Expectations of train drivers for innovative driving cabin

Jean-Valentin Merlevede, Simon Enjalbert, Frédéric Henon, Alexandre Pereda Baños, Stefano Ricci, Frédéric Vanderhaegen

Expectations of train drivers for innovative driving cabin

Jean-Valentin Merlevede*, Simon Enjalbert*, Frédéric Henon, Alexandre Pereda Baños***, Stefano Ricci****, Frédéric Vanderhaegen***

**Univ. Polytechnique Hauts-de-France, CNRS, UMR 8201 - LAMIH, F-59313, Valenciennes, France, (e-mail: simon.enjalbert@uphf.fr)*

*** UIC - International union of railways, 16 Rue Jean Rey, 75015 Paris, France*

**** Eurecat, Centro Tecnologico de Cataluna,08290 Cerdanyola del Vallès, Spain*

***** Dipartimento di Ingegneria Civile Edile e Ambientale, Sapienza Università di Roma, 00185 Roma, Italy*

Abstract: This paper aims at identifying the expectations of train drivers or other railway staff about Human-Machine Systems (HMS) in future cabins. The identification of the best technical solution needs surveying preferences and efficiencies of possible new information technology configurations of Human-Machine Interfaces (HMI) considering human factors as input and output sensors. Technical recommendations about the train cabin of the future are provided. They consider results from a state-of-the-art on HMI in transport systems, from technology maturity issues, and from two large scale surveys realized during project. Recommendations are then proposed to train manufacturers for deeper investigation or for innovative driving cabin implementation.

Keywords: Human-Machine Systems, Rail Transport, Drivers-Cabin Interfaces, Information Technologies, Large scale survey.

1. INTRODUCTION

In this paper, authors present activities realized during CARBODIN project which has received funding from the European Union's Horizon 2020 research as part of Shift2Rail initiative. One of the project objectives was to look at train drivers or railway staff preferences that can be an interesting way of exploring how interaction design of rail operations might be improved. The goal of the depicted work consists in proposing recommendations for future train cabin based on the results of a state-of-the-art about innovative Human-Machine Systems (HMS) in transport domain and on two experimentations involving more than 1700 professional railway staffs and drivers from 15 European countries.

First of all, a state-of-the-art regarding all technologies already existing and implemented in transports industries, or studied in simulators in research laboratories, has been realized (Enjalbert et al., 2021a, 2021b) in order to identify technologies to be tested on surveys. Already matures technologies are considered but are less interesting with regard to the CARBODIN project purpose study: it consists in determining expectations about new HMI devices on future train cabin by considering technologies that are rarely implemented by now in rail industry. Consecutively, the objective of the first experimentation was to collect a large number of contributions from drivers and staffs all over Europe. So, it was decided to organize an online survey and to collect the maximum feedback from people in railway transport services from different countries in Europe. Results about their statistical analysis supported the design of the second survey that included a driving phase with train drivers on a simulator. This second experimental protocol aimed to test gesture and voice control devices and to have feedback about relationships between groups regarding new sounds and driving tasks. Drivers were invited to answer to two lists of questions: same questions from the survey done on the first experimentation and other questions from a second survey about their feelings on gesture and voice control devices and on sounds that may be introduced in future cabin train.

With regard to results from both experimentation and surveys, guidance on Human-Machine Interfaces (HMIs) to be included in future train cabins are retained. It reminds conclusions from actual technologies already in use in transports, and from results of both experimental protocols. Finally, it gives a list of technical recommendations for future train cabins.

2. FIRST EXPERIMENTATION

2.1 Survey implementation

The first round of experimentation had two main goals. The first objective was to collect feedback from drivers on the technologies and sounds they might encounter in their cabin. A sufficient number of people were expected in order to make statistical analysis possible with dedicated tools. The second goal was to select technologies for the second experimentation.

In order to obtain meaningful statistics, a maximum number of answers was needed during the first experimentation. To achieve this high completion rate, a short online survey (called survey 1 or SV1) was designed. It should not take more than 20 minutes to be completed. The survey was open from February 2021 to April 2021 Train drivers were targeted, but the survey was also open to other professional railway staff. Participants had to answer to relevant information that have been categorized later by statistical analysis with regard to professional groups. The investigation of the first survey concerns feedback about the use of sounds, touch screen, gesture, visual and haptic technologies. Four input solutions and four output solutions are assessable by using these technologies, Table 1.

