In collaboration with
Deloitte

Global Governance Toolkit for Digital Mental Health:
Building Trust in Disruptive Technology for Mental Health

WHITE PAPER

APRIL 2021

Contents

"Our ambition for mental health is a world where every person realizes their desired emotional, psychological and social potential."

Stephanie Allen
Deloitte Global Health Care Sector Leader

Foreword

Disruptive technology in mental health provides an opportunity to create breakthrough solutions that improve mental health and well-being outcomes on a greater scale than ever before.

Stephanie Allen
Global Health Care Sector Leader, Deloitte Global, Australia

Arnaud Bernaert
Head of Global Health and Healthcare, World Economic Forum, Geneva

It could be argued that the greatest discoveries for humanity, the next frontier, will not be in quantum technology or space, but will instead be those that help us to better understand our own minds. The COVID-19 crisis, alongside the extraordinary political, economic and social disruptions of 2020, have exposed an enduring silent epidemic and greatly accelerated the need for a properly funded and functioning global mental health ecosystem. The technology required to unlock that next level of research in mental health, and scale access to treatments for 8 billion minds, is only just starting to be developed. Yet while disruptive technology in mental health has the potential to drive enormous health, social and economic change, it also presents the opportunity for misuse and mistreatment. To meet the challenge, these tools will need to be safe, clinically validated and trusted.

Essential considerations for governments, health system owners and civil society are how to make sure innovations and technologies are harnessed and regulated to accelerate growth and improve outcomes. How can governments and stakeholders approach governance in a way that will account for the unique challenges of AI in mental health, navigating culturally different approaches, dealing with product agility and ensuring they are incentivizing, not punitive? How can health system

owners and innovators work together to inspire trust in the next wave of digital services that will play an important role in how society, economies and businesses reset in a post-COVID world?

These questions provided a need and an opportunity for an impact-driven partnership between the World Economic Forum and Deloitte, focused on creating a policy framework to guide leaders towards crafting successful policies on technology and mental health, and an ethical framework to assist innovators in working with their consumers to solve moral dilemmas. The collaboration is part of a larger World Economic Forum initiative, the Great Reset, that explores how, as the world undergoes transformation, the ability to harness and disseminate disruptive technology will play a vital role in ensuring our recovery from the pandemic and the avoidance of future crises.

This partnership's ambition for mental health is a world where every person may realize their desired emotional, psychological and social potential. In mental health, trust is more than the mitigation of risks of unethical and malicious uses, it is working with communities to act responsibly. Not only is this the start of that journey – which will not be easy – but we have a clear medical, moral and economic imperative to do better.

Executive summary

Mental health and behavioural health disorders impose a significant and under-recognized burden on the global healthcare system, affecting social needs, basic human rights and the economy.

Between a quarter and a half of the global population is affected by a mental disorder at some point in life.[1] The human cost is immense: 800,000 people commit suicide every year. Suicide is a leading cause of death among young people. People suffering from untreated mental health disorders are unable to realize their desired potential and are all too often exposed to a wide range of human rights violations.[2] Between 2011 and 2030, the cumulative economic output loss associated with mental disorders is projected to be $16.3 trillion worldwide.[3] The secondary consequences of mental ill-health are estimated to cost employers $2,000 per employee per year from presenteeism (i.e. being present but unable to focus), absenteeism and unnecessary turnover.[4] The direct and indirect costs can amount to 5% of a country's GDP.

In recent times, major health, climate and political crises, such as the COVID-19 pandemic and the 2020 bush fires, have highlighted and exacerbated the world's mental health challenges. Public health agencies have warned that a wave of depression,

suicide and other mental ill-health issues is on the horizon.[5,6] COVID-19 has also accelerated digital innovation at an unprecedented rate, transforming services and offering a huge opportunity to move forward with changes in the mental health system.

Disruptive technology – such as artificial intelligence and machine learning (AI/ML), digital reality (DR), blockchain and the cloud – is ushering in a new era of productivity and operations for consumers, industries and organizations.[7] Disruptive technology also offers a tremendous opportunity to improve global mental health systems, making them more affordable and more easily scaled, particularly for countries and people without adequate access to such services. Technology has transformed finance, transport, tourism, education and media industries beyond recognition over the past 20 or even 10 years, and it is set to transform the mental health sector as well.[8] There are more than 10,000 apps related to mental health in the Apple App Store and the Google Play Store.[9] The overwhelming majority of those 10,000 apps, it should be noted, are not currently evidence-based.[10]

Guiding purposes

Trusted, strategic and safe digital mental health and well-being services

Ensure that digital mental health and well-being services can be accepted as reliable and safe by consumers, promoted by clinicians and trusted by governments and healthcare-system owners so they can be supported and strategically integrated

Principles and standards

Cultivate trust through common ethical principles and standards that safeguard the needs of consumers, clinicians and healthcare systems

Drive action

Demonstrate results by piloting with innovators, governments and health-system owners and seeing widespread adoption of our principles and standards

Much disruptive technology in digital mental health solutions, however, raises new ethical questions about safety, efficacy, equity and sustainability. Who is deciding whether a psychologist chatbot is trained and optimized for a consumer's mental-health outcomes or for the use and profitability of the service? Who is ensuring that a person's mental health-related information is not being used unscrupulously by advertising, insurance or criminal justice systems? Questions such as these are troubling in the light of current regulatory structures.

The primary purpose of this toolkit is to provide governments, regulators and independent assurance bodies with the tools to develop, adopt and engage standards and policies that address major ethical concerns, thereby protecting consumers, enabling them to assess quality mental healthcare more easily, helping them make more informed choices about their own mental health and encouraging the strategic growth of safe, ethical and effective digital mental health services. The toolkit is also intended to provide governments, businesses and service vendors with a guide to operating ethically and safely. It does this by exploring the current and future opportunities and ethical issues presented by disruptive technology in mental health, and offering a framework of governance principles, standards and processes that may be adopted by stakeholders as a code of ethics, as regulatory standards or simply as a kitemark of compliance, with a means for adapting these to the cultural, legal, medical and clinical situations in different jurisdictions.

This toolkit is unique in, first, promoting broader ethical, clinical, technical and human mental health domains based on shared goals and values; second, addressing a broad scope of disruptive technology in mental health and well-being; and, third, being designed to have a global effect with practical and implementable recommendations. There are early pilots already in play. The toolkit enables stakeholders to validate and develop digital mental health services that are safe, strategic and ethical, by focusing on the means of developing trust in digital mental health services through assurance schemes and transparency. Trusted services are defined as those upholding the values of being effective, ethical, equitable, safe, sustainable, trustworthy and usable. Encouraging consumer trust in the services will support higher engagement and better outcomes. If clinicians can trust in these services, they will recommend or prescribe them more frequently and may integrate these services into their practice, potentially increasing their own practice's reach, affordability and effectiveness. Building trust in these services will empower businesses and communities to invest in them and incorporate them into health systems, to be more holistic in approach, offer services based on consumer needs, scale services to more people, provide more affordable options and enable improved data-driven decision-making for better population health and well-being outcomes.

Guiding principles

Beneficial intent

We imagine a world in which everyone can access services that support them to reach their desired emotional, social and psychological potential

Community openness

This toolkit is shaped with opinions from a broad spectrum of stakeholders across countries and cultures with differing levels of maturity, including consumers, innovators and those in the public sector, business, healthcare sector and academia. The intention is for wide, open release for everyone. We recognize that perspectives will diverge and we seek bipartisan support

Peer review

We validate our work with experts in the field, encouraging respectful debate to strengthen our collective thinking

What is the role of this toolkit?

Building on the World Economic Forum report Empowering 8 Billion Minds: Enabling Better Mental Health for All via the Ethical Adoption of Technologies,[11] this toolkit seeks to help improve the accessibility, quality and safety of services that support all members of society to meet their desired emotional, social and psychological potential by helping users to:

– Understand the potential for digital mental health services in improving the mental and behavioural health of all people

– Develop principles and standards for the safe, ethical and strategic implementation of digital mental health services

– Adapt, pilot and adopt these standards and principles in countries, jurisdictions and health systems across the globe

– Improve access, effectiveness, quality and safety of digital mental health solutions by adopting better practices and standards

– Make strategic investment and incentivization decisions in the global digital mental health ecosystems to encourage its growth

– Make informed decisions to incorporate digital mental health tools into a health system, workplace, community, product or service seamlessly

Who will find this toolkit useful?

Government and regulators
The primary goal is to persuade government health departments and health regulators to pilot and adopt principles and standards that encourage the safe, ethical and strategic implementation of digital mental health services in their respective jurisdictions. The potential impact is the adoption of services that offer low-, middle- and high-income populations access to scalable, effective, affordable and needs-based services. The toolkit also provides foundational insights vital to understanding the levers available to develop incentives for growth, make more strategic commissioning decisions, and augment and integrate with existing healthcare models for a more holistic, seamless mental and behavioural health system.

Healthcare and insurance organizations
This toolkit also provides healthcare and insurance organizations with rubrics to understand digital mental health services and the core components to be considered in order to assess, build and integrate high-quality digital mental health services. The principles are fundamental to making successful commissioning and investment decisions, creating effective combined models of care and encouraging safe innovation in approaches to consumers' mental and behavioural health.

Digital mental health innovators and vendors
By bringing together the views of consumers, clinicians, innovators, government and healthcare providers, this toolkit offers a roadmap for digital mental health innovators and vendors to create trusted services. The standards detail the potential critical interests of those stakeholders. When these standards are applied to their services, they are the means to exceed consumers' expectations and accelerate opportunities for adoption, investment and integration.

Businesses and employers
Organizations have an opportunity to increase productivity, attract better personnel and reduce insurance risks by adopting improved employee mental health and well-being programmes, with a return on investment of between 2.3 and 5.7 to 1.[12,13] This toolkit provides businesses with the tools needed to understand the core attributes used to assess and integrate innovative, efficient and effective digital mental health apps and services into employee well-being programmes and practices.

Consumers and communities
The ultimate purpose of this toolkit is improving access to services that will enable all consumers around the world to meet their desired emotional, social and psychological potential. Communities may use these principles to advocate for, and make the best use of, state-of-the-art services to pursue their mental health goals.

FIGURE 1 | Navigating the toolkit

Executive Summary
Who will find this toolkit
useful and why?

1. Digital mental health
What is disruptive technology in
mental health?

- Major challenges to global mental health
- What is disruptive technology and digital mental health?
- Benefits vs. barriers, the trust challenge for digital mental health

2. Goals, values and standards
What does quality, safety and
effectiveness look like?

- Our shared goals and values for digital mental health
- Standards for digital mental health

3. Ethical AI in mental health
How can we solve ethical dilemmas, in
partnership with our communities?

- Operate ethically in our communities and solve new moral challenges

4. Policy governance
How do we transform standards
into a system of governance?

- Creating a policy framework for governance
- Enforcing regulation and standards
- Organizational structures for governance

5. Incentivizing innovation
What about scaling and growth?

- Target operating model
- Incentivizing innovation

6. Governance pilots
Where has it been done before?
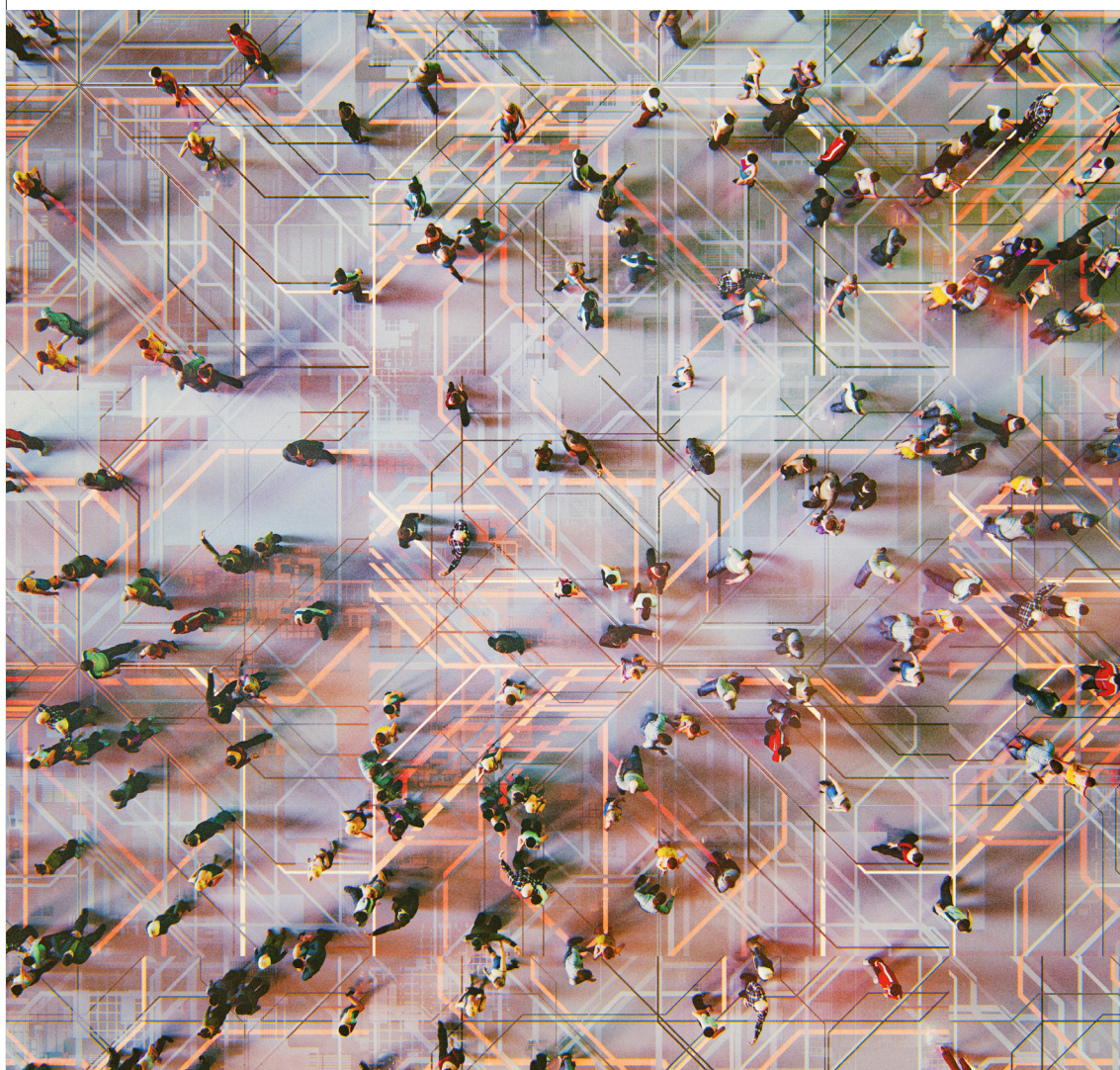
- Other examples of governance efforts and lessons learned

(1) Digital mental health

What are the challenges to improving global mental health? Here we define disruptive technology in mental health and how it will shape our future.

Disruptive technology in mental health provides an opportunity to find breakthrough solutions that improve mental health and well-being outcomes on a greater scale than ever before. This section provi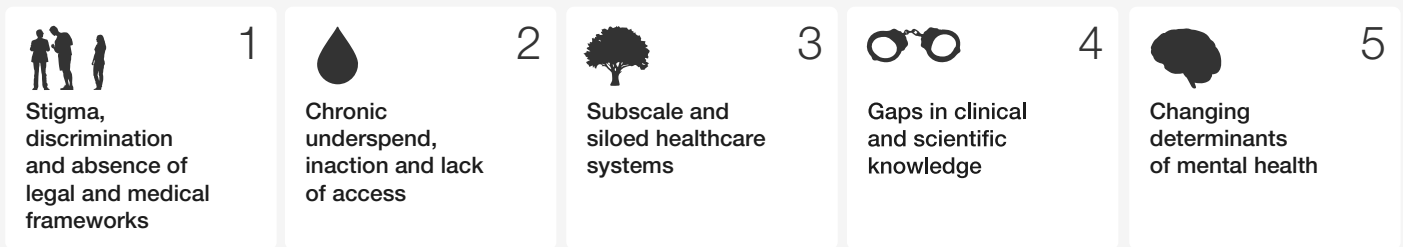des insights into the challenges to improving global mental health and defines disruptive technology in mental health, describing how this new mental health technology will shape the future and explaining the ethical concerns being raised about these new tools and services.

1.1 | Challenges to achieving better global mental health outcomes

Although some governments, employers, insurers and care providers are starting to invest in addressing mental health crises, there are significant global and local barriers to positive change. Below we explore some of those challenges before discussing how digital mental health approaches can play their part in addressing global mental health and well-being.

FIGURE 2 | Major challenges faced in achieving better global mental health outcomes

1	2	3	4	5
Stigma, discrimination and absence of legal and medical frameworks	Chronic underspend, inaction and lack of access	Subscale and siloed healthcare systems	Gaps in clinical and scientific knowledge	Changing determinants of mental health

1. Stigma, discrimination and absence of legal and medical protections
Around the world, individuals struggling with mental ill-health commonly face a far greater degree of stigma and discrimination than those with physical health conditions.[14] In many cases, the stigma means people are less willing to seek treatment or share their personal information with mental health professionals. In some cases, the discrimination leads to isolation, unemployment, criminalization, abuse of fundamental human rights through forced restraints, physical violence, sexual violence and torture and even premature mortality.[15] While the United Nations[16] and the World Health Organization[17] have recognized mental health as a universal human right, in places where mental health conditions are not universally recognized as a legitimate condition, current legal and medical frameworks do not effectively protect the best interests, or even the basic human rights, of individuals suffering from mental distress. This is especially so where their mental disorders profoundly interfere with their capacity to make informed decisions, resulting in unproved medical interventions, neglect, punishment, violence and criminal incarceration.

