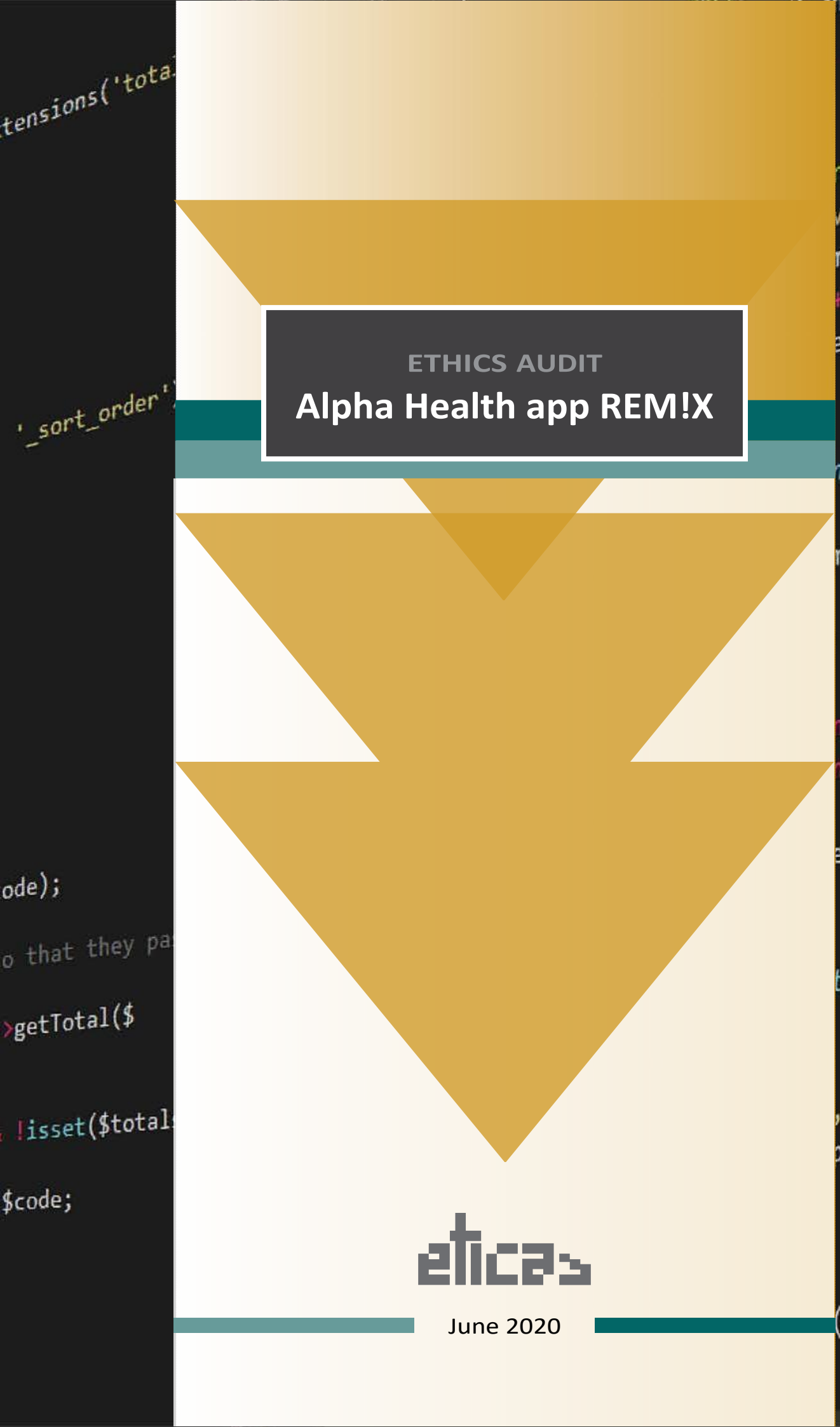
ETHICS AUDIT

# Alpha Health app REM!X



June 2020

# Table of Contents

# Annexes



# Tables



# Research Team

- Mariano Martín Zamorano;
- Gemma Galdon Clavell;
- Victoria Peuvrelle;
- Miguel Valbuena;
- Sara Suárez Gonzalo (Eticas R&C);
- Carlos Castillo (UPF).

# Foreword

At Alpha Health we believe that being trusted and trustworthy is critical to our mission to help people improve their health. To deliver this we have developed an ethics strategy which you can find on our website, and in which you will see that we are committed to undertaking and publishing an external, annual audit of progress against our ethics strategy. It is therefore with great pleasure that I write this foreword to our first external ethics audit, carried out by Eticas Research and Consulting (Eticas R&C).

The focus of this audit is on one of Alpha Health's prototypes – an app called REM!X. This app no longer exists as its purpose was to test various hypotheses relating to helping people improve their health and happiness, and to help us build our capabilities. Indeed, part of that capability building was itself to test how we could implement our ethics strategy, and consequently evolve and improve it. In this regard, the audit is of huge importance to Alpha Health.

The audit largely occurred over the course of 2019. It has taken us some time to publish it because it involved multiple analyses, and because we changed some elements of REM!X as the audit progressed, presenting somewhat of a moving target for the auditors. I also think that it's fair to say that the concept of an ethics audit remains fairly novel in the world of digital technology. Therefore, we have been learning a lot about what information is relevant and how to structure the analysis as we have progressed.

I am very grateful to the team at Eticas R&C, and their colleagues at Pompeu Fabra University, for all of their hard work, their unflinching ability to call us out when we weren't meeting our aspirations, and their patience with our agile development model.

Needless to say, what follows is the view of the auditors, but on behalf of Alpha Health, I wholeheartedly welcome this report and its conclusions. We aim to do even better in our next audit, later this year.


Ollie Smith

Strategy Director & Head of Ethics

# Introduction

This document presents the results of the Ethics Audit of the REM!X mobile application, developed by Alpha Health. Alpha Health is one of the teams within Telefónica Innovación Alpha, which is a company within Telefónica S.A. that focuses on long-term, disruptive innovation. Alpha Health aims to create technology that supports people to improve their health, with a portfolio of products that aims to prevent; predict; and treat mental illness. Throughout this audit, Alpha Health is referred to as Alpha as a shorthand.

The audit, carried out by Eticas Research and Consulting (Eticas R&C) jointly with the Universitat Pompeu Fabra of Barcelona[1] (UPF) in 2019, was mainly oriented towards examining the ethical standards behind the development of this app, and identifying technical and organizational issues related to its design and implementation. In this context, the analysis has been particularly focused on detecting unfair forms of bias and discrimination derived from the algorithmic processing involved in this application. This aspect was addressed explicitly through an Algorithmic Impact Assessment, which examined the technical specifications and social implications of such algorithmic processing. Based on these assessments, the Algorithmic Audit of the REMIX app allowed the Eticas team to recommend measures and practices aimed at improving the acceptability, desirability and proper management of personal data within the app, as well as minimising and/or preventing bias and discrimination.

The assessment consisted of two phases: The first one, aimed at collecting and analysing information regarding the practices, legal grounds and organisational aspects of the Alpha team in charge of the project. The ethical, acceptability and desirability dimensions of the REM!X app were addressed in light of data protection legal standards and principles (according to the European General Data Protection

[1] Dr Carlos Castillo (UPF) was only involved in the Algorithmic Impact Assessment.

Regulation, 2016/679), respect of users' integrity, and notions of accountability and technological design best practice. This phase took place over a number of months, which enabled the Alpha team to react to some of our initial recommendations. This led us to re-review certain aspects of REM!X; these adaptations to our recommendations are reflected in the analysis set out in this document.

The second stage of the Ethics Audit specifically concerned the Algorithmic Impact Assessment (AIA). The system was analysed, and indirect evidence of bias based on interviews with the Alpha team was examined. Moreover, the analysis was complemented by desk research and other fieldwork activities. The results of these research activities were examined to decide whether the system should be tested using quantitative methods and collecting data about the identified protected groups. Since no direct evidence of bias or discrimination was found in REM!X algorithmic processing, both Alpha and Eticas teams decided to reformulate the scope of the second of the AIA, and organise three training sessions for the Alpha team. These sessions addressed strategies for preventing algorithmic bias and were conducted between December 2019 and March 2020.

This report summarises all of the stages of the Ethics Audit. It provides an overview of its results, including the main findings, the exchanges of information between the Alpha and Eticas teams, and the recommendations offered by the audit team. After briefly presenting the REM!X app in Section 2, we summarise the ethics strategy of Alpha Health in Section 3; describe the legal and broader social analysis of the REM!X app in Section 4; detail the main organisational aspects examined through fieldwork in Section 5; recap the results of the AIA in Section 6; and provide overall conclusions in Section 7. The Audit outcomes allowed Alpha Health and Eticas to use the REM!X technologies to set standards and protocols applicable to all Alpha projects, and provided ethics best practice for activities concerning technological development in other, similar domains.

# The REM!X app

REM!X was a recommender app, developed by Alpha, whose algorithmic system was based on artificial intelligence techniques and aimed at offering customised advice on healthy habits through small exercises and wellness challenges. The application was therefore presented as one aimed at improving users' well-being, facilitating their personal development goals, helping them to overcome anxiety and stress, and ultimately, making them happier. In order to do this, the system was able to identify the current emotional state of the user by means of processing a set of data actively or implicitly provided by him/her, to then use it to provide personalised recommendations. REM!X was created by Alpha as a prototype to test various approaches to supporting people to manage anxiety and stress. Whilst REM!X itself no longer exists, it served its purpose as a test-bed for Alpha and it has informed the creation of Alpha's current product portfolio. The ethical audit of REM!X was a core part of this learning.


Drawing from data on users' sentiments and their desired emotional state, REM!X offered simple tricks, brief tutorials and challenges (recommendations) that sought to help them feel better. These included exercises that lasted for a few minutes, or programmes to follow over the course of several days, until they became healthier habits (e.g., going to the beach, reading a book, going to the theatre, etc.). With this purpose, the app asked users to indicate how they feel, how they want to feel, and what they are doing at the moment (e.g.: resting, working, studying, etc.). Other input data fed the REM!X algorithm in order to make recommendations, including: the interactions by users with the suggested activities and the app's interface, such as views, bookmarks, likes, and activities 'done'. Moreover, the system gathered data on users' tags, ambient light, activity, pedometer, battery, device, phone, call duration and social graph, ambient noise, Wi-Fi, Bluetooth, and screen lock / unlock.


## REM!X presentation:

*"According to many experts in the field, the secret to happiness is to find the balance between things that give pleasure and that give meaning to your life.*

*All the tricks and activities that we propose in REM!X are scattered to help you find your balance to be happier and overcome your anxiety and stress.*

*Therefore, to achieve your personal development goals and improve your well-being, it is good that you try new things so that you discover and learn more about yourself, because after all, isn't that life?"*

Additionally, the application included a feature that allows users to "save" challenges previously proposed and completed, as well as to store the most popular challenges that they had chosen. Based on their activity, REM!X scored users in five different categories: self-esteem, fun, productivity, relationships, and vitality. Once users set a goal, they could accumulate points by completing challenges and exercises. If users changed their mood or interests, or simply wanted to receive new recommendations, they could ask the app to help them in doing so.

The above characteristics and aims of REM!X, suggest that the app's customising features could pose certain challenges concerning the accuracy and fairness of its implementation. To further understand how this system is examined, section 4 covers the main concepts behind REM!X's legal and ethical aspects, technical specifications, and frames the more detailed Algorithmic Impact Assessment that followed.

# Alpha Health's ethical framework[2]

In this section, we will summarise the review of the Alpha Health ethical framework as it was when the audit of REM!X commenced. Alpha Health has subsequently iterated its Ethics Strategy, taking into account Eticas' insights throughout the audit. The latest version of Alpha Health's Ethics Strategy can be found on their website.

In our assessment of the Ethics Strategy, we started with the ethical principles set out in the strategy, commented on how to expand on these, and proposed indicators of success. The table below summarises our findings.

| Alpha Health's Principles | Detailed Commitments | Eticas Assessment for How to Measure Commitments |
|---|---|---|
| Accountability | Focus: On health, no other purposes. | Understanding and knowledge of users about the system purposes, capabilities and their rights upon it. |
| | Funding: Your money pays for the service, not your data. No reuse: We will no reuse or monetise your personal information. | App payment system. Use of app features to collect data and sell or sharing with it third parties (advertisers, other companies, etc.). |
| | Community impact: We will continually assess the impact of our app on your health and the broader community. | Trials on Rem!x usability and acceptability. Social impact assessment. |

[2] Alpha Health Ethics Strategy - as of 02/05/2019.

| Alpha Health's Principles | Detailed Commitments | Eticas Assessment for How to Measure Commitments |
|---|---|---|
| Accountability | Responsiveness: You will always have someone to contact in case of data concerns, to rectify or delete your data or to ask questions about how we use your information. | DPO contact included in privacy policy. Access, rectification, cancellation and objection forms are available to users. Technical and managerial resources to conduct these procedures and interact with users. |
| | External monitoring: We will have regular algorithmic audits, a community board to assess the social impact of our app and an ethics board to ensure all research undertaken follows responsible, ethical and data protection research standards. | External ethics review of Rem!x.<br><br>Algorithmic Impact Assessment (AIA) of Rem!x. |
| Control | Preferences: You will always be able to set your preferences in the app. | Rem!x functions and features for users to set up targeted preferences both in terms of usability and privacy. |
| | Consent: We will always ask for consent for specific uses of your data. If we come up with new services or purposes, we will ask you again for consent. | Explicit and informed consent provided for using Rem!x. DPO contact in privacy policy. Description of special categories of data for users. Updating of consent when data collection (new categories of data) or processing (other purposes) conditions change. |
| | Access: You will always have access to the information we hold about you, your digital identity in the system and how we make decisions to improve your health. Also, you will be able to talk to a person tasked with answering your questions and reply to your access requests. | DPO contact in Privacy Policy and information about the possibility of sending an access request. Technical and logistical mechanisms to exercise ARCO rights. Information on the user digital identity and his/her activity provided by the app. |

| Alpha Health's Principles | Detailed Commitments | Eticas Assessment for How to Measure Commitments |
|---|---|---|
| Transparency | Privacy Policy (PP): We will develop a complete PP. | Adequate information on algorithmic processing and the logic followed by the algorithms to make recommendations<br><br>Include in the PP the typologies of third parties involved in data processing and their segmented role within data processing and the app functioning. Consider the possibility of publishing the full list of third parties in the Privacy Policy.<br><br>Include the third states to which data are likely to be sent and the relevant adequacy decisions issued by the European Commission (in the case of Rem!x, referencing the decision for the US should be enough).<br><br>Include the right to opposition of users to the processing of their data for direct marketing purposes in the Privacy Policy.<br><br>Include a simplified taxonomy of the categories of data processed by the app in order to aid the understanding of the privacy policy. |

| Alpha Health's Principles | Detailed Commitments | Eticas Assessment for How to Measure Commitments |
|---|---|---|
| Transparency | Explainable AI: Our AI systems will remain explainable for users. | Implement AI for the general public - explaining how our products get to their recommendations.<br><br>Such explanation should be reviewed by the legal team and included in the Privacy Policy. |
| Security | Anonymization and pseudonymization: Data will be properly (pseudo)anonimyzed when applicable. | Anonymisation or pseudonymisation should be explained to users. |
| | Special categories of data: concrete safeguards will be taken to protect sensitive data. | Collecting potential biometric data without meaning to - we might be able to assess a person's gait from accelerometer data, and this is considered biometric data- is a risk.<br><br>Gender could also be removed if it is not essential for the functioning of the system |
| | Proportionality and minimization: purpose limitation will be respected and the minimum amount of data needed for these purposes will be collected. | Avoid using those categories of information that are considered as less useful and still present risks for privacy or integrity (i.e. light).<br><br>Only integrate extra sources of data, third parties or historical, at the point where they start to deliver real user value. |

| Alpha Health's Principles | Detailed Commitments | Eticas Assessment for How to Measure Commitments |
|---|---|---|
| Security | Data breaches: measures will be taken to avoid data breaches and users/authorities will be informed about these events in due time. | Encrypt all data before it is processed.<br><br>Undertake regular data security testing.<br><br>Define proactive and reactive protocols data breaches. This should include:<br><br>1) "Just-in-time' mechanisms for alerting users about potential privacy risks.<br><br>2) Communication channels between different teams in order to boost the response capabilities in the event of a data breach. |
| Governance | Internal capabilities: governance tools to ensure secure and ethical treatment of personal data will be established. | Establish better protocols of communication between the teams. Have training sessions for the different teams on core issues:<br><br>- Exchange of information<br><br>- GDPR by design<br><br>- Privacy by design<br><br>- Ethical research<br><br>- Discrimination |

| Alpha Health's Principles | Detailed Commitments | Eticas Assessment for How to Measure Commitments |
|---|---|---|
| Governance | Internal Protocols: governance mechanisms to ensure secure and ethical treatment of personal data will be established. | Develop a data management plan for the full spectrum of using data, from research through to an in-market product.<br><br>Develop an impact assessment to allow Alpha to measure the health impact alongside broader societal impact.<br><br>Integrate socio-cultural background as a variable when recruiting participants for usability testing.<br><br>Include accessibility testing in the research process.<br><br>Test the privacy policy in terms of intelligibility and acceptability. Conducting a Privacy Impact Assessment would be useful at this point.<br><br>Tests features that might lead to "addiction by design". |

Source: Alpha Health and Eticas

# Contextual analysis of REM!X: analysis of ethical and legal frameworks

Eticas' role in this assessment is to establish the ethical grounds of REM!X and its automated processing mechanisms, rather than providing legal advice on whether REM!X is compliant with GDPR. Nevertheless, the principles and requirements embedded in this regulation have been used to analyse the system and develop recommendations for improvement.

