Quantitative assessment of IA-assisted sound-based interpretation of neonatal EEG

A Degree Thesis
Submitted to the Faculty of the
Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona
Universitat Politècnica de Catalunya
by
Montserrat Anglès Cabau

In partial fulfilment
of the requirements for the degree in AUDIOVISUAL SYSTEMS ENGINEERING

Advisors: Andriy Temko and Emanuel Popovici
Co-advisor: Climent Nadeu Camprubí

Barcelona, October 2018
Abstract

The goal of the project was to study the usability of Neurobell, an Android App designed to help healthcare professionals without EEG interpretation expertise. The application, which is currently being developed, detects neonatal brain injuries using sonification and artificial intelligence methods.

This work also aims to assess the performance of the neonatal brain stethoscope by comparing it with the current method of visualising EEG signals. To accomplish this, a web-based survey was previously developed. In this work, the usability of the survey was studied, and the website was redesigned to be more appealing to users by forming a game based scenario.

The improvements of the Android app and website were discussed with a user interface expert and healthcare professionals in order to tailor both for clinical needs and settings.

To conclude, the results of the survey are analysed and it is proven that the neonatal brain stethoscope is improving the detection of brain injuries in neonatal patients.
Resum

L'objectiu del projecte era estudiar la usabilitat de Neurobell, una aplicació d'Android dissenyada per ajudar els professionals de la salut sense experiència en interpretació d'EEG. L'aplicació, que actualment s'està desenvolupant, detecta lesions cerebrals en nounats mitjançant mètodes de sonificació i intel·ligència artificial.

Aquest treball també pretén avaluar el rendiment de l'estetoscopi cerebral neonatal comparant-lo amb el mètode actual de visualització de senyals EEG. Per aconseguir-ho, es va desenvolupar prèviament una enquesta basada en web. En aquest treball, s'ha estudiat la usabilitat de l'enquesta i s'ha redissenyat el lloc web perquè fos més atractiu pels participants convertint-lo en un joc.

Les millores de l'aplicació Android i la web es van discutir amb un expert en interfícies d'usuari i professionals de la salut per adaptar-se tant a les necessitats clíniques com a les tècniques.

Finalment, s'han analitzat els resultats de l'enquesta i s'ha demostrat que l'estetoscopi millora la detecció de lesions cerebrals en pacients neonatales.
**Resumen**

El objetivo del proyecto era estudiar la usabilidad de Neurobell, una aplicación de Android diseñada para ayudar a los profesionales de la salud sin experiencia en interpretación de EEG. La aplicación, que actualmente se está desarrollando, detecta lesiones cerebrales en neonatos mediante métodos de sonificación e inteligencia artificial.

Este trabajo también pretende evaluar el rendimiento del estetoscopio cerebral neonatal comparándolo con el método actual de visualización de señales EEG. Para ello, se desarrolló previamente una encuesta basada en web. En este trabajo, se ha estudiado la usabilidad de la encuesta y se ha rediseñado el sitio web para que fuera más atractivo para los participantes convirtiéndolo en un juego.

Las mejoras de la aplicación Android y la web se discutieron con un experto en interfaces de usuario y profesionales de la salud para adaptarse tanto a las necesidades clínicas como las técnicas.

Finalmente, se han analizado los resultados de la encuesta y se ha demostrado que el estetoscopio mejora la detección de lesiones cerebrales en pacientes neonatales.
We were hiking through Picos de Europa and the boy in front of me was complaining all
day about EVERYTHING. ‘It is a tough route but you are the one who has decided to come
here’ you told him. To Berta, you inspired me since the first day we met and you keep
doing it even now that I cannot hug you. My whole life changed because of you.

Also, to all the “Corkians” I met while I was doing this project, you made me realise how I
feel when I am truly happy. It was the best craic!
Acknowledgements

I would like to specially thank Dr. Andriy Temko, Climent Nadeu and Emanuel Popovici for giving me the opportunity to work on this project in INFANT Centre and Embedded.Systems@UCC, University College Cork, in Ireland, and expand my knowledge in fields I probably would not have explored otherwise.

My sincere thanks to Sergi Gómez for being supportive and helpful with any doubt I had during the realization of the project.

Thanks to Prof. Gealdine Boylan (INFANT Centre director and healthcare professional) and Mike Cunneen (User interface expert) for being so receptive and spend some of their time discussing and giving ideas in order to improve both apps.

Many thanks to Kevin (back-end developer of the website) and Edu (Android developer) for their patience and for making the changes and improvements in their respective apps as fast as possible.

Also many thanks to Valentina, Paola, Mark, Alison and Oksana. It has been a pleasure to work in your team.

Thanks Atenea for always being there and gràcies family for your unconditional support.

