Mobile Health App Database - A Repository for Quality Ratings of mHealth Apps

Michael Stach[1], Robin Kraft[1,3], Thomas Probst[2], Eva-Maria Messner[3], Yannik Terhorst[3], Harald Baumeister[3], Marc Schickler[1], Manfred Reichert[1], Lasse Bosse Sander[4], Rüdiger Pryss[5]
[1]Institute of Databases and Information Systems, Ulm University, Germany
[2]Department for Psychotherapy and Biopsychosocial Health, Danube University Krems, Austria
[3]Department of Clinical Psychology and Psychotherapy, Ulm University, Germany
[4]Department of Rehabilitation Psychology and Psychotherapy, Albert-Ludwigs-University of Freiburg, Germany
[5]Institute of Clinical Epidemiology and Biometry, University of Würzburg, Germany
{michael.stach, robin.kraft, eva-maria.messner, yannik.terhorst, harald.baumeister, marc.schickler, manfred.reichert}@uni-ulm.de
thomas.probst@donau-uni.ac.at, lasse.sander@psychologie.uni-freiburg.de, ruediger.pryss@uni-wuerzburg.de

***Abstract*—The utilization of mobile technology in the field of medicine and healthcare has become a decisive aspect. The entire field is denoted as mobile health (mHealth). For mHealth, the development and use of mobile applications are crucial. The purposes and goals of mHealth apps, in turn, are manifold. As a consequence, a plethora of mHealth apps can be found in the app stores. Interestingly, for patients, users, and health care providers that consider to use mHealth apps one aspect has been less pursued so far: Systematic and standardized ways that help about the quality of an app or its medical evidence are mainly missing. The Mobile App Rating Scale (MARS) is a standardized instrument that aims at the systematic and comparable evaluation of the quality of mobile health apps as well as categorizing their goals and functions. It comprises 23 items, which are utilized to calculate a rating scale. Having *MARS* in mind, a database was developed that is called Mobile Health App Database (MHAD). The latter offers technical features to systematically utilize the *MARS* for researchers as well as clinicians and end-users that (i) want to evaluate apps as well as (ii) want an interactive and easy-to-use web interface that shows the results of the rating procedure. *MHAD* comprises a rating platform that supports the conduction of *MARS* ratings and their release process. With the information platform, a web application was developed that prepares the data stored in the rating platform for being freely viewed and studied by users, patients, and health care providers. The goal of *MHAD* constitutes to be an open science repository that encourages researchers to release their *MARS* ratings to a broader audience. Such repositories become more and more important in many fields, especially in the field of mHealth.**

***Index Terms*—mHealth; mobile health; Mobile App Rating Scale; MARS; app quality; medical database**

I. Introduction

Mobile apps and their data collection capabilities are only one direction that has garnered a lot of attention in the last years [1]. Especially in healthcare and medicine, mobile apps are a basis for new data sources and medical insights [2]–[4]. In addition to the sophisticated collection possibilities of data, mobile apps can be utilized to guide and inform patients about health conditions and questions as well as health-related day-by-day issues [5], [6]. Interestingly, both in research and industry, less efforts have been undertaken to evaluate the quality and evidence of mobile health apps. When monitoring the field of mHealth in the major app stores (i.e., Apple App Store, Google Play Store), it can be observed that currently over 300,000 apps are available, with over 200 additional apps being added on a daily basis [7], [8]. Although the dramatic increase of mobile health apps is ongoing, efforts to systematically evaluate them are still rare. From a broader perspective, the following aspects are particularly apparent at the moment:

1) Although recent regulations (e.g., the General Data Protection Regulation or the Medical Device Regulation in the EU) have been pursued to guide and protect users in the digital world better, no **general standards** in terms of quality have been presented for mHealth apps so far.
2) The number of users that will utilize mHealth apps will constantly increase [9]. Hence, the demands of the users for offered apps increase in the same magnitude of orders. Therefore, mHealth apps are introduced in a way that makes it very complex for interested stakeholders to **reliably decide** which app shall be used.
3) The development of mobile apps in general and for mHealth in particular is **technically challenging** due to the frequent update cycles of the mobile operating system vendors and the associated API changes. This includes the frequent hardware advancements of smart mobile devices (e.g., sensor features are offered with a new smartphone version).
4) For the quality as well as for the evidence, still less instruments exist that provide **guidance** in the jungle of offered and newly introduced mHealth apps.