## 4.1 DATA PROTECTION PRINCIPLES AND REQUIREMENTS

### 4.1.1 Personal data

#### 4.1.1.1 Anonymisation and pseudonymisation

The anonymisation of data forms part of REM!X's data management plan. In our initial review of REM!X, we identified some potential concerns concerning re-identification risk, particularly concerning the combination of different data obtained by the app, such as location. During the course of the audit, Alpha Health evolved their approach, in part in response to our initial findings, and therefore minimised the type and amount of data collected and processed.

In the initial stage of the audit we found that the type and amount of data to be gathered by the app could easily lead to the identification of users, because the app was, in its initial version, collecting personal data according to GDPR, although it should be said that this collection was properly documented in the REM!X privacy policy. Two direct identifiers would have been used by the system: user ID; and location (GPS data). We identified location data points as the main source of re-identification risks within REM!X. In the app, location data was obtained through the mobile phone sensors and used to provide the user with tailored recommendations. Third parties were not to be given access to this data, which was pseudonymised during data processing.

The accuracy of location data was initially considered to be vital for REM!X's performance, but it was also acknowledged that it involved multiple risks for the users' privacy. Indeed, the e-Privacy Directive defines location data in its article 2(c) as "any data processed in an electronic communications network or by an electronic communications service''. Recital 2 provides more information on the different modalities that location data can adopt. More concretely, it indicates that location data may refer to:

*''the latitude, longitude and altitude of the user's terminal equipment, to the direction of travel, to the level of accuracy of the location information, to the identification of the network cell in which the terminal equipment is located at a certain point in time and to the time the location information was recorded.'' It has been proven that it takes only four location points to accurately identify an individual (De Montjoye, et al., 2013).*

Location information from mobile phones has also shown to be very useful to accurately identify individuals' behaviours and relationships (Dong, 2011). These important findings from the literature show evidence that location data must be taken seriously in matters of privacy. This is particularly important also considering that users of similar recommender systems have shown to be "slightly greater than moderate concern" about the use and retention of their location information by these systems (Hersh and Leporini, 2017).

After receiving our initial advice, Alpha chose to remove location data from REM!X. This was a positive step, however, we were clear that not including this information did not render the information anonymous[3]. Hence, the processes followed within REM!X were framed as pseudo anonymisation[4].

Considering the above, we recommend following the guidelines attached to this report as a general reference for pseudonymisation, but also to consider the measures described below in this document with respect to improving the security of data access, including encryption, as mechanisms to minimise re-identification through the aggregation of data corresponding to one data subject.

[3] Re-identification can also be potentially achieved by aggregating data collected by REM!X. In this regard, it is important to understand that simply riding a database of a name, address and other obvious identifiers does not make a database anonymous if other quasi-identifiers are present in the database. Moreover, complete anonymization is technically impossible. Furthermore, the benefits of anonymization versus the costs it represents for usefulness must be taken into account when anonymizing a database.

[4] Recital 26 GDPR establishes the definition of anonymized data and Article 4 the same for pseudonymisation.

### 4.1.1.2 Special categories of data

The inclusion of sensitive categories of data was addressed as part of the ethics assessment. According to Article 9 of the GDPR, on "Processing of special categories of personal data", special categories of data include:

- Racial or ethnic origin.
- Political opinions.
- Religious or philosophical beliefs.
- Trade union membership.
- Genetic data.
- Biometric data for the purpose of uniquely identifying a natural person.
- Data concerning health or a natural person's sex life and/or sexual orientation.

In the case of REM!X, we identified a possible risk of the collection of biometric data and the classification of users according to their sexual orientation.

In the first case, the data derived from the accelerometer, the location, etc., were to be used to determine the physical and behavioural characteristics of the user and thus fall within the category of biometric data. It is important to understand whether the data to be gathered are biometric data because these are listed as "sensitive data" in the GDPR, which means that their processing is prohibited under article 9(1), though exceptions listed in Article 9(2) apply. Moreover, this type of data requires special safeguards in its processing.

As far as REM!X was concerned, it seemed that the biometric data involved was not meant to identify a person uniquely, but rather to link the user's behaviours with those of others, which could be segmented as algorithmic patterns. Nevertheless, even if data was not being collected to identify a person uniquely, could this still have happened? For instance, the measurement of a person's gait is listed explicitly as an example of the behavioural-based technique. What if REM!X had been able to determine a person's gait? Would it have sought to determine a person's gait? If the purpose was less intrusive than measuring the gait -for instance, determining the pace of the user-, then REM!X should have made sure that no unnecessary biometric data was created. Otherwise, both consent and security mechanisms should be developed accordingly.

In the second case, when signing up for the service, the user was to be required to indicate their gender: Non-binary, male or female. However, this information translated into the initial draft of the privacy policy as “sex”, instead of “gender”. In general, whether it is gender or sex that is being collected must be clarified. Furthermore, the relevance of knowing this information appeared minimal with respect to making recommendations and we proposed that this was removed.

On both of the above points, during the course of the audit, Alpha Health amended REM!X to address these concerns, leading us to conclude that there was no risk of special categories of data being collected.

#### 4.1.1.3 Data proportionality and minimization

An effort was made to minimize the amount of data needed to meet REM!X goals. Still, the initial system we reviewed used multiple information derived from a variety of phone sensors, such as Gyroscope, GPS, or microphone. During the research process we identified that Walk Activity (e.g. running, walking, cycling – using labels from the Google API or iOS), Location data (using the location provided by the operating systems) and data from the Accelerometer were the sources of data with the greatest risk of identifying. Though very useful for the functioning of REM!X, there is always a risk that such data can lead to privacy risks.

We also noted some perhaps less obvious examples of re-identification risk. The main example is light, which showed an imbalance between low relevance in terms of REM!X performance and medium/high risk in terms of tracking/misuse. Though light can be useful to guess whether the person is inside or outside of a closed space, the analysis of other data can lead to a similar result, such as a combination of the gyroscope and GPS data.

In response to our initial recommendations Alpha removed data collection on location and light. Our final assessment was that data minimisation principles were followed by REM!X. In general we recommend that Alpha continue to pay close attention to data minimisation and security principles, and to avoid using those categories of information that are considered as less useful and which present risks for privacy or integrity. This includes only integrating additional sources of data, third parties or historical, if it is vitally needed for the functioning of the system.

## 4.1.2 Users' data protection rights

### 4.1. 2.1 Privacy Policy

The Privacy Policy (PP) is the main instrument with which data controllers inform data subjects about the circumstances that the processing activities involve. We reviewed the REM!X's privacy policy in order to suggest some changes that could foster trust and favour the spread of the app. We have condensed our suggestions in Table 2:

**Table 2. Privacy Policy recommendations**

| Automated data processing |
| --- |
| The only reference to automated data processing was made as a right (not to be subjected to a solely automatic decision), but there was no information about the processing/profiling that is carried out or the logic that it follows. An intelligible and most accurate possible explanation of the role and goals of algorithmic processing should be included in the Privacy Policy. It was recommended to cover the following aspects as part of this explanation:<br>• Problem to be solved by the formula/procedure.<br>• User data used by the algorithm to solve this problem.<br>• Description of the algorithmic processing results. |
| **Data sharing and third parties** |
| In terms of data sharing, the guarantees and safeguards that were provided were very vague and did not even specify the countries to which data could be sent. A far as the guarantees, the PP stated the following: "Alpha siempre certificará la seguridad de tus datos exigiendo alguna de las garantías previstas en la legislación europea a los proveedores con los que se relaciones" (Translation: "Alpha will always certify the security of your data by demanding some of the guarantees provided in European legislation to the suppliers with whom it is related"). The countries are not specified and there is no mention of the adequacy of decisions on the part of the European Commission. According to the list of third parties that was made available to Eticas, data was stored mainly in the EU and the US by third parties, such as external providers storing personal data, which should have been made explicit. |

Regarding the adequacy of the data protection legislation currently in force in the US, the European Commission has issued a decision that should have been referenced in the Privacy Policy[5]. Furthermore, there was no mention of the existence of binding corporate rules that could shed light on the data sharing activities that take place within the company either.

We suggested to include in Privacy Policies the typologies of third parties involved in data processing and their segmented role within data processing and the app functioning. According to the list of third parties, for REM!X these categories could have been:

- data storage,
- communication/customer support,
- app functionalities/interaction,
- tracking/monitoring of users' activity.

Alternatively, the list with all the third parties could be published in the Privacy Policy. This would constitute a very significant act of transparency on the part of Alpha Health since legal compliance can be achieved merely by disclosing the categories of recipients of the personal data (article 13.1.e GDPR).

### Marketing purposes

Marketing purposes: "In addition, we can use your data to communicate and be in contact with you, in order to inform you about our products or services, as well as to improve the knowledge we have about the impact of our advertising through the App, social networks (e.g. Facebook or Instagram) and online banners."[6]

No mention was made here (nor later in the document) of the right of data subjects to oppose the processing of their data for direct marketing purposes (article 21.2 and 21.3 GDPR). Besides making explicit mention to the opposition right, we suggested explaining the existing mechanisms for exercising it, following our guide for ARCO rights (Annex 2).

[5] Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2016.207.01.0001.01.ENG

[6] 'Además, podemos utilizar tus datos para comunicarnos y estar en contacto contigo, con el objetivo de informarte acerca de nuestros productos o servicios, así como para mejorar el conocimiento que tenemos sobre el impacto de nuestra publicidad a través de la App, redes sociales (p. ej Facebook o Instagram), banners online, etc.''

| Types of data and risk assessment |
| --- |
| We also recommended including a simplified taxonomy of the categories of data processed by the app in order to aid the understanding of the privacy policy. The means through which data is gathered could also be stated more clearly. This could have been done by rearranging the structure of the Privacy Policy as suggested in Annex 3, by adding explanatory Icons or illustrations, by providing a simplified Privacy Policy (layered approach to privacy policies recommended by the AEPD) or by combining all of those at the same time. Our proposal for the PP indicates how its layered implementation would work. The objective is for the user to have a clear understanding of the terms and conditions to which they are giving consent. In essence, this implies that whatever means are chosen to inform users, they should be clear on the personal data that is going to be processed, the purposes of the processing and on the legal grounds that will legitimise it. |

Source: own elaboration.

We note that, whilst some aspects of REM!X were updated during the audit process, the privacy policy was not changed before REM!X was discontinued. However, we were able to review the draft privacy policy of another Alpha Health prototype app during the audit. LULL – an app to help improve people's sleep – included a much improved privacy policy – see Appendix 4 for more details.

#### 4.1.2.2 Informed consent

We have addressed the informed element of consent for the processing of individual categories of data. In this section we will review the app from the perspective of consent, a broader concept that encompasses the following dimensions: free, specific, informed and unambiguous[7].

##### Free

According to Guidelines on consent under Regulation 2016/679 from the Article 29 Working Party, ''if consent is bundled up as a non-negotiable part of terms and conditions it is presumed not to have been freely given''. The app stated 3 different purposes in its Privacy Policy: giving recommendations to the user, improve the app and marketing. To our knowledge, users could not refuse the use of their personal data for any of those individual purposes. It was not clear to us that

[7] Source: See Guidelines on consent under Regulation 2016/679 from the Article 29 Working Party.

marketing was necessary for the performance of the contract, and certainly using the performance of a contract as the consent basis for marketing is against best practice and the guidelines. This approach may be problematic in terms of article 7.4 GDPR, in which consent for something that is not necessary for the performance of a contract should not be required. Therefore, we recommended providing the option to opt-out to this purpose. Moreover, the Opinion from the Article 29 Working Group, ''When data processing is done in pursuit of several purposes, the solution to comply with the conditions for valid consent lies in granularity, i.e. the separation of these purposes and obtaining consent for each purpose.'' This provides further arguments for the inclusion of differentiated consent forms for each of the different processing purposes, although this should not be implemented in a way that creates complexity and confusion. We indicated that this could take the form of an unchecked box.

### Informed

Broadly, the GDPR considers that consent has been given in an informed fashion if the user has been provided with information regarding all the aspects present in its articles 13 and 14. This has been tackled in the section above devoted to the privacy policy. A layered privacy policy helps to be accurate and comprehensible when providing information by electronic means as the guidelines establish.

In order to increase transparency in this framework, it was recommended to integrate a tool for users to analyse the list of data shared with the system beyond their profile. This dashboard could include data about what the data shows about the user. This tool could be available in the app.

### Unambiguous

This implies that the user must give consent actively through a clear affirmative act. In the context of the internet, this usually means that the user can only give consent by ticking a previously unticked box. To our knowledge, this was not an issue in REM!X.

### Specific

This dimension of consent is closely linked to the information that is provided to users and to the principle of granularity. According to the guidelines, consent will be specific when the following conditions are met:

- Purpose specification as a safeguard against function creep,
- Granularity in consent requests, and
- Clear separation of information related to obtaining consent for data processing activities from information about other matters.

Therefore, it would be advisable follow best practice and to ask for consent for the different processing activities in a different form and in a way that allows the user to opt out from the ones that are not strictly necessary for the performance of the contract.

### Opt out from one or more types of data

When giving consent, the user had the opportunity to prevent the app from accessing their phone's microphone, location, and photos. However, the privacy policy indicated the use of more data points than microphone, location, and photos. Therefore, we indicated that it would be better for the PP to give more control to the user, by giving them the opportunity to reject more data points.

