

# CHARACTERIZING MEDICAL ANDROID APPS

Raina Samuel, Iulian Neamtiu, Sydur Rahaman, James Geller  
*New Jersey Institute of Technology*

## ABSTRACT

There is a proliferation of medical mobile apps: Google Play alone has thousands of apps in the “Medical” category. Many such apps perform critical tasks (e.g., are used with a medical device or in lieu of a device); handle sensitive patient related information; perform diagnosis; or treat diseases. However, there are wide gaps between an app’s claims and users’ expectations as well as between app implementations and regulatory frameworks’ mandates. We perform the first study, based on analyzing more than 4,000 Android apps, that characterizes medical apps. We begin by introducing an automated classification scheme that integrates textual information extracted from multiple sources to establish the purpose and target audience for an app, based on fine-grained traits and high-level categories; we found that the most common functionalities involved connecting to medical devices (e.g., hearing aids, glucometers), offering tele-health services, or patient management. We then dive deeper into app nature and characterize according to the function and domain of the app. We reveal actionable findings found in various facets of medical applications, regulatory frameworks and user privacy and safety.

## KEYWORDS

Mobile computing, characterization, mobile health applications, Android, health informatics

## 1. INTRODUCTION AND BACKGROUND

Over the past decade, the digital/mobile health area has grown substantially, as devices have become more advanced and more ubiquitous (Statista, 2021). On the Android platform alone, this pervasiveness has led to thousands of health-related apps. Furthermore, virtually all hospitals have enabled patients to access their health information via portal apps in both the outpatient and inpatient setting (Johnson, 2021).

Many users of these medical apps are unfamiliar with the app landscape and unsuspectingly trust that the apps are safe. Users should not be expected to question the legitimacy of a medical app or “dissect” an app to understand what it is doing with personal data. Medical apps can be valuable tools, but there is no universal standard that defines what is effective and does not put personal data at risk. A 2015 study of apps that evaluate symptoms for self-diagnosis and triage revealed that many deficits exist in both aspects (Semigran, 2015). Such lapses are potentially a public health issue, as apps are often used to make healthcare decisions.

Moreover, with the widening scope of medical apps, their capabilities and intended audience remain unclear. For example, the app **Instant Heart Rate: HR Monitor & Pulse Checker** has over 10,000,000 installs; the app's description states that it is the “most accurate” heart app and has been used in research. While its functionality is legitimate, the disclaimer states that “Instant Heart Rate should be used for entertainment purposes.” In order to triage potential app abuses or misleading claims, a clear consistent classification scheme of apps and their functionalities is essential.

We present a characterization of medical apps in Section 2. We begin by categorizing apps using a multifaceted analysis employing three main sources from an app's metadata: app description, XML (extensible markup language) assets, and image assets. By this process we extract relevant medical keywords. The keywords are used to determine orthogonal fine-grained traits that describe app functionality (e.g., sending patient data, found in 775 apps, or handling insurance, found in 376 apps). Combining these traits leads to a higher-level categorization scheme that allows us to better understand the app's intended audience. We establish six categories: virtual visit, patient portal, medical device, professional, reference, and patient. Virtual visit apps allow users to interact with medical professionals remotely. Patient portal apps

allow users to access information regarding visits or book appointments. Medical device apps interface with an external device, such as a glucometer. Professional apps are intended for medical professional uses in office management or patient care. Reference apps provide study and reference material. Finally, patient apps are intended for general personal health reminders. We find that the most popular category is *patient*, with 1,993 apps, followed by *reference* with 1,590 apps. Our classification scheme shows that the most common functionalities of medical apps involve connecting to medical devices, tele-health, and medical calculators.

Section 3 discusses actionable findings from our research. We investigate possible lapses found in the way regulatory agencies approve and determine medical apps and their functionalities. We discuss privacy implications of handling user data and how developers and marketplaces should be more transparent in how sensitive data is handled.

Our paper makes several contributions:

1. An automated approach and study that characterize medical apps into sub-categories to better understand their purposes and functionalities.
2. A discovery of the most common functionalities of medical apps, such as connecting to medical devices, providing telehealth management, or patient management.
3. A discussion of regulatory frameworks and user privacy practices and how these can be improved for the benefits of both developers and users.