2. Chronic underspending, inaction and lack of access
The financial resources allocated by governments for mental health are alarmingly low. Development assistance for mental health has never exceeded 1% of global health development assistance,[18] and it does not come close to the expected $16 trillion negative economic impact that mental ill-health will have on the global economy between 2010 and 2030.[19] In the European Union (EU), mental health expenditure ranges from 2% to 14% of total health expenditure,[20] which falls universally short of the 19.5% total Disability Adjusted Life Years (DALY) health burden accounted for by mental disorders in the EU.[21] As a consequence, the majority of the world's population lack access to affordable, effective, timely and appropriate mental healthcare. In low-income countries, the number of mental health workers can be as low as two per 100,000 people[22] and in the United States, more than 100 million people live in communities designated as "health professional shortage areas" for mental health professionals.[23] It is not just a matter of a gap in access to treatment; experts also call attention to the quality gap, where lack of oversight leads to the promotion of unproven treatments; and the prevention gap, where there is a lack of investment across the board in awareness, early interventions, preventative and protective measures.[24]

3. Subscale and isolated healthcare systems
The traditional responses to mental ill-health – including professionally delivered talking therapies, prescription of pharmaceuticals, voluntary and

involuntary hospitalization and surgical interventions – often fail to meet the mental health demands being presented in high-income countries.[25, 26] When considering low-income countries that lack any existing mental health infrastructure, this model presents a high financial and time investment barrier.

As with physical health, systems of data gathering and sharing are imperfect across mental healthcare systems, making it harder to gather the required data to make clinical mental health decisions or refer consumers, and more difficult to create the links with other health initiatives, all leading to mental disorders being underreported or misreported.[27]

Advocacy for global mental health is threatened by the fragmentation of opinion in constituencies and scientific perspectives, exasperated by diverse cultural and social views of mental ill-health. Views differ in relation to, for example, the medicalized and specialized care model promoted by clinical practitioners; promotion of the study of the human brain by neuroscientists; the focus on fighting discrimination and social determinants by civil social activists; and holistic care and alternative approaches promoted by Indigenous populations. Differences of opinion lead to contradictory messages, lack of coherent plans and the risk of mental health becoming an isolated area within global health initiatives.

4. Gaps in clinical and scientific knowledge

Research into understanding mental health disorders is at an emerging and ongoing phase. Mental health disorder classifications often deal with multiple causalities, diagnoses have less discrete categorizations with unclear thresholds between disorder and non-disorder markers, and there are frequently comorbidities with other mental and physical health ailments that make treatment more complicated. Often, consumers are recommended treatments on a trial-and-error basis, and their awareness of, and the actions needed to address, mental health issues are less clear. Even well-defined disorders are difficult to diagnose and treat; for instance, consumers with bipolar disorder were initially misdiagnosed in 69% of cases, and more than one-third remain misdiagnosed for 10 years or more.[28]

5. Changing determinants of mental health

Major determinants of mental ill-health, such as poverty, pandemics, climate change, conflict, population demographic changes and political uncertainty, continue to contribute to both the number of people at risk of mental disorders and the dramatically increasing global disease burden represented by mental disorders. While reductions in the proportion of people living in absolute poverty is beneficial to overall mental health, other social determinants such as increasing global income inequality, rapid urbanization, economic and political uncertainty and the stress of increasing natural disasters from climate change, alongside improving capabilities to measure population health, contribute to the increasing risk and prevalence of mental disorders. The recent COVID-19 pandemic, for example, was a major contributor to severe mental distress tripling among Americans in 2020.[29]

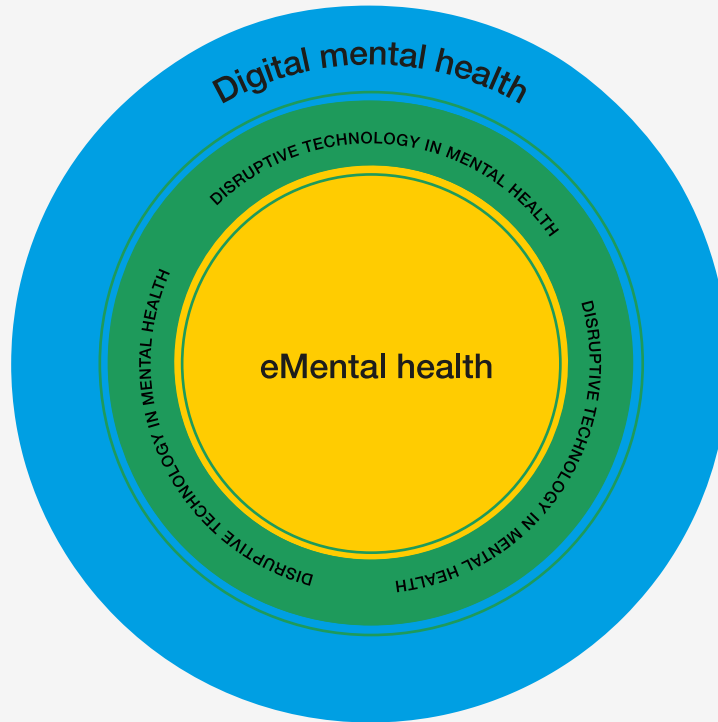## 1.2 | The role of disruptive technology in mental health

Disruptive technology in mental health refers to innovative technology-based solutions that significantly change the way societies prevent, identify, diagnose, treat and support mental health and wellness. Disruptive technology in mental health is broad and diverse and includes:

– Emerging technology, such as computational psychiatry, virtual assistants, voice tracking of mood and implantable neurological nano devices, which are creating completely novel breakthrough solutions in mental health

– Mature digital mental health and eMental health solutions, such as guided self-help online treatments, telehealth and online information material, which are evolving traditional mental health models of care

"eMental health" is a phrase that grew alongside the usage of the term "eHealth" and is typically used to refer to more clinically based mental health solutions. Using the WHO definition of eHealth as a basis,[30] we can define eMental health as the use of information and communication technologies for the provision of mental health services.

The term "digital mental health", or "digital behavioural health", has grown in popularity recently and is used more broadly to encapsulate eMental health while including other technologies that help improve a consumer's mental health and overall wellness. Digital mental health refers to any services, accessed through a digital platform, that aim to prevent or treat mental health disorders or promote the well-being of people.

Digital mental health (digital behavioural health)

A broader term that encapsulates eMental health while including other technologies that help improve consumers' mental health and overall wellness.
e.g. Self-guided treatment applications, virtual assistants and chatbots, virtual reality to facilitate training and awareness, gamified treatment experiences

eMental health

Using the WHO definition of eHealth as a basis,[1] eMental health can be defined as the use of information and communication technologies for the provision of mental health services.
e.g. Online support groups, forums, blogs podcasts, online assessments

Disruptive technology in mental health

Disruptive technology in mental health refers to **innovative technology solutions that significantly change** the way mental health and wellness are identified, diagnosed, treated and supported. This could include emerging technology or mature technology that could change the traditional way of doing business. **e.g. Digital phenotyping, computational psychiatry, neurological interventions**

Our work in disruptive technology in mental health gathered and analysed 190 use cases of information and computer technology being applied to mental health and well-being services, tools and solutions where there was evidence of more than 10,000 users. This research has shown the diversity of currently available digital tools, ranging from: resilience and prevention services to targeted diagnosis and treatment tools; from simple educational websites to virtual-reality avatars delivering cognitive behavioural therapy; from Excel spreadsheets guiding consumers in a local community to effective services to globe-spanning electronic mental health records that support both integration of services and computational psychiatric research efforts; and from simple phone applications developed by small agile teams to implantable neurological nano devices that affect addiction and brainwave-measuring headsets, delivered at consumer-appropriate prices to optimize personal stress and relaxation for enhanced well-being and productivity.

Novel solutions are being built on the back of a range of types of disruptive technology that have already transformed many other industries. Mobile devices and computers are enabling consumers to access services anywhere, at any time, and provide a means of providing intervention at the point where it is most required, rather than waiting for a clinical
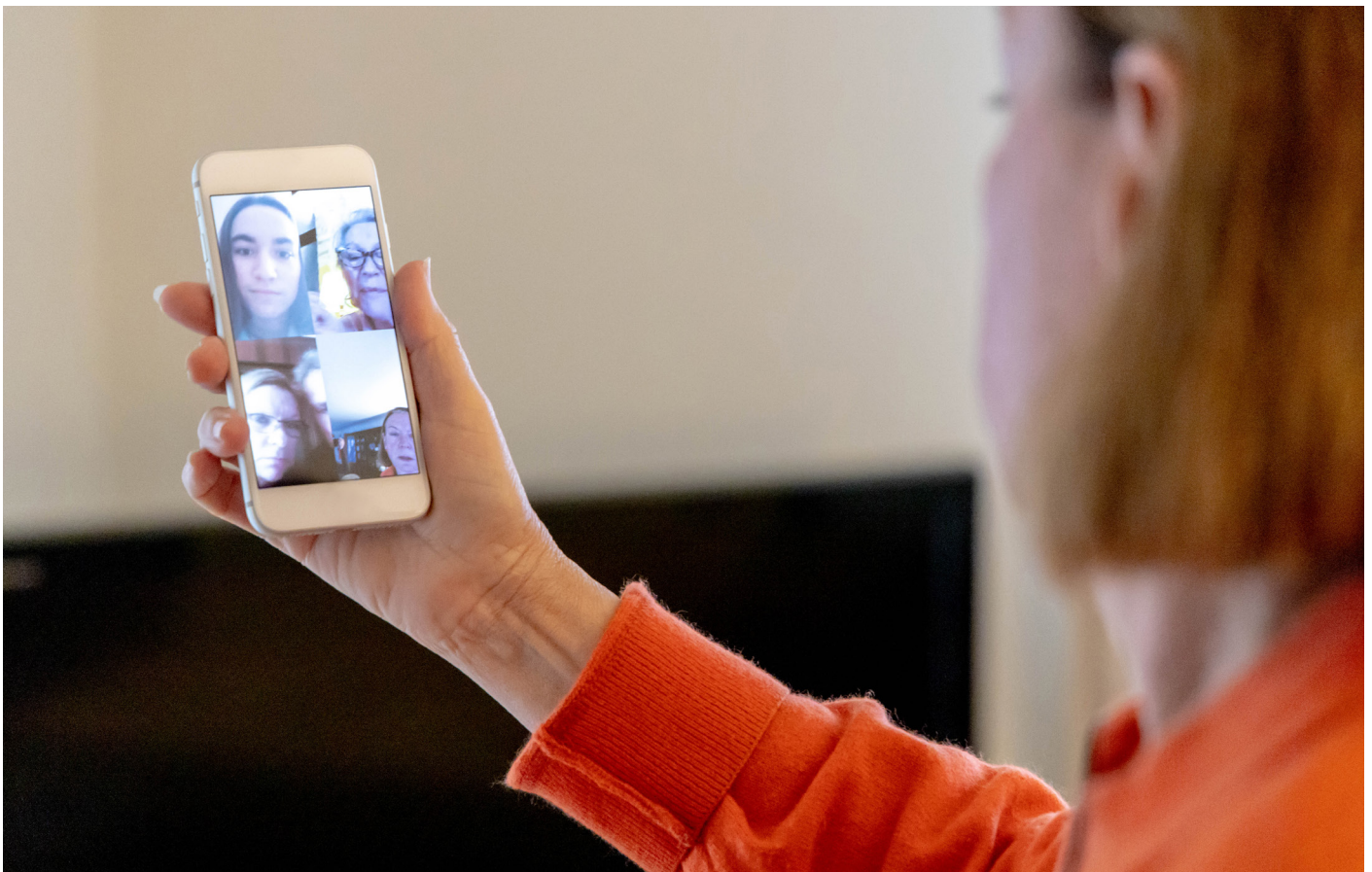
appointment. Digital phenotyping, where phones, watches and computers measure everything from behaviour, language, facial expressions and voice tonality to political and cultural views, is being combined with big data to unlock new understandings of psychosocial cause and effect, recategorize and discover new psychiatric conditions and bring greater accuracy to treatment options. Artificial intelligence/machine learning (AI/ML) and natural language processing have been deployed to make diagnostic decisions based on more holistic patient data, assess a variety of new possible scenarios for early intervention,

such as suicidal posts on social media, or come up with novel therapeutic chatbots and avatars. Augmented and virtual reality are enabling consumers to experience and practise real-life scenarios within the safe confines of a visit to their psychologist. Games have been developed that can more accurately diagnose specific conditions by observing behaviour and reactions, and gamified treatments are more engaging for children and adults alike. The diagram below highlights some of the categories of innovation identified in these 190 use cases, with the proportion in each category indicated around the circle.

FIGURE 4 | Novel technologies in mental health

Deloitte has collected **190 examples of novel technology use in mental health**, from apps with a reach greater than 10,000 downloads to the state-of-the-art tech coming from universities, pharma, healthcare, NGOs and mental health start-ups.

Self-guided treatment   28.4%

Delivering mental health and wellness management tools for stress, anxiety, depression and addiction in an accessible and scalable interface

Neurological interventions   4.7%

Innovative surgical, ingestible and non-invasive interventional tools observing and affecting brain and body function

Virtual assistants   4.2%

Facilitating natural interaction and learning without judgement, without stigma, at scale

Computational psychiatry   13.7%

Cognitive science, computational neuroscience and AI improving psychiatric diagnosis and treatment, and better explaining causes

Telehealth and peer-to-peer   16.3%

Finding and connecting with support networks, therapists and doctors and recommending supportive options

Physical, augmented and virtual reality   6.8%

Delivering more interactive awareness, management and training services with robots and simulated realities

Digital consumer experiences   10.0%

Managing the patient experience, emotion and mood journalling and recommending supportive options

Diagnostic support   6.3%

Combining new phenotypes, insights and tools to facilitate more accurate self-assessment and professional diagnosis

Novel digital treatments   5.3%

Gamified treatments for brain injury, neurocognitive decline, occupational therapy or mental health

Digital phenotyping   4.2%

New biomarkers from IoT devices and platforms enabling early, accurate detection and delivering new services

An individual's psychosocial well-being and mental health are influenced by a wide range of factors including the individual's strengths, beliefs and values, their connections, relationships and support networks, how their culture and value systems interact with their relationships and a great variety of environmental conditions. Unlike treatments for physical health conditions, many of the treatments and programmes used to improve mental health and
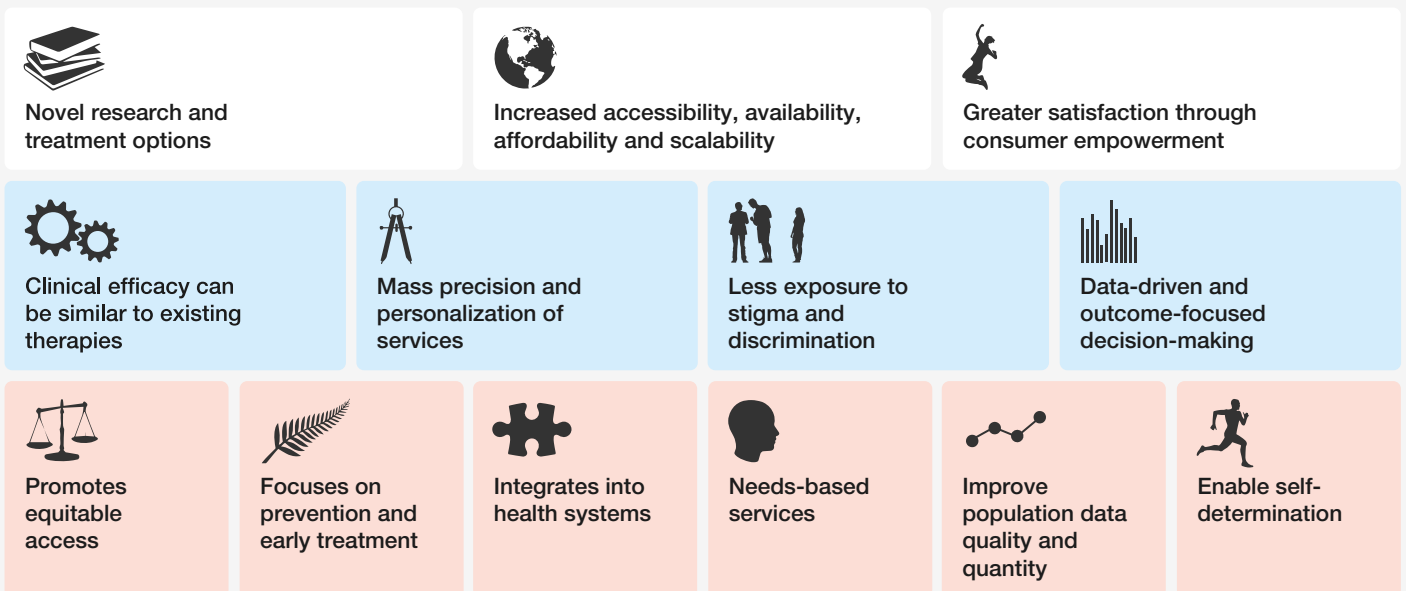
well-being are immaterial exchanges; they do not require physical interventions such as medication. Instead, they rely on forms of talking therapy and emotional or behavioural support and adjustment. Digital technology's ability to deliver virtual experiences independent of a brick-and-mortar site – such as telehealth, guided self-help treatments and virtual assistances available in the safety of the home – make it uniquely suited to delivering mental

health services. Services and specialist treatments, such as psychopharmacology, surgical interventions and in-person therapy, also have proven positive outcomes for consumers and are a vital part of a functioning mental health ecosystem, often used in conjunction with non-material therapies. And here, again, digital services may offer support to traditional, material services – to, for example, increase adherence through the gamification or tracking of medication – or provide improved consumer experiences to make the service easier to organize and undertake.