In general, when a certain type of information is deemed absolutely necessary for the performance of the activities included in the contract, the privacy policy should inform, according to article 13 and 14 of GDPR, ''whether the provision of personal data is a statutory or contractual requirement, or a requirement necessary to enter into a contract, as well as whether the data subject is obliged to provide the personal data and of the possible consequences of failure to provide such data''. For instance, if location data plays a vital role for the functioning of the app (which implies that the user cannot opt out), this should be made explicit in the Privacy Policy.

In the case of REM!X, when looking into the permissions the phone gave to the app, other data were included which were not explicitly part of the consent process: the phone specifications, and data on SMS. Although these were requested before the app asked for consent on personal data processing, these were not specified in the consent tool. These two points should have been added when the user first consented to the use of the app. The user could only react negatively if they later found out that more information than they explicitly accepted was collected.

Furthermore, in the settings for the app, the accessing of the phone's photos was called "storage". This could have led to confusion for

the user. A remedy for that would have been simply to explain what "storage" means exactly before the user consents. For instance "We will only collect the timestamps of your pictures, this will help us to XXXX. Though your phone tells you we access "storage", which includes photos, files and others, we only access the picture timestamps". If this changes, asking for permission again will help the user trust the app more.

To summarise, although REM!X included a reasonably thorough explanation of personal data processing in its consent mechanisms in line with the GDPR, there was some room for improvement concerning best practices. In this regard, the app should have asked for consent for the different processing purposes in a form that is separate from the privacy policy and the terms and conditions. Also, consent should have been given individually for each of the purposes of personal data collection. Users should have been able to opt-out from the processing of their data for purposes that were not essential for the performance of the contract. This included those data that did not put in danger the basic functionality of the app, such as the one connected with marketing purposes.

### 4.1.3 Security and data breaches

Alpha had already developed a set of security systems and protocols for the management of REM!X. These included the appointment of a security officer (DPO), mechanisms for secure data access, including identification and authentication procedures and systems for access control. Security measures also included systems for backup and recovery of data, protocols for ensuring that any support media and document generation are used, security systems for telecommunications, and a set of protocols for security incidents, data management audit and data deletion. Even though the goal of the ethics assessment is not to set the computational mechanisms needed in order to secure data, in this section we provide some inputs about risks and security standards, and will describe further measures required in order to reach best practice in the security standards related to the ethical principles, societal challenges, and legal requirements described above.

In this framework, the main security issues found can be summarised as follows:

- Risks for re-identification of users: Data breaches, misuse. Risks of breaches are always present. It must be considered that the app was able to capture data from multiple sensors and online and offline filters were applied.
- Risks for unfair profiling and discrimination of users: Data breaches, misuse. In this regard the literature notes that "Medical information stored on devices that are lost or stolen may be accessed by malicious users, particularly if information is not secured using encryption."(Huckvale et al, 2015, p. 7).
- Risks for data protection and limitations in the exercise of ARCO rights through design capabilities and technical mechanisms.

It should be noted that unencrypted data storage (of any data) has been identified as a security vulnerability of health apps (Huckvale et al, 2015, p. 7). A study analysing 79 different apps revealed that 73 of them (92 %) presented this problem. But the study also identified that many apps were allowed to send information without encryption and a minor amount of them had unencrypted usernames/passwords or unencrypted personal or sensitive information. Moreover, other less common security vulnerabilities were identified, such as:

- sensitive information sent without encryption,
- username/password captured in network cache or log,
- health-related information sent to third parties,
- fixed device identifier used as user identifier,
- unencrypted access to server-side API,
- access to user data without authorization.

Considering the technical information about REM!X we analysed during the assessment and the above findings of the literature, security risks posed by the App were regarded as low. Still, we recommended the following best practice privacy by design approach, and assessed REM!X against its elements:

1. Data encryption: Since personal identifiers are used by the app, it is recommended to encrypt data before it is processed for machine learning purposes. Data should therefore be encrypted within storage. In the above framework, interoperability with Amazon cloud and security standards of data storage must be confirmed. The database should be protected by

a firewall and data is pseudonymized at the point of transference. Encryption should also be applied to communication of users' data, usernames and passwords as well as all personal data. All personal identifiers should be removed from the data and only aggregation can lead to identification.

In line with this recommendation the Alpha team confirmed the use of encryption.

2. Data security testing: We recommended assessing the effectiveness of the technical and organisational measures for ensuring the security of processing activities. This includes the performance of an Algorithmic Impact Assessment, the on-going studies on usability and the testing of the technical infrastructures around the ARCO step by step (in Annex 2).

This recommendation was addressed through the usability studies conducted by Alpha, the integration of protocols for addressing ARCO rights and the AIA detailed in section 6.

3. Proactive and reactive protocols: Security protocols should be established, such as authorised access to data by users using a Secure Shell (SSH), which authenticates server access with digital certificates and encrypted passwords. Communications between users (phones) and servers should also be encrypted. Moreover, when designing the above safeguards, it is necessary to consider how data aggregation can lead to data subject identification and account for the fact that apps (including REM!X) can access location data through the use of alternative sources which are not GPS, such as Wi-Fi networks. So, we also recommend focussing on the oversight and governance of data location management.

Furthermore, during the audit we pointed out that the software used for locating mobile phones should be regularly subjected to safety checks in order to confirm that they have not been fraudulently accessed. Here, two protocols could be put in place:

- a notification of access could be automatically sent to the user notifying of possible unauthorized access to his or her location data.
- a dashboard could be developed for allowing data subjects to check if their location has been accessed.

In a similar vein, we recommended ensuring the integration of other 'just-in-time' mechanisms for alerting users about potential privacy risks

(Martinez-Perez et al, 2014), which could be aimed at alerting users about unauthorized access attempts.

Following the above recommendations, data used for location was removed, as explained above.

### Data breaches

Protocols for the communication of possible events (such as data breaches) were included in the security form (describing GDPR requirements and protocols for their compliance) used by relevant actors involved in REM!X. As we will describe in the next section, it is advisable to expand the channels of communications available to allow the different teams working on Alpha's products to stay up to date on issues having to do with GDPR compliance. This should be complemented with activities aimed at the training of the staff.

## 4.1.4 Data retention period in REM!X

The retention period for the REM!X app was 12 months after the last use. The same retention period applied to the data collected through the website. REM!X's privacy policy stated that the personal data is to be deleted after that period, unless certain legal requirements arise.

The GDPR states that the retention period must to be kept to a strict minimum. Though the privacy policy complied with the GDPR, the explanation of cases requiring data retention for a longer period was not clear; the way in which the policy was written had the potential to leave the determination of the retention period completely at the will of the legal team or whomever was tasked with determining the retention period in REM!X. In order to tackle this point in general, we recommended including in the Privacy Policy something along the lines of ''In the event of your data being kept for longer than 12 months due to some of the exceptions listed above, we will notify you about it. The notification will include information about the retention period for your data''.

However, the privacy policy also stated that in certain cases, personal data may be anonymised in order to be able to use it for a longer period of time. There were certain issues arising from this statement; firstly, as we explained earlier in this document, complete anonymisation is impossible. Secondly, we understood from our interviews with the Alpha team that anonymisation meant ridding the dataset of obvious identifiers such as name and location data. If the extent of the anonymisation was simply that, then this could have

been defined as pseudo-anonymisation in the privacy policy/consent. Whilst we acknowledge that, under GDPR, pseudo-anonymised data can be described as anonymised if the risk of re-identification is very low, it was not clear to us that such a risk-based analysis had been undertaken to allow this claim to be made. Thirdly, the purpose for which this data would have been used was not stated, neither was the time period for retention. Though this last point is not mandatory provided that the data can genuinely be considered anonymised, it would have been best practice to provide this information. We were told that the anonymised data would be used to conduct research to improve the app; this is a legitimate aim which should have been mentioned as the purpose for the further use of data.

### 4.1.5 DPO (Data Protection Officer)

The DPO must be appointed by those organisations that are either public or that process personal data on a large scale. We believe that the characteristics of REM!X made it necessary to appoint a Data Protection Officer, since ''the core activities of the controller or the processor consist of processing operations which, by virtue of their nature, their scope and/or their purposes, require regular and systematic monitoring of data subjects on a large scale.'' (Article 37.1.b).

In line with the above, the Telefonica DPO acted as DPO for the app. His/her contact details were available in the Privacy Policy. Since this requirement was already addressed, we make the following observations to facilitate compliance with this particular aspect of the legislation in the future:

- The appointment must be made in observance of the relevant articles of GDPR, such as Recital 97, article 27, article 37, article 38.6, article 39. These articles affect matters such as the professional credentials that the appointed DPO must hold, certain incompatibility rules and other respects that we assume have been taken into account at the moment of appointing the current DPO.
- There is mention of a Security Officer in a document that was sent to Eticas. Such a figure does not exist in GDPR, which led us to think that it might have been an error. If 'Security Officer' has been intended as another way to refer to the DPO, this should be corrected in order to avoid confusion. Additionally, the duties would have to be reviewed in order for them to match those established by the GDPR. The Alpha team addressed this point following Eticas recommendation.

## 4.2 CONSIDERATIONS ON BROADER SOCIAL IMPACT

The ethical and legal compliance of technical systems and data management protocols of REM!X does not mean the absence of externalities at the societal level. In this section we comment briefly on possible negative effects that could potentially have been caused by the app in order for the Alpha team to be more aware of their existence and, thus, be able to take a proactive approach to prevention in future apps.

Firstly, as a result of our study, we found that the scope of the REM!X could have been clarified: was it a preventive or a management tool? Was it both? Moreover, the interrelations between happiness and health could have been further developed and shared among the different teams, as it had already been scientifically justified in project documentation. A shared view on this could have helped align the work of Alpha teams and better coordinate their tasks.

The first of the problems that we hypothesised as potentially to have been engendered by the app is addiction. The user could have developed the habit of resorting to the app more than they would have considered necessary or appropriate because of the potential addictive properties of the app, which could have potentially turned the app into a negative influence on the user's wellbeing. For instance, phone addiction has been linked to depression, mainly among teenagers (Young and Rogers, 1998; Bian, and Leung, 2015, Elhai et al. 2017). We cannot know whether this would have been the case or if, conversely, the app would have helped users to get rid of their own pre-existing addictions. It seemed to us that an appropriate solution would have been to test the functioning of the app in regards to this dimension and, if the app was found to be addictive, steps could then have been taken to address this, and the risks made known to users so that they were aware of it and could manage their behaviour accordingly.

Secondly, users could have lost their personal autonomy in terms of deciding how they would like to feel or the means to achieve such a mental state without the constrictions imposed upon them by the app's features and functionalities. It could have also prevented individuals from trying activities or solutions to their problems that might not be recommended by the system, creating a sort of ''echo chamber'' effect in their lives. We do not know if this outcome was likely to come about, but it would have been a good idea to account for this dimension when testing the system. Considering that this dimension is closely linked

to the previous one, they could have both been assessed at the same time.

Lastly, self-tracking apps entail a risk for encouraging a whole array of dysfunctional behaviours, as Gross et al. (2017) show. They argue that the possibility for self-tracking opens up a new performative space for self-tracking, which can wind up not only causing an inappropriate use of the app but also the appearance of conditions such as bulimia, anorexia, or other mental health issues stemming from body image distortions (Idem, p. 331). Since REM!X was not a fitness tracker, its potential for producing those outcomes was probably negligible. However, given that the end goal of the recommendations provided by the app was to improve the wellbeing of the users, they might have been encouraged to hold unrealistic expectations with respect to the degree of control that they have over their emotions and overall satisfaction, or with the desire to feel good all the time. In this regard, self-tracking and behavioural assessment have been shown to produce anxiety in many cases, causing in some instances the reproduction or worsening of the behaviour that the user sought to modify (Calvo and Peters, 2013; Gross et al., 2017).

## 4.3 CONCLUSIONS OF ETHICAL REGISTER OF THE AUDIT

To conclude, we would like to point out that the development of this app, and similar apps by Alpha can be an excellent opportunity to test the effectiveness and suitability of passive sensing-based apps, which present many opportunities but also some issues, among which privacy occupies a unique position (Cornet and Holden, 2017). Moreover, even though this app was not conceived to be part of medical treatment or as a substitute for medical advice, the development team could have explicitly drawn from the lessons that have been extracted from the scientific literature on medical applications (Bakker et al., 2016).

In summary, this audit shows that REM!X followed high ethical standards. Efforts made by the Alpha team in the design of REM!X, which included significant data minimisation, mechanisms for securing data exchanges and a comprehensive informed consent policy, led to an appropriate ethical framework for the app. In terms of improvement, the following recommendations were made to Alpha whilst REM!X was still operating. The status of these issues and recommendations are also detailed.

**Table 3. Summary of recommendations for REM!X tool**

| Issue | Recommendations | Status |
|---|---|---|
| Anonymisation and pseudonymisation | In terms of improvement, it is recommended to follow the guidelines attached to this report as a general reference for anonymisation (Annex 1). | Addressed in final version of the app. |
| | In terms of improvement, consider the security measures in this document oriented to secure data access, including encryption, as mechanisms to avoid re-identification through the aggregation of data corresponding to one data subject. | Addressed in final version of the app. |
| Special categories of data | Although data minimisation has properly been applied, as best practice, REM!X should make sure that no unnecessary biometric data is created. Otherwise, both consent and security mechanisms should be modified accordingly. | Addressed in final version of the app. |
| | Consider refraining from collecting data on sex/gender. We consider that this data is not required for the app to function well. | Gender was removed as data collection/ processing category. |
| Proportionality and minimisation | Avoid using those categories of information that are considered as less useful and still present risks for privacy or integrity. Here we recommended removing the data points “light” and “location” | Addressed in final version of the app. Light and location data were removed. |

| Issue | Recommendation | Action |
|---|---|---|
| Privacy Policy (PP) | Include adequate information on algorithmic processing and the logic followed by the algorithms to make recommendations according to the model suggested in this document. | Not addressed, although lessons applied to another Alpha app, Lull |
| | Include in the PP the typologies of third parties involved in data processing and their segmented role within data processing and the app functioning. Consider the possibility of publishing the full list of third parties in the Privacy Policy. Include the third states to which data are likely to be sent and the relevant adequacy decisions issued by the European Commission (in the case of REM!X, referencing the decision for the US should be enough). | Not addressed, although lessons applied to another Alpha app, Lull |
| | Mention the right to opposition of users to the processing of their data for direct marketing purposes in the Privacy Policy. | Not addressed, although lessons applied to another Alpha app, Lull |
| | We also recommend including a simplified taxonomy of the categories of data processed by the app in order to aid the understanding of the privacy policy. | Not addressed, although lessons applied to another Alpha app, Lull |

| Issue | Recommendation | Action |
|---|---|---|
| Consent | Ask for consent for the different processing activities in a different form and in a way that allows users to give consent in a specific way. In particular, users should be able to not give consent for data that are not essential for the performance of the contract. We do not believe that marketing should be considered as part of performance of a contract, and so should be treated differently. | Not addressed, although lessons applied to another Alpha app, Lull |
| | Add SMS and the phone to the sources of information to which the user agrees. | Addressed in final version of the app consent procedure. |
| | Explain what ''storage'' means before the user gives consent. | Addressed in final version of the app consent form. |
| Security and data breaches | Proactive and reactive protocols, including the already established data encryption measures but also security testing trough attacks. | Addressed in final version of the app. |
| | Ensure the integration of other 'just-in-time' mechanisms for alerting users about potential privacy risks. | To be considered for other apps. |
| | Improve communication channels between different teams in REM!X in order to boost the response capabilities in the event of a data breach. | To be considered for other apps. |

| Issue | Recommendation | Action |
|---|---|---|
| Retention period | The way in which the policy is written had the potential to leave the determination of the retention period completely at the will of the legal team or whoever is tasked with determining the retention period in REM!X. In order to tackle this, we recommended including in the Privacy Policy something along the lines of ''In cases that your data will be kept for longer than 12 months due to some of the exceptions listed above, we will notify you about it. The notification will include information about the retention period for your data''. | Addressed in final version of the app Lull PP. |
| | We have understood from our interviews with the Alpha team that anonymisation meant ridding the dataset of obvious identifiers such as name and location data. If the extent of the anonymisation is simply that, then it this could be defined as pseudo-anonymisation in the privacy policy/consent. | Addressed in final version of the app Lull PP. |
| Data Protection Officer | Make sure that the figure of the ''Security Officer'' is different from the DPO. | Addressed in final version of the app PP. The Security Officer reference was removed. The DPO correspond to Telefonica. |

| Issue | Recommendation | Action |
|---|---|---|
| Social impact | Test the system for addiction. | To be considered for other apps. |
| | Test the system for its effects over personal autonomy. | To be considered for other apps. |
| | Test the system for its possible effects over the promotion of dysfunctional behaviours or unrealistic expectations concerning happiness or mental health. | To be considered for other apps. |
| | Review the lessons coming from the literature on medical apps. | Addressed for REM!X and the other Alpha apps. |

# Research processes underpinning REM!X

As mentioned above, REM!X involved empirical and theoretical research at different levels and from different disciplinary domains (including engineering, psychology, marketing, law, etc.). These research activities were carried out at different stages of the app's development. The diversity of perspectives and actors participating in this process made it necessary to deploy a great effort in coordination and required an effective exploitation of interdisciplinary work. During our fieldwork different concerns about the degree of inter-team sharing of conceptual, legal and normative definitions and requirements to be followed by the system were found. In addition, some doubts and enquiries on the methodologies and best approaches to deal with the project development from an ethical standpoint were raised by different interviewees. This section briefly addresses these two issues by proposing specific recommendations aimed at fostering compliance with the managerial and privacy by-design requirements described in previous sections.