*Prior Work.* Safety concerns regarding medical mobile apps have been a prominent subject of study. Magrabi et al. (Magrabi, 2019) (1) argue that it is difficult to regulate healthcare apps due to the fact anyone can develop an app and (2) confirm that there is little to no monitoring of use or formal evaluation of such apps, which is exceptionally concerning as more and more medical apps are being produced and recommended by physicians to patients to help track and monitor symptoms. Mobile mental health apps are a growing field aimed to help patients manage their mental health conveniently on their mobile devices. However, Terry et al. (Terry, 2018) revealed a lack of clear regulations of mobile mental health apps and created a typology of mobile mental health apps. Terry et al. also discuss that it is very difficult to judge the quality and efficacy of mental health apps, especially since many of the apps were developed outside of traditional healthcare spaces, revealing deficiencies in current regulatory frameworks. However, none of the aforementioned studies provide an overview of current regulatory frameworks globally found in our study.

Developing safe medical apps and understanding the risks that apps entail has been a topic of discussion and research. Yet there is a lack of discussion on the primary audience(s) apps are intended for. For example, Lewis et al. (Lewis,) evaluates and creates a risk framework of medical apps based on functionality. Additionally, Wicks et al. (Wicks, 2015) provide methods on how one can develop a medical app safely and securely. Akbar et al. (Akbar, 2020) performed a series of meta-analyses on 74 app studies; they exposed a variety of safety concerns in medical apps and grouped apps based on functionality. In our work, however, we determined app functionality on thousands of medical apps currently on Google Play.

Tangari et al. (Tangari, 2021) discovered severe privacy issues in 88% of medical apps used in their study, i.e., medical apps could potentially share user data with third parties, namely advertising and tracking services. Despite conducting a study with thousands of medical applications, the authors did not provide a characterization scheme as we have.

Tools assessing mobile medical app quality have been developed, namely Stoyanov et al.'s MARS (Medical App Rating Scale). However, they focus on iOS apps and do not rate apps based on overall intended usage and audience as well as potential risks (Stoyanov, 2015).

## 2. CHARACTERIZATION

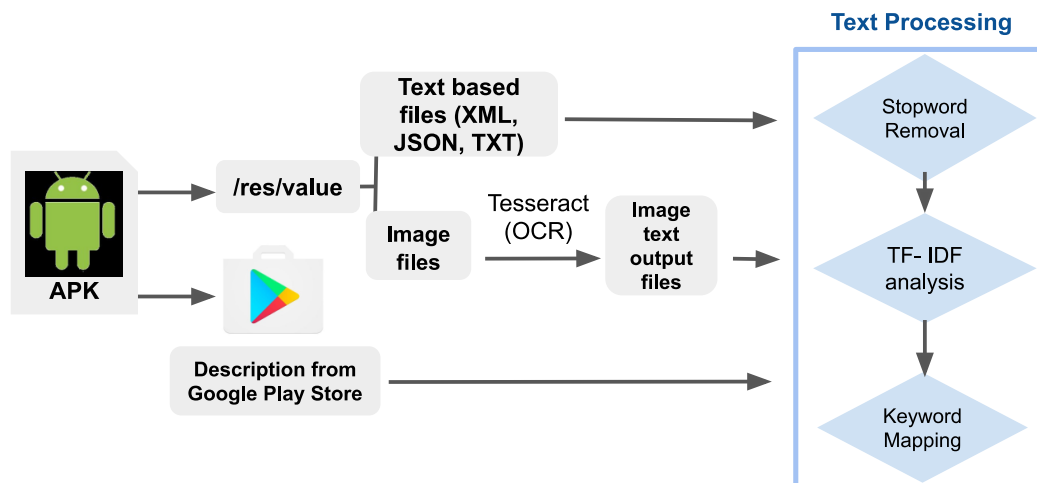
App characterization -- understanding the nature, purpose, and target audience of an app -- is challenging, as detailed next. To address these challenges, we use multi-source information along with a multi-rater human approach. First, we use information retrieval to extract terms of interest, and then define first-order low-level *traits*. Building upon traits, we then establish high-level *categories*.

*Challenges.* Characterization is a major challenge for several reasons. First, apps may serve more than one purpose, e.g., an app may manage a patient's prescriptions, help locate the nearest emergency room, and support video chats with the provider. Second, app features are hard to detect automatically (e.g., video chat software can be home-made as opposed to using a video chat library); similarly, location/mapping services

can serve several purposes, thus the presence of such a library simply indicates that the app provides location-relevant services. Third, the app description on Google Play is at the developers' discretion, and can be incomplete, inaccurate, or downright misleading. Fourth, actual app functionality can only be reconstituted from heterogeneous sources via a multi-faceted analysis of app description, embedded images, app bytecode, etc.