Finally, many many thanks to Aleksandra, Anika, Frifri, Rourou, Daniel, Raji, Júlia, Eoin, Rodolfo, Mario, Anjali, Hannah, Brigit, Peter, Jenny and Nora. I probably would have run away without you.
# Revision history and approval record

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>13/08/2018</td>
<td>Document creation</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>Document revision</td>
</tr>
</tbody>
</table>

## DOCUMENT DISTRIBUTION LIST

<table>
<thead>
<tr>
<th>Name</th>
<th>e-mail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Montserrat Anglès Cabau</td>
<td><a href="mailto:rat.angles@gmail.com">rat.angles@gmail.com</a></td>
</tr>
<tr>
<td>Andriy Temko</td>
<td><a href="mailto:atemko@ucc.ie">atemko@ucc.ie</a></td>
</tr>
<tr>
<td>Emanuel Popovici</td>
<td><a href="mailto:e.popovici@ucc.ie">e.popovici@ucc.ie</a></td>
</tr>
<tr>
<td>Climent Nadeu</td>
<td><a href="mailto:climent.nadeu@upc.edu">climent.nadeu@upc.edu</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Written by:</th>
<th>Reviewed and approved by:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>Date</td>
</tr>
<tr>
<td>10/09/2018</td>
<td></td>
</tr>
<tr>
<td>Montserrat Anglès Cabau</td>
<td>Andriy Temko</td>
</tr>
<tr>
<td>Project Author</td>
<td>Project Supervisor</td>
</tr>
</tbody>
</table>
Table of contents

Abstract ..................................................................................................................................................1
Resum ..................................................................................................................................................2
Resumen ...............................................................................................................................................3
Acknowledgements ............................................................................................................................5
Revision history and approval record ................................................................................................6
Table of contents ..................................................................................................................................7
List of Figures .......................................................................................................................................8
List of Tables: .....................................................................................................................................9
1. Introduction .......................................................................................................................................10
  1.1. Project overview and goals ......................................................................................................10
  1.2. Project requirements and specifications .............................................................................12
  1.3. Work plan, Gantt Diagram and milestones ........................................................................13
  1.4. Deviations and incidences .....................................................................................................14
2. State of the art of the technology used or applied in this thesis ......................................................15
  2.1. Usability concept ....................................................................................................................15
  2.2. Creation of a webpage ...........................................................................................................19
  2.3. Analysis of the results ............................................................................................................19
3. Methodology / project development: .................................................................................................20
  3.1. Web application ......................................................................................................................20
  3.2. Android application ..............................................................................................................30
4. Results ..............................................................................................................................................32
  4.1. Web application ......................................................................................................................32
  4.2. Android application ..............................................................................................................37
5. Budget ..............................................................................................................................................41
6. Conclusions and future development: .............................................................................................42
Bibliography .........................................................................................................................................43
Glossary ................................................................................................................................................44
List of Figures

Fig. 1 Scheme of the full project ................................................................. 11
Fig. 2 Scheme of the project ........................................................................ 11
Fig. 3 Summary of the work plan and Gantt ................................................. 13
Fig. 4 Scheme of the project detailed ............................................................ 14
Fig. 5 Clear and busy composition ............................................................... 16
Fig. 6 Paletton.com website ......................................................................... 17
Fig. 7 Logos ................................................................................................. 17
Fig. 8 Previous registration interface ........................................................... 21
Fig. 9 New registration page proposal ........................................................... 22
Fig. 10 Webpage flow .................................................................................. 23
Fig. 11 Final characters ................................................................................ 24
Fig. 12 Different displays of the probabilistic output .................................... 25
Fig. 13 Interface of the first test ................................................................. 27
Fig. 14 Preferences database ....................................................................... 29
Fig. 15 Results database ................................................................................ 29
Fig. 16 One of the menus in the analysis app .............................................. 30
Fig. 17 Training phase before ................................................................. 32
Fig. 18 Second test interface before ............................................................ 32
Fig. 19. Training face after ......................................................................... 33
Fig. 20 Second test interface after .............................................................. 34
Fig. 21 Main interface before ................................................................. 38
Fig. 22 Main interface after ........................................................................ 38
Fig. 23 Review mode before ........................................................................ 39
Fig. 24 Review mode after ........................................................................... 39
Fig. 25 Patients list ....................................................................................... 40
List of Tables:

Table 1 Milestones ........................................................................................................14
Table 2 Legibility level ..................................................................................................17
Table 3 Results ignoring 'not sure' answers ...............................................................35
Table 4 Results treating 'not sure' answers as errors ..................................................37
Table 5 General budget ...............................................................................................41
Table 6 Personnel budget ............................................................................................41
# Introduction

## Project overview and goals

Although electroencephalography (EEG) is the most useful tool to assess neonatal brain health, it is difficult to find healthcare professionals with expertise in EEG interpretation or the necessary equipment in the majority of hospitals in need, especially in developing countries. Since neurological injury in neonatal patients can lead to life threatening experiences, the following work aims to contribute to the cause and help to avoid such situations.

The Embedded Systems group in the Electrical Engineering Department of the University College of Cork (UCC) and the Irish Centre for Fetal and Neonatal Translational Research (INFANT) Centre are developing a Neonatal Brain Stethoscope called Neurobell [1], a solution which aims to detect seizures in new-born babies, through a system assisted by sonification [2] and deep learning AI (artificial Intelligence). This innovative handheld and low-cost tool aims to help all neonatal health care professionals, particularly those without EEG interpretation expertise.

The whole Neurobell project is divided in three parts: hardware (signal acquisition), communication (data compression, Bluetooth transmission and reception, and data decompression) and software (signal processing and graphic user interface).

In the hardware part, the EEG signal is acquired through a handheld device with dry [3] electrodes which can record up to eight channels [4] In the first prototype just one channel is used. The EEG signal is compressed and transmitted to a tablet through Bluetooth connection and the data is finally decompressed.