Table 1. List of technologies for survey 1

Technologies	Action(input)	Information(output)
Audio	Voice control	Audible notification
Touch Screen	Tactile Control	Screen reading
Gesture	Gesture recognition	N/A
Visual	N/A	Head-Up Display (HUD)
Haptical	Manipulator	Feedback force

Yet haptic technology is mainly relevant regarding the field of speed regulation in the driving cabin of a train. Thereby three input solutions and three output solutions assessable for all tasks and one additional input solution and one additional output solution for the speed regulation task. Each technology is described by a short explanatory text and a visual support in the "Presentation of technologies" section of the survey. The visual support was a pictogram or video clip. As a matter of fact, the respondents had at their disposal all the elements to understand the proposed technologies before starting the survey. Various technologies are being studied for integration into the train cabin of the future. These technologies can be applied to the means of Action that drivers would have to act on their train, but also to the means of Information that would enable them to keep informed about the train state. Possible relationships between technology and Action or Information are as follows:

- AUDIO: Voice control of the system (Action) and audible notification (Information)
- TOUCH SCREEN: Use of a touch screen (smartphone, tablet, etc.) for entering commands (Action) and reading information (Information).
- GESTURE: Gestures recognition technologies allow to act on the system without contact (Action)
- VISUAL: A system allows information to be projected into the driver's field of vision. Thus, the desired information can be accessed through a Head-Up Display (HUD), without leaving the lane of sight (Information).
- HAPTICAL: Haptic technology allows you to use the manipulator you are used to and adds force feedback functions. The manipulator is then able to assist the driver; accompanying or resisting the driver's movements. It is therefore used by the driver to control traction (Action) and to be informed of his speed (Information).

Due to the huge number of actions or information regularly performed or checked during railroading by drivers, a functional analysis was realized in order to define categories and groups. Categories were organized according to the criticality and time constraint of the interaction, and groups were defined according to similar interactions in each category. The identified categories were as follows:

- Category 1 includes the KEY elements: Actions and/or information having a high impact on the safety of the train and passengers with a high time factor.
- Category 2 contains AUXILIARY systems: Actions and/or information related to driving but which are carried out either at a standstill or in specific but limited circumstances.
- Category 3 includes SUPPORT devices: Actions and/or information related to the comfort, putting into service or parking of the train.

Then 3 groups were created in each category (Table 2), bringing the total number of groups to 9. Each group is assessed in terms of information and action by using the same layout for each group:

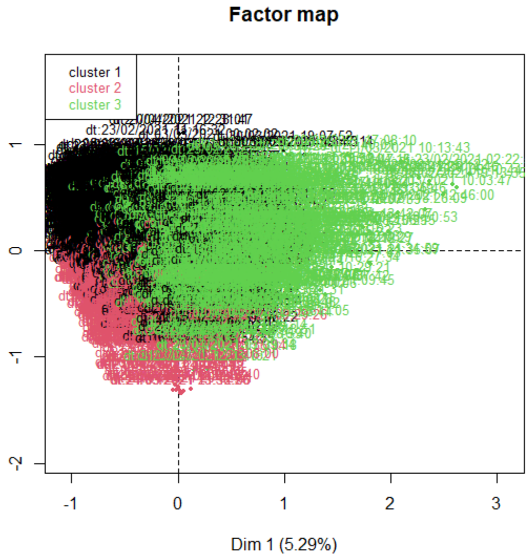
- How relevant would you consider the following technologies for: Action on "group's name"?
- How relevant would you consider the following technologies for: Information on "group's name"?

To facilitate survey understanding, each group of actions was also accompanied by a short explanatory text. Regarding the 18 questions, participants had to select the correct evaluation of their own feeling on a 5-degree scale from 0 "Very unsuitable" to 4 "Very suitable". Of course, there were no limitation in participants answers and they could evaluate all technologies as very suitable (important to be included in next generation cabin design) or all unsuitable.