We started by retrieving all apps (APK files) from Google Play's *Medical* category along with their descriptions. We only retained those apps that had English descriptions and at least 1,000 installs, for a total of 2,215 apps. Our approach for extracting relevant text is shown in Figure 1 and described in detail next.

Figure 1. Trait extraction



## 2.1 Sources

An APK is essentially an archive of directories and files that include bytecode, resource files, assets, and libraries. Among all these files, we find that text relevant to app functionality and user interactions with the app appears in three main locations: XML assets, images, and app descriptions, as shown on the left of Figure 1. We describe each of these and provide evidence why all three sources are needed.

Table 2. Location and frequency of relevant keywords

App	Frequency (%)		
	XML Assets	Image Assets	Description
SimplePharmacology	31	69	-
MyNM by Northwestern Medicine	100	-	-
AnthroCalc	-	-	100

*XML Assets.* Layout XML files, stored in the app's *res/* directory, define the user interface by storing all text views, buttons, and other UI (user interface) elements. We are interested in these features as we can discover what information the app is requesting from the user and what kind of information the user provides to the app by interacting with it. String XML files store strings accessed by the application, which constitute another key location of medical terminology that can be extracted.

*Image Assets.* These assets, also stored in the app's *res* directory, are relevant as well. For example, some apps may use an image as a button, rather than defining the button's text string as an XML asset. In general, images may contain relevant text. We use the Tesseract OCR (optical character recognition) package (Tesseract, no date) on image files to extract the English text present in images.

*Descriptions.* App descriptions are found on Google Play and not within the app itself. As a result, the description of an application allows a user to understand what an app does prior to installation. As the description is the first impression a user has of the app, its functionalities should be clearly defined in a way that help users establish an app's purpose confidently and securely. However, as the description is usually

written by the app developers themselves, the app is often portrayed in a flattering and overly positive way to attract users. Therefore, app descriptions can be inaccurate, misleading, or incomplete, which we found to be another essential aspect of an app that should be considered in our analysis.

*Why all three sources are necessary.* Table 1 illustrates why using only one of the three sources is insufficient: the table shows, for three apps, where the relevant keywords are located. For the app **SimplePharmacology**, 69% of the relevant keywords are in the image assets while the description contains no relevant keywords whatsoever. In contrast, for the app **MyNM**, all the keywords are found in the XML assets; images and the description contain no relevant terms. Finally, for the app **AnthroCalc**, all keywords are in the app description. Therefore, we need to analyze and integrate information from all three sources.

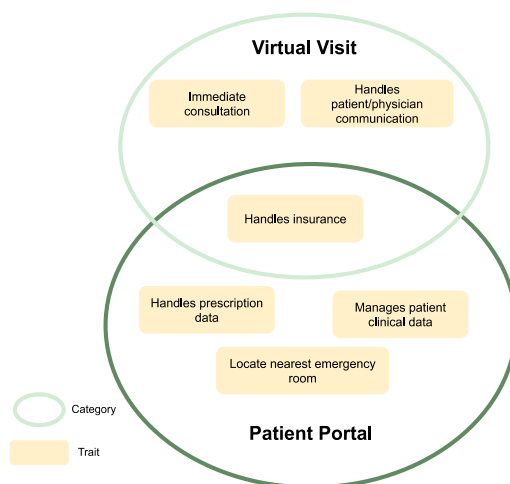
## 2.2 Methodology

In order to create a clear classification scheme based on app functionality, we referred to ICD'11 (International Classification of Diseases) and PHI (Protected Health Information) terms. ICD codes provide a reliable established standard of diseases and health conditions. PHI terms allow us to obtain a broad idea of what data is required from users in certain apps to determine their functionalities. We began our text processing with extracting the descriptions.

First, the descriptions had stop words removed to focus on conceptual information in the text, followed by a TF-IDF (term frequency – inverse document frequency) analysis (tf-idf, no date) based on ICD keywords and PHI terms. As a result, we could provide a preliminary classification of each application based on keyword matches and frequencies. The process was repeated for XML and image assets. With the resulting keywords extracted, we observed which resources provided the most relevant results. More keywords were found in the XML files of apps as opposed to image files (via Tesseract), or app descriptions. This evidences that descriptions do not paint a complete picture.