Once the data is in the tablet, the signal is interpreted using three methods: visual, as currently used in hospitals; sonification, converting the signal into sound using signal processing techniques [5]; and obtaining the probability of having seizures using a deep convolutional neural network (CNN) with 6 layers or 11 layers [6]. Finally, the results are shown in an Android app user interface.
The purpose of this work is to study all the aspects related to the usability of the software. It covers the usability study of the Android Application used in a tablet and the usability study and improvement of a web survey implemented in order to determine the accuracy of the novel software interpretation system. The author of this work thought that the aforementioned web survey could be converted into a game to attract more participants, so this work also covers the process of the design and implementation of a game-based survey, its population, and the analysis of the obtained data (coloured in purple in the following scheme).

Fig. 1 Scheme of the full project

Fig. 2 Scheme of the project
The project main goals are:

- Develop a thorough understanding of the application that is being developed.
- Study the usability of the previous survey webpage
  - Redesign the webpage to make it more popular.
  - Conduct a survey that involves personally talking to neonatologists and other healthcare professionals.
  - Analyse the results and determine the accuracy of the system.
- Usability study identifying the weaknesses and strengths of the Android app in order to better match the clinical needs.

1.2. Project requirements and specifications

The project requirements are:

- Present a detailed usability study of the current webpage
- Redesign the webpage in order to attract more participants and gather more results
- Implement the changes in the webpage before the 5th of May so that can be populated at the annual Paediatric Academic Societies Meeting in Toronto.
- Study of the Android app ensuring the ease of use, efficiency, user satisfaction, safety and effectiveness for the intended users.

The project specifications are the following:

- The implementation of the website game is done from scratch using HTML, CSS, JS, JQuery and PHP. It will have a simple interface and it will keep the core of the already-done webpage.
- The usability study follows the principles and process of international standard IEC62366-1:2015 Application of usability engineering to medical devices [7].
1.3. **Work plan, Gantt Diagram and milestones**

A detailed version of the Work packages and Gantt diagram is attached in the Annex I. Moreover, the differences between the actual development of the project and the first plan are explained in section 1.4.

**Work packages and Gantt Diagram**

![Gantt Diagram](image)

Fig. 3 Summary of the work plan and Gantt
Milestones

Table 1 Milestones

<table>
<thead>
<tr>
<th>WP#</th>
<th>Short title</th>
<th>Milestone / deliverable</th>
<th>Date (week)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PP&amp;WP</td>
<td>Project Proposal and Work Plan</td>
<td>05/03/2018</td>
</tr>
<tr>
<td>2</td>
<td>CR</td>
<td>Critical Review</td>
<td>15/06/2018</td>
</tr>
<tr>
<td>3</td>
<td>P</td>
<td>Presentation</td>
<td>10/05/2018</td>
</tr>
<tr>
<td>4</td>
<td>FR</td>
<td>Final Report</td>
<td>8/10/2018</td>
</tr>
<tr>
<td>5</td>
<td>FP</td>
<td>Final Presentation</td>
<td>15/10/2018</td>
</tr>
</tbody>
</table>

1.4. Deviations and incidences

In the beginning, the plan was to study the usability of the Android Application and make the necessary improvements to match the clinical needs. It was then decided that it was more convenient to continue developing a webpage in order to get data and evaluate the algorithms used during the seizure detection. The plan ended up as follows:

After experiencing low participation on the web surveys first launch, the webpage was adapted to mobile phones and tablets and the content of the survey was reviewed (text, examples…). For that reason, the analysis of the webpage was delayed some weeks.
2. **State of the art of the technology used or applied in this thesis**

2.1. **Usability concept**

The accurate definition of the word ‘usability’ can be found in the official ISO 9241-11 [8] as: “*the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*” However, in this project, at the time of using it as a tool, it was decided to define the usability of the system by taking five attributes into consideration:

- Ease of learning: how easy is for the user to learn how to use the product/app.
- Efficiency of use: the user should be able to do the tasks as fast as possible.
- Facility to remember: after a while, it should be easy to use the app again without too much effort.
- Degree of errors: the product/app have to have a low error rate.
- Satisfaction of the user: the product has to be attractive for the user so they keep using it.

There is no general method that applies to all apps, a usability study has to be adapted to the application done at hand. For webpages it is usually helpful to do tests on a large cohort of people while observing the users voice and movements while doing specific tasks. Since users of this webpage are mostly busy clinicians, and moreover, the fact that it was unfinished, a different method was used in this project.

For medical devices it is necessary to use the International Standard IEC 62366-1:2015: Medical devices Part 1: Application of usability engineering to medical devices which covers the most important necessary aspects in an application intended for medical needs.

2.1.1. **Webpage usability**

Undertaking a usability study of a webpage means to make it more usable and this directly results in multiple benefits: costs savings, user satisfaction, brand prestige and, higher productivity of the employers.

Certain attributes are taken into consideration when analysing the usability of web pages, the most important are:
- Interface design
- Navigation
- Architecture of the information

In the following sections, these three attributes are explained in detail.

2.1.1.1. Interface design

While analysing the interface design, the aesthetic part of the website is considered. It is very important because it is what the user sees in first place and it depends on it that the user stays or leaves the site. The main factors that have to be taken into account while designing a webpage are listed below:

- Composition of the webpage: it should be attractive and not too busy so that the user has a good impression. For good design, it is fundamental that the webpage adapts its content depending on the size of the screen.