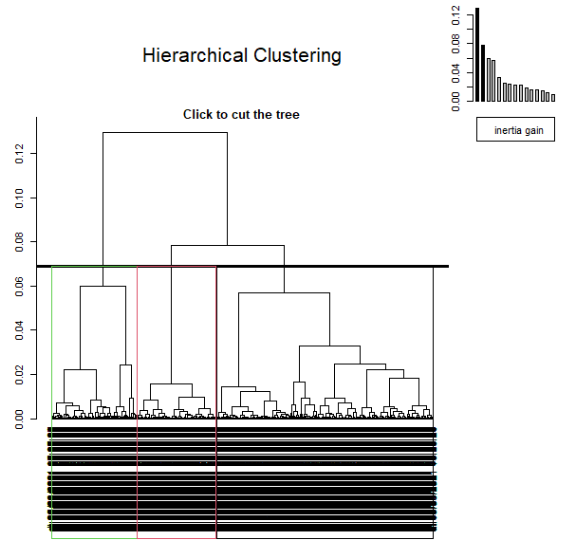
2.2 Methodology

Due to the large number of respondent (more than 1700 from 15 different European countries, project data available on demand), statistical analysis tools are required to determine relevant tendencies on the results by considering data like the age of drivers, the homeland of the drivers, their living country, or their experience, their perception of sounds or of technologies. The implementation of these tools is done with R. It is facilitated by R studio. Multiple Correspondence Analysis (MCA) and Hierarchical Clustering on Principal Components (HCPC) are data analysis techniques. The first one aims to find links between several variables. It identifies independent groups of data by considering respondent profile (age, experience, job, nationality) and scores given for sounds and technologies assessment. It prepares data for the second analysis technique. The second technique aims to determine clusters that link members of the same group and differentiate them from other groups. It will reveal which respondents are the most representative and the most divergent of each group with regard to factors from data analysis for each cluster. It also sorts the different groups according to their size.

These tools will create groups in our pool of participants. To do so they will use the answers to the survey question to place each participant in a multidimensional space. Closer are the answers, shorter is the distance between the point. Next it will start making groups, for example 3 groups for KEY systems analysis visible in factor map (Fig. 1) or in hierarchical clustering (Fig. 2). It will find the best breakdown. where for each group, groups member is close to each other and distant from any no group members. After it will start to look if there are any criteria which could be typical of the created groups. It will look it will look for criteria in the group whose distribution is different in the group compared to its distribution in the whole population. If there is a relevant breakdown, it is going to find it, but it will also try to create and add a meaning to unrevealing breakdown. Moreover, each criterion is studied individually.



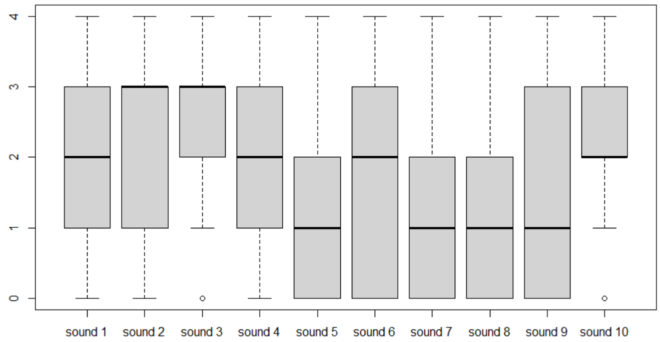
Factor map for KEY systems Category during the experimentation 1 (Fig. 1).



Hierarchical clustering for KEY systems Category during the experimentation 1 (Fig. 2).

For example, for technologies, a group whose criteria are Italian, trainers do not necessarily include Italian trainers, but Italians who can practice any profession and trainers of all nationalities. If a group comes up regularly during the grouping and is associated with a different response from the mass of respondents, then the results must be checked by looking in the clusters where it appeared if the individuals who correspond to the criterion have indeed responded differently from the rest of the participants.

In addition to technologies to be tested regarding group of tasks, different sounds have been tested during first experimentation and classified into three main appreciation levels according to their level of acceptance: the appreciated ones, the rejected ones and the so-so in the middle. A 5-degree scale from 0 "Not at all" to 4 "Absolutely" was used to determine how enjoyable were sounds. An example of box plotting is given for the first experimentation (Fig. 3). Finally, participants have been asked if the sounds should be dedicated to a category of sounds in particular between the 6 proposed (Action, Alarm, Driving Mode, Home, Notification, None). If 50% of respondent identified the same category, sound will be retained for the next step of implementation.