Some tools were developed for the systematic evaluation of the quality of mHealth apps. For example, the app evaluation model of the American Psychological Association [10], the Evaluation Tool for Mobile and Web-Based eHealth Interventions (ENLIGHT) [11], [12], or the Mobile App Rating Scale (MARS) [13]. The *MARS* [13] is an instrument based on 23 items, which provides standardized expert ratings of mobile health apps.

The items were developed based on an intensive literature study and evaluation of existing apps. The instrument revealed to be reliable in evaluating mHealth apps (i.e., in terms of internal consistency, inter-rater reliability and confirmatory factor analysis of its structure [14]). Notably, *MARS* has a global score based on the following subscales: *engagement*, *functionality*, *aesthetics* and *information quality*. In addition, there is a *subjective quality* section, a *perceived impact* section and a *classification* section for the theoretical background and the app functionality.

As the *MARS* has proven its feasibility and psychometric quality in existing works and evaluations [14]–[16], the authors of the current work decided to develop a database solution that shall support making the *MARS* expert ratings for mHealth apps available to the public more easily and comprehensively. This is necessary, since consumers are particularly vulnerable right now, since the mHealth market is not well regulated and most of the currently available mHealth apps are not scientifically evaluated [15]–[17]. Thus, there is an urgent need to inform users, patients, and healthcare providers about the quality of mHealth apps [17]. The project presented in the paper at hand is called Mobile Health App Database[1] (MHAD) and is online since the end of 2018 and currently comprises *MARS* ratings of 1112 mHealth apps (see Fig. 1), stemming from nine categories; i.e., *mindfulness*, *anxiety*, *depression*, *support children and young people*, *cancer*, *PTSD*, *pain*, *support the elderly*, and *sports*.

Technically, *MHAD* consists of a powerful rating platform to manage *MARS* ratings as well as an information platform[1]. The work at hand delineates the important requirements of *MHAD*, its technical features, as well as its use in practice. In particular, it will be shown how the technical procedure from obtaining relevant apps from the apps stores up to their rating release on the information platform was conceived and implemented. As less comparable solutions exist, the *MHAD* repository shall help to release results of instruments like the *MARS* for the overall guidance of users, patients, and healthcare providers in the jungle of mHealth apps more efficiently and systematically.

The remainder of this work is organized as follows. In Section II, related works are discussed. As *MHAD* is based on *MARS*, the latter is presented in more detail in Section III. Based on this, Section IV discusses the overall concept of *MHAD*. In particular, it is shown how the flow of the overall rating procedure was chosen. The technical requirements that are addressed will be discussed in Section V, while Section VI presents a selected choice of impressions of the platform and how the technical requirements have been implemented. In addition, this section discusses the current statistics of the platform. Finally, Section VII summarizes this work and provides an outlook on future work.

II. Related Work

Three kinds of related research are relevant in the scope of this paper. First, Internet-based platforms that offer guidance for mHealth apps. Second, approaches that deal with the quality and evidence of mHealth apps. Third, general considerations on the use of mHealth apps by researchers and institutions. Regarding the first category relevant for the work at hand, several other platform exist. In Germany, three other platforms are prominent [18]–[20]. In contrast to these platforms, *MHAD* is the only one that is based on a standardized and scientifically proven instrument. In addition, the mentioned other platforms use their own criteria for the evaluation of mHealth apps. In research, also works can be found that establish database sources for the evaluation of mHealth apps. For example, the authors of [21] have made their data source of the paper available for other researchers. However, they only provide the crawled app store information and not calculated results. Regarding the second category, summaries that can be used as a very good starting point can be found in [22]–[24] or works that address particular aspects [25]–[27] or the development of rating tools [28], [29]. Furthermore, works can be found that deal with the user-friendly development of mHealth apps, thus increasing their quality by design [30]–[32]. Research that are utilizing the *MARS* are also all in this second category [16], [33], [34]. The number of works based on the *MARS* show its general applicability. Related to *MARS*, other established instruments (e.g., from the computer science field) have been adjusted to rate the quality of mHealth apps [11], [35], [36]. It should however be kept in mind that the latter instruments cannot replace controlled trials. Only rigorously tested mHealth apps can be described as evidence-based. The *MARS* includes an item in the information quality scale regarding evidence-base of the app. In general, the evidence-base of most of the mHealth apps is low at the moment [14], [37], [38]. Regarding the third category, on the European level, for example, first considerations were done very early (e.g., [39]). Also in the USA, such regulations were released very early [40]. Recently, in Germany, a new recommendation was released