![Fig. 5 Clear and busy composition](image)

- Typography: for webpages, is better to use Sans Serif fonts (without ornaments). The size of the font and the colour helps to give more or less importance to the text. The contrast with the background colour is essential for a good legibility.
Table 2 Legibility level

<table>
<thead>
<tr>
<th>Combination</th>
<th>Legibility level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hello world, how are you?</td>
<td>Low, there is no contrast at all.</td>
</tr>
<tr>
<td>Hello world, how are you?</td>
<td>Medium, it is readable but could be better</td>
</tr>
<tr>
<td>Hello world, how are you</td>
<td>The highest contrast possible.</td>
</tr>
</tbody>
</table>

- Colours: the combination of colours is probably one of the interface elements that has more influence in the design of a webpage. There are resources like: paletton.com or coolors.co that may help while choosing the palette of colours.

Fig. 6 Paletton.com website

- Images: Pictures, illustrations and icons are very important in webpages. Not just to decorate but to give information.

Fig. 7 Logos
2.1.1.2. Navigation

The users have to know where they are in every moment and it has to be easy to change the page where they are. The links have to be surely working.

There is four types of navigation: the main, the local and the footer navigation, and the search bar. While it is true that the webpage has to have a flow, sometimes it will not be useful to have too much paths to get to the same pages.

2.1.1.3. Architecture of the information

The information has to be sorted and well-defined. Each element of the webpage has its place in the interface and it has to be coherent with the information that has to be given. For example, important things on top should be on top.

2.1.2. Usability of medical devices (Android App)

For medical applications and devices, a rigorous engineering usability process is mandatory. Since humans’ lives are concerned it is necessary to follow the International Standard IEC 62366-1:2015: Medical devices Part 1: Application of usability engineering to medical devices.

Basically, it consists in eight steps:

1) Define the use specification
2) Identify user interface characteristics related to safety and potential use errors
3) Identify known and foreseeable hazards and hazardous situations
4) Identify and describe hazard-related use scenarios
5) Establish user interface specification
6) Establish user interface evaluation plan
7) Perform user interface design, implementation and formative evaluation.
8) Perform summative evaluation of the usability of the user interface

These steps are detailed later.
2.2. **Creation of a webpage**

The webpage has been implemented from scratch. Templates were not an option since it was necessary to have the control of all the content and data. During the creation of the webpage some programing languages have been used:

- **HTML**: is the standard markup language for creating Web pages. It Stands for Hyper Text Markup Language and describes the structure of Web pages using markup.

- **CSS**: stands for Cascading Style Sheets and describes how HTML elements are to be displayed on screen, paper, or in other media.

- **JavaScript**: It is basically used to add interaction in the webpage. It can change HTML, HTML attributes values, HTML styles (CSS), can hide and show elements, among other things.

- **jQuery**: it is a JS Library that simplifies its programming to make JS easier to use.

- **PHP**: used for making dynamic and interactive Web pages.

Moreover, the AUC (area under the ROC curve) is applied to figure out the accuracy of the survey. An area of 1 represents an excellent test and an area of 0.5 represents a worthless test.

2.3. **Analysis of the results**

Some parameters [9] defined as follows are used in order to assess the system:

**Accuracy**: it is the ability of the test to differentiate the unhealthy and healthy cases correctly. To estimate the accuracy of a test, the proportion of true positive and true negative in all evaluated cases is calculated. Mathematically, this can be stated as:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

Eq. 1
Sensitivity: it is the ability of a test to determine the unhealthy cases correctly. To estimate it, the proportion of true positive in unhealthy cases is calculated. Mathematically, this can be stated as:

\[ Sensitivity = \frac{TP}{TP + FN} \]  
Eq. 2

Specificity: it is the ability of a test to determine the healthy cases correctly. To estimate it, the proportion of true negative in healthy cases is calculated. Mathematically, this can be stated as:

\[ Specificity = \frac{TN}{TN + FP} \]  
Eq. 3

Where in Eq. 1, 2 and 3,

**Patient:** positive for disease

**Healthy:** negative for disease

**True positive (TP)** = the number of cases correctly identified as unhealthy

**False positive (FP)** = the number of cases incorrectly identified as unhealthy

**True negative (TN)** = the number of cases correctly identified as healthy

**False negative (FN)** = the number of cases incorrectly identified as healthy

3. **Methodology / project development:**

3.1. **Web application**

In this section all the aspects related to the web application are detailed. The initial usability study, the idea of converting it into a game and the development process and finally, how the analysis of the survey results was done.

3.1.1. **Usability study**

As it has been said before, the goal of creating a webpage was to get data from clinicians and determine the accuracy of the system comparing the performance at detecting seizures with the proposed android app system (Visual+ Sonification+AI) versus the conventional method (Visual).
At the beginning of this project, an essential interface and the core of the survey were already done but it was necessary to study its usability and to improve it to attract more participants in order to get more data.

![Previous registration interface](image)

Fig. 8 Previous registration interface

The usability study, which summary can be found in Annex II, led to the idea of converting the website into a game to make it even more appealing.