## 5.1 COMMUNICATION AND COORDINATION OF RESEARCH

The main issues found in terms of research methodology related to miscommunication of the output and approach by each team involved in the project. This was revealed during our individual meetings with members of each team, where we were exposed to different views on what REM!X features and functionalities consisted of, or to different levels of knowledge about relevant legal requirements or technical specifications.

We recommended establishing clear protocols of communication between the different teams participating in the development of apps, which would have improved the coherence of the project to create REM!X - and possibly even its functionality. Following this line of argument, the already existing methodologies for data sharing and project planning and development could have been improved by establishing formal and consistent data flows between the research, legal, AI and engineering teams. Regular reporting should have included information on:

- what each team had done during the period,
- what had been completed/concluded,
- instructions on how could these results have been helpful for each of the other teams.

The list of legal requirements for data protection and the Privacy Policy, Terms & Conditions and Consent of the app, based on GDPR analysis and elaborated by the legal team, should have been shared with all teams participating in the project. The same applies to the security form which detailed what information was collected, where it was stored, the risks associated to such data, etc., developed by the compliance team. It should have been regularly updated as well. More positively, the updated versions of the security document (including an exhaustive analysis of GDPR compliance) were accessed by all teams involved, which helped to avoid discrepancies in their understanding about the changing normative definitions. Moreover, according to our interviews, all personnel accessing personal data related to REM!X received training (an introductory explanation) on the latest version of Alpha's GDPR compliance document as well as on the responsibilities and prohibitions of each user and of the Security Officer. Furthermore, prior to every data access, a banner displayed the following text:

AVISO: para acceder a este sistema necesita estar previamente autorizado, estando usted estrictamente limitado al uso indicado en dicha autorización. El acceso no autorizado a este sistema o el uso indebido del mismo está prohibido y es contrario a la legislación vigente. El uso que realice de este sistema puede ser monitorizado.

According to this document "Any personnel outside of Telefónica is STRICTLY FORBIDDEN to be included in the authorised access list, and to access personal data unless a Data Protection Agreement has been signed and the Security Officer has authorised such access. Any personnel outside of Telefónica who has access to any Telefónica Innovacion & Desarollo (TID – Alpha is part of this group within Telefónica) resource involved in REM!X (other than personal data) must be subject to the same security conditions and obligations as Telefónica's own personnel." These protocols, along with the identification and authentication procedures stated in the security form, ensure the efficient monitoring of data access.

Results from each of the tests conducted by the research team, both scientific research about the relations between app indicators and health/happiness variability and also usability of the app, could have been better shared with the rest of the teams. This would have helped to establish a workflow based on shared views on the technical and normative specifications and requirements.

At the same time, the teams should have ensured that the approach to marketing established by the department in charge was in line with the technical capabilities and goals of the system, which required an on-going exchange of information. We recommended improving the approach to the communication of the system goals and capabilities, so they are better aligned with the concepts of purpose limitation and fairness defined in the Privacy Policy.

## 5.2 UNDERTAKING RESPONSIBLE RESEARCH: CONCEPTS AND TRAINING

Besides the requirements of technological design, the empirical research behind REM!X, and indeed any other system development, must be aligned with basic ethical principles on Responsible Research and Innovation (RRI). The Ethical Approval of the research test conducted by Alpha team with the London School of Economics and Political Science on the Reflections app [a different app designed to gather data on wellbeing through surveys], revealed that relevant aspects of responsible research are currently considered as part of Alpha team tests, including informed consent, right to withdraw from research, confidentiality, freedom from harm and human participants (with focus on vulnerable groups). However, with respect to REM!X, other aspects should have been considered. A framework of societal issues that could involve ethical concerns should have been integrated into Alpha methodology for technological design. Embedding concrete societal values in REM!X and putting in place safeguards for the consequences of technological development was a crucial consideration. Based on our experience, besides the already addressed GDPR compliance, we considered the following four dimensions in the analysis of REM!X:

### Ethics

Ethics relates to the values and moral standards guiding the project. These include both societal and individual values. It also refers to the social contract between the state and citizens, which may be reinforced or threatened by technological innovation. We recommended that, from the start, an explicit social impact analysis for REM!X should have been undertaken. Based on Alpha's ethical framework, hypotheses should have been established to guide the development of REM!X through an attempt both to grasp the underlying values that guided a specific piece of existing legislation and to forecast all potential ethical issues that may have been raised by the new project. In REM!X, these ethical issues involved: privacy of users (operationalised in Privacy by Design and by Default), ARCO functionalities and mechanisms, prevention of harm, consideration of social groups, adaptability to vulnerable groups, gender equality and non-discrimination.

## Desirability

The second pillar analysed as part of this roadmap for REM!X development refers to the very need for REM!X. An assessment of desirability should have been required because the raison d'être of new technology is often forgotten, assuming that the incorporation of technological advances in the society is invariably good. This is, however, not always the case: any technological solution must be proportionate to the problem it aims to address. Desirability, therefore, can be achieved through a clear definition of the problem (health and mental diseases) and the solution (recommender system), a careful planning of its implementation (with indicators, maintenance needs, etc.) and a cost-benefit analysis of the system. The costs to be considered are not only economical but also societal. While this type of analysis will not always quantify costs, it is critical decision-making support for designers, which can anticipate possible disruptive exclusionary effects of REM!X or issues related to equity of access to REM!X, within targeted populations. In this regard, APHA Health strategy focuses on the fact that many people have unmet health needs as the primary rationale for this product.

## Acceptability

This element of the analysis accounts for the crucial issue of how citizens perceive, consent to, and adopt REM!X. Perfectly legal apps sometimes need to be withdrawn because they have not been accepted by society due to a variety of reasons, which may range from risks to health, to distrust, privacy, or cultural concerns. Acceptability therefore requires a public debate with an informed user base and the broader public, as well as ensuring that choice, consent and control are accounted for. The different testing stages developed in the next section address this issue. In addition, informed consent must be fully provided during the conducting of interviews, trials, surveys or any other fieldwork activity conducted within the project, and not on an opt-in basis. In this context, the teams and researchers involved must ensure that the value of conducting the research is explained to both participants and researchers. All of this appears to be normal practice within Alpha Health, therefore we consider that acceptability is addressed by Alpha Health through user testing.

## Data management

Finally, many of the concerns highlighted in the previous sections could have been addressed by a socially sustainable data management policy. In other words, data management is where, in practice, often both the problems and the solutions reside. It includes the legal framework for privacy and data protection, data management systems and protocols, as well as broader considerations relating to individual control and consent, methods of anonymisation, and how privacy enhancing mechanisms can be designed into technologies and projects. A Data Management Plan, prior to the development of fieldwork, should have been established for the testing of REM!X, covering specific protocols for data collection, exchange and deletion. This document should have been reviewed be the legal team and by the Ethical Committee, each providing their respective approval.

In order to address the above issues, training for all staff on these requirements and specifications as well as on the ARCO rights protocols is recommended at different points of system development, including inception and conclusion.

## Social impact and acceptability testing

As part of this training and data management programme, we recommended including modules on ethical principles to be integrated "by design" into REM!X, and other Alpha apps, and proposals for their translation into functionalities or features. This training should also address the societal implication of recommender systems, particularly analysing the impact of automation on human rights, accuracy and accountability. This should include three main strategies for integrating ethics into all phases of the Research and Innovation process (agenda setting, project definition, and implementation):

- Instruments promoting research integrity and which establish the ethical values to be followed by the project (codes of conduct, legal requirements and principles, train research integrity...)
- On-going discussion on current security protocols embedding the above ethics values and concerns as well as measures for improvement
- Structures for reflection (ethics board, ethical committees, community advisory boards...)

It is important to note that Alpha Health conducted training with its teams on what can and can't be done from a legal standpoint. Alpha worked with ElevenPaths (the Telefónica Group's global cybersecurity unit), which advised on compliance, and helped with preparations for the GDPR. Alpha still receives advice from ElevenPaths. This is a very good initiative and also included meetings of Eticas and the Alpha legal team to discuss the forms of implementing some of the audit recommendations.

Usability was tested for all the systems connected to the REM!X app. Concept testing, onboarding and degree of acceptance of the app features were examined in this context. Individuals were recruited by gender (male/female), occupation (studying/working), living conditions (at home/not at home), and age. In order to address the possible lack of representativeness within usability, the integration of different sampling methods during different moments of the system development is suggested for future apps. Within this framework, it is recommended to integrate relevant sociocultural variables (gender, disabilities, ethnic/ cultural groups, others) into the usability (and acceptability) tests.

Furthermore, we suggest:

- Testing those features which could potentially lead to "addiction by design": though addiction to an app would increase user retention and be economically preferable, and it could be perceived as helpful in the short-term', this practice is starkly criticised as being unhealthy for users, and we acknowledge is highlighted within Alpha's ethical framework as something that Alpha wishes to avoid
- Include testing of privacy policy, both in terms of intelligibility but also considering acceptability.

## Recruitment

Even though recruitment for REM!X was externalised (using an agency) and Alpha monitors sample bias in this process, it is recommended to assess stigma or discrimination by the contracting party (Telefonica Alpha). The above aspects concerning socio-cultural background in the recruitment variables should have been addressed depending on sampling design.

The analysis of social impact should have been widened to include broader issues beyond functionality for target groups. It is recommended to expand the scope of the societal analysis to adequately address ethical concerns related to the social impact of apps such as REM!X in terms of privacy, health and integrity. Besides reducing the risks of unexpected social effects, such as bias, this task can also contribute to reinforcing the validation process. This means that the actual efficiency and efficacy of technology can be ensured by addressing users' perceptions and interests from the very beginning and integrating outcomes of this assessment into the technological design.

The research team conducted tests to confirm the hypothesis about the different correlations between the measurements obtained from the gathered data and its expected performance as proxies. However, these assessments focused on usability with small sample groups and by considering users' feedback without paying significant attention to broader societal aspects. These aspects include the potential of the system to create stigma, and its adaptability to users belonging to particular cultures, minorities or protected groups. Accessibility and adaptability of the apps concerning people with disabilities may also be assessed in this framework. Along these lines, we recommended including a series of tests aimed at determining the acceptability of the app: expectations of the apps vs users' satisfaction, whether the recommendations were also relevant to people of different cultures or background, or whether the app fostered the inclusion of people with different disabilities (deafness, visually impairment, or certain mental diseases).

**Table 4. Summary of recommendations for Alpha's research approach (for REM!X and in general)**

| Issue | Recommendations |
|---|---|
| Communication and coordination of research | Establish better protocols of communication between the teams. |
| | Have training sessions for the different teams on different issues:<br>• Exchange of information<br>• GDPR by design<br>• Privacy by design<br>• Ethical research<br>• Discrimination. |
| Responsible research methods | Develop an on-going societal impact assessment<br>• Informed consent<br>• Training on GDPR, ethics, desirability and acceptability |
| | Develop a Data Management Plan for each project and address all data protection principles in Article 5 GDPR within it. |
| Methodological aspects in usability and inclusion of social impact test | Integrate socio-cultural background as a variable when recruiting participants for usability testing in a comprehensive manner. |
| | Include accessibility testing in the research process. |
| | Test the privacy policy in terms of intelligibility and acceptability. |
| | Include "addiction by design" in usability testing. |

Source: Own elaboration.

# Algorithmic Impact Assessment: Framing the REM!X AIA: Description and conceptual framework

The main algorithm used in REM!X was a recommender systems based on popularity. Although later versions used recommender systems based on collaborative filtering, the popularity-based algorithm was the subject of this AIA. This system aimed at suggesting activities to people. The output was a ranked list of activities that a user (a) was likely to undertake, and (b) was likely to value positively or have a positive impact on a user's wellbeing. The inputs fed to this algorithm were the different activities that a user had undertaken in the past[8].

REM!X used different algorithms for different submodules of data. The different purposes for which the algorithms were designed and implemented include tailored recommendations (recommender engine) for users and timing (right time engine). As has been thoroughly assessed by the literature (O'Neil, 2016), the use of algorithms with these purposes can involve different risks for the integrity and privacy of users and entail significant challenges in terms of accountability and transparency.

In REM!X one main issue needed to be addressed in this regard: Algorithmic unexpected bias needed to be addressed before and after the app is implemented. Algorithms are biased when "systematically and unfairly discriminate against certain individuals or groups of individuals in favour of others" (Friedman and Nissenbaum 1996).

[8] However, it is possible to design new algorithms in the future that take as input other elements from a user's profile such as gender or age.

Below we develop these issues.

## 6.1 PRELIMINARY ASSESSMENT OF ALGORITHMIC BIAS

### 6.1.1 Gender unfair bias

It is very important to make sure when designing the app not to introduce gender bias in the algorithms. Recommendations should be based on interest, not on what is likely to interest based on what gender the user is.

The following scenario exemplifies gender bias: The female user inputs "I am sad, I want to feel relaxed and I am at home" and receives as a recommendation "Put on some nail polish".


This recommendation is based on the biased assumption that women enjoy painting their nails. Though this may be true of certain women, it should not be a "standard" recommendation for females. Gender-based recommendations can be reductionist and could harm the app's popularity.

The recommendations provided by REM!X were not dependent on sex: No activities required skills that are biologically defined, such as strength. Furthermore, most of the proposals were gender-neutral (e.g., meeting friends, sleeping well or going for a walk). In this context, asking the user for their gender/sex was not required. Gendered recommendations could have been given once a complete profile of the person had been established, and once it could have reasonably been assumed that the person would want a gendered recommendation (e.g., paint your nails).

### 6.1.2 Socio-economic discrimination

While directly asking for economic means could be too intrusive for the user, it is important to take into account the user's economic means when making suggestions. Not everyone can afford to take a Pilates class to relax or eat extra-healthy, as suggested by REM!X.