Level of enjoyable sounds box plotting for survey 1 (Fig. 3).

There is no significant group that emerge from the MCA and HCPC for both sounds and technologies. This result means two things. Firstly, it seems that sound design in train cabin and the use of new technology depending on tasks will be universal for European people, and secondly the factor influencing sound design or technologies to implement was not in the spectrum of data collected during first experimentation.

3. SECOND EXPERIMENTATION

This second experimentation is a follow-up to the first one and uses the results to determine the relevant technologies to be used for future train cabin. It aims to test and assess innovative HMI technology by drivers. A three-steps experimentation was designed. In the first stage, participants will have to fulfil the first survey from experimentation 1 (SV1) then pass an experiment (Fig.4) and lastly, they will have to answer a second survey (called survey 2, SV2) based on the same 5-degrees scale depicted on SV1 but only for technologies tested during designed experimental protocol. The purpose of this protocol is to verify the relevance of the European sample by comparing the SV1 result during Experimentation 1 and those obtained during Experimentation 2 and to measure if differences exist in the responses after using a technology during a test protocol on simulator by comparing results from the Survey 1 vs Survey 2 both surveys coming from experimentation 2. Because there were no significant group related to nationality, 20 French drivers (due to covid-19 pandemic limitations) that had not taken part into first experiment has participated to the second one.

Train Sim World 2 was used for the experimental part of experimentation 2. It is a railway simulation video game launched on August 20, 2020. It provides a sufficient level of details to allow a good immersion and it is particularly fast to implement. RailDriver is a Desktop Train Cab Controller which was designed to evoke an operation panel in the cabin of a locomotive. It features throttle, brake, reverser, and switch controls, plus 34 programmable buttons. In conjunction with Train Sim World 2, it should establish a high level of immersion for drivers. Some of the train driving tasks are not triggerable from this support. A numpad was used to allow drivers to achieve these tasks when needed. The Leap Motion device is used to implement gesture recognition. The Leap Motion tracks users' hand in space, allocating the tracking data process to GameWAVE. With the tracking information, GameWAVE runs AutoHotkey scripts when it detects the corresponding gesture. AutoHotkey scripts simulate keyboard press which activates function in the simulator. Leap Motion is a reliable prototype device, but interferences may occur in degraded conditions due to light or heat levels. Due to the kind of risk of interpretation in such conditions, a wizard of oz approach is used to recover it. Regarding voice capture, Google Assistant is used to implement reliable and fast speech recognition. The full operating protocol is based on IFTTT and Assistant Computer Control. Those features consist in creating text file in a cloud storage area. In order not to introduce any bias during the experimentation because of response delays, lack of reaction and interpretation concerns, voice recognition feedback was implemented to the system. Thus, during a vocal command, the voice system acknowledges to the drivers with the command it understands. This aims to verify the voice command and identify possible errors of recognition and delays.

Experimental protocol of experimentation 2 (Fig. 4).

Experimental protocol has been divided into three parts. The first part is a training time. Drivers can get into the simulation and they should use this time to familiarize to the experimental set-up. A first sub-part aims to familiarize the participants with the hand tracking device and voice input system. learning about the use of gesture control device and the reaction of the system to acknowledge the required command. During the second sub-step, drivers used the simulator while driving the train in a scenario containing 3 stops. During each stop, one interaction way was studied. During the first test, they used the rail driver. During the second stop, they used the Leap Motion. During the last stop, they used the voice command. The second part of the scenario was about testing the gesture and voice control during predefined stops. The drivers had to perform the actions to be tested. Each driver had therefore tested each technology at least 3 times. Most of the time, the actions on the innovative interfaces were carried out during station stops. This way, drivers had the time they needed to find the right command and execute it safely. Finally, the last part consisted in driving by using freely the gesture and voice command. During this part drivers were free to use the technologies when and how they want.

4. RESULTS

4.1 Experimentations results

Regarding the global results, the more relevant technologies are indicated for Action and Information related to categories and groups of tasks (Table 2). When brackets are proposed, this means that a secondary relevant technology is also possible. Divergences of opinion between results of both experimentations are highlighted in grey and are detailed (Table 3). They should lead to new tests and/or decisions from train manufacturers.