Digital mental health services offer a range of potential direct and indirect benefits that promise to augment, supplement and scale existing mental health models of care. Far from being a replacement for existing models, these tools are most effective when they are used either in combination with current care models or to access populations that cannot or do not currently access traditional medical services. Below we explore these potential direct and indirect benefits, before examining the ethical and clinical concerns.

1.3 | Potential benefits of digital mental health

FIGURE 5 | Summary of the direct benefits of digital mental health

Novel research and treatment options

Increased accessibility, availability, affordability and scalability

Greater satisfaction through consumer empowerment

Clinical efficacy can be similar to existing therapies

Mass precision and personalization of services

Less exposure to stigma and discrimination

Data-driven and outcome-focused decision-making

Promotes equitable access

Focuses on prevention and early treatment

Integrates into health systems

Needs-based services

Improve population data quality and quantity

Enable self-determination

Novel research and treatment options
Disruptive technology is improving existing treatment options and unlocking novel approaches to enhancing mental health and well-being. The application of big data and AI is unlocking research to better understand cause and effect and increase the efficacy and effectiveness of treatments through data analysis and the use of more holistic data. In addition, disruptive technologies are enabling researchers to explore novel modalities such as computer games, virtual reality, new medications and neurological brain-machine interfaces as ways to unlock new potential in mental health services.

Increased accessibility, availability, affordability and scalability
People may access digital mental health services at their own convenience, removing traditional barriers of time, availability and travel. Plus, the inherent scalability

of technologies can drive both a reduction in cost and an increase in the accessible market for services.[31]

Less exposure to stigma and discrimination
Using digital mental health services can reduce the likelihood of being exposed to the stigma and discrimination associated with mental disorders. This lowers one of the main entry barriers for mental health services and encourages people to start their mental health journey earlier. Research has shown a greater willingness to open up about personal issues to a computer, which is perceived as being free of judgement.[32]

Clinical efficacy similar to, or greater than, existing therapies
When tested for clinical efficacy, properly designed digital mental health services have demonstrated that they can be as effective as and in certain

specific situations more effective than traditional therapies.[33] Furthermore, in many situations integrative approaches have been shown to be more effective than therapy as usual or a solely digital therapy modality.[34,35]

Mass precision and personalization of services
By more readily assessing situational markers such as language, culture, history or phenotype, digital services have a greater capacity than traditional methods to deliver precision mental health services, offering the right treatment for the right person at the right time, improving the focus of service delivery and outcomes. Personalizing services in terms of the cultural, social, emotional

and psychological context can lead to greater engagement with those services.[36]

Greater satisfaction through consumer empowerment
With digital services, people can more readily drive their own mental health and well-being by choosing their own treatment. This increases the likelihood of engagement and long-term positive outcomes.[37]

Data-driven and outcome-focused decision-making
Digital mental health services provide the opportunity to track outcomes and use that data to improve product features, enhance treatment effectiveness and advance clinical research.

TABLE | **Other potential indirect benefits of digital mental health**

Promotes equitable access	Focuses on prevention and early treatment	Integrates into health systems
Promotes equitable access, particularly for rural, remote and low-income markets	Allows for increased focus on prevention, resilience and early treatment	Expanded access to integrated mental health services across the community and in primary-care settings
Needs-based services	**Improve population data quality and quantity**	**Enables self-determination**
Access to mental health services in concert with people's preferences and needs	Improved quality of individual and population data to drive research, innovation and strategic decision-making	Increased self-determination for people to drive their own well-being

1.4 | The trust, evidence and data challenges for digital mental health

Despite the potential of digital mental health services, interviews conducted with consumers, clinicians and healthcare system owners often revealed safety and efficacy concerns, as well as ethical questions related to the use of data and AI/ML in mental health services. Trust is fundamental to the uptake, therapeutic outcomes and sustainability of mental health services. An empathetic relationship with one's mental healthcare professional is a primary factor in positive therapeutic outcomes. That relationship is linked to the likelihood of consumers feeling understood, talking honestly about their problems and accepting advice. Research has shown that the empathetic relationship accounts for approximately 9% of variance in therapeutic outcomes.[38] Similarly, trust in a digital mental health service is required for prescription by professionals, uptake by consumers and investment from system owners. A systemic assessment of depression management and suicide

prevention apps discovered that only 7% of the apps reviewed provided comprehensive and holistic support with evidence-based strategies for suicide prevention.[39] Furthermore, the mental health apps that are readily available rarely use an evidence-based approach or are comprehensively tested for medical treatment of mental health issues.[40]

There are two main factors to consider in seeking to improve consumers', clinicians' and health-system owners' trust in digital mental health services. First, the technology must be safe – it must not inflict harm on its users – and it must be effective in leading improvement for consumers, as is the case when developing any tool or service for clinical practice. Second, the technology must be trustworthy: consumers must trust the tool and feel safe when interacting with it. Specifically, trust in a tool involves consumers being confident: that the tool uses

proven, effective means to achieve its advertised outcomes; that their information is protected and is not used in a harmful manner; and that the tool has been designed to uphold their best interests. Although the relationship between the consumer and a digital tool is not usually intended to match the empathetic relationship that exists between patient and psychiatrist, "trust in the tool" still plays an important role in its effectiveness. The effectiveness of a digital mental health service has its basis in medical ethics and in the ethical use of data and AI/ML, as well as in providing meaningful and clinically verified therapies.

Several risks of real harm act as barriers to the adoption of disruptive technology in mental

healthcare and to high-quality consumer outcomes. Developers wishing to produce technology that can be readily used and trusted in mental healthcare need an in-depth understanding of such risks. Below we have summarized, and categorized into six core areas, the main ethical and practical risks associated with disruptive technology in mental health services. These have been identified through interviews with leading innovators and healthcare professionals around the world, and through reviews of the current literature. There are also second- and third-order risks that have not been highlighted in this summary because the total universe of risks associated with disruptive technology in mental health is large and continuously evolving.

1.5 Ethical risks of disruptive technology in mental health

1. Untested efficacy or misleading claims of mental health benefits

Making unsubstantiated or misleading claims about the health benefits of mental health services can cause unintended harm to the consumer, as illustrated in case study 1. A digital product may use aspects of clinical treatment that have not been trialled or validated in the delivery mode of that product. For example, it has been clinically proven and verified in the real world that cognitive behavioural protocols, delivered by a qualified practitioner, are effective for anxiety disorders, somatoform disorders, bulimia, anger control problems and general stress.[41] However, the use of cognitive behavioural therapy (CBT) in other modalities such as self-guided online training, or delivered autonomously via a chatbot, must also be validated as there are many fundamental differences, such as not performing the same level of pre-treatment assessment, or not having the same ability to respond to adverse consumer scenarios during treatment. As such, it would be an unsubstantiated claim to conflate the effectiveness of CBT delivered by a trained professional to an individual with CBT on a computer screen or with a

bot, until both the efficacy and effectiveness have been shown through gold- standard scientific trials.

Additionally, the effectiveness of a service in diagnosing or treating mental health disorders may be misconstrued or miscommunicated to consumers. Products may be advertised in such a way that consumers are not fully informed as to the purpose of the services – for example, they may be misled into believing that a product can manage mental health issues for which it was not designed.

While it is not necessarily the intention of the vendor to provide unsubstantiated or misleading claims about the health benefits of their products, misalignment between advertised claims and services, or not mentioning associated risks or conditions that should be precluded, mean consumers are not able to provide informed consent as they are not properly informed. Consumers may be at risk of unknowingly engaging with treatment that unintentionally harms them or worsens their mental condition.

2. Privacy breaches of health data

The collection, storage and use of individuals' healthcare data poses several important risks. Data that is collected and stored in the context of disruptive technology in mental health is highly sensitive. Video and audio data pose a particularly high risk for exposing sensitive and personal information. Currently, frameworks to protect video and audio data of this nature are not widely used and would need to be further developed to ensure optimal efficacy.[42] Collecting and using video and audio data poses the risk of data leakage and hacking due to

the underdeveloped technical and legal privacy protections that exist to protect this type of data. Consumers are often afraid that their data may be sold, leaked or hacked. They are particularly concerned that their private information will be used by insurance companies to deny or affect their coverage or for punitive reasons by government agencies; or that it might be sold to third parties and used for non-therapeutic purposes. This is a prevalent concern in the context of sharing information on mental health, where consumers may be divulging personal and/

or incriminating information when engaging with services. This concern is a barrier to the effective adoption of technology in mental healthcare as it affects consumers' trust in the tool to protect their private information and work in the service of their best interests.

3. Harm to patient through malfunction, incorrect advice or misuse

A service's digital capability can create new risks for the consumer, as a service that malfunctions, is misused or misinforms consumers can negatively affect their experience and health. Most pertinently, bots, such as those being trialled for use as chatbots in mental health apps, may malfunction in unpredictable ways, which could cause more harm than benefit, particularly to consumers with a severe mental health condition that clouds their critical judgement.

For example, in the case of predictive or diagnostic modelling, the efficacy of the model is largely dependent on the dataset on which it is trained, and the outcome for which it is being optimized. Models that are improperly trained on "unbalanced" datasets, where one outcome is significantly less likely than the other, will more often predict the more common outcome.[43] Additionally, predictive models are not always able to fully contextualize the data they are given. For example, predictive models reading the word "depression" may not be able to differentiate between a subject referring to a transient depressive state or to a chronic condition.[44] Furthermore, the increased use of AI/ML across numerous industries has identified the key detriments of overlapping automated decisioning networks, where the combination of multiple models amplifies the inaccuracy, specificity and sensitivity of the results, or amplifies the impact when one model makes a mistake. These examples highlight the limitations of predictive models and the way in which they can issue incorrect information to the consumer.

Without proper oversight, a digital service may provide information and advice that has not been validated either by clinical evidence or experienced professionals, as seen in case study 1. Services may also provide advice that is not matched to the specific social, cultural or relationship situation of the consumer, resulting in poor or negative outcomes. Finally, without oversight, a model that is designed to use psychologically influential techniques to modify a consumer's behaviour and thought processes may be used to affect political, abusive, criminal and other non-beneficial outcomes.

4. Lack of accountability or incentives for prioritizing consumers' best interest

There is currently limited regulation of technology in mental healthcare, and vendors are not held accountable to an appointed or competent regulatory body or to an overarching set of standards for their services. This issue is amplified by the cross-border nature of digital services. Companies may train their machine-learning models for outcomes that are not in the best interest of the demographic of their consumers, such as increasing the addictiveness and profitability of the service or product rather than aiming to improve the therapeutic outcome. The well-being of consumers is at risk because there are insufficient safeguards in place that hold vendors accountable for possible unethical outcomes in relation to the development of technology in mental healthcare.

5. Widening the digital divide through inaccessible or inequitable design

Technology in mental health can increase the global inequality and inaccessibility of mental healthcare in two ways. First, the premise that technology in mental health increases global access to mental healthcare rests on the assumption that there is equal access to a digital device that can provide such care. However, we know that this is not in fact the case: for example, 184 million fewer women own mobile phones compared with men;[45] additionally, there is a significant technology literacy gap worldwide among both consumers and clinicians, and this limits consumers' ability to access digital mental healthcare and reduces clinicians' capability to deliver it.[46] Such discrepancies in the ability to access and use technology globally have the potential to widen the gap for mental healthcare access, particularly for minority groups and those of low socioeconomic status, even in high-income countries. The situation is, of course, exacerbated in low-income countries or for those in the developing world.

Second, AI-enabled mental health devices may contain bias arising from the way in which they are trained and tested, which has the potential to harm consumers in unintended ways.[47] One form of algorithmic bias occurs when models

are trained on a dataset that does not reflect the demographic of the people for which it will be used. Additionally, unsupervised models may develop bias by associating certain features, events or words with certain demographics or socioeconomic groups as it learns on a dataset, especially if the dataset is seeded with biased labelling.[48] The integration of AI into mental healthcare potentially risks harming or detracting from the care of disadvantaged groups of people.

6. Long-term social and behavioural consequences

The implementation of technology in mental healthcare poses the risk of affecting the long-term behavioural and social capabilities of consumers. Worsening mental health symptoms have been associated with prolonged screen time, particularly in children and adolescents, which may highlight an unintended risk of integrating technology into mental healthcare.[49]

Additionally, recent publications have posited that technology developed for "self-care" or self-treatment of mental health lacks the therapeutic relationship that a consumer would have with a healthcare professional, and may exacerbate the behavioural traits of a person experiencing issues with mental health, including social withdrawal and isolation.[50,51] Research has also identified that disclosing personal health information online, where it may be accessed by relatives, friends, employers or others, may have unintended social and behavioural consequences for the consumer.[52]

1.6 Case studies of ethical challenges in digital mental health

These case studies, illustrating some of the more difficult challenges of technology in mental health, were developed by the World Economic Forum's Global Future Council on Mental Health in 2020 to illustrate the many ethical dilemmas that are bought up by disruptive technology in mental health.

CASE STUDY 1 | **Proactive mental health intervention by a social media company**

A social media start-up offers users a popular feature to livestream moments in their life. Some users elect to use the livestream feature to share their self-harm experiences, food-restriction strategies for people with anorexic behaviour and even some suicide events. After several such cases received bad publicity, the company launched an internal AI program to screen for self-harm behaviours. Anything flagged by the AI is reviewed by a trained staff member. The program has not been reviewed by a health regulatory agency but, according to the start-up's internal research team, may have saved 450 lives in the first two months alone.What is the responsibility of the social media company to address these cases? Is the company liable for harm to the user if it does not report the prediction to the police or emergency services?

1. What kind of consent is needed? Should users be allowed to opt out or should users only opt in to this function? What level of privacy can be promised to users, given the possibility that the algorithm may alertfriends, family and emergency authorities?

2. If the company has an internal AI program that predicts self-harm, what is the optimal accuracy for these predictions? What are best practices for training staff members to evaluate the AI prediction? What are the dangers of an incorrect prediction?

3. What are best practices for training the AI – such as an AI that would address biases that could exacerbate existing disparities?

CASE STUDY 2 | **Proactive mental health intervention by a social media company**

An app claims to improve mood and reduce stress during COVID isolation and secondarily boost the immune system. The app-store profile has 45,000 5-star and 4-star reviews. The press has celebrated the app's developer as a brilliant engineering student with no medical background who is on a mission to increase access to mental health for all. There are no publications from any randomized trials of the app, but the app's profile cites 25 publications from others (some peer-reviewed and others not) linking stress and depression to increased risk for viral infections. Recent peer review studies have also demonstrated that some placebo versions of popular mindfulness apps (such as symptom trackers) resulted in similar gains to the app: a digital placebo effect. This digital placebo effect further complicates the evidence needed to regulate the efficacy of any (medical) apps. What are the risks and harms of an app that offers digital placebo relief compared to an app considered by scientists/regulators to be medically efficacious? Is the digital placebo effect more helpful than having no access to mental health providers?

1. How can mental health providers, researchers and advocates help the public distinguish effective as against ineffective apps or even high product performance versus low performance?