The economic means of the user might have been established by using data about the district where the user lives, for instance. Based on a rough estimate of the user's economic means, recommendations could have been much more targeted.

It is important, if socio-economic profiling is done, to make the user aware of this, explain the reasons for profiling, and to give them the opportunity to block such profiling or correct it. The idea is to build a relationship of trust, like saying "We are simply trying to target our recommendations for you, this is what we have learned based on X, Y and Z data about you. You can always decide not to give us this data, just know that the recommendations will not be as tailored".

#### Summary of risks related to bias and initial recommendations

Besides the above initial analysis of the app and the two main risks within the development of algorithmic processing for REM!X, related to gender and socioeconomic bias, other (focus, interpretation, transfer or processing) biases may have potentially affected it. These forms of bias related to data input, with the specific learning context of the algorithms or with their outputs, and concerned all sensitive or legally protected attributes, such as race, gender, ethnicity, age or income. Such biases could also lead to wrong or prejudicial recommendations for users, with possible inappropriate/unexpected psychological effects on them, such as the producing an adverse health or emotional condition.

Even though algorithmic bias remains difficult to address, particularly when manifesting through proxies as was the case in REM!X, different actions can be taken to prevent it. In particular, we recommended:

- Defining protected groups for the app from the design phase and evaluating their chances of becoming the subject of bias during the algorithmic process, within each of the potential purposes of the algorithm. Protected characteristics correspond to attributes of people that anti-discrimination law mentions specifically, such as "sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation[9]." Individuals are grouped based on their protected attributes to evaluate the extent to which an algorithm may treat or impact a group differently from another. On the basis of the predefined categories of processed data, protected groups within REM!X were only defined by age (tool design and testing focused on the 16-24 group) and gender. So, risks of algorithmic bias were considered as low. Still, the app could make recommendations on the basis of other uncontrolled variables, so when testing to find models of prevention measures for an algorithmic design we recommend using integrate pre-processing and in-processing methods. It should be noted that in order to evaluate whether unwanted proxy measures are embedded into an algorithm, very sensitive categories of information, such as race, gender, age, ethnicity, will have to be identified and compared to know if the biases are sufficiently minimised.
- Conducting an Algorithmic Impact Assessment (Section 6.2) where religious, race, origin, disability and health condition variables can be tested. This assessment will measure fairness, justice, due process, and disparate impact. The purpose of an evaluation of algorithmic bias is first to detect discriminatory situations and practices, and second, to mitigate these algorithmic behaviours through pre-processing, in-processing, or post-processing methods (Hajian et al., 2016).

### 6.1.3 Explaining Algorithms

In general, it is recommended that the engineering team, jointly with the research team, develop an understandable and as accurate as possible explanation of algorithmic processing for the general public.

[9] Article 21 of the EU Charter of Fundamental Rights.

As indicated above, this explanation should define each of the inputs and outputs of automated decisions and predictive processing. As a method to conduct this, we suggested framing the explanation around the purpose or problem to be solved by the algorithm.

Such explanations should be reviewed by the legal team and included in the Privacy Policy in order to comply with the right to be informed of the data subject, as it is stated in the Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 adopted by the Article 29 Data Protection Working Party. The user has the right to be informed of the existence and logic followed by this type of processing as well as the significance of the envisaged consequences of such processing according to article 13.2.f GDPR.

A further detailed version of this text should be turned into a separate document to be published in an app such as REM!X and accessible for users online, which will contribute to solving problems derived from algorithmic opacity and will help to reduce the number of ARCO requests related to algorithmic processing.

## 6.2 ALGORITHMIC IMPACT ASSESSMENT

As described earlier, the REM!X recommender algorithm was popularity-based, i.e., the most popular activities were recommended first. There was also an element of randomness to generate serendipitous recommendations and for avoiding feedback loops in which popular activities became even more popular. This type of system works by establishing a set of preferences for items by users (Resnik, et al. 1997). In this way, these kinds of algorithms are able to match one user to others by identifying those who have historically had a similar taste or followed similar patterns. This process was at the root of REM!X's algorithmic design.

The scalability of these systems, namely their capacity to process large data sets and the quality of their recommendations constitute two of their major challenges (Sarwar, 2001). The literature has shown how recommendation algorithms tend to introduce biases that favour the most popular options (Abdollahpouri et al., 2019). More often than not, the capacity to recommend options beyond a certain "band of popularity[10]" will determine if a recommender system can introduce users to new options, given that a limited number of choices are likely to be highly popular among many users. In the case of REM!X, users

[10] For a complete description about how the band of popularity work, see: Abdollahpouri, Himan & Burke, Robin & Mobasher, Bamshad. (2017). "Controlling Popularity Bias in Learning-to-Rank Recommendation". RecSys.

were presented with a big range of recommendations, although they were suggested on the basis of how a series of categories correlated, such as current and expected status, practice or mood.

These suggestions could be unfair[11], not only inaccurate. This could be the result of the association of specific social groups to certain tastes, practices and attributes, which can lead to discriminatory or biased recommendations. For instance, to take a non-REM!X example of movie recommendations, if a group has a strong preference for a certain type of movie, whereas another group would rather go for a different genre, the movies favoured by the group with the strongest preference will be recommended more often to all users (Tsintzou et al., 2018). Machine learning can establish such associations based on the previously mentioned reproduction of historical selection trends by protected groups or collectives. Such processes may be based on discriminatory assumptions and stereotypes. If certain social groups are more inclined towards specific recommendations or are more likely to have certain preferences (for instance, women looking for certain jobs), this may determine the kind of recommendations that the system suggests to them (worse-paying jobs than men).

Bias in REM!X was defined by Eticas as any kind of unfair or discriminatory recommendation made on the basis of protected attributes such as race or gender. We define algorithmic discrimination or algorithmic bias as disadvantageous differential treatment of (or impact on) an already disadvantaged group. These disadvantaged groups can be defined in relation to the above indicated attributes mentioned in Article 21 (Non-discrimination), of the EU Charter of Fundamental Rights. Groups defined by these attributes are therefore potentially subjected to algorithmic discrimination when the systems "systematically and unfairly discriminate against certain individuals or groups of individuals in favour of others" (Friedman and Nissenbaum 1996).

These protected groups can be either legally protected (e.g., people with disabilities) or not, for instance in the case of the participation of women or minorities who might be underrepresented in certain positions. It should be noted that depending on the system, its

[11] Fairness in recommender systems is an on-going line of research (see, e.g., the FATREC Workshop at RecSys'18 https://piret.gitlab.io/fatrec2018/). In all cases, fairness begins with awareness, meaning that it is necessary to know which are the recommendations that different groups of people are receiving to evaluate if biases or discrimination have been introduced.

characteristics and goals, protected groups might be defined by the intersection of two or more variables, such as “gay people with a minority ethnic background”. The criteria according to which bias is defined also need to be framed from a social and ethical standpoint, since some attributes may be legal and considered legitimate for differential treatment, but still considered discriminatory in some social contexts due to cultural or ethical reasons. For instance, religious belief could be used as a variable for the purposes of assigning schools by means of an algorithm used by the social services. This could be done by following the GDPR’s security standards concerning the treatment of special categories of data when the individuals whose data are being processed are in the European Union. However, that does not mean that such processing does not present ethical and political challenges, such as the fact that it could contribute to the perpetuation of socioeconomic inequality.

### 6.2.1 AIA methodology

In order to establish whether algorithmic decision-making is based on unfair grounds and can lead to discriminatory outcomes, pre-processing, in-processing, and post-processing methods can be applied (Hajian et al., 2016). During the algorithmic design stage, developers should minimize risks either by eliminating categories involving protected groups, when they are not needed for achieving the purposes of the system, or removing possible discriminatory links between recommendations and protected groups. This can be done, for instance, by ensuring that the training data contains enough examples involving members of protected groups, and does not contain discriminatory associations.

So, one form of reducing such risks would be eliminating data corresponding to protected attributes. Going back to the example of gender, the gender of users could be removed as a data collection and processing category. However, this option presents two main problems. First, many systems cannot deliver precise and useful outputs without such information. Therefore, some categories of data concerning protected groups must be collected and analysed in order to produce targeted outputs. Second, not collecting data on gender or other protected attributes can make it challenging if not impossible to identify bias once the system has been implemented and machine learning has been deployed, as was in fact the case with REM!X. REM!X teams eliminated data, data points and other data

sources complementing user's interactions in order to comply with the principle of data minimisation and in observance of Eticas' ethical recommendations. These included some data gathered by sensors as well as information related to gender, education, age or geolocation. As a result of that, examining the algorithm to find bias became a very challenging endeavour.

Taking the above into consideration, the AIA methodology that was planned to be applied to REM!X in order to analyse the above aspects of algorithmic processing consist of the following four main steps.

1. Assign the data about individuals into groups:

In this phase, the data processed by the system is classified into groups that can be overlapping ("soft" assignment) or non-overlapping ("hard" assignment). Such groups will concern different categories of people or social groups which will be formed according to certain individual characteristics, especially those considered to be protected attributes. While in a soft assignment an individual can belong to several categories, a hard assignment will be characterised by the fact that individuals are categorised into closed categories.

2. Define a protected group(s):

In the second part of the assessment, the groups defined as protected are clearly determined and selected for their monitoring.

3. Determine a set of metrics aimed at measuring bias

The third step consists of determining the set of metrics to be used in the analysis. In general, these metrics quantify the extent to which an algorithm treats people differently (disparate treatment) and the extent to which an algorithm has a different impact on different segments of the population (disparate impact). There are multiple and often conflicting definitions of metrics that should be used to evaluate algorithmic bias. Therefore, it is necessary to choose one that is coherent. The two most common approaches are the following: the first approach assumes that there are only two possible outcomes (a positive or favourable outcome, and a negative or unfavourable outcome), while the second one attempts to order outcomes from most positive to most negative (e.g., in the case of salaries).

In the case of systems used to allocate a benefit such as REM!X, the proportion of people that receive negative/unfavourable outcome across groups, which is a measure of risk, should be equal if we want

to claim that the algorithm entails equal risks for protected and non-protected populations. In addition, error rates should be similar across groups, which mean that they should not be concentrated in the protected group. Additionally, in the case of systems used to allocate a benefit, measures aimed at ensuring consistency in the treatment of people should result in similar people being treated similarly. Similarity is defined exclusively on the basis of non-protected attributes. Hence, any difference in the way they are impacted can be attributed to their protected attributes, which is to be avoided as it constitutes an instance of algorithmic bias.

### 4. Measure and compare across groups.

In the fourth step, the data is analysed to obtain values and confidence intervals for these measurements. If the data goes through several steps in a system (such as data collection and data analysis), which is normally the case, the analysis needs to be carried out for each step separately. The computation of metrics is done by using a combination of existing libraries, which are general-purpose, and custom code for a particular purpose. The existing libraries used for assessing REM!X included Aequitas, developed by researchers at the University of Chicago, and AI Fairness 360, developed by IBM. After the measures and confidence intervals are computed, any disparity is noted, analysed, and reported; they constitute potential discriminatory situations in the data. When differential treatment follows a pattern, this structure may constitute a potential discriminatory practice.

If discrimination is established through this procedure, the disparities may be addressed through various mitigation measures. These measures are context-specific. To choose the mitigation measures to be applied, first a metric to affect and a target value for such a metric must be decided (e.g., the proportion of people receiving the beneficial or positive outcome should not differ by more than 20%). Then, a pre-, post, or in-processing modification of an algorithm is applied to obtain the desired outcome. An algorithm that does not discriminate on the basis of sensitive categories of data is both more desirable and likely to be accepted by society than one that does. Thus, a report including measurements evidencing this non-discriminatory behaviour is a key element of algorithmic accountability.

## 6.2.2 Fieldwork of the REM!X algorithm

In this section a brief explanation of the fieldwork activities conducted as part of the REM!X algorithmic impact assessment will be provided. Second, on the basis of the collected information a contextual analysis of the relevant social factors that could potentially lead to bias will be developed. Finally, in the third part of this section we will explain the extent to which the analysed information can be considered as indirect evidence of bias within the App.

### 6.2.2.1 Fieldwork activities

The qualitative analysis conducted for this assessment was mainly based on four data collection tools, i) the analysis of the recommendations provided by the App, ii) desk research consisting of the assessment of documents, iii) a series of interviews with different Alpha teams involved in the design and development of REM!X, and iv) a digital ethnography using the Messages of users in Google Apps, which were categorised and analysed to improve the understanding of its performance and social implications.

Concerning the interviews with the Alpha teams, the following meetings took place:

- Meeting held on the 14th of June 2019: This inception meeting was attended by the Eticas team and experts from the research, project and engineering areas from Alpha. The research grounds behind algorithmic processing and the methods and aims of the AIA were discussed in this occasion, as well as the next steps to be taken in the future.
- Meeting held on the 20th of June 2019: This second meeting was attended by Eticas R&C, Carlos Castillo from the Pompeu Fabra University and Alpha. Issues of a technical nature were discussed, such as the description of the deployed algorithms, the identification of possible proxies and data processed by the system and how they could lead to discriminatory outcomes for individuals belonging to protected groups, and future AI developments in REM!X. In this regard, the team developing emotion detection technology (including machine learning algorithms to analyse and interpret emotional content) explained how their outcomes could be embedded into REM!X in the future. Only one specific research document was examined in this context, concerning the how long users used REM!X. However, this study only differentiated data

retention periods concerning gender and age groups and its findings did not show any relevant difference which could lead to infer bias in a valid way.

- Meeting held on the 19th of July 2019: The meeting was attended by Martin Zamorano (Eticas) and Sarah Shepherd (Alpha). On this occasion, the work conducted by the user research team concerning REM!X was discussed, focusing on its aims, the structure of the sample used for the focus groups and its possible findings concerning protected groups. The different tests of the App were conducted with fewer than 8 participants and the variables of analysis did not include any information concerning protected attributes. Tests included usability/concept testing using a script or "ethnographic" research based on interviews conducted in 2017. Participants were recruited by an agency on behalf of Telefonica and, even though Alpha always asks for an even gender split when doing user research, this is not always achieved (people don't show, etc.). Other criteria for participation were age - between 16 and 24 years old - and occupation. The research activities were not intended to find out whether the app had gender bias and information indicating this was not found.

- Meeting held on 20th of August 2019 and follow-up communication. Finally, an iterative dialogue with Miquel Ferrer, Oliver Smith (both Alpha) and Carlos Castillo (UPF) was established in order to examine other possible sources of information for analysing bias in the system, including:

    - Previous studies about the system, including user experience concerning different social groups. We found that there were no useful studies or documents reflecting this on the basis of which we could make verifiable inferences, test hypotheses or derive estimates. We also found that there were no studies about the system analysing the requested categories of users or analysing social impact related to human-machine interaction.

    - Information collected and processed by REM!X during its first stage of development (prototyping), when gender and possibly other protected attributes (age) were integrated. The idea behind this was to access these data and analyse the relations between the social groups and their interaction with the App. However, due to changes in the design of

the app, none of the collected training datasets included protected attributes.

- Information about the data and sample used for training Algorithms. Any documents explaining in detail the sociodemographics and the criteria applied to the selection were sought in this context. Rather than the ability of these data to accurately represent a target population, we wanted to examine potential bias introduced into the algorithm and develop hypotheses about the above-mentioned "band of popularity".

Overall, the only available sources of information that were useful to analyse the impact of algorithmic processing in terms of bias or differential impact were the App data retention information and the messages sent by users using Google App. On these bases, we will now address the societal factors that might lead to bias in REM!X.