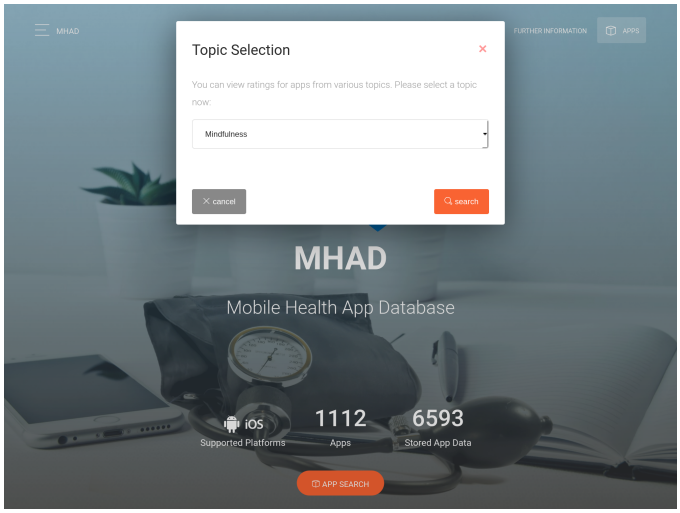

Fig. 1. Screenshot: the category filter function.

[1]MHAD information platform: `www.mhad.science`

[41]. Moreover, quality criteria based on existing models for the systematic assessment of telemedicine applications were suggested by researchers for Internet-based and mobile mental health interventions [42]. Altogether, for all existing works, to the best of the authors knowledge, *MHAD* is the biggest platform for mHealth apps using a reliable instrument of high psychometric quality. However, also many related works have to be mentioned, the goals partly differ significantly. This is especially the case for the third category. However, these works must be considered for platforms like *MHAD*.

III. *MARS* Background Information

As already introduced, the *MARS* was developed as an instrument to efficiently and reliably classify and assess the quality of mHealth apps [13]. It comprises five categories: *engagement*, *functionality*, *aesthetics*, *information quality*, and *subjective quality*. These categories were extracted and selected based on 372 criteria from 25 published papers. This emphasizes the broad perspective of the *MARS*. Note that the general quality rating for the 23 items utilizes the following 5-point scale: (1) inadequate (2) poor (3) acceptable (4) good (5) excellent.

For *MHAD*, the *MARS-G* is utilized [34]. It is a German translation of the original *MARS* and was developed by the authorization and help of the *MARS* authors. As the original version, the *MARS-G*, has the four sections *engagement*, *functionality*, *aesthetics* and *information quality* (see Table I, A-D). Based on the latter, a global scale can be calculated. It also has the *subjective quality* section and the *perceived impact* section like the original *MARS* (see Table I, E&F). One more section was added to the *MARS-G* that is not part of the original version, i.e., a therapeutic gain section (see Table I, PT) to evaluate the usefulness of the mHealth app in psychotherapy. To obtain *MARS* ratings for *MHAD* the following procedure is applied: Two independent raters that are trained in *MARS-G* (including an online training [43]) and supervised (if necessary) by a licensed healthcare specialist (e.g., licensed psychotherapist) perform the *MARS-G* ratings on apps that shall be evaluated. Their individual *MARS* ratings are summarized to final *MARS* scores for an evaluated app. On the *MHAD* website, the global *MARS* score as well as the scores of the four *MARS* subscales are displayed.

TABLE I
MARS-G: Sections and Items

Section	Title	Items
A	Engagement	fun, interest, individual adaptability, interactivity, target group
B	Functionality	performance, usability, navigation, gestural design
C	Aesthetics	layout, graphics, visual appeal
D	Information Quality	accuracy of app description, goals, quality of information, quantity of information, quality of visual information, credibility, evidence base
PT	Therapeutic Gain	gain for patients, gain for therapists, risks and side effects, ease of implementation into routine healthcare
E	Subjective Quality	recommendation, frequency of use, willingness to pay, overall star rating
F	Perceived Impact	awareness, knowledge, attitudes, intention to change, help seeking, behavioral change

IV. Mobile Health App Database Approach

In this section, the overall procedure to accomplish the rating process, which was introduced in the last section, is presented and discussed, while in the next section, the concrete technical features and their implementation are presented. In Fig. 2, the phases and their sequence are illustrated.

In the **Identification Phase**, an app store crawler was implemented for the two major app stores, i.e., the Apple App Store and the Google Play Store. In addition, the implementation and maintenance of the crawler is challenging because, for example, the Play Store provides no official API and its website is changed frequently by Google. However, the development is a decisive pillar of the entire solution. If the crawler would not exist, a manual procedure would become necessary. As thousands of apps could be relevant, such procedure is less feasible. Note that we implemented filtering techniques for the crawler, so that, for example, only sports apps are found. The technical obstacles for such a crawler are also one more reason why only the two predominant app stores are currently provided.