2. Some app developers have stated a personal view that all users should own their own data. However, the data often becomes a more valuable asset than the app itself. How can small independent app developers be given an incentive to help protect user data?
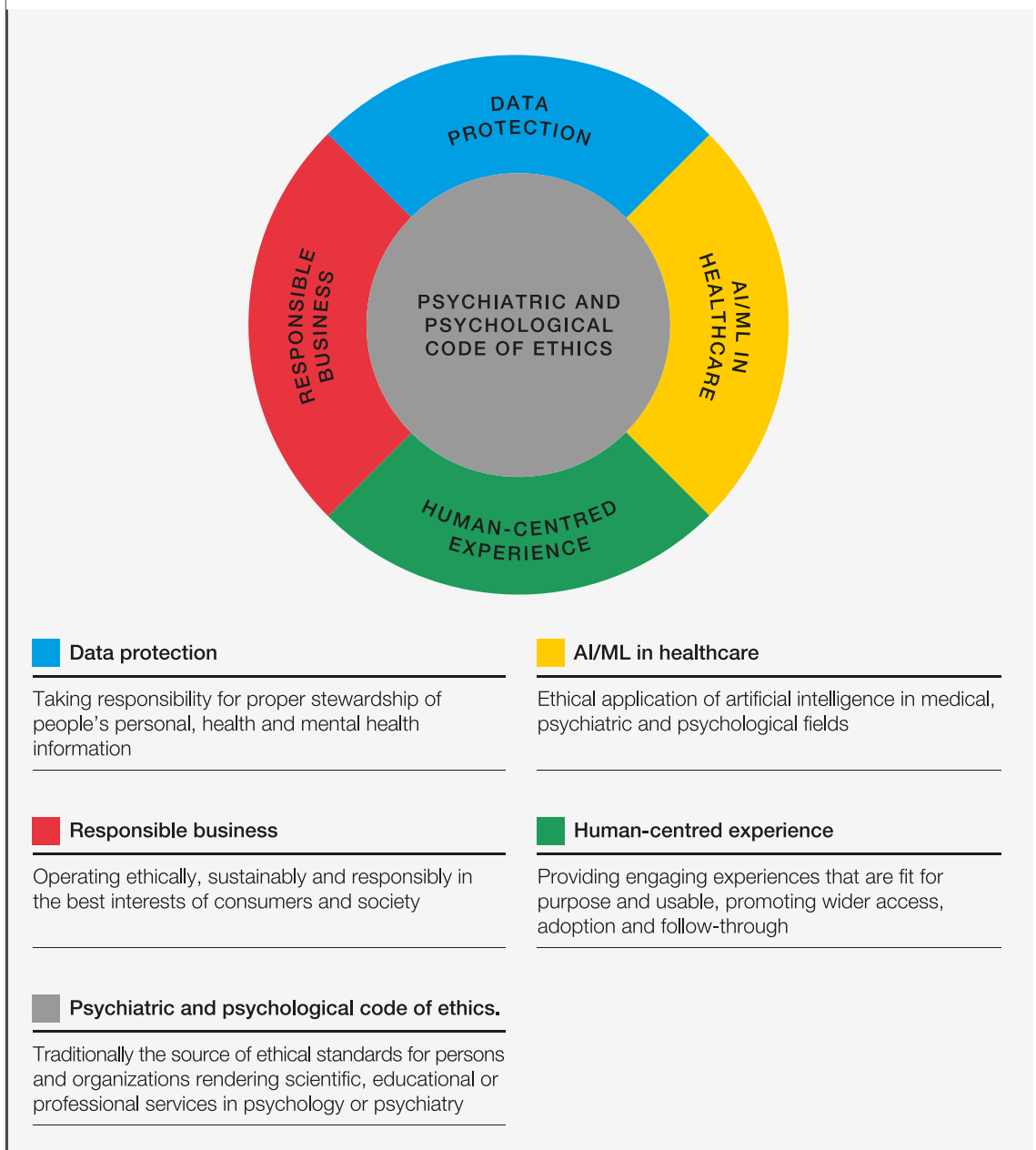
Additional ethical domains for consideration

Current best practices for ethical guardrails for persons and organizations rendering scientific, educational or professional services in psychology, psychiatry and mental health include the code of ethics governed and administered by organizations such as the American Psychological Association, standards of conduct, performance and ethics administered by the Health and Care Professions Council (HCPC) and Software as a Medical Device (SaMD) by organizations such as the Food and Drug Administration (FDA). These guardrails are not necessarily fit for purpose for tackling the particular ethical challenges of data and AI in psychology and mental health, since they were designed when these risks did not exist. Further, these existing guardrails do not always cover the vast range of services that are blurring the lines between consumer products, self-improvement products and mental health treatments. Considering the primary risks and central principles that are recognized as important for digital mental health services, as discussed in the next chapter, these guardrails could potentially be strengthened through detailed considerations of the four ethical domains shown in Figure 6.

FIGURE 6   Ethical domains in digital mental health covered and not covered by the current psychiatric and psychological code of ethics[53]

Data protection

Taking responsibility for proper stewardship of people's personal, health and mental health information

AI/ML in healthcare

Ethical application of artificial intelligence in medical, psychiatric and psychological fields

Responsible business

Operating ethically, sustainably and responsibly in the best interests of consumers and society

Human-centred experience

Providing engaging experiences that are fit for purpose and usable, promoting wider access, adoption and follow-through

Psychiatric and psychological code of ethics.

Traditionally the source of ethical standards for persons and organizations rendering scientific, educational or professional services in psychology or psychiatry

2 Goals, values and standards

Standards for digital mental health must be developed in line with the cultural, medical and social principles and values of the consumers and clinicians who are using the services and implemented to continuously and flexibly maintain high levels of safety and ethics on behalf of those stakeholders.

2.1 | Setting up robust policy foundations

There are many ways to create and enforce governance systems that ensure individual solutions in the digital mental health community are safe, strategic and trusted by all involved. This chapter focuses on two core methodologies for ensuring a safe, strategic and trusted digital mental health solution. First is a standards-based approach to regulation that can be applied widely to ensure that a service meets the ethical expectations of its stakeholders. Second is an ongoing governance model that a business or digital mental health service vendor could employ to continuously resolve new ethical dilemmas that arise from innovation and integrate themselves within their targeted community. Chapter 5 provides the framework for flexible implementation and ongoing administration of these standards for regulators and policy-makers.

Standards or policies may derive from a variety of sources, including: (1) identifying the ethical way an organization should behave and establishing the rights and values that it should uphold; (2) identifying particular risks involved for an organization's outcomes and establishing specific required mitigations; and (3) identifying systems that should be in place for an organization to ensure safe and good-quality outcomes.

Our approach to developing a robust policy foundation, as set out below, is to be purpose-driven, goal-orientated and flexible. To those ends we have used a methodology in which we start by identifying the goals of our framework and develop the shared values that define the purpose and principles for digital mental health. We have based these goals, values, attributes and principles on consultation with consumers, innovators and health-system owners around the world, as well as a research and literature review of current sources, as laid out below. We have included our pilot-project work in developing the Digital Mental Health and Addiction Services Evaluation Framework with the New Zealand Ministry of Health.

The framework for implementing these principles is outlined in the following chapters. Of special note is the intention to be flexible in the application of individual standards depending on the assessed risk of a service and whether a particular standard applies to that service's functionality. There are many applications, from simple informational services to software that supports secondary psychiatric diagnosis. The level of validation and application of standards should take this level of therapeutic intent into account. In some cases, such as an information website without any service workforce, the standards covering a service workforce simply do not apply according to the service's features. The framework for implementation in the following chapters includes a capacity for flexible assessment of services.

The final application of a policy in a jurisdiction will have to consider existing legislation for health devices, health software, data privacy and health information privacy, as well as responsible business practices, equality and justice. It may also have to consider jurisdictionally different views on social, cultural and medical practices. Provided below is a set of principles that offer a policy foundation and a methodology to further refine the standards in a particular jurisdiction.

Key terms used through out these principles

Term	Description
Disruptive technology in mental health	Innovative technology-based solutions that significantly change the way in which mental health and wellness issues are prevented, identified, diagnosed, treated and supported
eMental health	The use of information and communication technologies for the provision of mental health service
Digital mental health	Any services, accessed through a digital platform, that aim to prevent or treat mental health disorders or promote the well-being of people. Encapsulates eMental health while including other technologies that help improve consumers' mental health and overall wellness
Consumer	The consumer expected to be interacting with the digital mental health service
Digital service vendor (vendor)	The organization or business providing the digital mental health tool service to consumers in a jurisdiction; may also be called the provider
Tool	The actual technical solution that is used to render the service
Service	The application of a technical tool to provide services to a target market
Service workfoce	Portion of the workforce associated with – but not necessarily employed by – the provider or service and responsible for delivering services to the consumer. They may be counsellors, peers, clinicians or administrative workers

2.2 | Goals

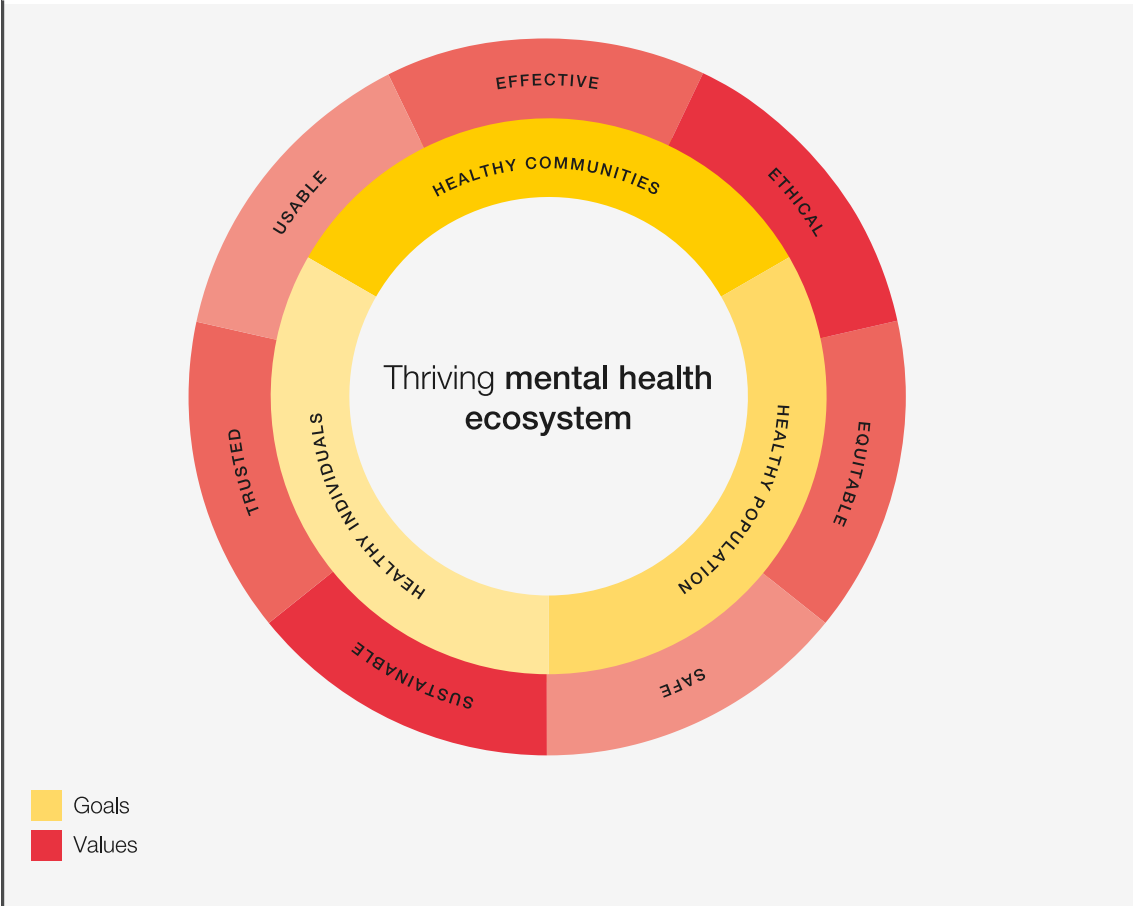**Ideals of the future digital mental health ecosystem that we envisage and commit to achieving**

HEALTHY COMMUNITIES

HEALTHY POPULATION

HEALTHY INDIVIDUALS

Thriving **mental health ecosystem**

Goals

Primary goals

Goal	Description
Healthy individuals	Individuals empowered to reach a state of well-being in which they may realize their own abilities, cope with the normal stresses of life, work productively and be able to contribute to their community
Healthy communities	Communities that support each other to enable individuals to reach their social and mental well-being goals and fully participate in the community
Healthy populations	Populations in which everyone has access to the means to detect and prevent mental illness early; receive effective and appropriate treatment and support; and enable recovery
Thriving digital mental health ecosystem	An integrated mental health and well-being ecosystem that supports and enables people and communities; integrates with health, economic, social and cultural systems; promotes pragmatic progress; and flourishes for the long term

2.3 | Values and attributes

Values are the fundamental beliefs that guide and motivate us towards our goals; our attributes describe our desired ecosystem. Together they help us in defining our purpose and principles

Core values

Legend:
- Goals
- Values

Center: Thriving **mental health ecosystem**

Outer ring (values): EFFECTIVE, ETHICAL, EQUITABLE, SAFE, SUSTAINABLE, TRUSTED, USABLE

Inner ring (goals): HEALTHY COMMUNITIES, HEALTHY POPULATION, HEALTHY INDIVIDUALS

| Value	What is it?
Effective	**Does the solution really work for consumers?** – The tools and services need to achieve their intended goals, as opposed to making false promises. We want to promote the testing and trial of tools, where appropriate, as well as continuously measuring outcomes to ensure that tools effectively contribute to people's mental health and well-being
Ethical	**Is the purpose aligned with consumers' interests?** – We need tools and services to be ethical and transparent in terms of their purpose and their agenda, including their treatment of consumers, models of care, use of data and technology, and integration into our health system

Equitable	**Does it promote fairness, justice and equality?** – Our tools need to contribute to reducing the gap in access and services in vulnerable communities by championing the provision of services and tools that deliver equity in mental health and well-being outcomes for those communities. This requires us to be explicit in our targeting and focus on vulnerable and underserved communities and population cohorts. Solutions must be accessible, appropriate and free from discrimination, for those who will need them most
Safe	**Is the solution safe to use for its intended purpose?** – We need solutions that are safe and as a minimum "do no harm". Many digital solutions show great promise, but we need to be certain that they do not put people in harm's way, potentially putting consumers at risk or even exposing others to harmful or adverse effects
Sustainable	**Will this solution be around to meet consumers' needs over time?**– Because the digital mental health landscape is evolving so rapidly, there is a high risk of "boom-to-bust" scenarios. At the same time, digital innovation does not always come with a ready-made business case and financial security. Therefore, there is a need for sustainable practices for ongoing development and refinement and for assured long-term business models, as well as transition plans to make sure a solution does not suddenly fail, leaving consumers without support when they might have become dependent on a particular tool
Trusted	**Can consumers trust the solution with their data and trust it will meet their needs?** – In the fast-moving world of online sites and apps, trust – and, in particular, digital trust – has become a complex issue: consumers sometimes have a false sense of trust in their interactions through social media and other such tools, while at other times there can be a very rapid and angry backlash, with consumers feeling that their trust has been violated. We therefore need to ensure transparency and clarity on how a given tool manages trust
Usable	**Does the tool make it easy for consumers to achieve their goals?** –One of the most common reasons for a digital service's lack of effectiveness is not the therapeutic basis of the content but the interface failing to meet basic expectations for usability or aesthetics. How many times do consumers delete an app because of an error in the first minute of use? While usability is a foundation of efficacy of digital services, aesthetics and ease of use adapt as people's expectations for usability and aesthetics change over time and a redesign may have little effect on the therapeutic content, other than seamlessly facilitating the use of that content

Core attributes
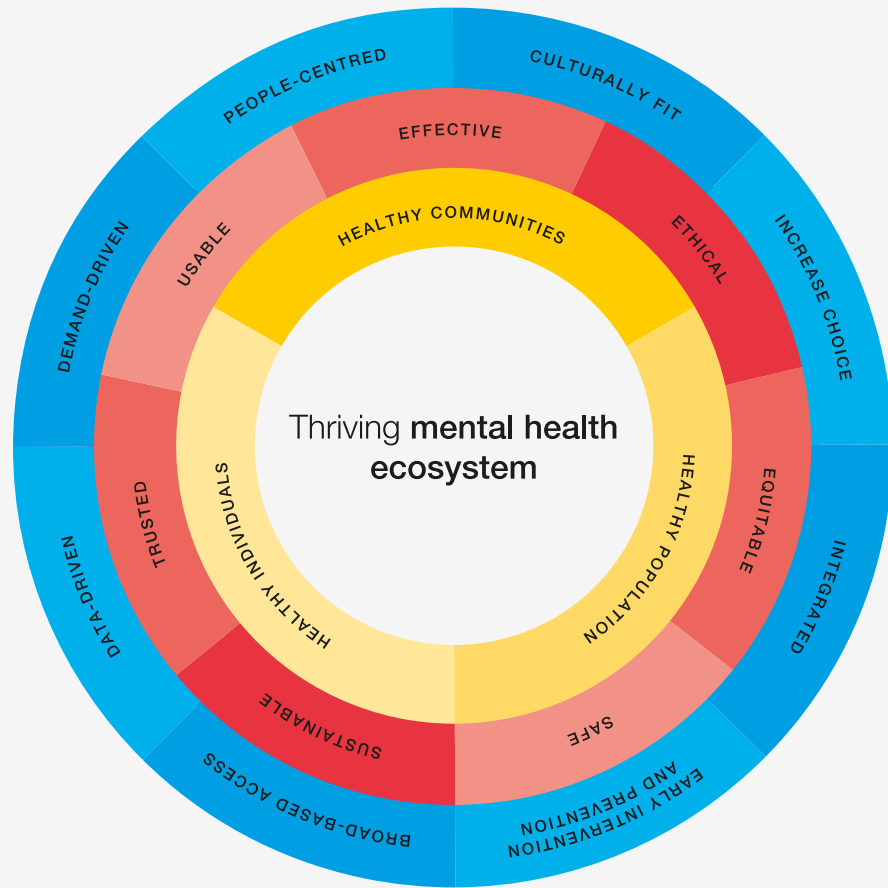
Legend:
- Goals
- Values
- Attributes

Attribute	Why is it important?
People-centred	We need tools that are easy to use, convenient, engaging and empowering for consumers. Our digital mental health tools must be first and foremost about enabling people to live lives that are healthy and emotionally, vocationally and socially satisfying – not about provider convenience. This also means that our tools must be culturally aware and sensitive, ensuring that we actively promote solutions that respond to the needs of diverse and vulnerable communities
Culturally fit	We need tools that are responsive and tailored to the beliefs, attitudes and circumstances of groups of people who share a common cultural heritage. We need human-centred principles in the design and deployment of these tools. This means that cultural groups need to have a voice in the design and implementation of tools, as well as in determining which needs are being targeted

Increase choice	We need to offer an array of options and choices that suit individual needs and preferences for support in times of mental distress or illness. There must be a continuum of options as well as a broad range of services available from which consumers can pick
Broad-based access	Our tools need to be universally available, not creating digital divides. They need to support access to services for vulnerable communities who often have barriers to access due to their location, economic status, technology literacy, language, culture or other factors. This is particularly important considering COVID-19 and the need for a cohesive psychosocial response
Early intervention and prevention	We have a clear mandate to address mental illness and distress. By intervening earlier and more frequently, we can reduce the incidence and impact of more severe mental health distress and illness, while reducing the cost of intervention
Integrated	We need to link people to communities and peers for support, as well as to other providers in the broader ecosystem of health and social services. We do not wish to create further islands of information or "digital fortresses" where information cannot be shared, or collaboration is hampered because of digital differences
Demand-driven	We need to prioritize tools that focus on addressing the gaps and needs in our current networks of support and that offer relevant support for our communities
Data-driven	Our digital tools need to be able to use data to enhance the degree of tailored support consumers receive. This data needs to be able to inform our decision-making to drive quality outcomes, characterized by safe, ethical and impactful data collection and analysis

2.4 | Standards for disruptive technology in mental health

Here we have compiled a full list of standards to audit and govern a digital mental health service and provide assurance of its safety, quality and efficacy. They are designed to be outcome-focused, i.e. rather than prescribing how to achieve a standard or requirement, they describe what the outcome of achieving that standard would be, allowing for greater flexibility in implementation options and future-proofing of requirements.

