
Robot model: LBR iiwa  
Year of robot: 2021

	Yes	No
The person and the robot have a shared workspace.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
The person and the robot have direct contact.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
The person and the robot have the same task.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
The person and the robot have simultaneous process.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
The person and the robot have sequential process.	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Choose the types of interface the robot has:

☐ Visual ☒ Voice ☒ Haptic ☐ Gesture

Choose the context the robot has:

☐ Laboratory ☒ Real

Before starting the assessment, do the following tasks:

☒ Make an observation of the workspace.  
☒ Execute all robot functionalities.  
☒ Make an observation during task execution.

4

Then, we go point by point evaluating the system, while indicating whether the system complies with the heuristic. To do this, in the squares on the right, the user fills in the corresponding square, "Yes", "No" or not applicable "Na".



HEUROBOX

2

ERGONOMICS - Cognitive ergonomics

COGNITIVE ERGONOMICS

Cognitive Ergonomics focuses on how well the use of a product matches the cognitive capabilities of users. It draws on knowledge of human perception, mental processing, and memory.

Basic evaluation

	Yes	No	Na
1. The system makes components well identifiable and distinguishable.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system makes components easy to orient, locate and fasten.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3. The system makes the work intuitive (supports the formation of a mental model, reduces the choice reaction time, facilitates the learning transfer, promotes similarity).	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system maximizes the user's trust.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The UI is designed with ergonomics and accessibility in mind, ensuring the robot UI is comfortable to work with for the necessary duration.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
6. The system maximizes operator psychological wellbeing and satisfaction.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
7. The system avoids high-speed motions.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The robot is as low as possible.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The system informs operators about the robot speed.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Advanced evaluation

	Yes	No	Na
10. The system makes menu items that need attention visually salient, do not attract attention unnecessarily.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
11. The system provides fused sensor information to avoid making the user fuse the data mentally.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

14

HEUROBOX

2

ERGONOMICS - Cognitive ergonomics

15

Once each category has been completed, the indicators obtained are summarised in the Summary section at the end of each category. In this example, physical ergonomics has obtained a score of 15 out of 26. This indicates that a total of 26 items have been evaluated, of which 15 are satisfactorily fulfilled, and therefore, there are still another 11 heuristics to be improved.

In addition, there is a space to add up the total of the categories and a section where comments can be made.

HEUROBOX

2

ERGONOMICS - Summary

02 ERGONOMICS

02.2 Physical Ergonomics: 15 / 26

02.2 Cognitive Ergonomics: 12 / 19

TOTAL SCORE: 27 / 45

Comments:

We need to improve the indications. Could we add symbols?

The posture of the workers needs to be improved.

16

## Expert evaluation on heuristic list

### General assessment

We use the Simplified System Usability Scale (SUS)<sup>2</sup> to evaluate the usability of HEUROBOX. We really appreciate your evaluation of HEUROBOX with the following questions:

#### QUESTIONS

**ASSESSMENT SCALES:** STRONGLY DISAGREE (1); DISAGREE (2); MODERATELY (3); AGREE (4); STRONGLY AGREE (5)

**(TICK "X" ON THE EVALUATION SCALE)**

**ASSESSMENT**

I WOULD USE "HEUROBOX".	
"HEUROBOX" WAS TOO COMPLEX FOR ME.	
"HEUROBOX" WAS EASY TO USE.	
I REALLY NEED HELP FROM SOMEONE TO USE "HEUROBOX".	
THE VARIOUS PARTS OF "HEUROBOX" WERE WELL INTEGRATED.	
"HEUROBOX" WAS CONFUSING FOR ME.	
LEARNING TO USE "HEUROBOX" WAS QUICK FOR ME.	
"HEUROBOX" WAS HARD TO USE.	
I FELT CONFIDENT USING "HEUROBOX".	
I WILL NEED TO LEARN A LOT BEFORE USING "HEUROBOX".	

### Assessment on categories

We defined 6 categories to be evaluated using the new heuristic list. Hence, the following survey is useful in practice and research by answering relevant research questions, including:

- WHICH category is more important?
- In which categories should time and resources be invested?