### 6.2.2.2 Hypothesis of bias in REM!X

Bias in REM!X was understood as the provision of recommendations which unfairly discriminated against people on the basis of the protected attributes described above. Disparate impact on a disadvantaged group should, therefore, have been explained by the protected characteristics of the users who were being discriminated against. Even though we should keep the above legal definitions as a reference, what can be considered as a disadvantaged group and disadvantageous treatment must be contextualised within each technological project and social setting in order to properly consider the norms or representations that can lead to algorithmic discrimination. Moreover, it should be noted that most recommender systems are based on the associations between individuals as part of (protected) groups and use collaborative profiling to construct targeted outcomes. Therefore, our qualitative analysis addressed the ethical implications of each individual recommendation and also their proportionality in terms of contextual factors, such as norms, symbolic considerations or ethical grounds, defining whether a recommendation is socially acceptable in itself.

In REM!X, different elements were considered in this regard. It was an App used to inform Alpha's work internationally, which required balancing universal ethical standards, which could be assimilated with human rights, with the recognition of national or cultural differences[13].

On a different note, as we mentioned above, protected categories were not collected nor processed by the app. This, combined with the lack of indirect and quantitative evidence, led us to examine the above information and the recommendations primarily from a qualitative standpoint. As part of this analysis, we defined four hypothetical domains that might be problematic in terms of algorithmic discrimination as explained above.

### 6.2.2.3 Socioeconomic barriers and implications

Even though the app was 100% free[14] and REM!X was not collecting information about the socio-economic status of users, all recommendations provided by REM!X had certain socio-economic implications in terms of access and/or inequality reproduction. Access related to the capacity of users to effectively carry out the recommended activities, since they could certainly be beyond their means. For instance, activities such as "Cómprate un despertador"("buy an alarm clock"), "Improvisa una escapada" ("improvise a getaway"), or "Al paso, al trote o al galope", ("stepping, trotting or galloping") as a recommendation to go horse riding, or others, could have been unaffordable for various users. This had the potential to lead to property-based discrimination, which could also have been deepened and expanded by machine learning, since some racial, gender or religious groups could have preferences that are correlated with socio-economic factors. Additionally, distributing recommendations based on popularity could favour the reproduction of socioeconomic status of certain groups.

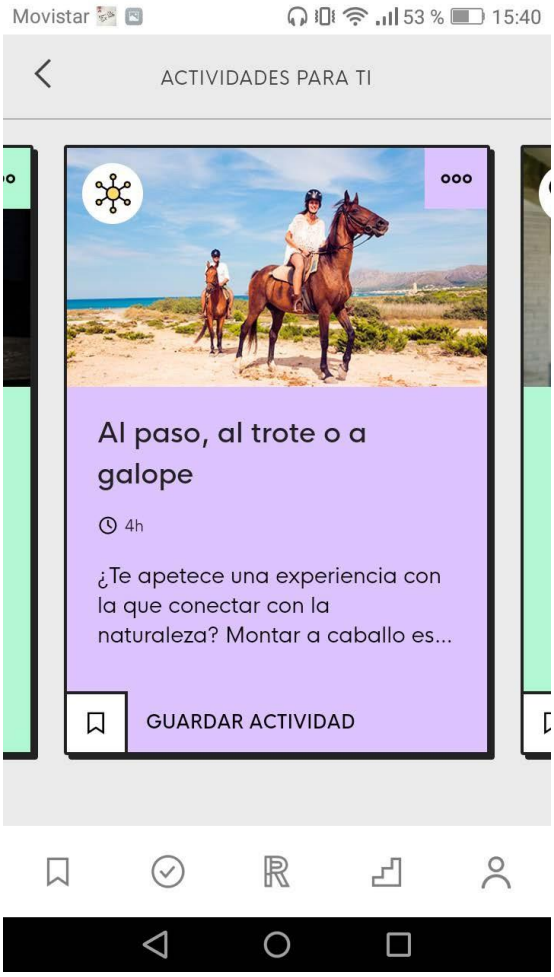
As Eubanks (2018) and O'Neil (2016) have noted, algorithms have the potential for increasing economic inequality and discrimination because of socio-economic status. Social practices and power relations are embedded into algorithmic decision-making systems in different

[13] For instance, should images of women used in the App be adapted to Islam in order to be used in Muslim majority countries?

[14] Even though the app is free, some comments on Google Play show some trends such as criticisms concerning the use of the App to make profit: "La app está hecha con mucho gusto y la idea es buena. Mi opinión es que se nota demasiado que está hecha para hacer publicidad. Es un poco contradictorio porque al principio piensas que te van a dar ideas para afrontar cosas serias como la ansiedad, la tristeza, la soledad... Y en realidad es una app que te ofrece sitios de los que ha recibido dinero para dar publicidad y que poco tiene que ver con tus problemas mentales." ... "The app is made with good taste and the idea is good. My opinion is that it evident that it is made to advertise. It is a bit contradictory because at first you think they will give you ideas to deal with serious things like anxiety, sadness, loneliness ... And in reality it is an app that offers you sites from which you have received money to advertise and that has little to do with your mental problems."

ways and end up amplifying biases present in human decision-making. In Eubanks' book, the case studies show how algorithms can be used to enforce decisions that take an actual toll on these people's ability to make ends meet. Even though this was not the case in REM!X as recommendations were voluntary by definition, it is true that they could have had an influence in a way that discriminates against people on the basis of their socio-economic status in the following ways:

- If recommendations are tailored to the user's economic status, they are likely to be either more affordable than the ones shown to other individuals or more attuned to the tastes and preferences of people who are in the same socio-economic category, which can contribute to perpetuating their situation by means of not exposing them to alternative realities.
- If they are not, that could also be negative as recommendations could encourage people with financial problems to live beyond their means or make them aware of their lack of financial resources, thus causing psychological suffering.



Having said all that, it seems evident that we are at a crossroads. On the one hand, customising recommendations on the basis of socio-economic status can lead to a sort of economic segregation. On the other hand, its absence can also have financial and psychological negative consequences. In this case, the solution would seem to involve striking a balance based on the principle of proportionality. Pros and cons need to be weighed, and once the most beneficial approach has been chosen, mitigation measures need to be put in place in order to further minimise negative impacts or even to contribute to lower the pain caused by a lack of financial means.

Even though these hypotheses were considered, they could not be properly proven either with the available set of user experience data, nor with the datasets and categories processed by the algorithms.

### 6.2.2.4 Cultural barriers and implications

Cultural, religious or ethnic factors have been shown to be problematic for some algorithmic models. Just to describe one example, the algorithms used by Google and Facebook discriminated against people on the basis of anti-Semite inputs. In the case of Facebook, it was revealed that advertisers were able to target Facebook users by taking their view of Jews into consideration. According to ProPublica[15], Facebook advertisements included organizations or content related to the SS, the Nazi Party, and Germany's far-right National Democratic Party. The company has implemented a policy to remove ads related to hate speech (Schindler, 2017).

REM!X introduced users to several activities with cultural implications. In particular, the App suggested activities, habits or goods that were linked to race, religion, or ethnicity. For instance, it suggested enjoying Christmas, but it did not do the same with other religious celebrations, which clearly made it less sensitive to the expectations of religious minorities. Likewise, we could entertain the idea that the algorithm was more likely to offer these kinds of recommendations to individuals belonging to groups that tend to prefer them over other options. Even though this sort of discrimination cannot be considered as derived from algorithmic processing it should have been considered in order to frame the app's approach to these issues, especially if REMI!X were to be offered in countries where Christianity is not the main religion.

[15] Detailed information at: https://www.propublica.org/article/facebook-enabled-advertisers-to-reach-jew-haters

In addition, the use of certain dialects as opposed to others could have alienated certain users, or at least worsened their user experience. One example of this was found in a comment by one user in Google Apps on the issue of adapting the app to the Spanish spoken in Latin America.

> *"I love this application, I am from Argentina and I was able to download and use it. The only thing that I would change is to modify some words and adapt it more to Latin America"*
>
> *"Me encanta esta app, soy de argentina y la pude descargar y usarla. Lo unico que le cambiaria serían algunas palabras y adaptarla más a latam"*

However, even though the barrier posed by dialects might be seen as discriminatory, it was not defined or boosted by the use of algorithms.

### 6.2.2.5 Gendered recommendations and gender inclusiveness

In terms of gender, some of the activities recommended by REM!X could have been considered as gendered since they followed dominant social practices and stereotypes in specific social contexts, e.g. women painting their nails. In order to analyse this issue, it is important to distinguish between historical bias in algorithmic processing (which involves models that reproduce gender bias existing in society), and other models based on normative interference. An example of historical bias happened with Google's algorithm for advertising that proved to offer jobs with higher wages to men than to women (Moore, 2018). Normative interference means the introduction of recommendations that challenge cultural practices, structures or representations, and that are based on human rights or ethical considerations. In order to do this, a social contextually-based analysis must be conducted.

In this regard, even though the "painting nails" recommendation with the picture of a woman's hands could have been considered as a sexist recommendation, we could not tell whether this recommendation was actually shown disproportionately more commonly to women than to men since we lacked the necessary information. To illustrate the importance of weighting these issues it is worth noting an example within REM!X of what could have been considered a normative interference (the use of a more inclusive i.e. non-gendered version

of Spanish), which, whilst it sought inclusiveness, actually fostered negative feedback in the form of comments about REM!X on the Google App Store, where up to 8 comments included negative statements concerning such inclusive language as opposed to a single comment in which inclusive language was evaluated positively. Although these comments would have been representative of a certain population Alpha could still have decided to keep the inclusive model in place, and indeed this was the case until REM!X was discontinued.

On a different note, after performing an ethnographic analysis of the comments left by users on Google AppStore, Eticas realised that women could have been overrepresented among users. Up to 61 comments were made by users identified as female out of a total of 206 comments. Comments made by users identified as male were less numerous (about 23). This may suggest that the app was more popular among women, which could have caused recommendations to become attuned to female preferences, which in turn could have made the application even less attractive to male users. Potentially, this feedback loop could have turned REM!X into an app used mainly by women in practice. This would have eliminated issues of gender discrimination within the app as differential treatment received by men would cease to be a major concern, but it would certainly have come at the cost to the level of inclusiveness of the app.

### 6.2.2.6 Accuracy of the recommendations

The amount and type of data processing categories can also affect the quality of the recommendations made in terms of accuracy. This is also related to all preliminary studies conducted during the design of the application, during the training of algorithms, and during their testing in the lab or with sampled data. If at any of these steps a group was excluded or not represented sufficiently, it would not be possible to claim that the app worked accurately for members of that group.

In the case of REM!X, data minimisation involved both the reduction of data points and of the categories of data to be used for profiling users and suggesting recommendations. This minimisation of personal data collection was recommended in Eticas's initial analysis of REM!X,

[16] The criteria used to classify users between female and male were their name and the picture. Comments made by users whose accounts have a name that is not commonly given to men or women and with no pictures were not taken into consideration, although in some cases we took into account the way in which they decided adjectives with which users described themselves. For instance: 'estoy cansada' (I am tired) agrees in gender with the subject in the Spanish language.

in general and in particular for data on gender and location. As mentioned above, while not having certain information can minimise the risk of discrimination, it can also reduce the capabilities of the system to provide customised recommendations to users. In this way, some comments made on the Google App Store underlined the need to enhance the accuracy of the system in order to recommend useful activities. In particular, the lack of data about location seemed to have negatively affected the accuracy of the system:

*"La aplicacion promete más de lo que da. Pones según un listado tu estado de ánimo, dónde estás y cómo te quieres sentir. A partir de eso te recomienda varias actividades, nada prácticas para realizar en el momento. Por ejemplo, le pones que estás en el transporte público y te recomienda que ordenes tu cuarto o hagas yoga. No le encuentro el sentido a la aplicación, me esperaba algo mejor y sobre todo, posible de realizar con el móvil."*

*"The app promises more than it gives. You list your mood, where you are and how you want to feel. From that, it recommends several activities, nothing practical to do at the moment. For example, you put that you are on public transport and it recommends tidying your room or doing yoga. I do not see the sense of the application, I expected something better and above all, possible to perform with the mobile phone."*

Still, the trade-off between data protection rights and reputational risks derived from unfair discrimination, on the one hand, and accuracy on the other hand, seemed to be balanced, since one of the aims of the system was to be based on ethical grounds and, at the same time, the app was able to reach most of its aims without having to integrate extra categories of data.

Furthermore, Eticas found three comments on Google AppStore that were critical of the amount of personal data collected by REM!X as part of the analysis performed on the comments section of the app. This may indicate that a certain minority of informed users had concerns about privacy that should have been addressed in some way. Maybe that could have been done by further explaining the rationale behind data collection to users, as well as by reassuring them in terms of how their data was being processed and managed.

## 6.3 CONCLUSIONS OF THE AIA

The conclusion of the first part of the AIA for REM!X is that the system

was able to achieve good performance while reducing the amount of data (and particularly sensitive data) to a minimum. No data on gender, race, religion or other protected attributes were collected or processed by the algorithm. These variables were neither used to train the system (training data) nor to assess its performance and social impact once in operation. As was said above, this fact had two main implications in terms of machine learning. On the one hand, it minimised the potential discriminatory consequences of the algorithmic processing outcomes. On the other hand, it boosted opacity in terms of the capacity for identifying differential impact, particularly considering that it was a collaborative system.

We were not able to undertake the second element of the AIA (definition of the protected group) since even though protected groups could be inferred, the information about them could not be correlated to other variables or proxies. This is why most of the fieldwork activities were oriented towards finding indirect evidence or sources of bias to determine the strategy to be followed. Based on these indirect sources of information, we developed a series of hypotheses about the risk of bias in order to consider the need for conducting a trial to collect direct evidence of bias and test it based on further (protected) data provided by a restricted sample of users.

As can be seen in Table 5, three main hypotheses were developed based on the available data, but no significant concerns were identified:

- As far as property discrimination is concerned, risks were low since only a few recommendations were relatively expensive for users and the system was not designed to distribute benefit, which means that was not allocating or limiting material resources to specific groups.
- In terms of religious or cultural discrimination, we could say that, although it could certainly have happened, it was not expected that the algorithm was capable of recommending items according to a "band of popularity" based on linguistic or religious grounds. Religious implications within the recommendations were very few and the issue concerning the dialectal variants of the languages used did not represent a major barrier for users.
- Finally, the potential for gender bias was only detected in the way in which certain activities or challenges were recommended to women based on historical bias. But again, since REM!X performed well, and it did not seem to be encountering major ethical challenges, the risk of reputational losses seemed to be low, especially in the case of gender discrimination.

**Table 5. Hypothesis about bias in REM!X based on indirect sources of information**

| Protected attributes | Sources | Societal implications | Bias potential implications |
|---|---|---|---|
| Property | • Recommendations analysis | • In general, activities were accessible for a general public<br>• Perpetuate socioeconomic differences<br>• Psychological suffering | • Property-based discrimination<br>• Inequality reproduction<br>• Low risk |
| Religion<br><br>Ethnic grounds | • Recommendations analysis<br>• Comments online | • Only one case was identified (Christmas)<br>• Adaptation to dialectal form might not be proportional | • Religious discrimination<br>• Language based discrimination |
| Gender | • Recommendations analysis<br>• Comments of users online<br>• Interviews with Alpha teams<br>• Report on the App retention | • Gendered practices based on historical distribution of gender roles (did not involve explicit male chauvinism)<br>• Possible majority of female users | • Historical bias<br>• Low acceptability and reputational risks<br>• Feedback loop |

Source: Own elaboration.

The last aspect considered was the impact that the lack of personal data belonging to special categories of information could have had on the app's performance and its capacity to provide targeted and customised recommendations. In this regard, we consider that the right balance between the risks associated with collecting sensitive data (such as location data) and providing more accurate recommendations was struck.