Some of these requirements are considered mandatory to achieving safety for consumers, such as utilizing clinically validated therapeutic models of care or monitoring clinical safety and quality, and some may be considered as merely important for consumer choice, such as ensuring high usability. In addition, there are some requirements that may not be applicable to every service – for example, a purely informational website may not gather personally identifiable healthcare data and as such not need to implement privacy controls related to healthcare records. Thus, during application of these requirements, flexibility needs to be considered for the following drivers of exclusion of specific requirements: therapeutic claims, functional scope, data use and workforce. These are further explained in the following chapter.

The intended outcome of auditing a service against these requirements is a categorized grade where the service is scored according to its ability to meet or exceed various requirements in those categories. This is discussed in the following chapter.

FIGURE 7 | Goals, values and standards for digital mental health

Legend:
- Goals
- Values
- Attributes
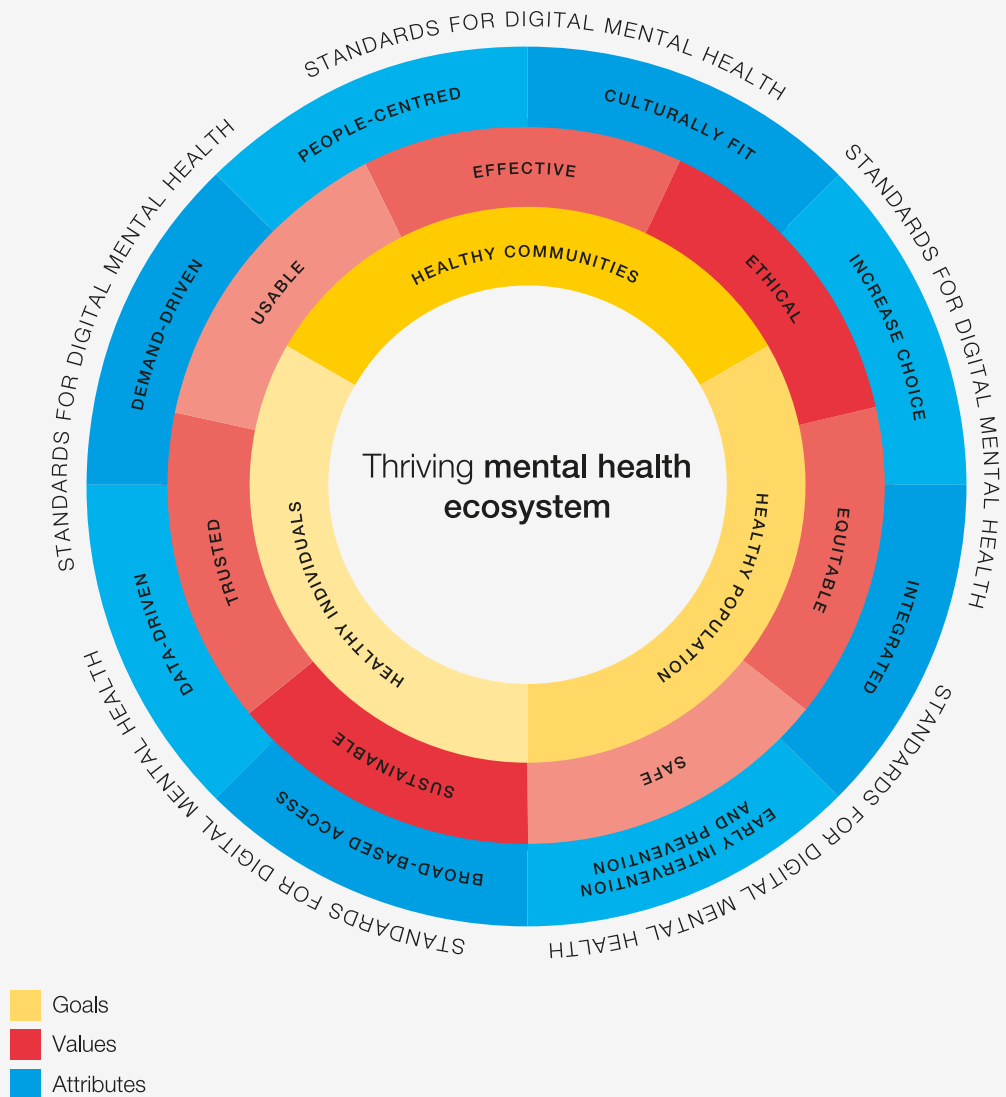
Standards for digital mental health

1. Lead practice with evidence
2. Do no harm
3. Establish a risk management culture
4. Ensure clinical safety and quality
5. Commit to a robust service workforce
6. Promote data privacy and transparency
7. Maintain data and information security
8. Ensure interoperability of healthcare records
9. Orient around person-centred design
10. Promote social and cultural respect
11. Commit to equity and justice
12. Advertise responsibly and accurately
13. Enact a responsible business model
14. Operate with accountability and accept feedback and complaints
15. Provide a sustainable platform and continuity of care
16. Govern AI/ML responsibly

1. Lead practice with evidence

Use clinically validated therapeutic models of care	Ensure that the specific model of care provided has been validated through clinical trials for both efficacy and effectiveness, that trials have been academically verified and accepted by professionals as best practice and are not outdated or disproven
Offer content and therapeutic approaches based on science	Ensure that the content provided in each service is based on the best available evidence and best practice, current know-how and is not outdated or disproven, relevant to the therapeutic goals, academically verified and accepted professional practice
Support ethical research and publication	Where the services capture or store information for research purposes, ensure that consumers are made fully aware of the research through the consenting process and that the research follows ethical principles

2. Do no harm

Promote ethical decision-making	Strive to benefit those with whom the work is undertaken and take care to do no harm; be alert to and guard against personal, financial, social, organizational or political factors that might lead to misuse of their influence; allow individuals to report ethical concerns and attempt to resolve these concerns in a responsible manner. When conflicts occur among obligations or concerns, attempt to resolve these conflicts in a responsible fashion that avoids or minimizes harm
Incorporate processes to do no harm	Have processes in place to analyse and document the potential risks surrounding the proposed service; identify mitigating factors and controls to manage the identified risks; implement fail-safes and control points to monitor these risks; regularly report on these and review them; and link reports to corrective action

3. Establish a risk-management culture

Govern responsibly with processes for risk management	Have governance in place for risk management to establish accountability for risk management from the top in the leadership team; define and document the assurance landscape and scope; set out, review and maintain the currency and effectiveness of policies, procedures and protocols against this landscape; monitor and take action to improve adherence to policies, procedures and protocols; review compliance with legislation, regulations and jurisdictional requirements; establish and maintain a systematic governance process for clinical safety and quality; and monitor these processes to drive improvements in clinical safety, quality, performance and effectiveness for consumers
Govern and manage risks proactively	Identify and document service risks; use clinical, technical and other data collections to support risk assessments; have a governing body accountable for risk management; maintain an active risk and issues register; act to reduce risks; regularly review and act to improve the effectiveness of the risk-management system; plan for and manage internal and external emergencies and disasters, including cybersecurity risks and threats; and have a business continuity plan (BCP) in place

Create a risk-management culture	Where services are delivered in conjunction with service workers, promote an appropriate risk-management culture by: creating awareness of risks among the service workers; training service workers on how to identify and report risks; making service workers and service consumers aware of risk governance and reporting channels; reporting on risks to the workforce and service consumers; proactively notifying service consumers and workers of relevant risks; allowing service consumers to escalate from self-management to getting additional help

4. Ensure clinical safety and quality

Monitor clinical safety and quality	Where the functional scope captures relevant consumer data and/or involves service workers, have processes in place to communicate when critical information about a service consumer's care emerges or changes, through the use of the tool, to ensure the safety of the consumer; use defined parameters to recognize acute deterioration in mental state that requires care to be escalated; have protocols that specify the criteria to call for emergency assistance; enable service consumers to communicate critical information and information on risks to either their vendor or an alternative provider
Operate with a defined model of care	Where services are delivered in conjunction with service workers, ensure the implementation follows a defined model of care that monitors the delivery of services to ensure they are consistent with the model of care; have a process for assigning responsibilities to a member of the workforce (where applicable) for the overall accountability of the care of each service consumer; develop the goals of care and actions for treatment in partnership with the service consumer; clearly communicate the care plan to the service consumer; enable the involvement of support people to the extent that the consumer chooses; have a process for referral to follow-up services that is consistent with the model of care
Manage the clinical safety and quality performance	Where services are delivered in conjunction with service workers, ensure that performance is managed: assign improvements to members of the workforce with clear responsibility for safety and quality; implement identified improvement initiatives for quality and safety; monitor implementation; formally document and verify corrective actions for their impact; undertake proactive testing at regular intervals
Manage adverse events in terms of clinical safety and quality	Where services are delivered in conjunction with service workers, use an open disclosure programme that is consistent with the jurisdiction in which it operates; monitor and act to improve the effectiveness of open disclosure processes; identify sentinel events or adverse events, which are issued to a cross-functional governing body for review and action; include feedback to the workforce with appropriate performance management; provide disclosure to affected consumers, notifying the appropriate authorities
Continuously improve clinical safety and quality	Have in place a quality and safety improvement system that identifies customer satisfaction and usage patterns as well as safety metrics; define and capture quality and safety measures; monitor and report performance and outcomes against these measures; feed these into a governance group and continuous improvement process; maintain a quality-improvement register to log initiatives to improve quality and safety

Prevent harm and suicide	Have systems to identify service consumers who are at risk of harm, including self-harm and suicide; effectively respond to service consumers who are distressed, have expressed thoughts of self-harm or suicide or have self-harmed; identify service consumers whose healthcare needs are beyond the scope of the service; have protocols that specify the criteria to call for emergency assistance; and provide information to service consumers with healthcare needs beyond the scope of the service on where to access services appropriate to their clinical need

5. Commit to a robust service workforce

Operate with an appropriately qualified service workforce	Where services are delivered in conjunction with service workers, have processes to ensure clinicians involved in the design and delivery of services have the necessary skills, experience and qualifications for these roles and work within a defined scope of clinical practice; and that technicians involved in the design and delivery of services have the necessary skills, experience and qualifications for this role
Ensure clarity in roles and responsibilities	Where services are delivered in conjunction with service workers, have processes to assign safety and quality roles and responsibilities for service to the workforce; and support the workforce to understand and perform their roles and responsibilities for safety and quality
Ensure the service workforce is skilled	Where services are delivered in conjunction with service workers, have training systems that assess the competency and training needs of the workforce; implement a training programme to meet its requirements arising from these standards; provide orientation training that describes roles and responsibilities for safety and quality for new members of the workforce; provide access to training to meet its safety and quality training; and monitor the workforce's participation in training
Manage service workforce performance	Where services are delivered in conjunction with service workers, have valid, reliable and fit-for-purpose performance review processes that require members of the workforce to regularly take part in a review of their performance; identify needs for training and development in safety and quality; and incorporate information on training requirements into training systems
Ensure cultural and safety training for the service workforce	Where services are delivered in conjunction with service workers and target consumers with different cultural and linguistic backgrounds, have strategies to improve the cultural safety and cultural competency of the workforce to meet the needs of culturally and linguistically diverse and other non-mainstream population groups; partner with a service workforce that is competent in servicing the cultural and linguistic needs of the diverse groups

6. Promote data privacy and transparency

Use informed consent for data privacy	Have mechanisms and easy-to-use functionality for consumers to ensure they understand exactly what data will be captured, what the purpose of capturing this data is and who this data will be shared with – all presented in plain and easy-to-understand language
Have system controls for data privacy	Where a service collects individual consumer data, have systems in place for the collection, use, disclosure, storage, transmission, retention and destruction of data that is provided: information on the types of data collected and how the information is used or shared; information on any interoperable healthcare services; information on who has access to the data, including through data-sharing agreements, provision or sale to third parties; timely information if requests to access data by external parties are granted by the provider; protection of data that was provided anonymously or using a pseudonym; prevention against the reidentification of anonymous or deidentified data and notification if the service ceases to operate or changes ownership
Allow data management by the consumer	Where data is captured and/or shared, have mechanisms for consumers to: control the sharing of information where applicable such that consumers can withhold consent for the collection, storage or distribution of their personal data; opt in or out of sharing their data with third parties, including for research purposes; access in-app controls and checkpoints before data is shared with another party to deliver the service (e.g. sharing personal contact details for a referral); access in-app or overall consent controls that permit consumers to opt in or out of sharing their data for research or other purposes not directly linked to the service; have disclosure where their data is shared to get revenue (e.g. advertising or paid contributions to analytical datasets; control the level of personal data shared with other consumers/peers or the public throughout the use of the tool
Gain consent to capture data	Inform consumers about any data that is captured through use of their service and ask for their consent to the data capture; ensure they are adequately informed as to what data is captured and how this data is used, and that they have the opportunity to control any data capture or information sharing resulting from the use of the service
Gain informed consent for data usage	Provide information on the model of care, purpose, evidence and risks, which complies with legislation and best practice; ensure such information is easy to understand and meets the service consumer's needs; explain the purpose of the solution; explain the rationale for the proposed model of care or service; adequately inform service consumers about potential risks

7. Maintain data and information security

Have secure data and information through systems for data encryption	Where a service collects personal information, maximize the safety and quality of care through the design of services and the use of internal access controls; have a system for data encryption and use a risk-based approach to assign responsibility and accountability for data encryption; develop a key management plan to detail key encryption protocols and allocate appropriate encryption permissions; support systematic audit of encryption keys and protocols
Have secure data and information through systems for data security	Where a service collects personal information, have information security management systems and use a risk-based approach to assign responsibility and accountability for data security; complete and maintain an information and data inventory; protect data in transit and at rest; protect against interruption, damage or disconnection of the service; assess the size and extent of threats to information assets; consider and mitigate vulnerabilities and threats; conduct regular updates, reviews and audits of information security; and detect, respond and report to the governing body, workforce and service consumers on information security incidents and technical faults
Have secure data and information through systems for consumer authentication	Where a service collects personal information or clinical data, have processes that authenticate service consumers and match them to their care; protect the anonymity of the service consumers where this is part of the model of care; use appropriate identifiers for service consumers according to digital services best-practice guidelines

8. Ensure healthcare records interoperability

Make responsible use of healthcare records	Where a service captures or stores healthcare information, have healthcare record systems that comply with healthcare data security legislation and regulations; implement technical and operational procedures that also comply with the above; comply with cloud-based security and privacy standards when using cloud-technologies; support the systematic audit of clinical information; ensure information is secure from unauthorized access/changes/edits/deletions; allow consumers to update, verify and change information held about them
Provide interoperable healthcare records	Where a service is delivered in conjunction with service workers and involves named referrals, have healthcare record systems that comply with healthcare data privacy legislation and regulations; integrate multiple information systems where they are used for referrals or integration with other systems; use national patient and provider identifiers and standard national terminologies and ontologies
Ensure accuracy of healthcare records	Where a service captures or stores healthcare information and personal information, have healthcare record systems that ensure the creation and maintenance of accurate healthcare records and personal information using clear editing and access permissions; tag all healthcare information and personal information against a unique identifier; routinely audit records for accuracy
Enable bidirectional management of care	Where a service is delivered in conjunction with service workers, have processes to effectively communicate when all or part of a service consumer's care is transferred; determine minimum information content to be communicated when care is transferred; set out the process for a transfer of care in line with the model of care; assess risks relevant to the service's context and the particular needs of the service consumer when a transfer of care occurs; support service consumers to be involved in the transfer of their care; refresh the relevant information-sharing agreement and informed consent with the consumer