Therefore, we use AHP<sup>3</sup> (analytical hierarchy process) to design and analyse the following survey with the 9-point evaluation scale on each side of pairwise comparison. The fundamental question is:

- **How much a criteria category is more important to another category in the following 9-point scale evaluations:**

<sup>2</sup> Holden, R. J. (2020). A Simplified System Usability Scale (SUS) for Cognitively Impaired and Older Adults. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 9(1), pp. 180–182.

<sup>3</sup> Al-Harbi, K. M. A.-S. (2001). Application of the AHP in project management. *International Journal of Project Management*, 19(1), pp. 19–27.

Rating	Journal
1	A category is equally important
3	A category is slightly preferred than another the category on the other side
5	A category is moderately preferred than another category on the other side
7	A category is strongly preferred than another category on the other side
9	A category is absolutely preferred than another category on the other side

**For example:**

Which of following criteria category prefers to be used for “Evaluation and Testing Human-Robot Interaction” more than the other criteria category on the same comparison row?

(tick “X” ON THE EVALUATION SCALE)

A category is preferred than the left-hand category									Equally	A category is preferred than the right-hand category								
	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
Safety			x															Ergonomics

- the tick “x”—of 7 on the left—indicates that, according to an expert’s perspective, “Safety” should be strongly more important than “Ergonomics” to evaluate and test Human-Robot Interaction.

**Another example:**

Which of following criteria category prefers to be used for “Evaluation and Testing Human-Robot Interaction” more than the other criteria category on the same comparison row?

(tick “X” ON THE EVALUATION SCALE)

A category is preferred than the left-hand category									Equally	A category is preferred than the right-hand category								
	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
Functionality									x									Assistance

- the tick “x” of 1 above indicates that, according to an expert’s perspective, both “Functionality” and “Assistance” are equally important to evaluate and test Human-Robot Interaction.

The following evaluation questionnaire is designed to evaluate which category is more important:

Which of following criteria category prefers to be used for “Evaluation and Testing Human-Robot Interaction” more than the other criteria category on the same comparison row?

(tick "X" ON THE EVALUATION SCALE)

[illegible]

Which of following criteria category prefers to be used for “Safety” more than the other criteria category on the same comparison row?

(tick "X" ON THE EVALUATION SCALE)

A category is preferred than the left-hand category									Equally	A category is preferred than the right-hand category								
	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
Motion planning																		Robot systems
Motion planning																		Organizational measures
Robot systems																		Organizational measures

(tick "X" ON THE EVALUATION SCALE)

A category is preferred than the left-hand category									Equally	A category is preferred than the right-hand category								
	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
Physical ergonomics																		Cognitive ergonomics

(tick "X" ON THE EVALUATION SCALE)

[illegible]

## Information

## Assistance

[illegible]

Which of following criteria category prefers to be used for “Interfaces” more than the other criteria category on the same comparison row?

(tick “X” ON THE EVALUATION SCALE)

A category is preferred than the left-hand category									Equally	A category is preferred than the right-hand category								
	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
Visual																		Voice
Visual																		Haptic
Visual																		Gesture
Voice																		Haptic
Voice																		Gesture
Haptic																		Gesture

## Possible comments and advice

We appreciate your comments and advice that can be written here:

---



---

## Thank you

We strongly thank for your contribution. Your responses to our survey will be kept as anonymity and we also update to you the analysis result in convenient time.

If you have any questions or discussions, we are willing to arrange a meeting for a discussion. We are looking forward to your responses that can be sent to us the following email:

- [aapraizi@mondragon.edu](mailto:aapraizi@mondragon.edu); [jose.alberola@stiima.cnr.it](mailto:jose.alberola@stiima.cnr.it); [glasa@mondragon.edu](mailto:glasa@mondragon.edu); [mmazmela@mondragon.edu](mailto:mmazmela@mondragon.edu)

Regards,

Ainhoa Apraiz Iriarte