Bearing the above in mind, Eticas did not regard it as mandatory to conduct a trial aimed at testing bias in human-machine interaction, since this would not have been proportionate considering the low level of risk that was found, and that no evidence pointing to high risks was identified. Such a trail would have meant gathering sensitive information from a representative sample of users in order to measure disparate treatment and impact, which would itself have created unnecessary risks concerning privacy and data protection.

Concerning the methodology, one of the main lessons drawn from this part of the assessment is that the analysis of algorithmic bias necessitates the use of a series of methods and fieldwork activities from the very beginning (design and pre-processing phases) (Galdon et al., 2020). This is in line with recent findings regarding the needs of industry practitioners regarding algorithmic fairness (Holstein et al., 2019).

Additionally, even in cases where protected categories of data are not meant to be collected and processed on a long term basis, it is important to gather relevant information about the training data sample and human-machine interaction, including protected attributes, which can then be used to infer potential bias before the system's deployment.

On a slightly different note, the analysis of bias in recommender systems can also concern the analysis of the actual recommendations; which means that their narrative, assumptions, and integrated stereotypes should be examined. This analysis is in general context and domain-dependent.

Taking this into account, three main measures have been recommended for future projects:

1. The Alpha team could establish a set of pre-processing mechanisms for testing bias and develop a set of cross-checks and milestones for each project.

2. In the same manner and depending on the project at hand, it is recommended to carry out user experience tests with small-scale samples (more than 20) and algorithmic trials with large-scale samples (more than 200) in order to test bias during the system's development and once the system is in operation.

3. Lastly, we recommend that in future projects, recommendations be analysed by an anthropologist/sociologist, and assessed separately through empirical studies in order to inquire on their possible biased assumptions properly.

# Conclusions of the Ethics Audit

This report summarises our audit and analysis, in which was found that the REM!X app proactively and adequately addressed data protection practices and principles through data minimisation techniques and by applying proportional security standards. These include encryption, pseudonymisation and the locking of logging systems. REM!X also addressed relevant questions on accountability and users' rights through a comprehensive informed consent and privacy policy, which later integrated Eticas inputs aimed at making opt-in more targeted. The app adopted a proactive approach towards minimising potential negative impacts on users' rights by tackling the risks of discrimination and biases throughout the reduction of sensitive data collection, while keeping a high level of accuracy in recommendations.

Overall, it is our judgement that through this audit that REM!X demonstrated an adequate application of ethical standards. Moreover we have been able to consider our findings against Alpha's final published set of five principles and ten commitments. This assessment can be summarised as follows:

| Principles | Alpha Commitments | Compliance in REM!X |
|---|---|---|
| Improving your health and happiness | We aim to support you to have the greatest possible health and happiness and we will never try to increase one if it will significantly reduce the other | Aims and capabilities of the app were properly presented to users. It was recommended to further explain to users the associations between happiness and health. |
| | We will ensure that our recommendations are not based on gender, race, ethnicity, sexuality, age, belief, or other characteristic protected under anti-discrimination law, unless there is evidence that demonstrates that such a characteristic is an important driver of the desired outcome of the recommendation | The AIA did not find direct evidence of unfair discrimination regarding users' sensitive attributes. Measures such as data minimisation were implemented so as to achieve this goal. |

| Principles | Alpha Commitments | Compliance in REM!X |
|---|---|---|
| Putting you in control | You will always know and control how we use your personal data | Users control over the data they provide was considered adequate. The app's Privacy Policy included control access and opt-out mechanisms, as well as the accomplishment of consent legal requirements. Some improvements were introduced on the basis of the Audit recommendations |
| | We will deploy the best available techniques to prevent any user from becoming addicted to any of our services | Usability tests had been conducted and this variable (addiction) had been considered. |
| Being understandable and transparent | We will explain how our services work to support you in having the greatest possible health and happiness. We will ensure that such explanations are comprehensible | Informed consent was required so as to use the app and the information provided was clear about the purposes of the app and its data management aspects. |
| | We will publish the ethical approvals of our research and external audits of our work, although we may redact commercially sensitive information | This publication responds to this criterion. |

| Principles | Alpha Commitments | Compliance in REM!X |
|---|---|---|
| Securing your data | We shall share only the minimum data about you on our internal networks, and when we do this we will use the highest industry standards of encryption to protect your data | Data minimisation practices were followed, and were improved after Eticas' initial recommendations. Sensitive data collection was reduced. |
| | We will create services that preserve as much privacy as possible, for you and your community | Overall the app was privacy friendly, due to its security mechanisms and transparency. |
| Being accountable | We will never generate revenue through advertising | No revenue was obtained from the app, although this was not completely clear informed to users. |
| | We will hold regular external audits of our algorithms [developed past a certain point] to test that it is doing its job and avoiding biases and other unintended consequences | This audit and its publication respond to this criterion. |

Source: Alpha Health and Eticas.

Overall, we believe that the audit process has allowed Alpha to improve its own Ethics Strategy by integrating both new concepts, new forms of operationalisation of certain ethical principles and also detailed protocols for its organization. This last point concerns, for instance, the establishment of an ARCO rights policy based on Eticas recommendations. This fruitful collaboration has, therefore, helped to develop best practices for technological design, improve the ethics strategy of Alpha Health and integrate protocols on data governance based on the audited app into the Alpha organization. Findings have

also been disseminated through an open publication derived from the audit, by Galdon Clavell, Gemma; Mariano Martín Zamorano, Carlos Castillo, Oliver Smith and Aleksandar Matic (2020), entitled: "Auditing Algorithms: On Lessons Learned and the Risks of Data Minimization". In: Proceedings of the AIES 2020 conference. ACM Press.

Other materials were developed by Eticas on behalf of Alpha, integrating an algorithmic bias preventative strategy for recommender systems. Besides summarising the main findings of the AIA, these documents use REM!X as a case study to develop a Manual for Alpha focused on how to minimise risks of bias when developing recommender systems. This document includes detailed information about how to avoid and monitor bias in the context of technological developments. Lastly, a series of training sessions for a group of Alpha employees was delivered by Eticas and the UPF.

# References


Abdollahpouri, Himan; Robin Burke, and Bamshad Mobasher. (2019). "Managing Popularity Bias in Recommender Systems with Personalized Re-ranking". In AAAI Florida Artificial Intelligence Research Society (FLAIRS '19), May18–22, 2019, Sarasota, Florida, USA.ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/XXXXXXX

Bakker, D., Kazantzis, N., Rickwood, D. and Rickard, N., (2016). "Mental health smartphone apps: review and evidence-based recommendations for future developments". JMIR mental health, 3(1).

Bian, M., & Leung, L. (2015). "Linking loneliness, shyness, smartphone addiction symptoms, and patterns of smartphone use to social capital". Social Science Computer Review, 33(1), 61e79.

Calvo, R. A., & Peters, D. (2013). "The irony and re-interpretation of our quantified self". Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration. New York, NY: ACM.

Cornet, V.P. and Holden, R.J., (2018). Systematic review of smartphone-based passive sensing for health and wellbeing. Journal of biomedical informatics,77:120-132.

Elhai, J. D., Dvorak, R. D., Levine, J. C., & Hall, B. J. (2017). "Problematic smartphone use: A conceptual overview and systematic review of

relations with anxiety and depression psychopathology". Journal of Affective Disorders, 207, 251e259.

Eubanks, Virginia (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press.

Friedman, Batya and Helen Nissenbaum (1996). "Bias in computer systems," ACM Transactions on Information Systems (TOIS)" 14, no. 3. 330-347. https://www.nyu.edu/projects/nissenbaum/papers/biasincomputers.pdf.

Galdon Clavell, Gemma; Mariano Martín Zamorano, Carlos Castillo, Oliver Smith and Aleksandar Matic (2020). "Auditing Algorithms: On Lessons Learned and the Risks of Data Minimization". In Proceedings of the AIES 2020 conference. ACM Press.

Gross, S., Bardzell, J., Bardzell, S. and Stallings, M., (2017). "Persuasive anxiety: Designing and deploying material and formal explorations of personal tracking devices". Human–Computer Interaction, 32(5-6), 297-334.

Hajian, Sara, Francesco Bonchi, Carlos Castillo (2016, August). Algorithmic bias: From discrimination discovery to fairness-aware data mining. Tutorial at KDD'16.

Hersh, M. and Leporini, B. (2017). "Mobile recommender apps with privacy management for accessible and usable technologies". Studies in Health Technology and Informatics, 242, 630-637.

Holstein, Kenneth; Jennifer Wortman Vaughan, Hal Daumé, III, Miro Dudik, and Hanna Wallach. (2019). "Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?". In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). ACM, New York, NY, USA, Paper 600, 16 pages.

Huckvale,Kit, José Tomás Prieto, Myra Tilney, Pierre-Jean Benghozi and Josip Car (2015). Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment, BMC Medicine, 13:214.

Martinez-Perez B, de la Torre-Diez I, Lopez-Coronado M. (2014). "Privacy and security in mobile health apps: a review and recommendations". J Med Syst. 39:181.

Moore, M. (2018). “How the online business model encourages prejudice”. The Guardian. [online] Available at: https://www.theguardian.com/technology/2018/oct/28/how-target-ads-threaten-the-internet-giants-facebook [Accessed 26 Nov. 2018].

O’Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Noew York: Broadway Books.

Resnick,P.,Iacovou,N.,Suchak,M.,Bergstrom,P.,andRiedl,J.(1994).” An Open Architecture for Collaborative Filtering of Netnews”.In ProceedingsofCSCW’94,ChapelHill,NC.

Sarwar, Badrul; George Karypis, Joseph Konstan, and John Riedlfsarwar, “Item-Based Collaborative Filtering Recommendation Algorithms”, Research Group/Army HPC Research CenterDepartment of Computer Science and Engineering University of Minnesota, Minneapolis, MN 55455.

Schindler, M (2017): “Google and Facebook allowed advertisers to target “Jew Haters” In Jerusalem Post. Available at https://www.jpost.com/International/Google-and-Facebook-allowed-advertisers-to-target-Jew-haters-505356

Tsintzou, V., Pitoura, E., & Tsaparas, P. (2018). “Bias Disparity in Recommendation Systems”. arXiv preprint arXiv:1811.01461.

Young Kimberly and Rogers, Robert (1998).”The Relationship Between Depression and Internet Addiction”, CyberPsychology & Behavior, 1.25.

# Annex

## ANNEX 1. DATA ANONYMIZATION

### Definitions:

- Identifier: An attribute that identifies the individual to which it refers directly. Examples: passport number, fingerprint, name.
- Quasi-identifier: An attribute that in itself does not lead to re-identification but may do so if combined with other attributes. Example: ZIP code, birthdate, gender.
- Non-identifier: Attributes that are neither identifiers, nor quasi identifiers and do not enable the re-identification of an individual.
- Pseudonymization: The replacement of a value, normally an identifier, by another value to render it more difficult to re-identify.
- Anonymization: Process that transforms a dataset in order to ensure that an adversary cannot recover information about individuals.
- Anonymized dataset: A dataset where no individual can be identified, where no information can be linked to an individual and that cannot be used to infer information about an individual.
- Attack: A process that takes as input an anonymized dataset and outputs information related to an individual.

### Risks:

- Consider pseudonymised data to be anonymised data. Pseudonymity likely leads to identifiability and stays within the scope of data protection.
- To think that anonymized data deprives individuals of safeguards. The e-Privacy Directive prevents the storage and access to information (which includes non-personal information) on terminal equipments without the subscriber's/user's consent.
- Neglect to consider the impact on individuals of properly anonymised data (especially in the case of profiling).

**Attacks on anonymity can take the following forms:**

- The singling out of an individual by isolating records identifying that individual. For instance, a researcher found that in the U.S. with only sex, ZIP-Code and date of birth, an individual is re-identifiable 87% of the time.
- The linkability of two datasets, leading to the re-identification of an individual. If an attacker can use a public dataset or other available dataset to re-identify an individual through the correlation of both datasets. For instance if both sets include the attributes sex, ZIP Code and date of birth, re-identification would be quite straightforward.
- Inference as being the ability to infer the value of an attribute for a certain record.

**The success of one or a combination of all of these attacks may lead to the re-identification of an individual.**

Steps to anonymization:

- Remove identifiers
- Identify quasi-identifiers
- Based on the quasi-identifiers present, the level of risk of re-identification and the consequences of re-identification, apply the appropriate anonymization techniques (such as randomization, suppression and generalization)

Assess the utility and risks for privacy that the dataset constitutes.

# Annex 2 ARCO rights

## ARCO + 2 REQUESTS AND PROCEDURES: STEP-BY-STEP PROCESSES:

This is a step by step guide for Alpha management of the main procedures included within the ARCO rights.

### Access[17]

1. Request submission:

the template used (see below) for requesting access to the users' data will include the following information: biographic data, purpose of the

[17] Users can access their personal data being processed by the system and the other aspects listed in article 15 GDPR.

request (the right that the user wants to exercise), the personal data affected and the format in which the information should be provided.

2. Request reception and identity corroboration:

Upon receiving an ACCESS request from a Rem!x user -sent to the email of the DPO included in the Privacy Policy-, the request is passed to the Alpha Legal Department. The identity of the data subject submitting the access request is therefore confirmed by the Legal Department of Alpha, by matching data provided by the user within the request template to biographic data stored in the system.

*If the identity of the data subject cannot be confirmed by these means, he/she will be contacted to provide further evidence of his/her identity (passwords, etc.). The same process will be conducted in case a representative is asking for this information on behalf of the user.*

3. Request approval and information gathering/setting:

If the data subject or his/her representative has proven to be who he/she claims to be, then a message confirming the approval of the access request will be sent to the user.

The Legal Department will submit a formal request to the IT Manager of Rem!x to collect all/the specifically requested (personal) data corresponding to the person involved in the access request. The information will be organize in a user friendly manner so the user/representative can easily identify the set of records. Alpha will conduct this process within 1 month after the access request has been approved.

*In case the access request cannot be technically processed, the data subject will receive a written explanation about this, informing also about possible ways to pursue a legal remedy.*

4. Information provision:

Once the requested information has been put together, organized and received by the Legal Department of Alpha, it will be sent by the Legal Department to the data subject requesting access, or to his/her legal representative.

The dataset will be provided in paper or in electronic format, depending on the request. If the user requests a paper copy, the first shall be free of charge, but subsequent ones could be charged with a reasonable fee based on administrative costs.

The information will have to be structured in a succinct, clear language, as well as in a comprehensible and effortlessly manageable format.

## Rectification[18]:

- Request submission: same than for Access above.
- Identity corroboration: same than for Access above.
- Confirmation of inaccurate/incomplete data: If the data subject or his/her representative has proven to be who he/she claims to be, the legal department will submit a formal request to the IT Manager of Rem!x to confirm the data inaccuracy. Once this is confirmed, a message confirming the approval of the rectification request will be sent to the user.

*In case the rectification request is considered not binding or cannot be technically solved the data subject will receive a written explanation about possible ways to pursue a legal remedy.*

- Modification of data: Once wrong data or errors have been identified, the legal department will request to the IT manager to modify inaccurate data about the user held by Alpha corresponding to the person involved in the rectification request. Details will be amended within a month after the request has been approved and as indicated by the user. A note referencing the modification and its cause/s will also be included on the system by the IT manager.
- Notification on third parties disclosure: The data subject involved in the rectification request will be informed, within a month from the data modification, if his/her rectified data was disclosed to third parties legally involved in Rem!x.