Considering sounds, voice control has better results in post-experimentation survey (SV2). First thought was a possible problem with the simulation that was not as noisy as a real train cabin. However, during the second round of experience in Paris, this hypothesis was rejected because the experimental room was really noisy with the air ventilation, and the window open on a busy street. Moreover, voice recognition systems are now quite common and drivers would prefer giving commands using natural speaking. Audio feedback was also developed to let know train divers if the command has been correctly recorded and applied by system. So, the voice command obtained high scores during experimentations. During the tests, the drivers showed a very strong interest in replacing the beeps emitted by the train with voice interaction. Indeed, according to them, voice interactions, although longer, allow to avoid the time of analysis of the beep and the search for its meaning. Moreover, in the second test of sounds, some of them occurred regularly and their meaning was ambiguous for the drivers. This confirms the potential interest for clear and concise vocal messages emitted by the system. Finally, drivers complained about the huge amount of consecutive alarms generating a mental overload and high level of stress.

Gesture control has lower results in Survey 2. It mostly comes from technical problems with Leap Motion technology. Indeed, as gesture control is disruptive, it needs an adaptation time. Leap motion is only one medium, other gesture recognition system can be very different as using arms, full upper body or whole body to trigger commands. However, there is no tangible feedback with this device. In order to improve gesture commands, different solutions can be studied: trigger a pop up on screen, trigger a voiced feedback, display the movements of the hands of the driver, or last but not least create haptic feedback on gestural commands. BMW was used as an example for using voice command, and they also start using gesture for infotainment device.

Table 2. Result of the comparison between EXP1-SV1 et EXP2-SV2

KEY elements			AUXILIARY systems			ANNEXES devices		
Groups	Tasks		Groups	Tasks		Groups	Tasks	
	Action	Information		Action	Information		Action	Information
Situation awareness	Touch (Gesture)	divergence	Station stops	Touch (Gesture)	divergence	Lighting	Touch (gesture)	Screen (HUD)
Speed regulation	Haptic (Gesture)	divergence	Accessories	Touch (Gesture)	divergence	Comfort systems	Touch (gesture)	Screen (HUD)
Flow Management	Touch (Gesture)	divergence	Power management	Touch (Gesture)	divergence	Configuration	Touch (gesture)	Screen (HUD)

Table 3. Divergence cases between EXP1-SV1 et EXP2-SV2

Group	Interaction	SV1	SV2
Situation awareness	Information	Audible (HUD)	HUD
Speed regulation		HUD	HUD (Haptic)
Flow Management		Audible	HUD
Station stops		Equality for the three technologies	HUD
Accessories		Screen HUD	Screen (HUD)
Power management		Screen HUD	HUD (screen)

Human is multitasking but only two hand to move actuators. Nested situation can be improved by Artificial Intelligence (AI) with will choose which data and button to display on touch screen, which data send on HUD and using voice or gesture command when hands are already busy.

4.2 Other general recommendations

As a matter of facts, there is additional and complementary recommendations that have been collected during both experimentations thanks to discussions with participants:

- Some drivers wanted to have a better mutual understanding with their train because willingness to cooperate is linked to the understanding of the system. They wanted more information and a better troubleshooting system. One of the most feared events is of failures that cause degraded mode.
- The drivers have also complained about the driving cabin itself. They asked for some quality-of-life improvement. Mostly they are interested by equipment that becomes common in the automotive field as: more efficient automatic headlight and windshield wipers, a comfortable seat.
- Drivers are willing to use a HUD. HUD information must be carefully picked. Unlike haptics, it is easy to ensure that the information transmitted on a HUD will be well understood. The problem here is the possible overload of the user with unnecessary information. For example, tracks profiles could be included for safety and eco-driving purposes.
- Most drivers wanted to maintain the speed control with a physical manipulator but are strongly interested in haptics. Therefore, the traction manipulator should be enhanced with haptic feedback. Information relayed by haptics must be carefully picked. Haptics could be used for a lot of different combinations of stimuli and actions.
- Drivers want to use touch screen. Touch screen allows information gathering and action at the same time and on the same device. Several parameters can be handled in order to satisfy preferences of drivers: relevance of information, readability, colour, standardization.
- Some drivers asked that the layout of the controls be redesigned, and they would also like to see more consistency in the design of the controls, especially regarding location of some actuators (ergonomics). Indeed, they find that rarely used controls take up too much space on desk or cabin and that commonly used control devices are too small and poorly placed.
- Music appears as concentration improver, and some drivers asked to have radio or music in cab.
- AI oriented Advanced Driver Assistance Systems should be developed to predict relevant actuator to show up on touch screen to reduce workload from the drivers.