Then, in the next phase, which is denoted with **Screening**, it is decided which apps that the crawler found are actually rated. As the crawler uses the information that the app developers have provided in the stores, it might happen that these information do not fit to the criteria that were chosen for a rating (e.g., the category). Therefore, this screening is done for each identified app manually.

Afterwards, in the **Reviewing** phase, the remaining list of mobile apps is rated with the *MARS* by the reviewers. Due to the lack of space, not all features that are provided for the distribution can be discussed in detail. However, to mention one example, reviewers see their assigned apps in the rating platform, but it is possible to decline *MARS* ratings if a reviewer feels not comfortable or has not the time for a rating at the moment. By providing such features, *MHAD* must keep track that always two reviewers provide a result in time.

Finally, in the **Presentation Phase**, the ratings are released to the information platform. For the two phases *Screening* and *Reviewing*, a further aspect has to be briefly discussed. In the first technical stage of the *MHAD* solution, these phases were not technically supported (i.e., rating platform features weren't provided). Thus, for some app ratings of the information platform, these two phase were accomplished solely manually (i.e., using Excel and paper-based *MARS-G* questionnaires). For this manual process, *MHAD* provides an sophisticated import mechanism for Excel files that is able to manage the release procedure. As this manual procedure was very time-consuming, the current implementation of *MHAD* provides required functions for all of the four shown phases in Fig. 2. During the course of running *MHAD*, more features emerged that should be implemented. The complete list is therefore discussed and presented in the next two sections.

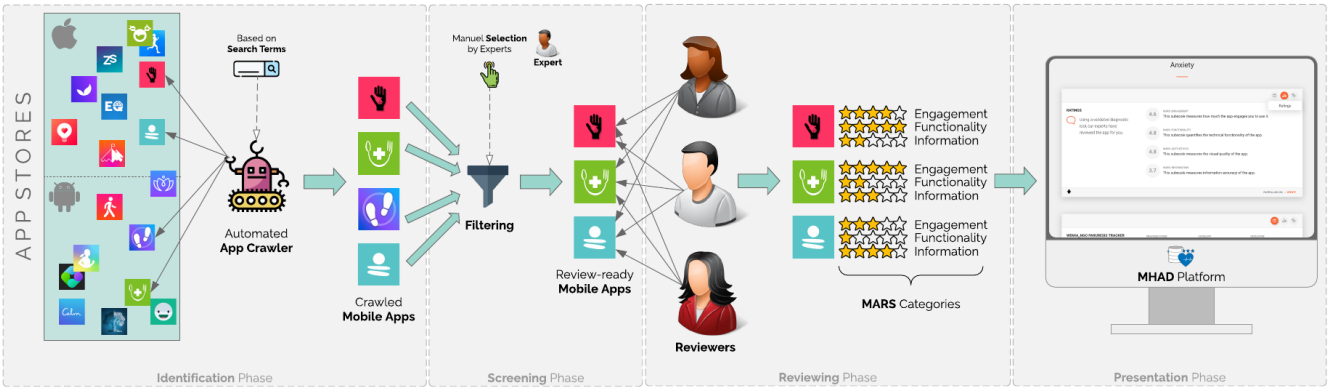

Fig. 2. Mobile Health App Database (MHAD) approach at a glance.

V. Technical Requirements

The entire *MHAD* endeavor started in 2018 and was initially intended to be a dynamic website displaying apps that were rated based on the *MARS*. After starting with the rating of apps, mainly based on pen-and-paper, many issues emerged that lead to the current development stage of *MHAD*. For example, it quickly emerged that a powerful platform is needed to manage the entire review and release process. With having a rating platform in mind, the idea came up to manage also dynamic content for the *MHAD* information platform. Following this history, a modular **crawler** component providing a RESTful API, a **rating platform**, and an **information platform** were conceived and implemented. Table II summarizes all requirements at a glance. Only one selected feature is further discussed, the management of groups. It emerged that it is helpful to assign reviews to groups of users instead of single users. This approach has advantages and technical drawbacks. The advantage is to better address the preferences of users to rate particular app categories or apps with particular characteristics. On the other, from a technical perspective, the management of groups harbors the risk of further considerations. For example, if assigning reviews to groups could purport that nothing has to be monitored as there should be always someone in a group that certainly takes over a review. Such situations must be therefore handled through proper monitoring features. Current considerations follow the idea of an open rating platform, like the review process of scientific publications.