3.1.2. Design and implementation of the game

Although the core of the webpage was kept, a whole new interface was necessary. In the proposed game, the users were asked to identify in two separate tests the presence of seizure events within a sample by using one of the mentioned methods in each test.
3.1.2.1. Flow

After some proposals and meetings with the user interface expert, the simplest flow was chosen. If the users are opening the survey from a tablet or mobile phone, they are redirected to a warning page explaining that the experience would be better using a computer. If they decide to not change the device, it is recommended to set the landscape mode and the adapted version for mobile phones and tablets is shown. If the survey is opened from a computer the users can enjoy the full experience.
In the first page, the participants can find an introduction of the project, the goals of the survey and the registration form. Once they are logged in, the game is explained and differences between seizures and no seizures are taught. The users can always go back to this information by opening a pop up in the test pages.

In the first test, there is one question per test and a bar with a character in the middle that goes backwards, forwards to a toy or stays at the same place depending on the answer given using visual methods. This process continues until the character gets the toy (congratulation page) or arrives at the beginning of the path (game over page). In both cases the survey continues the second test training where the sonification and AI differences between seizures and no seizures are explained.

The second test has the same structure than the first one but in this case the character is grown up and the sonification and AI methods are included. When the users finish the test, they get a game over or a congratulation page, however, this time they have the option to choose which methods were more useful in their opinion.
Moreover, there is four links where they can play again (directly to the test, avoiding the training phase), find more information about the team (www.infantcentre.ie), send the link of the game to a friend or contact the team (neurobell@ucc.ie) for queries or suggestions.

3.1.2.2. Characters

Nowadays, Neurobell is an application developed to help neonatal patients, therefore, the main character of the game had to be a newborn baby so that the participants are fully conscious the purpose of the survey, which is to help real patients.

In order to attract more participant, the main character was called Emily (as a baby and as a toddler) since this is the most popular baby name in Ireland and the users could feel identified.

First, Emily was drawn by hand and then, digitalised with a Huion graphic tablet. As the author of the project did not have expertise in digital drawing, two online courses were attended: ‘Introducción a la ilustración digital’ by Wacom and ‘Ilustración personalizada: el retrato oriental’ by Rod Feliciano.

In total there are twenty-eight different drawings. For each level: a way used as a progress bar, a character introducing the current level, a happy character playing with the toy if the user wins the match, a sad character if the user exceeds the maximum number of questions without getting the toy and, since the expression of the character depends on the distance to the toy, ten babies crawling or children walking towards the toy are also needed.
3.1.2.3. The address

The website is hosted at https://www.ue.ucc.ie/neurobell/ but it can be accessed from neurobell.com, http://neurobell.com and http://www.neurobell.com so that it is easier to spread the link and remember it.

3.1.2.4. The text

Up to five versions of the text were written before the final version, taking into account the audience, purpose, style and tone. And it is being improved periodically considering the healthcare professionals and user interface expert opinions.

3.1.2.5. The data: EEG files

There are 200 EEG files (converted into images) in the database that can be displayed in the survey: 100 containing seizures and 100 without any them. The name of the files are encrypted using the checksum method so that the user is not able to discover the correct answer beforehand.

To prove that the correct 200 file samples from the EEG database are used, the AUC is computed using the ground truth and the probabilistic outputs. Doing that, we ensure that after the AI representation there is room for the subjective interpretation using sonification and visual methods. Using a MAF (medium average filter) of 30 seconds, the epoch-level AUC is 93, 07%. This means that there is a 7% that depends on the subjective interpretation.

In order to represent the probabilistic output according to the medical needs, some options were proposed to healthcare professionals.

Fig. 12 Different displays of the probabilistic output
Other options such as using traffic light next to the EEG signal or displaying the probabilistic signal under the EEG were discarded. Finally, in the first test the files were represented in white on a blue background and in the second one the background colour changes to purple depending on its AI output, where the purple represents a high probability of being seizure and blue no seizure at all.

3.1.2.6. Interface

To avoid proportions issues, the game is showed in a white square on top of a purple background. INFANT colours are used as a palette to align the institution with the game.

The structure and the basic elements of the webpage were created using HTML and CSS. As the author of this project did not have web development expertise, an online course was attended: ‘Desarrolla una página web con HTML y CSS’ by Sergio Agamez Negrete. The animations and interactive elements were programmed in JavaScript and JQuery and it was joined to the backend part in PHP.

3.1.2.7. Miscellanea

- In order to not distract the participants, the links that are usually set in the header (‘About us’, ‘Contact us’, ‘Recommend to a friend’) are hidden until the end. The only links showed during the whole survey are the ones that refers to the institutions that have participated in this project.

- In the log in is taken into account if the participant is a healthcare professional. The second question is only activated if the answer is yes. Moreover, there is a question to make sure the users give permission to use the generated data, if they do not check it they cannot proceed to the survey.

- During the whole process there is a blue bar on the top of the page that allows the user to know in which stage of the survey he is.

- Once the users start the test, there is another progress bar that represents a path where the baby and the toddler move letting know if their answer is correct or not. Apart from the position and the face of the character, a sound celebrating or complaining the action is reproduced automatically.
- The image of the EEG signal scrolls automatically in both tests. Moreover, in the second test the sonification is reproduced automatically to force the participants to use it for the decision. Since the audio is ten times faster than real time, the velocity of the scroll and the sonification aren't synchronized. The image could not be seen properly if it was also accelerated.

- The user can select the scale of the image (10 or 20 seconds) as well the healthcare professionals do with the current apps. In addition, on top of the EEG signal image, there is a scale indicating also the amplitude.