*Defining Traits and Categories.* We employed a multiple-rater approach (Green, 1997) to determine traits and categories: three human raters had to come to 100% agreement on what constituted and differentiated the various traits. Raters had to agree first on what should be considered a unique trait of a medical app and which keywords should be used in determining that trait (traits essentially define low-level orthogonal functionality “facets” for apps). As a result, 19 traits were determined. Subsequently, raters would then agree on what combination of traits would dictate the category of an app. Once the baseline was set, the app was classified using traits into categories, as illustrated in Figure 2. In this way, some apps may have multiple traits and belong to various categories. However, such categorization provides a more nuanced view on the general functionality of certain medical apps.

Figure 2. An example of category determination based on traits



## 2.3 Traits

Traits are defined as single aspects of app functionality, orthogonal to other aspects. Apps can exhibit multiple traits, as apps can provide several functionalities. We determined traits by finding common ICD terms and medical keywords. As a result, we defined 19 unique traits; their definitions and frequencies are shown in Table 2. Many of the traits are self-explanatory and commonly used, such as 'Anatomy' and 'Locate nearest emergency room.' There are certain traits that needed further refinement, specifically those dealing with patient data management. Table 2 reveals that many common traits involve reference material for medical professionals. For instance, Medical Student Study Aids are found in the top five traits in medical apps, as well as medical calculators, which are often used by professionals.

Table 2. Frequency of traits found in apps

Trait	Description	% Apps	#Apps
Anatomy	Anatomy reference material	61	1351
Well-being reminder	Keeps track of patient habits	51	1129
Medical student study aids	Medical student study material	38	841
Medical calculator	Calculates unsaved patient readings	37	819
Sends acquired patient data	Sends inputted patient data to a provider	35	775
Handles prescription data	Patient info regarding prescriptions	28	620
Manages patient clinical data	Stores medical history of a patient	27	598
Visual guide	Visual reference material used by a professional	26	575
Medical procedures	Procedural references for professionals	22	487
Patient journal/diary	Patient behavior or progress with disease or nutrition	20	443
Disease name	Disease reference material used by a professional	20	443
Handles patient/physician comm.	Stores and transmits patient info between patient and provider	20	443
Patient symptom tracker	Keeps track of various symptoms, may lead to diagnosis	19	420
Drug name	Drug name and pharmaceutical reference material	18	398
Handles insurance	Stores patient medical history related to insurance policy	17	376
Immediate consultation	Virtual consultations with a provider	16	354
Dose calculator	Calculations for patients and providers to administer medication	16	354
Locate nearest emergency room	Using current location to find an ER	14	310
Device measuring patient data	Using an external device to collect readings, e.g., blood pressure	14	310

## 2.4 Categories

Table 3. Category determination from traits

Trait \ Category	Reference	Patient Portal	Professional	Patient	Virtual Visit	Medical Device
Sends acquired patient data			●			●
Handles prescription data		●		●		
Handles patient/phys. comm.			●	●	●	
Handles insurance		●		●	●	
Manages patient clinical data		●	●			
Device measuring patient data						●
Patient symptom tracker				●		
Well-being reminder				●		

Dose calculator			●			●
Patient journal/diary				●		
Immediate consultation			●		●	
Anatomy	●		●			
Medical student study aids	●		●			
Medical calculator			●	●		
Disease name	●					
Medical procedures	●					
Visual guide	●					
Drug name	●		●			
Locate nearest ER		●	●	●		
<b>Total number of apps</b>	1,590	327	509	1,993	1,269	609
<b>Total percentage of apps</b>	72%	15%	23%	89%	57%	27%

The various combinations of certain traits allow us to determine specific categories of medical apps as is evident in Table 3. We established six unique categories that apps may fall into.

*Reference.* These apps serve either as general references regarding medical terms or first-aid procedures. Some apps are study aids or provide quizzes for medical professionals in training.

*Patient Portal.* Users can schedule and make appointments with their medical providers and view their lab results or test results and data from their visits. In addition, users can search for nearby providers.

*Professional.* These apps are directed towards medical professionals ranging from medical staff to office assistants. Many apps help medical clinics with scheduling and handling patient data in a professional setting.

*Patient.* Apps in this category are aimed at patients to help them log their daily progress or daily habits such as sleeping or pill reminders.

*Virtual Visit.* These apps provide for virtual visits, e.g., via a video call with a medical professional. In doing so, users often provide personal information and discuss their symptoms.

*Medical Device.* These apps are considered as medical devices or work in tandem with devices, such as hearing aids, glucometers, or sphygmomanometers for hypertension. Apps in this category can be used to store device readings and be maintained as a log or can be used as a remote control for the device.