VI. *MHAD* in Practice

This section illustrates selected features of *MHAD*. First of all, some statistics are briefly mentioned. Currently, *MHAD* comprises freely available ratings of 1112 apps. That means, for all of the 1112 apps, the procedure including two reviewers (as described above or in [16]) for the final rating was accomplished. These ratings are categorized – as already mentioned – with the following numbers for each category maintained so far: mindfulness: 192, anxiety: 104, depression: 39, support children and young people: 13, cancer: 75, PTSD: 82, pain: 218, support the elderly: 77 and sports: 312.

TABLE II
Functional Requirements of the Mobile Health App Database

No	Title	Description
Crawler		
1	Scrape Play Store	The crawler enables authorized users to gather data from the Google Play Store.
2	Scrape App Store	The crawler enables authorized users to gather data from the Apple App Store.
Rating Platform		
3	App Search Feature	Two search features are provided. First, the crawler can be used to scrape the app stores. Second, scraping results can be stored in the internal database and scanned.
4	App Management	Found and stored apps can be managed (deleted, attributes changes, etc.).
5	Category Management	Features are provided to put apps into categories. This information, in turn, is utilized for the information platform to show the apps in categories.
6	Dynamic Questionnaire Views	Questionnaire input structures (e.g., for *MARS-G*) can be managed and dynamically rendered.
7	Review Management	Features are provided to manage reviewers, reviews, and the release of reviewed apps.
8	User and Group Management	Features are provided to manage users and also to manage groups. Users can be grouped. This feature is used with respect to the better grant of rights as well as to assign reviews to entire groups. The latter shall ease the review management.
9	Export Feature	Features to export data (e.g., all stored apps.)
Information Platform		
10	App Information Feature	A feature is provided that displays all created app information as well as the ratings information.
11	Category Feature	App categories from the rating platform are utilized and presented. More specifically, apps can be generally found and filtered with respect to the determined categories.
12	Content Management (News, Team, etc.)	A content management module was implemented to dynamically manage news, team information, publications, etc.
13	Search Feature (in development)	Currently, a search feature is under development, which shall enable the website users to find ratings of particular apps more easily.

New categories and ratings are currently in preparation. In the following, selected features presented in Table II will be briefly introduced. As the most important feature, Fig. 4 display the three tabs that are maintained for each app. In Fig. 4 (top), the final result of the *Pacifica - Stress & Anxiety* app [44] is shown, which is `4.5`. Fig. 4 (middle), in turn,

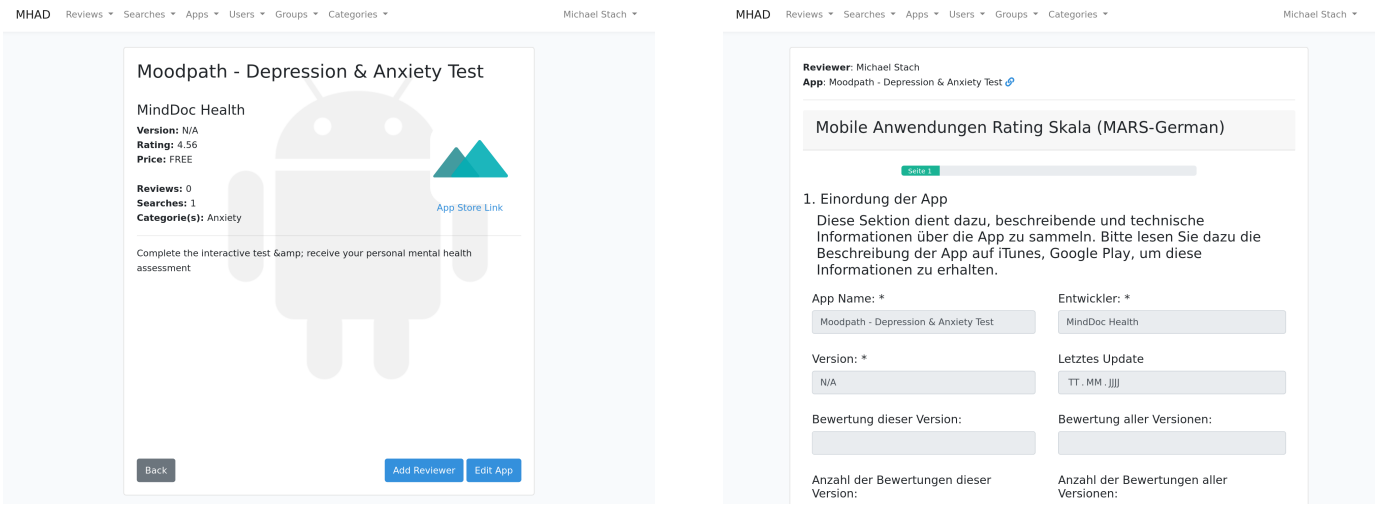

Fig. 3. Screenshots: Manage app information and its reviews (left) and *MARS-G* reviewer view (right).