Manage referrals responsibly	Where a service is delivered in conjunction with service workers and involves named referrals, use systems that optimize the safety and quality of care to service consumers; use national patient and provider identifiers; use standard national terminologies and ontologies; manage access to the system by the workforce to comply with legislative requirements; maintain the accuracy and completeness of the clinical information that the vendor uploads in the system

9. Orient around person-centred design

Promote accessibility and reach	Partner with service consumers to minimize the barriers to accessing services associated with hardware, software, data requirements and platform of the services; minimize the barriers to accessing the services relating to language, location, age, ability and technology literacy; ensure services are compatible with commonly used assistive technologies; meet relevant standards for web or app design; regularly review access to services and take action to improve access by the target consumers; use communication mechanisms tailored to the diversity of service consumers
Ensure high usability	Have processes to assess and optimize the usability of each service, including processes that meet service consumers' functional requirements according to the service's claims, and means of measuring and maximizing service consumer satisfaction for the service
Promote positive consumer experience	Where the service has a therapeutic purpose, consider processes to optimize consumer engagement, including providing options for consumers to actively engage with the service and measure adherence of consumers by tracking engagement and usage
Tailor communications and services to diverse groups	Support the communication of information to service consumers, and their support people, including products and services that meet the needs of the targeted consumers and are easy to understand and use; tailor content to the diversity of service consumers; communicate clearly and appropriately with the targeted consumer group, using language that is culturally appropriate
Co-design service and governance with service consumers	Partner with service consumers and carers from target consumer groups; have processes to involve a mix of people that are reflective of the diversity of service consumers and target consumers; provide orientation, support and education to service consumers, carers and service consumers who partner with it; incorporate the consumer's views and experiences into the governance, planning, design, measurement and evaluation of services; incorporate the consumer's views and experiences into the training and education for the service's workforce
Empower service consumers to make decisions relevant to their care	Where the vendor has processes to partner with service consumers to make decisions about their current and future care, identify a substitute decision-maker if a service consumer does not have the capacity to make decisions for themselves
Use a person-centred model of care	Where a service shows a therapeutic intent or claim, document the purpose and intent of the model of care for each service and the context in which it will operate; define the intended consumer demographic and their needs; match the model of care or therapeutic intentions to the target consumer's needs; monitor and evaluate the performance and effectiveness of the model of care; assign accountability for maintaining and improving the effectiveness of the model of care

Co-design services with vulnerable cohorts	Demonstrably incorporate a co-design approach using vulnerable cohorts whose design input influences the final service provision and supports the workforce to understand and perform their roles and responsibilities for safety and quality with identified vulnerable cohorts

10. Promote social and cultural respect

Ensure services are free from discrimination	Have processes to assess and minimize the stigma and discrimination for persons using mental health services, including removing barriers to service; provide similar levels of quality of care to all service consumers; include aids for people with a disability; treat service consumers with privacy and respect; treat service consumers fairly, equally and with dignity; implement necessary feedback and assessment systems, including risk assessment and editorial review
Show social and cultural respect	Identify the diversity inherent in targeted service consumers; ensure that a broad-based analysis is undertaken of consumers' cultural needs and norms; include people with diverse backgrounds in the design of the tool; identify groups of service consumers who are at higher risk of harm or exclusion; identify groups of service consumers who may have disadvantages or disabilities that affect tool use; incorporate information on the diversity of service consumers and higher-risk groups into the planning and delivery of the service

11. Commit to equity and justice

Promote equitable health outcomes	Actively enhance the tool's features for better cultural, economic, geographic and health outcomes by making good use of data evidence to tailor services to improve outcomes, avoiding the introduction of institutional bias from age, gender, ethnicity and social-economic status perspectives; have processes in place to apply and monitor progress towards closing the gap in mental health services for traditionally vulnerable communities, providing greater equity of service to vulnerable cohorts, and connecting service consumers' anonymized feedback to help resolve systemic inequities
Prevent racism and marginalization	Put systems and processes in place to surface and address underlying racism or marginalization for service consumers, address institutional racism within the mental health sector and eliminate internal service bias
Prevent coercion or exploitation in services	Have systems to minimize the risk of exploitation, including coercion, harassment, abuse, bullying and hate crimes against service consumers; minimize the risk of children and young people being harmed by the service; preserve the dignity of service consumers
Govern responsibly on ethical principles or values	Govern responsibly and comply with a charter of ethical principles or values
Govern responsibly on equity, ethics and diversity	Have a governance structure to implement the organization's equity, ethics and diversity strategies; have a diverse representation in the governance structure; have processes in place to monitor and address issues

12. Advertise responsibly and accurately

Make responsible therapeutic claims	Where the service makes a therapeutic claim, ensure that the claim is based on independent evidence, founded on sound therapeutic principles, can be objectively verified, does not set unrealistic expectations for consumers, is statistically significant and relevant and considers the placebo effect or counterfactual baseline
Advertise responsibly by avoiding false and deceptive statements	Ensure advertising related to services, products or functionality complies with local advertising standards, consumer laws and regulatory requirements and is appropriate for the target consumers; ensure therapeutic or clinical claims comply with local advertising standards, consumer laws and regulatory requirements
Ensure responsible and ethical in-product advertising	Ensure that in-product advertising complies with local advertising standards, consumer laws and regulatory requirements and is appropriate for the target consumers

13. Enact a responsible business model

Put in place a responsible business model	Operate a responsible business model, where the business results and incentives are aligned with the overall success of consumers, employee incentives are aligned with company/business incentives, consumers can expect long-term access to and support from the service, and there is a transition plan for consumers if the service is no longer viable
Be transparent about costs to service consumers	Provide service consumers with clear and transparent information on the direct costs to access the service, estimated data usage requirements for using the service, and other substantial indirect costs that may be incurred

14. Operate with accountability and accept feedback and complaints

Demonstrate accountability for the too	Have a feedback and complaints management process or system in place, and seek regular feedback from service consumers about their experiences of the service and outcomes of care, including how they have provided culturally appropriate care; encourage and supports service consumers to report complaints and feedback; involve service consumers in the review of complaints and feedback; have data-capture processes to address feedback on system-level inequities; resolve complaints and feedback in a timely way; provide timely feedback to the governing body, the workforce and service consumers on the analysis of complaints and actions taken
Uphold the rights of service consumers	Have a charter of rights in place that is consistent with local legislation and regulation, is easily accessible to service consumers and their support people, articulates any warranties, offers a mechanism to raise complaints or provide feedback to the vendor and outlines the claims process for breach of any warranties

15. Provide a sustainable platform and continuity of care

Keep systems and services up to date	Actively manage the continuity of services by maintaining platform and operating system updates and patches on an ongoing basis; regularly update the service as required for security or feature improvements; effectively communicate service changes or interruptions to service consumers; have a change control board and governance in place; conduct adequate testing and regression testing as part of the release management cycle; have adequate backup and recovery mechanisms in place; have a business continuity plan (BCP)
Ensure platform stability and availability	Where a service is delivered in conjunction with service workers, provide a stable platform that has the requisite level of availability and operating hours to meet business and clinical needs; monitor platform availability; specify hours of service/operations to consumers; ensure that issues are resolved in a timely manner; provide adequate helpdesk and support services; have a business continuity plan (BCP) in place

16. Govern AI/ML responsibly

Put in place quality systems and good AI/ML practices	Where a service uses artificial intelligence or machine learning (AI/ML), ensure that quality systems and good AI/ML practices are in place, including appropriate evaluation of the validity of the AI/ML's clinical, clinical association and analytical functions; establish appropriate data management, feature extraction, model training and model evaluation practices and practices to assess the relevance of available data for the clinical problem; ensure data acquisition that is consistent, relevant and generalizable; set out appropriate separation of training, tuning and test datasets; ensure appropriate assessment and clarity of the algorithm and its output
Establish pre-release review and modification governance	Where a service uses AI/ML, conduct an appropriate pre-release risk assessment of the models, including assessing the performance parameters, inputs, intended use and potential real-world impacts; develop an algorithm change protocol that contains allowable limits for changes in performance, inputs and intended use; ensure protocols for quality and good AI/ML practices, including data management, model retraining, performance evaluation, update procedures and triggers for updating the risk assessment
Monitor and publish real-world performance	Where a service uses AI/ML, monitor the real-world performance, safety and use of the AI/ML models and work with transparency by publishing model use and performance information appropriately to relevant internal and external stakeholders such as clinicians, regulatory bodies, partners and consumers

(3) Ethical AI in mental health

Our approach to ethical and moral dilemmas creates a social licence to operate with our communities and acts as a bridge to ethical innovation in mental health.
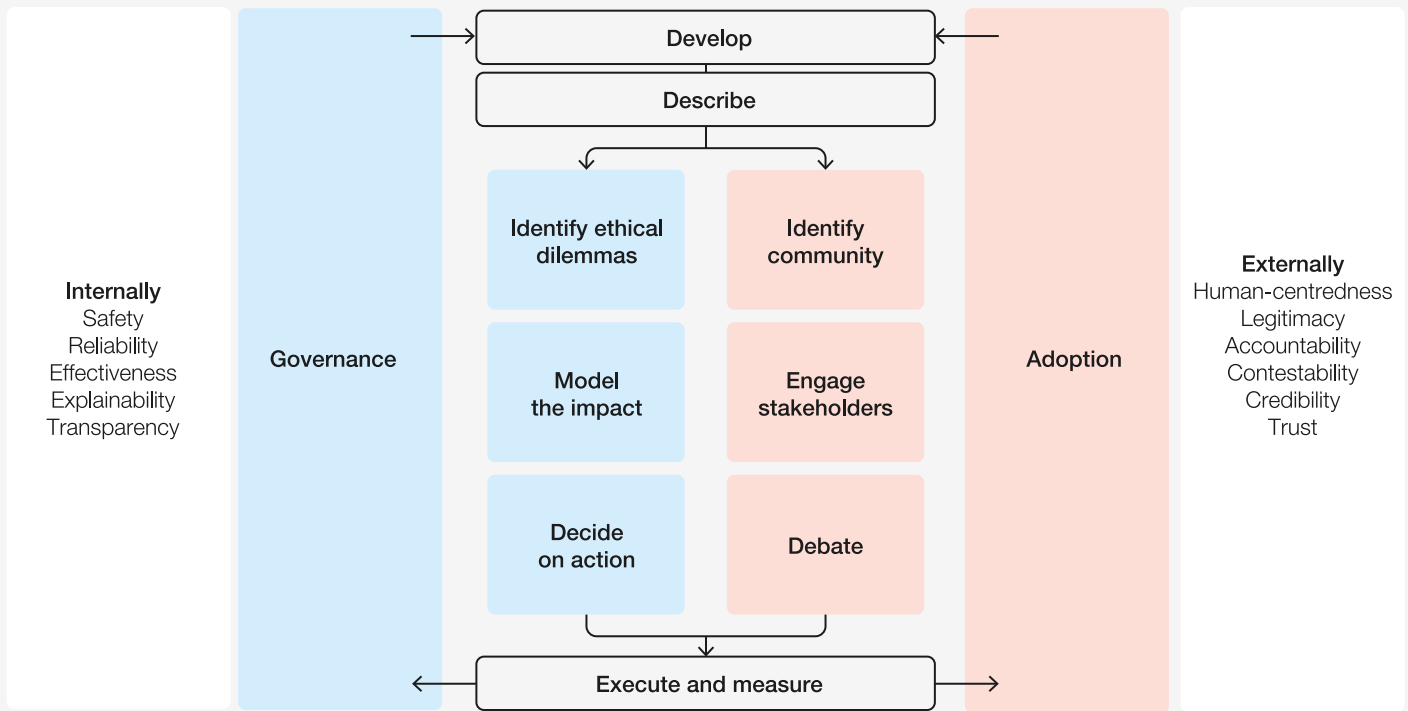
3.1 | Ethical primer

The application of an audit-based approach to digital mental health governance will not solve every ethical dilemma that presents itself because disruptive innovation introduces novel solutions and new moral questions, and only partially answers the question of how to work ethically within our community. This leaves questions such as: How to solve the new moral dilemmas presented by the increasing complexity of innovations that are specifically designed to affect behaviour, thinking and emotion, new modality options, different models of care, changes in social and cultural norms or a changing political and regulatory landscape? How to make ethical business decisions when technology may provide new means to harm consumers? How to stay current when the definition of what is ethical and moral is subjective and changing? How to show our community that we continuously operate ethically within our society, to earn its trust?

In determining how to create ethical AI in mental health, an ethical framework can be applied in order to interpret presented ethical dilemmas, involve stakeholders, consumer and community members to encourage debate and come to acceptable solutions, adapt that framework to AI/ML and disruptive technology and sustainably operate ethically in the long term. The possible detrimental effects of AI go beyond algorithmic bias and extend into the complicated relationships that AI systems have within human communities and the propagation of error through multiple layers in computer decision networks. When faced with an ethical dilemma, the outcome of not making a decision is effectively still making a decision to do nothing, and this can be especially detrimental when the associated outcomes could be harmful to individuals, communities and businesses, and have legal or reputational consequences. It is also important to note that seeing an ethical dilemma as a barrier may mean missing out on novel opportunities and breakthroughs that could improve our mental health and well-being.

In frameworks for making ethical decisions – such as those from Brown University[54] or the Markkulla Center for Applied Ethics[55] – we find a methodology for solving ethical dilemmas but no consideration of the unique challenges presented by data and AI, or how to involve the consumer, carer and clinical community in the process to develop greater social capital and a social licence to operate, a fundamental part of human-centred design for mental health. Many papers on the principles for ethical AI, such as those recommended in the World Economic Forum's *Empowering 8 Billion Minds* report,[56] or the OECD's *Principles on AI,*[57] propose a principles approach, such as: AI should benefit people and the planet; AI should be designed with respect to the rule of law, human rights, democratic values and diversity; AI should have transparency, explainability, responsible disclosure, robustness, safety, contestability and accountability. These principles leave the difficult practical implementation up to the developer. In the *Proposed Regulatory Framework for Modifications to Artificial Intelligence/ Machine Learning (AI/ML)-Based Software as a Medical Device,*[58] the US FDA proposes a risk-management, quality systems and good machine learning practices system approach to ensure ongoing safety, which is largely covered by the requirements set out above. In the paper *A Moral Licence for AI*, Deloitte, CSIRO and Data61 argue that "fair" or "ethical" algorithms or development methodologies are not necessarily achievable in increasingly complicated systems, and that the right solution is an agreement to operate with a community through transparency and shared decisions on the best course of actions, or a "social licence to operate".[59] Below we provide a practical approach for an organization to apply ethics to disruptive technology and AI in mental health, providing transparency, fairness, social capital and a way forward in an open universe of ethical situations.

FIGURE 8 | A framework for the ethical operation of AI in mental health

Internally
Safety
Reliability
Effectiveness
Explainability
Transparency

Governance

Develop

Describe

Identify ethical dilemmas

Identify community

Model the impact

Engage stakeholders

Decide on action

Debate

Adoption

Externally
Human-centredness
Legitimacy
Accountability
Contestability
Credibility
Trust

Execute and measure

1. Develop and describe the functional intent and the intended stakeholder communities

By accurately describing the intent of a tool, service or decision AI, we define the objective for which we intend to train or optimize the service; the outcomes against which we can measure the consequences of our actions; and the communities that may be affected by, and are affecting, those decisions and outcomes. An accurate description provides a basis for identifying risks and ethical dilemmas and a starting point to engage the community in solving these dilemmas. An accurate description may, while avoiding technical jargon or technology specifics, include the following:

–   **Intended consumers and stakeholders:** to determine our human sphere of influence, including demographic, geographic, social and cultural specificity. Other stakeholders may include clinicians, service workers, health-system owners, insurance providers and the cohorts used for data sources

–   **Intended state of the health situation or conditions of the consumers:** including the type of disease or condition, and the seriousness of the condition

–   **Significance of service to managing those health conditions:** describing the intended use of the information to prevent, diagnose, treat, manage or inform management of the consumer's conditions or situations

–   **Information and data:** used by both humans and machines to make decisions or the information that is being created and disseminated by the system or tool

–   **Decisions and actions:** judgements made by humans or machines while providing services or tools such as diagnosis, consumer management or treatment options

2. Identify the value conflicts that may generate ethical dilemmas in the context of your impacts within the community

An ethical dilemma may come from uncontrolled bias within the dataset or an inappropriate choice of algorithm or outcome measures, or from the application of a decision and actions to a consumer or community and their reactions – in other words, the feedback loop that is introduced into an environment with a new service or AI. For example, the introduction of an algorithm that selects the most relevant mental health information for a user may be unintentionally optimized to feed through the articles with headlines that are most picked by other users rather than to assess what is most relevant for that consumer's situation, raising the question of how we decide what to show the consumer.

– **Data:** Is the data used with consent, transparency and proper privacy, and are security controls in place? Do the datasets represent the target population demographically, culturally and clinically? Are there quality, consistency or bias issues?

– **Model:** Can we explain and interpret how the data is used to derive an answer? Have we introduced discrimination on grounds such as gender, sex, race, culture or medical background? Is the model clinically validated and more effective than current best practice? How is model performance measured over time?