- There is a popup where the user can check which channel corresponds to the signal.
- For each EEG file, the user can choose between three options: ‘Seizure’, ‘Not seizure’ and ‘Not sure’. If their answer matches the ground truth the character moves to the toy. If the answer is incorrect the character goes backwards. Choosing ‘Not sure’ lets the character in the same position and a new question appears.

- At the end of the survey the participants are asked to choose the most useful method on their opinion.

- The webpage is adapted to mobile phones and tablet although it is better to play in a computer or laptop so that the signals are better shown.

- The webpage code (HTML, CSS, JS, PHP) can be found in Annex III.

3.1.3. Conduct survey

The survey was populated mainly to healthcare professionals without expertise in EEG interpretation. That includes clinicians from the Cork University Hospital (CUH), medical students’ societies, clinicians and students from other countries... furthermore, some healthcare professionals trained in EEG interpretation and some layout people have also participated.

3.1.4. Database

The information obtained by the participants are stored in two different databases. In the first one, the user information is saved: the session ID, if they are healthcare professionals and if they are trained in EEG interpretation, their methods preferences and the timestamp.
In the second database the information is saved: the session ID, the file ID, the test number, the given answer, the ground truth and the timestamp.

3.1.5. Analysis of the results

Several metrics can be extracted from the results obtained in the survey, including the accuracy at detecting seizures for each method. This metric should allow to compare the improvement (if any) provided by the proposed Android app system among the traditional methodology.

The analysis of the result has been carried out in Matlab. The program can analyse the data according to the participants’ knowledge (healthcare professionals with or without expertise in EEG, or laypeople). It can also classify the data by the file ID, by the person ID or process all the data at the same time. The ‘Not Sure’ answer can be interpreted as an error or ignoring the question. All the Matlab codes can be found in the Annex IV.
3.2. **Android application**

3.2.1. **Usability study**

The usability study of the Android App has been carried out following the International Standard for medical devices previously commented. First of all, a general study of the app has been done detecting its strengths and weaknesses. Since the goal of the usability study is to develop the app covering the medical needs, the author of the project has had weekly meetings with the user interface expert and has occasionally consulted healthcare professionals during the whole process. The information extracted from those meetings has been given to the Android developer who has made the proposed improvements.

All the information collected is summarised in a Usability engineering file that can be found in the Annex V. In that file, the steps of the International Standard are followed.

Firstly, the use specification has been defined explaining the intended medical indication and patient population, the intended part of the body that interacts with the app, the user profile, the use environment and the operating principle.

Following that, the user interface characteristics related to safety and potential use errors have been identified, as well as the known or foreseeable hazards and hazardous situations.
After that, the hazard-related use-scenarios have been identified and described. The user interface specifications have been reviewed and a user interface evaluation plan has been designed and carried out.
4. **Results**

4.1. **Web application**

4.1.1. **Survey**

After the usability study and the redesign of the website, the survey has ended up being more appealing and usable in computers, tablets and mobile phones.

A page of the website before the redesign and after it can be compared in the following images:

![Fig. 17 Training phase before](image1)

![Fig. 18 Second test interface before](image2)
Background: What are Neonatal Seizures?

"Neonate" refers to a time period, neonatal seizures are seizures that occur in the first month of an infant's life. When newborn babies are monitored with EEG, the readings may be for a few hours or several days. A seizure is an abnormal electrical discharge in the brain and these are seen as a sudden and abnormal event in the readings. Usually these 'events' have a repetitive and evolving pattern with changes in frequency, voltage and morphology.

What do Neonatal Seizures Look Like?

- Seizures tend to last between 1 and 5 minutes, but can be shorter.
- Seizures show a repetitive periodic waveform. The time between 2 consecutive periods (repetitions) usually takes between 0.5 sec and 2 sec (1 sec corresponds to a single square on the grid).
- Amplitude (vertical distance from the centre) may change as the seizure evolves.
- The shape of the waves (known as the morphology) tends to change within seizures, and can differ between seizures. They can include spike and wave, sharp and slow waves, and repetitive spikes or rhythmic activity.

Features: Periodic signal, spikes & sharp waves, temporally evolving

Fig. 19. Training face after.
Fig. 20 Second test interface after
So far, 58 people have participated to the survey and it is expected to reach 100 participants in less than two months.

### 4.1.2. Analysis of the results

In the following tables the results of the survey carried out in the website are analysed. Of the 81 registrations in the website, a total of 58 participants had answered the survey on the 1st of September.

It is observable that, in general, the methods used in this project permit identify seizures significantly better. Analysing all the data at the same time, the accuracy increments up to four points and the sensitivity up to eight points. The specificity doesn’t improve as much in this case, although that, is greater (0, 2%) than with the methods currently applied in hospitals.