## Cancelation (Restruction of Processing)

According to GDPR, users shall have the right to restrict the processing of their personal data when at least one of the four circumstances listed in article 18 GDPR are happening. Therefore, the request form will allow users to select which among them serves as the basis for their request in order to facilitate the work of the legal team:Cancelation (Restruction of Processing)

[18] According to article 16 GDPR, users shall have the right to obtain rectification from the controller without undue delay, taking into account the purposes of the processing.

- "the accuracy of the personal data is contested by the data subject, for a period enabling the controller to verify the accuracy of the personal data;
- the processing is unlawful and the data subject opposes the erasure of the personal data and requests the restriction of their use instead;
- the controller no longer needs the personal data for the purposes of the processing, but they are required by the data subject for the establishment, exercise or defence of legal claims;
- the data subject has objected to processing pursuant to Article 21(1) pending the verification whether the legitimate grounds of the controller override those of the data subject."
- Request reception: same than for Access above.
- Identity corroboration: same than for Access above.
- Identification of data and confirm data subject rights upon (personal) data: If the data subject or his/her representative has proven to be who he/she claims to be, the legal department will analyze the rights of data subject to restrict the processing of the indicated information. If this is confirmed, this department will submit a formal request to the IT Manager of Rem!x to confirm the existence of data whose processing has to be stopped.
- Cancellation of data: Once this is confirmed, a message confirming the approval of the cancellation request will be sent to the user and the information will be processed appropriately by the IT manager. Also, in the processing restriction happens to be lifted, the user will be informed in advance.

In case the cancellation request is considered not binding or cannot be technically solved the data subject will received a written explanation about this, informing also about possible ways to pursue a legal remedy.

Notification on third parties disclosure: Data subject involved in the cancellation request will be informed within a month of the data modification if his/her data was disclosed to third parties legally involved in Rem!x

## Objection[19]

According to article 18 GDPR, users shall have the right to restrict the processing of their personal data when processing of personal data concerning him or her which is based on point (e) or (f) of Article 6(1), including profiling based on those provisions. Especially, users will be able to object the processing of their data when it's being processed for marketing purposes. Therefore, the request form will allow the user to specify which one of those circumstances are taking place and legitimizing his or her petition in order to facilitate the work of the legal team.

1. Request reception: same than for Access above.

2. Identity corroboration: same than for Access above.

3. Identification of data and confirm data subject rights upon (personal) data: If the data subject or his/her representative has proven to be who he/she claims to be, the legal department will analyze the rights of data subject to object the processing of the indicated information. If this is confirmed this department will submit a formal request to the IT Manager of Rem!x to confirm the existence of data whose processing must be stopped or restricted. Otherwise the rejection of the objection request will be automatically submitted to the user[20].

4. Modification of data processing: Once data whose must be changed have been identified, the legal department will request to stop the processing of the data corresponding to the person involved in the objection request. Processing will be amended within a month after the request has been approved.

Depending on the kind of objection this process can imply the erasure or suppression of data or the ceasing of the processing. A note referencing the modifications made and its causes will also be included on the system.

According to article 21 GDPR, the controller should demonstrate compelling legitimate grounds that override the interests, rights and freedoms of the data subject if he or she is to not accept the user's demands. Therefore, if the request was not accepted, the reasons should be made known to the user.

[19] For further information, see: https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-object/

[20] The list of cases where objection does not apply must be published by Alpha.

5. Notification on third parties disclosure: Data subjects involved in the objection request will be informed within a month of the data modification if his/her data was disclosed to third parties legally involved in Rem!x.

## Right to Data Portability

According to article 20.1 GDPR, users shall have the right the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format. They also have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided when one of the circumstances contained in that same article applies.

1. Request submission: same than for Access above.

2. Identity corroboration: same than for Access above.

3. Confirmation of the request's legitimacy: If the data subject or his/her representative has proven to be who he/she claims to be, the legal department will submit a formal request to the IT Manager of Rem!x to assess if the request is based on legitimate basis. The presence of the circumstances listed on article 20.1 GDPR and of the exceptions present in articles 20.3 and 20.4 of the GDPR will be accounted for by the legal team. Once this is confirmed, a message confirming the approval of the request will be sent to the user. According to article 20.2 GDPR, the data will be directly transferred to the selected controller when technically feasible.

In case the portability request is considered not binding or cannot be technically solved the data subject will received a written explanation about this, informing also about possible ways to pursue a legal remedy.

4. Portability of the data: Once wrong the data has been compiled, the legal department will request to transfer the data corresponding to the person involved in the data portability request to the designed controller if that is technically feasible. If not, the data will be facilitated to the user. Details will be transferred within a month after the request has been approved and as indicated by the user. A note referencing the procedure and its cause/s will also be included on the system.

## Erasure[21]

1. Request submission: same than for Access above.

2. Identity corroboration: same than for Access above.

3. Confirmation of the request's legitimacy: If the data subject or his/her representative has proven to be who he/she claims to be, the legal department will submit a formal request to the IT Manager of Rem!x to assess if the request is based on legitimate basis. The presence of the circumstances listed on article 17.1 GDPR and of the exceptions present in article 17.3 GDPR will be accounted for by the legal team. Once this is confirmed, a message confirming the approval of the erasure request will be sent to the user.

In case the portability request is considered not binding or cannot be technically solved the data subject will received a written explanation about this, informing also about possible ways to pursue a legal remedy.

4. Erasure of the data: Once data involved in the request have been identified, the legal department will request to erase the data on the user held by Alpha corresponding to the person involved in the erasure request. Details will be amended within a month after the request has been approved and as indicated by the user. A note referencing the erasure and its cause/s will also be included on the system.

5. Notification on third parties disclosure: The data subject involved in the erasure request will be informed within a month from the data modification if his/her rectified data was disclosed to third parties legally involved in Rem!x.

According to article 17.2 GDPR, where Alpha has made the personal data public and is obliged pursuant to paragraph 1 to erase the personal data, the DPO and the legal team, taking account of available technology and the cost of implementation, shall take reasonable steps, including technical measures, to inform controllers which are processing the personal data that the data subject has requested the erasure by such controllers of any links to, or copy or replication of, those personal data.

[21] According to article 17 GDPR, users shall have the right to have their data erased from the controller's database without undue delay when one of the circumstances contained in article 17.1 GDPR apply.

## ARCO Rights Request Form

As a data subject under the protection of the GDPR, you are entitled to a set of rights that are commonly known as ''ARCO rights''. The present form will allow you to exercise those rights as a user of Alpha. If you want to find out more about your rights, you can access our Privacy Policy, in which there is detailed information on this matter.

### 1- Identifying information.

In order for us to know who you are and assess your case, we need the following information[22]

Data Subject's Name:

Email:

Any other information that may help us to locate your personal data:

### 2. Representatives (only complete if you are acting as the representative for a data subject)

[Please Note: We may still need to contact the data subject where proof of authorisation or identity are required[23]]

Representative's Name:

Email:

### 3- The right you want to exercise:

Now we know who you are. It is time for you to tell us which right you wish to exercise from among the different rights recognized by GDPR:

- Right to access
- Right to rectification
- Right to the restriction of processing
- Right to object
- Right to data portability
- Right to erasure

[22] One important note: the information collected must be the minimum necessary to carry out the processing.

[23] This right may also be exercised through a legal representative, in which case, as well as submitting the interested parts' DNI (national identity document), the representative's DNI and the authentic document that proves the right of representation must also be submitted.

### 4- Reason or ground on which you base your request:

This will change according to the specific right that the user wants to exercise. For instance, if the user wanted to exercise his or her right to erasure, the different circumstances present in article 17 would appear in the menu.

### 5- Data affected:

Here the user would have to specify which data he wants to have affected by his or her request. That could be all the data in the hands of Alpha or just a part of it.

Format of the data: paper / digital

If you do not know the address of the file manager or you have any problem with this request, you may contact the Spanish Data Protection Agency by telephone on 901 100 099. Personal data that are provided by the applicant will be included in a file owned by the Alpha, the purpose of which is to process requests to exercise your ARCO rights.

# Annex 3 Privacy Policy

## REM!X PRIVACY POLICY

We have reviewed the current privacy policy. After having analyzed it, we have suggested a whole array of improvements (see implementation plan) in order to:

- Facilitate understanding

-Improve legality and ethics

-Increase transparency and improve trust

Aside from all the recommendations we made in the Implementation Plan, we suggest the introduction of icons and visual cues that aid the understanding of the privacy policy by as many people as possible. The chart down below is intended as a summarized version of the current version of the privacy policy. Given its simplicity, we suggest its use at the moment when the user is going to give consent instead of the long version. This would be in accordance with a layered approach to the right to be informed, which has been recommended by the AEPD[24].

[24] https://www.aepd.es/media/guias/guia-modelo-clausula-informativa.pdf

| | | |
|---|---|---|
| WHO?<br><br>(Responsible of the treatment) | TELEFÓNICA INNOVACIÓN Alpha S.L. | Social domicile: Ronda de la Comunicación, Distrito Telefónica, Madrid, España.<br><br>CIF: B87453643 |
| | | Contact details: hola@remix-app.com |
| | | DPO email address: DPO_telefonicasa@telefonica.com |
| WHY?<br><br>(Purpose of the treatment) | 1-Provide recommendations | Data will be processed by automated means to create recommendations aimed at improving the wellbeing of users, to improve the functioning of the app and to be in touch with the user for marketing purposes. |
| | 2-Improve the app | Retention period: 12 months after the last moment that the user logged into the app. To know more about the exceptions to this rule, go to the long version of the Privacy Policy. |
| | 3-Direct marketing purposes | We processed your data with algorithms that predict the recommendations that will be best suited for you according to a logic that we explain here. |
| ON WHAT BASIS?<br><br>(Basis for the processing) | Recommendations: The legal basis is the need for the processing in order to provide the service agreed in the contract.<br><br>Improvement: The legal basis is our legitimate interest in improving out service.<br><br>Direct marketing: The legal basis is your consent. | You have the right to withdraw consent for the processing of your data for direct marketing processing at any time and without any consequences. |
| | | Location data is necessary for us to provide you with our services. That is the reason why providing these data is not optional. Therefore, not providing us with your location data will cause the termination of the contract. |

| | | |
|---|---|---|
| With WHO?<br><br>(Recipients) | We will share your data with third parties. | These third parties will be out providers and collaborator, as well as some companies such as Facebook, which will use the data to provide us with service of data analytics and marketing. The roles of these parties are:<br>• data storage,<br>• communication/customer support,<br>• app functionalities/interaction,<br>• tracking/monitoring of users activity |
| | Some of these data transfers could go to countries that do not belong to the European Union, such as the US. | Decisions of adequacy, guarantees, binding corporate rules or specific situations applicable. |
| YOUR Rights | As a resident in the European Union you enjoy a set of rights over your data. | If you want to access, rectify, erase, object its processing or take it somewhere else, we explain to you how here. |
| | | You have the right to withdraw your consent at any moment, which you can do by letting us know through an email. |
| | | If you wish to lodge a complaint, you have to contact the Agencia Española de Protección de Datos. |

| | | |
|---|---|---|
| WHAT?<br><br>Data gathered in a passive way | We gather data coming from different sensors in your phone. | These data come from sensors in your phone, such as:<br>• Accelerometer<br>• GPS<br>• Wifi and Bluetooth<br>• Your configuration<br>To know more, go to the long version of the Privacy Policy. |
| | | This are some of the data we collect:<br>• Phone type<br>• Phone Operating System<br>• (Android or iPhone)<br>• When the screen is switched<br>• on/off<br>• When the screen is<br>• locked/unlocked<br>• When the screen is locked/ unlocked<br>• Current state of phone<br>• connectivity,<br>• The time when new photos<br>• are taken<br>• Step counter<br>• Activity (walk)<br>• Location data<br>• Movement of the phone |

The extended version of the Privacy Policy may include further details on each item pointed out below.

THE DATA WE COLLECT

WHAT

The information you provide

-Name

-E-mail address (if you enter it in our website)

-Phone number

-Occupation

-Age

-Sex

-Educational Attainment

-Interests and goals

-Personal traits (extroverted or introverted)

-Personal values (loyalty, security, success…)

-Activities and mood

-Information about the recommended activities

The information we gather

-If the phone is blocked or unblocked

-If the screen is turned on or turned off

-Estimations of the levels of physical activity based on the movement registered by the sensors.

-Number of steps

-Geolocation

-Environmental noise (not the content)

-The movement of the phone.

-Environmental light

-Wifi networks around the phone

-Bluetooth devices around the phone

-The state of connectivity of the phone (if you activate roaming abroad or if the type of connection changes, for instance)

-Data consumed

-Battery levels and information about if the phone is charging

-Type of phone, OS installed, used memory, number of processors, manufacturer and the name of the phone.

-Number and types of apps installed, as well as the versions installed and the time when they were downloaded.

-Proximity of the phone to other objects

-Earphones connectivity

-Time tags of pictures (not the pictures themselves)

-IP address.

-City, region, country and time zone (from syntactic analysis of the IP address).

-Language

-Information about the browser

-Type of mobile device, including the screen dimensions

-The URL that took you here, including the browser and the keywords.

-Sections you access in out website.

-The messages that you open in your REMIX account.

-Where you click on our website and on the emails we send you.

We suggest that the data that are being gathered should be reduced to the minimum amount necessary (principle of data minimization). We also suggest that users should be able to opt out more often.

HOW

Icon: Sensors

We are able to collect automatically many of the typologies of data above thanks to the sensors and features in your phone.

-Accelerometer

-Pedometer

-Microphone

-GPS

-Wi-Fi and Bluetooth systems

Icon: Questions and forms

We also get information directly from you when you. For instance, at the moment when you fill in the initial questionnaire in which you actively give us information about you.

Icon: ML

Icon: Security

We guarantee the security, secrecy and the confidentiality of your data, communications and personal information. We have adopted the most robust and strict security measures in order to guarantee the integrity, confidentiality and availability of your data.

We are committed to acting quickly and to inform you in the event of any situation with the potential to put in danger your data.

## WHY

Our app attempts to provide the user with recommendations and council on time management and personal wellbeing. This constitutes a service that is performed on the basis of a contract which also serves as the legitimizing purpose for the processing of your data.

We also want to learn more about how to optimize the app so its functioning is more accurate. This forms part of our legitimate interest in improving our services.

Your data can also be processed for marketing purposes. When your consent is required for that, we will ask for it through the app.

With regards to our cookies, they are used to improve the user experience in our website and to unlock additional features.

Personalise

Recommend

## FOR HOW LONG

We will not keep your data for longer that it is necessary. More concretely, we will keep the data until 12 months after the last time you logged into the app. There are some potential exceptions to this principle that you can consult in the extended version of the Privacy Policy.

Also, if you exercise your right to erasure, we will delete your data right away. Read the section down below to learn how you can exercise that right among others.

(we suggest removal of longer retention period)

## YOUR RIGHTS

According to GDPR, you have the right to ask us to give you access to your data, rectify it, erase it, restrict its processing, facilitate your right to data portability and to object to its processing. Also, you have the right to oppose being subjected to purely automatic decisions and to withdraw your consent at any time. In order for you to exercise those rights, you should get it touch with our team through the following email:

hola@remix-app.com

You should attach documentation that helps our team to identify you and what right do you want to exercise and to what extent. You will get a response in a month time since your email. If we take longer to reply, you will be informed about the reasons for the delay. In case you are not satisfied with our decision, you should contact the Agencia Española de Protección de Datos (Spanish Data Supervisor Authority).

Finally, here is the contact of our DPO:

DPO_telefonicasa@telefonica.com

## DOWNLOAD

Your data

Our privacy policy

Our conditions of use

## QUESTIONS?

We suggest the following sections of the privacy policy are reviewed:

- SHARING PART (particularly research, third countries/parties)

- COOKIES (if cookies collect further data, this should be specified above)