– **Action and reactions:** What is the mechanism of intervention? Is there a valid association between the clinical output of our model and the clinical condition of the user? How does the output action affect the consumer? How will the consumer react to the action? Does this system interact with other human and machine systems? How are our assumptions about the effects validated?

– **Feedback:** How does our system measure external actions? What effect will external actions have on the system and its data and model over time?

Two questions that may be asked when assessing a dilemma are: what are the objectives or rights that we are balancing in this situation, and what is the scale of worst-case to best-case scenario in this situation?

For example, if we have identified that our ethical dilemma is: "When should we intervene if we identify a potentially harmful situation on public social media posts?", we can say that we are balancing the natural right to self-autonomy with the objective of healthy individuals and healthy populations. The worst-case scenario is identifying suicidal or homicidal intent in a post, and the best case is, most likely, mild anxiety. Clearly, there is a personal rights issue here that makes this a difficult question, and perspective is needed because different situations may require different levels of intervention, or no intervention at all, depending on the balance of rights.

3. Model the clinical impact

Measuring the impact on the consumer and the environment of introducing our system is fundamental to assessing whether the intervention is appropriate compared to the possible damage. Measuring the impact also brings clarity and transparency to the scale of intervention and effect. Clinical outcome measures may be based on symptomatology (e.g. CORE-OM, OQ-45), social disability (e.g. HONOS), needs (CAN), service usage or, increasingly, recovery outcomes (CHIME). Measuring must also be considered for the non-clinical, but still important, outcomes such as environmental, financial and social impacts.

4. Engage stakeholders

To gain a social or moral licence to operate, an organization must engage with the affected community. Our goal in consulting with the community is to identify and explore the range of solutions available. Consultation will provide insights into what members of that community consider to be possible/impossible and acceptable/unacceptable solutions. Through debate we may find the best balance of solutions – and we introduce transparency and accountability through public consultation, which provides an added layer of trust. Our prior work in describing the solution, understanding the ethical dilemmas and modelling the impacts will directly inform this debate. Arriving at decisions to balance opinion about difficult moral questions may be informed by three ethical approaches:

Consequentialist utilitarian approach	Produce the greatest balance of good over harm over large groups of the populations, recognizing a degree of pain and pleasure for some. This is a pragmatic approach in which the end justifies the means – and, taken to its logical conclusion, no action may be wrong if there is a net positive outcome
Deontological approach (rights and duties)	Enshrining ethical principles into laws, government policies and regulatory frameworks; a system of rights and duties and responsibilities that has consistent standards and applies to everyone in every situation. The means justify the outcome, but what are our rights in this situation? Do individual rights, laws and regulations exist that cover this situation?
Virtue approach	How do actions reflect on one's character and traits? What action will make the action taker the most virtuous person or organization? The motivating intent is more important than the outcome

5. Decide on action

Many ethical situations make us uncomfortable because they may be novel and we can never fully know the answers. However, we must act. Our chosen course of action may not be the consensus view of the community we have consulted, yet it is informed by what the community may view as acceptable and unacceptable choices, narrowing down our final decision. We have provided ourselves with the greatest possible information to inform our actions and an audit trail to show how we arrived at our decisions. Be aware that our action may have intended and unintended outcomes, and as such we should be continuously monitoring those outcomes.

6. Execute, measure and iterate

Modelling the impact of our actions, as undertaken in step 4, must continue after implementation to confirm our assumptions, review the consequences and unearth any unintended consequences.

Being able to advertise the continuous impact of our system is also valuable in providing ongoing transparency and building trust with the community we are serving.

7. Ensure continuous evaluation

Ensuring ongoing performance and operation to prevent safety issues and operate within the guardrails of our intended solution when implementing and updating the AI system is essential. Here the *Proposed Regulatory Framework for Modifications to Artificial intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device*[60] provides a roadmap for proper continual AI/ML governance:

– An organizational culture of quality and excellence, with quality systems and good machine-learning practices

– Pre-market review and specification of safety, ethics and effectiveness

– Algorithm change protocols to control the risks as expected performance or input modifications are introduced

– Real-world performance monitoring and publishing of results.
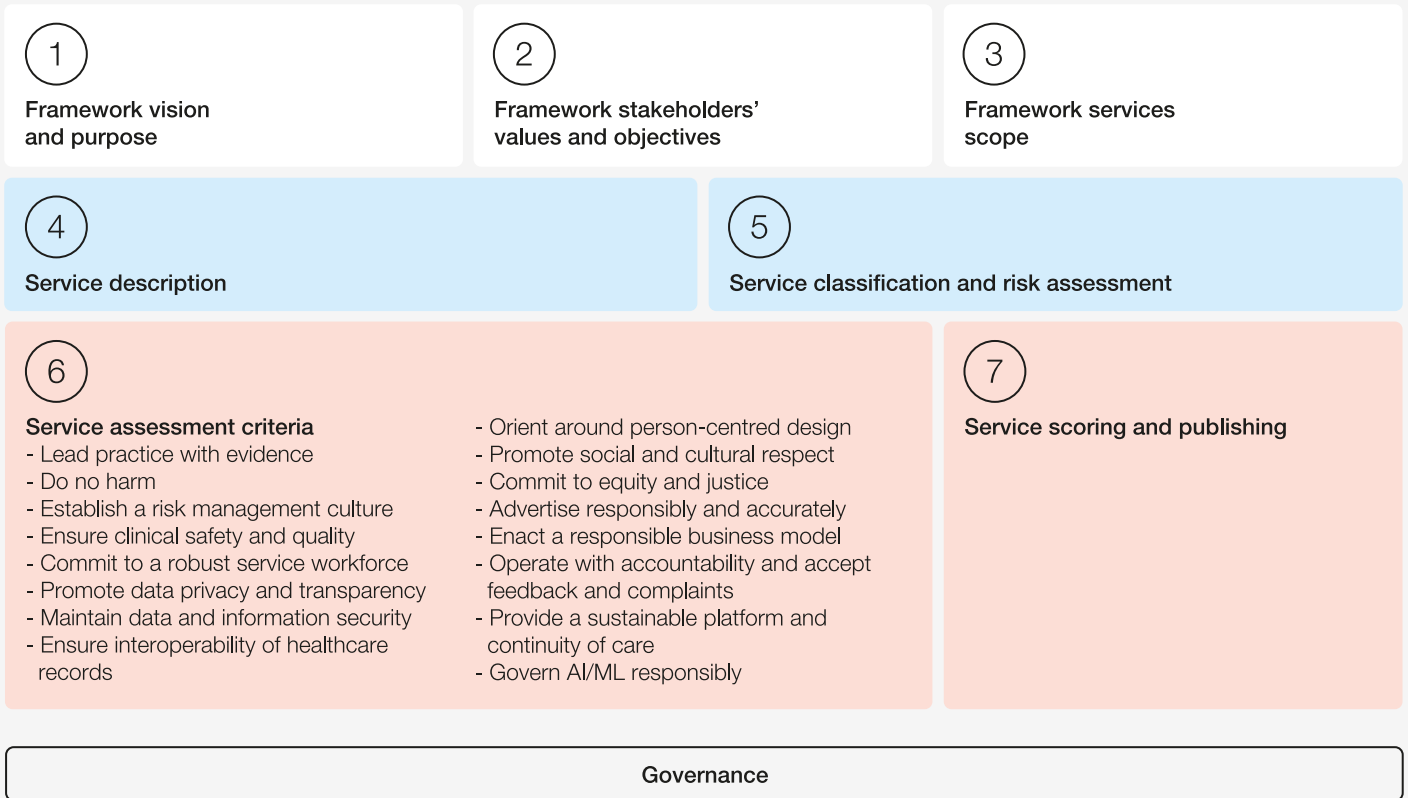
④ Policy governance

Functional, robust and flexible implementation of policy and regulations will drive the development of a safe, strategic and ethical digital mental health environment for the good of consumers and service providers.

4.1 | Elements of a functional governance system

Several key elements of an assessment framework must be considered to regulate the digital mental health industry ethically and effectively.

FIGURE 9 | Digital mental health evaluation framework

1
Framework vision and purpose

2
Framework stakeholders' values and objectives

3
Framework services scope

4
Service description

5
Service classification and risk assessment

6
Service assessment criteria
- Lead practice with evidence
- Do no harm
- Establish a risk management culture
- Ensure clinical safety and quality
- Commit to a robust service workforce
- Promote data privacy and transparency
- Maintain data and information security
- Ensure interoperability of healthcare records
- Orient around person-centred design
- Promote social and cultural respect
- Commit to equity and justice
- Advertise responsibly and accurately
- Enact a responsible business model
- Operate with accountability and accept feedback and complaints
- Provide a sustainable platform and continuity of care
- Govern AI/ML responsibly

7
Service scoring and publishing

Governance

1. Framework vision and purpose

Clearly setting out the vision and purpose of the assessment framework is the way to develop an overarching purpose and set of goals, and communicate those goals to stakeholders in the system. In this instance, the assessment framework must encourage products that are safe, trusted, effective, equitable, ethical and sustainable. The vision and purpose should be refocused by the regulatory body in its regulatory, clinical and cultural environment. Our vision and purpose are outlined in the previous chapter.

2. Framework stakeholders' values and objectives

The vision and purpose of an assessment framework influences, and is influenced by, the values, principles and objectives of stakeholders. A human-centred approach should be used to identify key stakeholders and what they wish to achieve from implementing this regulation in digital mental health. This approach is especially important in being recovery-oriented and involving consumers in their own outcomes, even at a policy level, rather than being paternalistic. In this context, an assessment framework would ideally optimize the well-being and outcomes of consumers while simultaneously supporting the development of novel technology in mental healthcare. By identifying the stakeholders and consulting with them throughout the development of the framework, we are more likely to achieve outcomes that are important to our stakeholders.

3. Framework services scope

The framework services scope is where the regulator clarifies the applications and services that lie within the jurisdiction of the regulation. The services included and excluded should be determined by the therapeutic mental health and well-being outcomes that the services intend to achieve and the use of information and communication technology in support of that therapeutic outcome. The framework will play an important role in enforcement actions and should consider services and tools currently covered, or not covered, by other clinical regulations and policies.

A definition of covered services will also have a degree of compulsion in its service coverage. For example, any tool being used to perform surgery clearly should be covered by consumer safety requirements. In well-being services, this may not be so clear-cut because there is a degree of self-autonomy and alternative therapies may exist, blurring the lines between clinical tool and consumer application. In balancing the clinical need for high-quality services in mental health with the right to self-autonomy in well-being, some regulatory bodies in the mental health and well-being regulatory domain currently offer a degree of voluntariness to being included in their assessment.

The framework services scope is predicated on the service being accurately described. A description template that is easily digestible should include the product description, provider information, clinical evidence, data usage, privacy and security information. These factors are particularly beneficial in the context of digital mental health services because they also provide transparency and clarity for users.

4. Service description

An accurate, easily digestible and honest description of the service, reliant on the vendor fairly describing the service, is essential to help in assessing the service classification and to give consumers and clinicians the information they need to compare and make an informed choice among digital mental health services. Providing consumers in a particular location with descriptions that follow a similar template greatly aids them in finding the service that will best meet their needs, rather than the service that has the biggest marketing budget. An accurate description, in line with consumer needs and based on the Transparency for Trust (T4T)[61] principles that are derived from experimental studies, will contain:

– **Provider information:** including business name and local contact details

– **Product description:** accurately describing the functional purpose, key product features, target audience and conditions

– **Clinical evidence:** describing the therapeutic model of care along with the level of, and evidence for, validation of efficacy and effectiveness of the service

– **Data usage and privacy:** what data is gathered, what is the purpose of capturing that data, how is it processed and to whom is it provided?

– **Security:** describing the key security features of the service

– **Business model:** showing how the service is funded and any third-party collaborations and partnerships

5. Service classification and risk assessment

The level of regulations put in place via the assessment framework should reflect the level of risk the product or service poses to the user. A risk assessment is important to determine the strictness of the regulations that the regulator must implement. In the context of digital mental healthcare, the regulations must mitigate the risks of the products or services such that they are safe, trusted and ethical for users, but must simultaneously promote innovation and growth in the industry. It should be noted, however, that in the context of mental health, risk is often a word associated with stigma and discrimination by employers and society, and with paternalistic and restrictive therapeutic choices that are not always in the best interest of consumers. A preliminary classification does, in this context, act as a risk assessment; it defines the level of examination to apply to services and is used to outline what criteria are relevant or not relevant for a particular service. Classification across the following lines should be used to define mandatory, non-mandatory and supplementary assessment questions:

– **Therapeutic claims:** what is the claimed level of evidence about the therapeutic effectiveness of the service? e.g. no therapeutic claims, evidence-informed, clinical validation in progress, or clinically validated and peer reviewed

– **Functional scope:** what is the functional intention of the service for the consumer? e.g. educational, measurement, self-assessment, referral, support, diagnosis, primary care or specialist care

– **Data usage:** what level of information does the service collect from consumers? e.g. anonymized data, personally identifiable information, basic health information or protected health information

– **Workforce:** does the service connect consumers with a service workforce as part of the model of care? e.g. volunteers and peers, trained counsellors or licensed mental health professionals

6. Service assessment criteria

The structure of the assessment criteria is crucial in providing the vendor with a clear set of requirements to achieve. The assessment criteria should be predicated on the ethical principles that underlie the regulations, thus communicating the importance of compliance to the vendors. Additionally, the assessme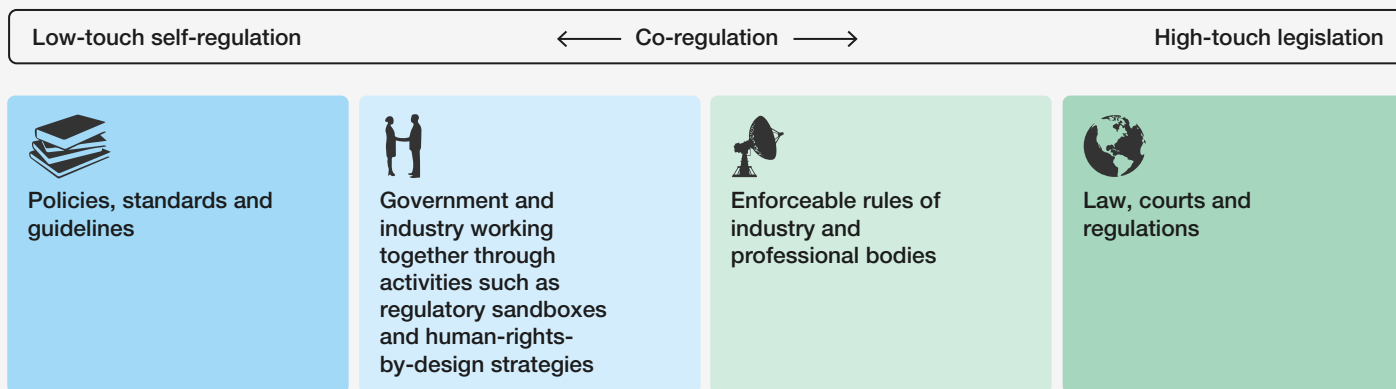nt criteria should provide pass/fail examples of how to comply with a particular requirement in that jurisdiction to enable the vendors to easily understand how to comply with regulations. Providing clear, easy and logical assessment criteria to follow means that vendors are more likely to comply and engage with regulations, thus improving the safety and effectiveness of their product or service.

7. Service scoring and publishing

The assessment criteria may be scored by marking the pass/fail of each standard or criterion and providing an overall score; each standard could be linked to a desired consumer, clinical or strategic outcome, in order to provide a score for that outcome that better communicates achievement to that stakeholder group for comparison purpose. For example, linking each assessment criterion to one of the key values – safe, trusted, ethical, effective, equitable or sustainable – will produce a scoring rubric that indicates levels of adherence to each of these outcomes and can be used to quickly assess and compare services by consumers.

| Low-touch self-regulation | ← Co-regulation → | High-touch legislation |

Policies, standards and guidelines

Government and industry working together through activities such as regulatory sandboxes and human-rights-by-design strategies

Enforceable rules of industry and professional bodies

Law, courts and regulations

Regulatory bodies can adopt varying levels of enforcement along the light- to heavy-touch continuum of regulation, within a principles-based or rules-based framework. Examples of the ways in which varying levels of regulation can be practically implemented and a discussion of the implications for a digital mental health assessment framework are detailed below. Local regulatory environments will inform the decision as to the level of enforcement.

- **Acts and laws:** acts and laws governing the regulation of goods and services are classified as "heavy touch". The Australian Therapeutic Goods Act is a set of regulations and orders that outlines the criteria for therapeutic goods to be included in the Australian Register of Therapeutic Goods. Failure to comply with the Therapeutic Goods Act when developing, selling or advertising a product is a civil offence that carries civil penalties for the provider. Acts and laws provide a high level of assurance to consumers but represent a large enforcement overhead and are less future-proof in design. As such, they tend to be made only to define the remit of the regulatory body, and not the specifics of the assessment criteria created by that body.

- **Pre-market authorization:** pre-market authorization of a product requires a vendor to meet quantitative criteria points before the product can be put on the market. Pre-market authorization can be enforceable by law, as is the case with the Therapeutic Goods Act; however, this is not always the case. Pre-market authorization is most common for products that pose a high risk to users because this kind of regulation requires a high level of logistics and administrative work. Pre-market authorization can also act as a barrier to market entry for tools that require a large randomized control trial.