5. CONCLUSION

Finally, facing the complexity of systems and situations, drivers ask for the development of more user-friendly HMI. Tactile, haptics, and voice command are the most appreciated innovative interfaces They must be studied and implemented on future train cabin. Gesture and HUD should be studied as they can improve the train driver comfort. Gesture control should not be used on action with high temporal constraints. However, due to current pandemic condition, touchless interactions via gesture and voice recognition device are welcome. Drivers asked for studying the security level of voice and gesture commands in case for instance someone who is not the train driver tries to use voice or gesture commands. Firstly, physical devices are not protected, so someone can already act on the train if they grad the manipulator. Moreover nowadays, voice assistant on GSM can recognize their owner voice and security software are able to recognize face and blur the screen if someone is spying on personal phone. Those technologies can be enabled in driving cabin. If there are enable, a workaround will be needed for drivers training.

Drivers also asked for more action on computer, and some of them stated that assistance tool have to be redesigned because as long as drivers are accountable, they must feel they remain in control. Moreover, such assistance needs the development of drivers' state assessment and monitoring devices. For instance, some drivers already complained about concentration loss and negative impact on workload for the dead man's switch verification. Due to human complexity, HMI can therefore become problem makers instead problem solvers. Correlation between information from HMI and behaviour of the train via automated or manual tasks has to be verified on field. Moreover, task allocation between a driver and a train has to consider human and technical advantages and weaknesses in order to take advantage of the benefits of one to compensate the weaknesses of the other

By the end, the most commonly mentioned problem concerning nested situation and cascading alarms. Drivers complained about the avalanche effect while driving, i.e. a lot of events come in a same time. As the Survey 2 has involved a limited number of participants (and most of them had more than 10 years of experience), a short-term perspective will consist as an extended study that could imply young drivers who may be familiar with innovative technology like voice and gesture recognition systems.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 881814. The present research work has also been supported by:

- The French Ministry of Higher Education and Research,
- And The Hauts-de-France Region.

The authors gratefully acknowledge the support of these institutions.

REFERENCES

Chen, L. L., Zhao, Y., Ye, P. F., Zhang, J., & Zou, J. Z. (2017). Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. 85, p. 279-291.

Collet C., Salvia E., Petit-Boulanger C. (2014). Measuring workload with ElectroDermal activity during common braking actions. *Ergonomics.* 57(6), p. 886-896.

Dehzangi O., Rajendra V., Taherisadr M. (2018). Wearable driver distraction identification on-the-road via continuous decomposition of galvanic skin responses. *Sensors.* 18(2), p. 503.

Dufour A. (2014). Driving assistance technologies and vigilance: impact of speed limiters and cruise control on drivers' vigilance. *Seminar on the Impact of Distracted Driving and Sleepiness on Road Safety*. Paris, 2014.

Enjalbert S., Gandini L. M., Pereda Baños A., Ricci S., Vanderhaegen F. (2021a). Human–Machine Interface in Transport Systems: An Industrial Overview for More Extended Rail Applications. *Machines*. 9 (2), p. 36.

Enjalbert S., Gandini L. M., Pereda Baños A., Ricci S., Vanderhaegen F. (2021b). Drivers' assisting Human Machine Interface: a review for future applications onboard trains. *Resource Efficient Vehicles Conference.*

JTSB (2008). Aircraft serious incident – Investigation report. *Report # AI2008-01.* November 28, 2008

Gandini L. M., Ricci S., Vanderhaegen F., Enjalbert S. (2020). State of the art of HMI in transport industries. *CARBODIN project, Delivrable D10.1.*

Merlevede J. V., Enjalbert S., Vanderhaegen F., Hassoun C., Henon S., Pereda Baños A., Ricci S. (2022a). Results of the innovative driving cabin survey. *CARBODIN project, Delivrable D10*.2.

Merlevede J. V., Enjalbert S., Vanderhaegen F., Henon S., Pereda Baños A., Ricci S. (2022b). Proposals of pertinent new HMI. *CARBODIN project, Delivrable D10.3.*