shows the second tab, the ratings for the four *MARS* subscales *engagement*, *functionality*, *aesthetics*, and *information quality*. The third tab finally shows the app classification information, here shown in Fig. 4 (bottom). The three latter figures have shown information platform features. This is supplemented by Fig. 1, which shows a screen that enables users to filter apps based on their category. The remaining two screenshots show rating platform features. In Fig. 3 (left), the app view is shown. As can be obtained, the gathered information can be edited or the app can be assigned to reviewers. Finally, Fig. 3 (right) shows a part of the *MARS-G* questionnaire and how it is displayed to a reviewer. As can be further obtained by the menu shown in Fig. 3, the rating platform currently summarizes features on: *Reviews*, *Searches*, *Apps*, *Users*, *Groups*, and *Categories*.

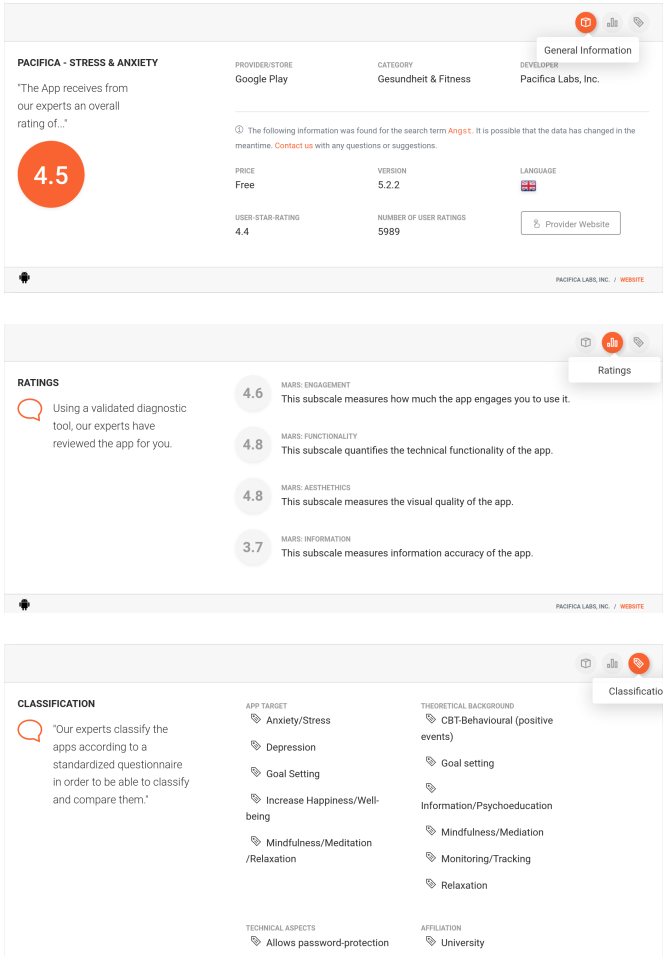

Fig. 4. Screenshot: Basic app information & overall app rating (top), detailed *MARS* rating results (middle) and app classification information (bottom).

VII. Summary and Outlook

This work has presented the Mobile Health App Database (MHAD) project, its background information as well as its use in practice. The overarching goal of *MHAD* is to provide a guide for users, patients, and health care providers regarding the quality of mHealth apps. As the market of mHealth apps is currently challenging to overlook and the ratings of apps in the app stores are not based on scientific instruments, solutions like *MHAD* can contribute to help various stakeholders in finding suitable apps for a healthcare question. Therefore, *MHAD* was conceived and implemented. The application of the *MARS* rating to systematically assess the quality of mHealth apps is time-consuming and challenging. For example, it must be determined how new releases of already rated apps are handled. Beyond this untackled aspect in *MHAD*, many more limitations could be mentioned, including the lack of evidence-based studies. However, *MHAD* is a first starting point to guide interested stakeholders in the field of mHealth apps based on qualified ratings. To conclude, many more goals must be addressed in future work.

REFERENCES

[1] F. Beierle *et al.*, "What data are smartphone users willing to share with researchers?" *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–13, 2019.