### 3. ACTIONABLE FINDINGS

Our categorization has revealed that medical apps serve a broad audience and variety of purposes. However, because many such purposes are sensitive or even critical, and not intended for a general audience, there should be barriers for app access control. Theoretically, as these apps are free, and found on a public app distribution platform, anyone can download and use them, even though the apps are meant exclusively for professionals. Generally, apps that are meant for professionals in a hospital or clinical setting usually require credentials to access such systems. However, there are professional apps which can potentially result in a diagnosis or interface with a medical device for a procedure; if such apps are available for general use, it can lead to possible user harm. Hence there is a need for strong regulatory frameworks protecting end-users.

We now describe actionable findings covering various aspects of medical apps. We review current regulations and definitions regarding mobile health and medical apps established by various legal entities throughout the world, while also finding certain lapses and difficulties in implementing these guidelines. From these definitions, we discuss potential privacy implications and user safety concerns.

#### 3.1 Regulatory Framework Enforcement

*Actionable finding: Regulatory frameworks should be clearer defined and more accessible for developers when creating medical apps managing user data.*

Medical apps can perform critical functions that involve patient data or other sensitive information. Overall, app users generally assume that apps are “certified” and trustworthy when making medical decisions. The question that arises is whether these apps are indeed approved by regulators and safe for use.

For example, in the United States, the FTC (Federal Trade Commission) provides definitions and guidelines for app developers. The guidelines indicate whether the app is a medical device, or a medical app; as well as whether the FTC will apply any regulatory oversight (FTC, 2019f). Additionally, the FDA (United States Food and Drug Administration) regulates functions of mobile devices that use device sensors (camera, light, vibrations) to perform medical device functions (e.g., measuring blood pressure), connecting a mobile device to a medical device and being able to manipulate it from the mobile device (e.g., alter settings of an implant), or active patient monitoring (e.g., acquiring signals from a cardiac monitor) (FTC, 2019d).

EU regulation of mobile medical apps focuses on potential privacy concerns that may arise. Mobile health apps must comply with data protection laws (Data Protection Directive) that were enacted, as well as ensuring that apps provide ‘clear and unambiguous information about processing to end users before app installation’ (Crossley, 2016).

Some Asian countries, such as China and Japan, regulate standalone medical software as medical devices, though depending on the overall software class, whose definition is based on functionality (Gross, 2017).

Overall, regulatory bodies have general guidelines on medical app behavior and functionality. However, there is no clear standard for app developers to easily refer to when developing a medical app. Having an accessible flowchart or a streamlined explanation of definitions would aid developers as well as app markets (Google Play, Apple's App Store) in managing the apps, especially apps handling users’ medical data.

### **3.2 User Security and Safety**

*Actionable finding: Medical apps should be more transparent regarding user data management prior to installation.*

App functionality plays a large role in determining whether the app falls under a regulatory framework. Medical apps often manage identifiable and private health information, that is, demographic information related to a user's health or condition that can be used to identify the user. For instance, in the US, if such apps work with health care providers or HIPAA entities, they are subject to HIPAA rules regarding security (FTC, 2019e) and privacy (FTC, 2019a) and what must be done when a breach has occurred (FTC, 2019b). However, not all data acquired by an app is considered identifiable health information. For example, an app measuring a user's weight and blood pressure is not considered a big security risk, compared to an app that tracks patient activity and prescriptions. Thus, certain apps pose lower risks to user privacy and would not need to be under scrutiny from regulatory bodies. An example would be apps that are general aids or of general purpose (e.g., magnifying glass); automate general office functions in healthcare and are not used for diagnosis; and educational apps (e.g., flashcards, encyclopedias, textbooks). These apps are neither regulated nor will have any discretionary enforcement exercised on them. However, as discussed previously, many medical apps handle patient data, and despite regulations and guidelines, users do not know how securely their data is managed or transmitted. App developers and markets must be more forthcoming and transparent about patient data management, by concisely explaining to users prior to installation what happens to their private health information. Potentially, these entities should be held accountable, should any leaks occur.

## **4. CONCLUSION**

Medical apps across many categories have been implemented and publicized that serve millions of users and provide a multitude of functions. To better understand the app landscape, our study categorizes medical apps based on stated and observed functionality. Overall, our research makes several contributions. First, we provide an automated approach and study that characterize medical apps into sub-categories to better understand their purposes and functionalities. Second, observe the most common functionalities of medical

apps. Third, we discuss regulatory frameworks and user privacy practices. By doing so, we are better equipped to undertake further studies into app behavior, app security, app claims, etc.; and ultimately improve the health and well-being of app users.