**Table 3 Results ignoring ‘not sure’ answer. ‘HcP’ means healthcare professionals and ‘T’ means trained in neonatal EEG interpretation. ‘AI’ is the probabilistic output of being seizure. The amount of participants is shown between brackets.**

<table>
<thead>
<tr>
<th>'Not sure': ignored</th>
<th>Test 1 (Visual)</th>
<th>Test 2 (Visual+sonification+AI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>Sensitivity</td>
</tr>
<tr>
<td><strong>Who</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All (58)</td>
<td>Overall</td>
<td>64.5%</td>
</tr>
<tr>
<td></td>
<td>By person ID</td>
<td>56.5%</td>
</tr>
<tr>
<td></td>
<td>By file ID</td>
<td>63.8%</td>
</tr>
<tr>
<td>Not HcP (37)</td>
<td>Overall</td>
<td>64.6%</td>
</tr>
<tr>
<td></td>
<td>By person ID</td>
<td>54.4%</td>
</tr>
<tr>
<td></td>
<td>By file ID</td>
<td>65.2%</td>
</tr>
<tr>
<td>HcP (21)</td>
<td>Overall</td>
<td>64.3%</td>
</tr>
<tr>
<td></td>
<td>By person ID</td>
<td>58.6%</td>
</tr>
<tr>
<td></td>
<td>By file ID</td>
<td>65.7%</td>
</tr>
<tr>
<td>T HcP (10)</td>
<td>Overall</td>
<td>66.0%</td>
</tr>
<tr>
<td></td>
<td>By person ID</td>
<td>57.9%</td>
</tr>
<tr>
<td></td>
<td>By file ID</td>
<td>77.3%</td>
</tr>
<tr>
<td>Not THcP (11)</td>
<td>Overall</td>
<td>62.9%</td>
</tr>
<tr>
<td></td>
<td>By person ID</td>
<td>58.8%</td>
</tr>
<tr>
<td></td>
<td>By file ID</td>
<td>59.6%</td>
</tr>
<tr>
<td>Not HcP and Not THcP (48)</td>
<td>Overall</td>
<td>64.3%</td>
</tr>
<tr>
<td></td>
<td>By person ID</td>
<td>56.6%</td>
</tr>
<tr>
<td></td>
<td>By file ID</td>
<td>64.4%</td>
</tr>
</tbody>
</table>
Doing the analysis accordingly to the profession of the participants some conclusions can be extracted:

**a) Healthcare professionals (21) vs. not healthcare professionals (37)**

Hypothetically, it is thought that healthcare professionals’ improvement would be higher, however, the non-clinicians participants have proved better skills in both tests. This could be due the fact that the first ones might be overconfident while doing the test.

**b) Healthcare professionals with expertise in neonatal EEG interpretation (21)**

Clinicians trained in neonatal EEG interpretation have improved the detection of seizures significantly. Although that, it is more difficult for them to determine healthy cases correctly with Neurobell.

**c) Healthcare professionals without expertise in neonatal EEG interpretation (11)**

Surprisingly, the intended users of the application, healthcare professionals without expertise in neonatal EEG interpretation, have done the worse performance. The number of participants in this case is very low, it ought to be necessary to wait until more users do the survey.

On the other hand, some minor analysis has been done classifying the data by files and by participants. In general, those analysis show worse results but it is probably because there is not a lot of information when the data is split.

The results of the previous analysis are interpreted ignoring the ‘not sure’ answers. In the following table, these answers are treated as errors and it is noticeable that, largely, the results get worse.
Table 4 Results treating 'not sure' answers as errors. HcP’ means healthcare professionals and ‘T’ means trained in neonatal EEG interpretation. ‘AI’ is the probabilistic output of being seizure. The amount of participants is shown between brackets.

<table>
<thead>
<tr>
<th>'Not sure': error</th>
<th>Test 1 (Visual)</th>
<th>Test 2 (Visual+sonification+AI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who</td>
<td>Accuracy Sensitivity Specificity</td>
<td>Accuracy Sensitivity Specificity</td>
</tr>
<tr>
<td>All (58)</td>
<td>Overall 62.7% 64.4% 60.8%</td>
<td>Overall 66.4% 72.5% 60.1%</td>
</tr>
<tr>
<td></td>
<td>By person ID 54.9% 57.2% 50.0%</td>
<td>By person ID 45.6% 49.5% 42.2%</td>
</tr>
<tr>
<td></td>
<td>By File ID 62.5% 64.9% 59.3%</td>
<td>By File ID 67.4% 71.9% 62.6%</td>
</tr>
<tr>
<td>Not HcP (37)</td>
<td>Overall 62.9% 62.8% 63.0%</td>
<td>Overall 68.5% 73.2% 64.0%</td>
</tr>
<tr>
<td></td>
<td>By person ID 53.0% 52.6% 51.0%</td>
<td>By person ID 54.3% 59.6% 50.5%</td>
</tr>
<tr>
<td></td>
<td>By File ID 64.1% 65.2% 63.0%</td>
<td>By File ID 70.2% 73.6% 66.6%</td>
</tr>
<tr>
<td>HcP (21)</td>
<td>Overall 62.3% 68.9% 58.6%</td>
<td>Overall 62.1% 72.4% 51.7%</td>
</tr>
<tr>
<td></td>
<td>By person ID 56.8% 64.5% 50.2%</td>
<td>By person ID 35.3% 40.8% 30.5%</td>
</tr>
<tr>
<td></td>
<td>By File ID 62.5% 71.8% 53.9%</td>
<td>By File ID 59.5% 65.2% 54.1%</td>
</tr>
<tr>
<td>T HcP (10)</td>
<td>Overall 64.0% 66.0% 62.3%</td>
<td>Overall 70.2% 69.6% 70.8%</td>
</tr>
<tr>
<td></td>
<td>By person ID 55.6% 54.1% 57.4%</td>
<td>By person ID 25.9% 24.5% 27.4%</td>
</tr>
<tr>
<td></td>
<td>By File ID 77.3% 80.0% 75.0%</td>
<td>By File ID 72.7% 80.0% 66.7%</td>
</tr>
<tr>
<td>Not T HcP (11)</td>
<td>Overall 60.9% 71.2% 52.2%</td>
<td>Overall 59.1% 73.4% 44.4%</td>
</tr>
<tr>
<td></td>
<td>By person ID 57.3% 70.9% 45.6%</td>
<td>By person ID 45.8% 55.4% 37.4%</td>
</tr>
<tr>
<td></td>
<td>By File ID 57.9% 68.1% 45.6%</td>
<td>By File ID 59.5% 54.3% 65.8%</td>
</tr>
<tr>
<td>Not HcP and Not T HcP (48)</td>
<td>Overall 62.5% 64.2% 60.6%</td>
<td>Overall 65.1% 73.8% 59.1%</td>
</tr>
<tr>
<td></td>
<td>By person ID 55.1% 54.6% 51.0%</td>
<td>By person ID 54.3% 55.5% 50.5%</td>
</tr>
<tr>
<td></td>
<td>By File ID 63.1% 65.6% 60.3%</td>
<td>By File ID 66.3% 71.2% 60.9%</td>
</tr>
</tbody>
</table>