[2] R. Kraft *et al.*, "Combining Mobile Crowdsensing and Ecological Momentary Assessments in the Healthcare Domain," *Frontiers in Neuroscience*, vol. 14, p. 164, 2020.

[3] T. Probst *et al.*, "Emotional states as mediators between tinnitus loudness and tinnitus distress in daily life: Results from the "TrackYourTinnitus" application," *Scientific reports*, vol. 6, no. 1, pp. 1–8, 2016.

[4] W. Schlee *et al.*, "Measuring the Moment-to-Moment Variability of Tinnitus: The TrackYourTinnitus Smart Phone App," *Frontiers in aging neuroscience*, vol. 8, p. 294, 2016.

[5] R. Pryss *et al.*, "A personalized sensor support tool for the training of mindful walking," in *15th Int'l Conf on Wearable and Implantable Body Sensor Networks*. IEEE, 2018, pp. 114–117.

[6] E. Messner *et al.*, "mHealth Applications: Potentials, Limitations, Current Quality and Future Directions," in *Digital Phenotyping and Mobile Sensing*. Springer, 2019, pp. 235–248.

[7] IQVIA, "IQVIA Institute for Human Data Science Study: Impact of Digital Health Grows as Innovation, Evidence and Adoption of Mobile Health Apps Accelerate - IQVIA," https://www.iqvia.com/newsroom/2017/11/impact-of-digital-health-grows-as-innovation-evidence-and-adoption-of-mobile-health-apps-accelerate/, 2017, [Online; accessed 05-March-2020].

[8] S. C. Mathews *et al.*, "Digital health: a path to validation," *NPJ digital medicine*, vol. 2, no. 1, pp. 1–9, 2019.

[9] C. Petersen, "mHealth: Don't Forget All the Stakeholders in the Business Case," *Medicine 2.0*, vol. 4, no. 2, 2015.

[10] American Psychiatric Association, "App Evaluation Model," https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model, 2017, [Online; accessed 05-March-2020].

[11] Baumel, Amit and others, "Enlight: A Comprehensive Quality and Therapeutic Potential Evaluation Tool for Mobile and Web-Based eHealth Interventions," *J Med Internet Res*, vol. 19, no. 3, p. e82, Mar 2017. [Online]. Available: {http://www.jmir.org/2017/3/e82/}

[12] J. B. Torous *et al.*, "A Hierarchical Framework for Evaluation and Informed Decision Making Regarding Smartphone Apps for Clinical Care," *Psychiatric Services*, vol. 69, no. 5, pp. 498–500, 2018, pMID: 29446337. [Online]. Available: https://doi.org/10.1176/appi.ps.201700423

[13] S. Stoyanov *et al.*, "Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps," *JMIR mHealth and uHealth*, vol. 3, no. 1, p. e27, 2015.

[14] Y. Terhorst *et al.*, "Psychometric evaluation of the Mobile Application Rating Scale (MARS)," *JMIR Preprints*, 01 2020. [Online]. Available: https://preprints.jmir.org/preprint/17689

[15] ——, "'Help from the App Store?': A Systematic Review of Depression Apps in German App Stores," *Verhaltenstherapie*, vol. 28, no. 2, pp. 101–112, 2018.

[16] L. Sander *et al.*, "'Help for trauma from the app stores?'A systematic review and standardised rating of apps for Post-Traumatic Stress Disorder (PTSD)," *European Journal of Psychotraumatology*, vol. 11, no. 1, p. 1701788, 2020.

[17] A. Carlo *et al.*, "By the numbers: ratings and utilization of behavioral health mobile applications," *NPJ digital medicine*, vol. 2, no. 1, pp. 1–8, 2019.

[18] HealthOn e. V., "HealthOn," https://www.healthon.de/, 2020, [Online; accessed 29-February-2020].

[19] Spirit Link GmbH, "digimeda," https://digimeda.de, 2020, [Online; accessed 29-February-2020].

[20] ZTG Zentrum für Telematik und Telemedizin GmbH, "AppCheck," https://appcheck.de/, 2020, [Online; accessed 29-February-2020].

[21] W. Xu and Y. Liu, "mHealthApps: A Repository and Database of Mobile Health Apps," *JMIR mHealth and uHealth*, vol. 3, no. 1, p. e28, 2015.

[22] R. Nouri *et al.*, "Criteria for assessing the quality of mHealth apps: a systematic review," *Journal of the American Medical Informatics Association*, vol. 25, no. 8, pp. 1089–1098, 2018.