## REFERENCES

- Akbar, S. et al. (2020) "Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review", *J. Am. Med. Inform. Assoc.* Oxford University Press (OUP), 27(2), bll 330–340.
- Crossley, S. (2016) "EU regulation of health information technology, software and mobile apps". Available at: [https://uk.practicallaw.thomsonreuters.com/2-619-5533?contextData=\(sc.Default\)](https://uk.practicallaw.thomsonreuters.com/2-619-5533?contextData=(sc.Default)).
- FTC (2019a) "Breach Notification Rule". Available at: <https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html>.
- FTC (2019b) "Breach Notification Rule". Available at: <https://www.hhs.gov/hipaa/for-professionals/breach-notification/index.html>.
- FTC (2019c) "Examples of Device Software Functions the FDA Regulates". Available at: <https://www.fda.gov/medical-devices/device-software-functions-including-mobile-medical-applications/examples-device-software-functions-fda-regulates>.
- FTC (2019d) "Examples of Software Functions for Which the FDA Will Exercise Enforcement Discretion". Available at: <https://www.fda.gov/medical-devices/device-software-functions-including-mobile-medical-applications/examples-software-functions-which-fda-will-exercise-enforcement-discretion>.
- FTC (2019e) "Health Information Laws". Available at: <https://www.hhs.gov/hipaa/for-professionals/security/index.html>.
- FTC (2019f) "MOBILE HEALTH APPS INTERACTIVE TOOL". Available at: <https://www.FTC.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool>.
- FTC (2019g) "MOBILE HEALTH APPS INTERACTIVE TOOL Glossary". Available at: <https://www.FTC.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool#glossary>.
- Gross, A. (2017) "Column - Medical Software Regulations in Asia 2017", *MedTech Intelligence*. MedTech Intelligence. Available at: <https://www.medtechintelligence.com/column/medical-software-regulations-asia-2017/>.
- Green, A. M. (1997) "Kappa Statistics for Multiple Raters Using Categorical Classifications", in *Proceedings of the 22nd annual SAS User Group International conference*, bl 4.
- Johnson, C. et al. (2021) "Hospital Capabilities to Enable Patient Electronic Access to Health Information, 2019". ONC Data Brief.
- Lewis, T. L. et al. (2014) "mHealth and Mobile Medical Apps: A Framework to Assess Risk and Promote Safer Use", *J Med Internet Res*. doi: 10.2196/jmir.3133.
- Magrabi, F. et al. (2019) "Why is it so difficult to govern mobile apps in healthcare?", *BMJ Health & Care Informatics*. BMJ Specialist Journals, 26(1). doi: 10.1136/bmjhci-2019-100006.
- Semigran, H. L. et al. (2015) "Evaluation of symptom checkers for self diagnosis and triage: audit study", *BMJ*. BMJ Publishing Group Ltd, 351. doi: 10.1136/bmj.h3480.
- Shuren, J., Patel, B. en Gottlieb, S. (2018) "FDA regulation of Mobile Medical Apps", *JAMA*, 320(4), bl 337. doi: 10.1001/jama.2018.8832.
- Statista Research Department (2021) "Healthcare apps available Google Play 2021", *Statista*. Available at: <https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/>.
- Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). "Mobile app rating scale: a new tool for assessing the quality of health mobile apps." *JMIR mHealth and uHealth*, 3(1), e27. <https://doi.org/10.2196/mhealth.3422>
- Tangari, G. et al. (2021) "Mobile Health and Privacy: Cross Sectional Study", *BMJ*. doi: 10.1136/bmj.n1248.
- Tesseract-Ocr (no date) "tesseract-ocr/tesseract: Tesseract Open Source OCR Engine (main repository)", *GitHub*. Available at: <https://github.com/tesseract-ocr/tesseract>.
- Terry, N. P. et al. (2018) "Regulating mobile mental health apps", *Behavioral Sciences & the Law*, 36(2), bll 136–144. doi: 10.1002/bsl.2339.
- "tfidf :: A Single-Page Tutorial - Information Retrieval and Text Mining" (no date). Available at: <http://www.tfidf.com/>.
- Wicks, P. et al. (2015) "'Trust but verify' – five approaches to ensure safe medical apps", *BMC Medicine*, 13(1). doi: 10.1186/s12916-015-0451-z.