Almost all the participants have stated that they prefer the second method using the sonification and the probabilistic output to make better decisions.

4.2. **Android application**

The Android app has also been upgraded and although there are some changes that are not implemented yet, the differences between the previous app and the renovated one are very remarkable.

The colour interface, contrast, information provided and vocabulary used in the main activity makes the app look more professional and credible. The review mode of the app, used to check and listen to the sonification again is also renewed and more intuitive and usable. Moreover, there is one new tab with the information of the patients.

Some screenshots of the comparisons are showed below.
Fig. 21 Main interface before

Fig. 22 Main interface after
Fig. 23 Review mode before

Fig. 24 Review mode after
Fig. 25 Patients list

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Birth date</th>
</tr>
</thead>
<tbody>
<tr>
<td>F3H86</td>
<td>10/7/1990</td>
</tr>
<tr>
<td>A3D6F5</td>
<td>10/7/1990</td>
</tr>
<tr>
<td>F1RT</td>
<td>10/7/1990</td>
</tr>
<tr>
<td>8R76</td>
<td>10/7/1990</td>
</tr>
<tr>
<td>T3HY</td>
<td>10/7/1990</td>
</tr>
<tr>
<td>45H16</td>
<td>10/7/1990</td>
</tr>
<tr>
<td>A5D5O6WE</td>
<td>10/7/1990</td>
</tr>
</tbody>
</table>
5. **Budget**

The budget of this project is:

<table>
<thead>
<tr>
<th></th>
<th>Cost (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students Matlab licence</td>
<td>0</td>
</tr>
<tr>
<td>International Standard IEC 62366-1:2015</td>
<td>248</td>
</tr>
<tr>
<td>Website host</td>
<td>10</td>
</tr>
<tr>
<td>Personnel</td>
<td>8806</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>9064</strong></td>
</tr>
</tbody>
</table>

It is not taken into account the energy cost because the project was carried on in a university and it is difficult to know the exact consume.

The personnel cost are detailed as follows:

<table>
<thead>
<tr>
<th>Personnel</th>
<th>€ / month</th>
<th>Total months</th>
<th>Cost (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior engineer</td>
<td>1000</td>
<td>7</td>
<td>7000</td>
</tr>
<tr>
<td>Three experts</td>
<td>258</td>
<td>7</td>
<td>1806</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
<td><strong>8806</strong></td>
</tr>
</tbody>
</table>


6. **Conclusions and future development:**

The goals of this project have been fully accomplished giving to the Neurobell App a better usability following the International Standard for medical devices and, at the same time, improving the ease of learning, efficiency of use, facility to remember and satisfaction of the user while using the application.

Moreover, a guideline to keep working on the improvement of the application interface was provided to the team members. The usability study of the application has to be completed periodically, even if the application is already on the market, so continuation of this work is essential.

On the other hand, a usability study of a web-based survey was done, and the webpage was fully redesigned converting it into a game that permits to get data from computers, tablets and mobile phones. The webpage is used to get data from different kinds of people, including our target: health professionals without expertise in neonatal EEG interpretation.

Additionally, the webpage is prepared to be easily used for different purposes such as an evaluation of different AI visualization methods (will be done soon).

Although the participation could be higher is expected to increment it during the following months since the website will be spread in conferences and events.

This work, as well as the other parts of the project, will be exhibited next month in the 7th Congress of the European Academy of Paediatric Societies in Paris, and it will be also presented in further conferences.
Bibliography


Glossary

AI: Artificial Intelligence
AU: AUdio
AUC: Area Under the Curve
CNN: Convolutional Neural Network
CSS: Cascading Styles Sheets
CUH: Cork University Hospital
EEG: Electroencephalogram
HTML: HyperText Markup Language
ID: IDentification
IEC: International Electronic Commission
INFANT: Irish Centre for Fetal Translational Neonatal Translational Research
ISO: International Organization for Standardization
JS: JavaScript
MAF: Moving Average Filter
PHP: Hypertext Preprocessor
ROC: Receiving Operating Characteristic
UCC: University College of Cork
VS: Visual