[23] B. Zapata *et al.*, "Empirical Studies on Usability of mHealth Apps: A Systematic Literature Review," *Journal of medical systems*, vol. 39, no. 2, p. 1, 2015.

[24] M. Yasini *et al.*, "mHealth Quality: A Process to Seal the Qualified Mobile Health Apps," *Exploring Complexity in Health: An Interdisciplinary Systems Approach: Proceedings of MIE2016*, vol. 228, p. 205, 2016.

[25] A. Powell *et al.*, "Interrater Reliability of mHealth App Rating Measures: Analysis of Top Depression and Smoking Cessation Apps," *JMIR mHealth and uHealth*, vol. 4, no. 1, p. e15, 2016.

[26] E. Jeon *et al.*, "Analysis of the Information Quality of Korean Obesity-Management Smartphone Applications," *Healthcare informatics research*, vol. 20, no. 1, pp. 23–29, 2014.

[27] L. Zhou *et al.*, "The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study," *JMIR mHealth and uHealth*, vol. 7, no. 4, p. e11500, 2019.

[28] M. Jin and J. Kim, "Development and Evaluation of an Evaluation Tool for Healthcare Smartphone Applications," *Telemedicine and e-Health*, vol. 21, no. 10, pp. 831–837, 2015.

[29] K. Scott *et al.*, "A Review and Comparative Analysis of Security Risks and Safety Measures of Mobile Health Apps," 2015.

[30] S. Berrouiguet *et al.*, "Fundamentals for Future Mobile-Health (mHealth): A Systematic Review of Mobile Phone and Web-Based Text Messaging in Mental Health," *Journal of medical Internet research*, vol. 18, no. 6, p. e135, 2016.

[31] R. Schnall *et al.*, "A user-centered model for designing consumer mobile health (mHealth) applications (apps)," *Journal of biomedical informatics*, vol. 60, pp. 243–251, 2016.

[32] B. Martínez-Pérez *et al.*, "Development and Evaluation of Tools for Measuring the Quality of Experience (QoE) in mHealth Applications," *Journal of medical systems*, vol. 37, no. 5, p. 9976, 2013.

[33] B. Kim *et al.*, "Consumer Mobile Apps for Potential Drug-Drug Interaction Check: Systematic Review and Content Analysis Using the Mobile App Rating Scale (MARS)," *JMIR mHealth and uHealth*, vol. 6, no. 3, p. e74, 2018.

[34] E. Messner *et al.*, "Development and validation of the German version of the Mobile Application Rating Scale (MARS-G)," *JMIR mHealth and uHealth*, 2019.

[35] J. Chen *et al.*, "The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment," *JMIR mHealth and uHealth*, vol. 3, no. 4, p. e104, 2015.

[36] W. Brown III *et al.*, "Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology," *Journal of biomedical informatics*, vol. 46, no. 6, pp. 1080–1087, 2013.

[37] E. Anthes, "Pocket psychiatry: mobile mental-health apps have exploded onto the market, but few have been thoroughly tested," *Nature*, vol. 532, no. 7597, pp. 20–24, 2016.

[38] K. K. Weisel *et al.*, "Standalone smartphone apps for mental health—a systematic review and meta-analysis," *npj Digital Medicine*, vol. 2, no. 1, pp. 1–10, 2019.

[39] European Commission, "Green Paper on mobile health ("mHealth")," https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth, 2020, [Online; accessed 29-February-2020].

[40] Food and Drug Administration, "Policy for Device Software Functions and Mobile Medical Applications," https://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications, 2020, [Online; accessed 29-February-2020].

[41] Federal Institute for Drugs and Medical Devices, "Guidance on "Medical Apps"," https://www.bfarm.de/EN/MedicalDevices/Differentiation/MedicalApps/_node.html, 2020, [Online; accessed 29-February-2020].

[42] J. Klein *et al.*, "Internetbasierte Interventionen in der Behandlung psychischer Störungen," *Der Nervenarzt*, vol. 87, no. 11, pp. 1185–1193, 2016.

[43] Ulm University, "MARS-G Online Training," https://www.youtube.com/watch?v=5vwMiCWC0Sc&amp=&t=1367s, 2020, [Online; accessed 29-February-2020].

[44] C. Moberg *et al.*, "Guided Self-Help Works: Randomized Waitlist Controlled Trial of Pacifica, a Mobile App Integrating Cognitive Behavioral Therapy and Mindfulness for Stress, Anxiety, and Depression," *Journal of medical Internet research*, vol. 21, no. 6, p. e12556, 2019.