- **Pre-approved library:** a pre-approved library provides users with a set of products that have been assessed by the regulatory body and meet the criteria it has established. Inclusion in a pre-approved library may be voluntary for the vendor; it is largely beneficial because it legitimizes its product and provides the product with a greater reach into the market. Head to Health is a free resource developed by the Australian government that offers a database of pre-approved mental health resources, as well as access to pre-approved phone, chat and email services. A pre-approved library also represents a legitimate resource for clinicians.

- **Independent body regulation:** independent bodies can regulate industries and professional services through licensing, membership and review panels. Independent bodies may be established as a separate entity to government regulation and may gradually legitimize themselves over time as an effective regulatory body. It is often the case that government bodies will collaborate with independent regulatory bodies to monitor or oversee an industry. This is the case with the American Psychological Association or similar regulatory entities such as the Royal Australasian College of Surgeons, where practitioners are required by the government to be a member of the entity in order to practise nationally.

- **Kitemark:** A kitemark is a product and service quality trademark that enables users to identify whether a regulatory body has assessed the product or service with which they are engaging. The key benefit of kitemarking as compared to completely voluntary regulatory standards is that it covers market education, where vendors must be audited in order to receive approval by the regulatory body. The Australian Heart Foundation Tick kitemark was established by an independent body and is approved and enforced by the Australian Competition and Consumer Commission: it provides consumers with information

on food and beverage nutrients and ingredients, enabling them to make better-informed purchase decisions. A kitemark for international digital mental health services would lend legitimacy to the services that are approved while enabling the market to continue to develop.

– **Voluntary standards:** regulatory bodies can also propose completely voluntary regulations, where recommendations are put forward regarding best business practice. In this instance, regulatory bodies run workshops or publish articles describing the regulations. Unlike kitemarking, vendors do not have to receive permission from the regulatory body in order to state that they have complied with their regulations. Additionally, because voluntary standards do not involve any level of audit, regulatory bodies must facilitate and encourage market education outside of their regulatory framework.

4.2 | Organizational structure options

Government-run regulation

A government entity can either develop a new regulatory body or extend and augment the remit of an existing regulatory body, to govern and enforce regulation in the digital mental health ecosystem. In the first instance, the government establishes a new regulatory body whose sole purpose is to regulate the industry in question. For example, the Healthdirect organization was set up to connect health districts in Australia with free Australian health advice. In the second instance, the government can establish new guidelines for regulating this industry within an existing regulatory body. As an example, the Food and Drug Administration (FDA) in the US established a novel set of regulations to oversee the use of software as a medical device. In this instance, the FDA was a suitable regulatory body for this industry because it focuses similarly on the approval of medical devices.

Independent body regulation

There is capacity for an independent body to also regulate the use of disruptive technology in mental healthcare, especially given the cross-border nature of digital technology. There are several ways in which this independent body may be put together, funded and operated.

Non-profit organizations can act as regulatory bodies. They are usually voluntary: donations and funding may be provided by philanthropists, industry donors or government bodies. PsyberGuide is a non-profit website that provides evaluations of digital mental health apps and a publicly accessible pre-assessed library of apps. It is aimed at clinicians and users, who can access the resources for free.

For-profit organizations that act as regulatory bodies may receive funds from industry bodies that benefit from the accreditation the organization provides, or from users. They often deliver additional, independent legitimacy for the accredited organization. ORCHA is a for-profit library of digital health resources, with users paying a fee to access its regulated collection of resources. ORCHA also works with app developers to review, rate and market products, thereby legitimizing the vendor's product.

Independent regulatory bodies may also be membership-based, using a system in which members pay a fee to support operations. The American Psychological Association (APA) is a scientific and professional organization whose members consist of psychologists, psychology students, teachers of psychology and other mental health professionals. Members of the APA are required to abide by its code of conduct and ethical principles. While there are no direct legal ramifications for psychologists breaching the APA ethical standards, government bodies and employers expect psychologists to hold themselves to these standards. In some cases, government bodies and employers do not permit psychologists to practise unless they are a member of the APA.

Finally, vendors can self-regulate, either internally or by forming an independent body consisting of industry stakeholders. Self-regulation often involves setting out a code of conduct or set of principles that a company must follow. Google has outlined a set of principles it aspires to meet when using AI/ML and other advanced technologies. These include aspirations such as to "be socially beneficial" and to "be accountable to people".

⑤ Incentivizing innovation

Our aim is not to restrict the global digital mental health ecosystem but to help it thrive – and integrate digital mental health into health and well-being operating models to deliver better opportunities and outcomes to consumers.

5.1 | A target operating model for digital mental health

The target operating model for digital mental health and well-being is a system integrated with current health and mental health models and strategies that promotes the needs of consumers. Psychosocial well-being is influenced by a wide range of factors that may include an individual's strengths and values, connections, relationships, support networks, culture, value systems and environmental conditions. The target operating model should be one that influences strategic decision-making on tools, services and options, in order to support the factors that influence the individual's and the population's psychosocial well-being and mental health.

The intervention pyramid below illustrates a system that aims to be integrated, multilayered, needs-based and broadly targeted, promoting well-being and psychosocial support while being focused on prevention. The system encourages services in the upper layers of the inverse pyramid and embraces individuals and communities, social considerations, basic services, risk-reduction efforts and the strengthening of resilience factors. Through early identification, readiness, resilience and support, we can increase well-being and reduce the incidence, impact and cost of mental health disorders, diverting mild cases from unnecessarily taking up specialist services.

FIGURE 11 | **Target operating model**
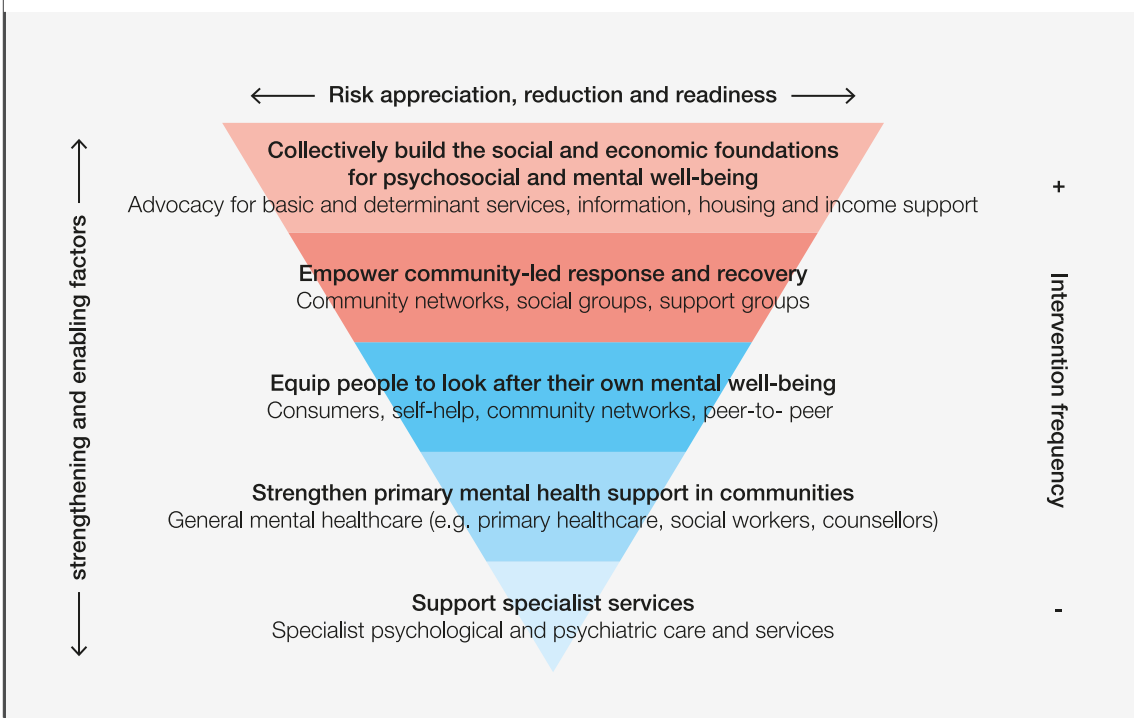
Adapted from the IASC (2007)[62] and the Framework for Psychosocial Support in Emergencies (2016)[63]

Individuals looking after their own mental health in this environment are likely to change their needs and desires in the course of their mental health journey and would be expected to use different supports dynamically, as their needs change over time, or multiple supports simultaneously. An individual is also likely to enter at any stage, using any service, with any level of severity of illness. Options at each intervention layer of the pyramid must provide needs-based services that range from broad-based services to more specialized or targeted interventions. It is expected that services rarely sit within a single layer of the pyramid and a single service may support consumers with varying intensity of interventions according to their needs. As people move away from self-care and closer to specialist services, the expectation of scrutiny of the service also increases.

Integration and interoperability of services are to be encouraged to facilitate services that can make better use of data across the ecosystem for data-driven decision-making and enable the seamless movement of consumers within the ecosystem. Digital and non-digital solutions should be integrated to take advantage of each other's strengths and offer the greatest number of options and opportunities to consumers.

5.2 | Providing incentives

The aim of a digital mental health governance framework is not to restrict new developments needlessly but to provide incentives for safe, strategic and trusted innovation. Focusing incentives, efforts and finances on mental health and well-being creates a huge opportunity to improve health and economic outcomes, and innovation brings great potential to accelerate and enhance those benefits.

The global wellness market was estimated to be worth $120.8 billion in 2020[64] and the global mental health market $207 billion in 2020.[65] At the same time, mental health is often seen as a social service, serviced by not-for-profit organizations, sidelined as a funding priority in health decisions and not a priority growth sector for investors. Although investment is changing, with $1.37 billion invested in mental health start-ups up to the third quarter of 2020[66] in the United States, in many countries it is still a major challenge to fund the services that are so badly required, and lack of access to funding is consistently listed as the number one barrier for innovators in the sector.

A governance system can and should be used to provide incentives for innovation in the ecosystem and for operations to scale with shared goals in mind. This may take the form of financial incentives; however, in order to nurture innovation and scale in mental health, access to resources such as consumers, data, research, knowledge and facilities can greatly ease the journey for providers. Below is a list of incentivization mechanisms that can be linked to adherence to governance, to be considered to develop the ecosystem.

Incentive	Description
Financial incentives	Lack of access to funding is still the number one barrier listed by innovators and providers in the digital mental health space, especially for operating in lower- and middle-income countries. Options include: – Venture capital funding – Grants, foundations and social bonds – Integration with public healthcare pricing mechanisms – Integration with public and private insurance pricing mechanisms
Operational education and support	Innovators and providers need to be supported with workshops and education programmes on how to improve their quality and safety, build capacity for sustainability and targeting of their services based on common service issues or common strategic goals for service targeting. Common example programmes and relevant topics include: – Health approval and commercialization – Human-centred and co-design methods for mental health – Data, privacy and security for digital mental health – Ethics of AI in healthcare and mental health – Cultural and social adaptation for vulnerable communities
Research, data and testing	Services require access to both consumers and high-quality data in order to perform efficacy and effectiveness studies, to advance research and to understand how to customize their service to meet the evolving needs of a population. This may be done through: – Consumer research groups – High-quality anonymized electronic mental healthcare records – Anonymized population and demographic data – University and academic research groups, platforms and networks
Access to and integration with the market	Access to consumers and the market is often a bidirectional relationship and many mental health consumers will present in existing health systems. Integration is vital and may be done through: – Integration with existing health and mental health systems as a new option for consumers – Recommendation to clinicians through a pre-approved library, clinical education material or clinician governance groups
Communities of expertise	Communities of expertise and storehouses of knowledge are vital to build on the work of others. These may be developed through the initiation or development of groups such as: – Universities – Research hospitals – Symposiums and conferences – Cooperative research centres – Collaborative research bodies

6 | Governance pilots

Lessons learned from our toolkit pilots and other governance model examples.

Pilot lessons from New Zealand

Digital Mental Health and
Addiction Service Evaluation
Framework
(DMHAS Evaluation
Framework)
New Zealand Ministry of
Health

Assessment approach and implementation model

In 2020, the New Zealand (NZ) Ministry of Health, in collaboration with the World Economic Forum and Deloitte, developed the initial Digital Mental Health and Addiction Service (DMHAS) Evaluation Framework as a part of its response to the *He Ara Oranga* report – the government inquiry into mental health and addiction – and the urgent need for remote mental health services presented by the COVID-19 pandemic.

The DMHAS framework covers mobile apps and online tools and is driven by guiding principles and goals that are intended to reflect the needs of all people in NZ.

The draft assessment approach was in the form of 55 wide-ranging individual standards that reflect the needs of consumers and clinicians for safe, trusted, ethical, sustainable and effective solutions. These standards wee divided into mandatory questions, for all services, and supplementary questions, which are required – depending on the functional scope, data collected, therapeutic claims and service workforce. Each standard contained an objective to meet, suggested evidence to show a vendor has met the standard and examples of not meeting it.

The initially trialled implementation model is for the vendor to self-audit against the standards and submit their results and evidence to a third party, which acts as the audit reviewer and keeper of a publicly accessible pre-assessed library. A score is then calculated by linking each standard to an output of either safe, trusted, ethical, sustainable or effective.

Discussion

The process was trialled in 2020 with three app vendors, with Health Navigator acting as the evaluator.

–   Vendors gained awareness of what the market desired and an understanding of potential weakness

–   The preliminary questions gave an effective advertising template to communicate to stakeholders

–   The goals and principles centred parties on what was important in New Zealand

–   Vendors found the initial self-audit process to be onerous, as well as having security concerns about providing information to a third party

–   It was suggested that the staged evaluation processes be redesigned to make better use of an initial lightweight preliminary screen followed by more detailed evaluation, remediation, certification and re-certification steps Some criteria were too broad or repetitive and the mandatory and supplementary linking and scoring rubrics were not easily understandable. More granular assessment criteria were suggested, along with further co-design of standards and an improved hosting system

–   It was suggested that multidisciplinary evaluators (clinical safety, data/security, equity, accessibility, cultural fit and user experience) be recruited

–   It was proposed that the applicability be broadened to all health conditions beyond mental health

–   An iterative, collaborative approach to refining the standards framework was suggested

–   The importance of communicating clarity and alignment of the consumer- facing scoring to all stakeholders was stressed, and revising and including user feedback suggested

The framework is currently undergoing assessment and public consultation, and a final version it is expected to be released in June 2021.

Other governance model examples

AMERICAN PSYCHIATRIC ASSOCIATION

American Psychiatric Association (APA) App Advisor[67]
United States of America

Assessment approach and implementation model

The APA is the main professional organization of psychiatrists in the United States and works to ensure humane care and effective treatment for all persons with mental illness.

The App Evaluation Model is a system for psychiatrists to self-rate apps based on a hierarchical rating system. The process involves a short question screen before a five-step evaluation is performed, with each step answering a set of specific questions in that domain:

1. Access and background 2. Privacy and security

2. Clinical foundation

3. Usability

4. Data integration towards therapeutic goal

The APA does not self-rate apps but offers the model for individuals to make their own decisions. The result is not a specific score, but a set of information that will inform a decision. The Division of Digital Psychiatry at Beth Israel Deaconess Medical Center (BIDMC) has published a publicly available list of apps that it has assessed with the APA Evaluation Model.

Discussion

– The model equips individual psychiatrists with the skills to both self-assess applications and learn to think about what makes apps effective

– The process encourages app developers to make information about effectiveness, privacy and security public and to be transparent

– The evaluation process is not verified as individuals will make decisions based on publicly available data

– While a distributed assessment model encourages free adoption and education for clinicians, it is an inefficient use of resources across an entire ecosystem

– The lack of a number score is an effective way to encourage understanding of the application purpose rather than judging from a "top ten" list. It is balancing ease of readability with a fundamental understanding of the app

---

AUSTRALIAN COMMISSION ON SAFETY AND QUALITY IN HEALTH CARE

National Safety and Quality in Digital Mental Health Standards (NSQDMH)[68]
Australian Commission on Safety and Quality in Health Care (ACSQHC) Australia

Assessment approach and implementation model

The ACSQHC is an Australian public health entity focused on contributing to better health outcomes and experiences for all patients and consumers and improved value and sustainability in the health system by leading and coordinating national improvements in the safety and quality of healthcare.

These standards were developed in 2020 through public consultation with a broad range of stakeholders. The standards consist of 59 individual actions that are outcome focused, divided into three categories:

– Clinical and Technical Governance Standard

– Model of Care Standard

– Partnering with Consumers Standard

At the moment, these are offered as a voluntary standard for digital mental health providers in Australia. ACSQHC has provided supplementary information to aid stakeholders such as: following a risk management approach, conducting self-assessment, and tips for consumers and clinicians to choose a digital mental health service.

Discussion

– The standards are an excellent example of applying a public consultation, or human-centered, approach to developing standards. As such they reflect the needs of a broad range of stakeholders and are relevant and transparent

– The format of the standards is largely in the form of outcomes and systems to ensure alignment with the consumer's needs, systems of governance and risk reduction and safety and quality. There is limited coverage addressing specific disruptive technology issues in terms of data and AI

– Without an implementation framework, the ability of the standards to affect the market is limited at this time

NICE National Institute for Health and Care Excellence

Evaluation and Assessment
for the Improving
Access to Psychological
Therapies Programme
(IAPT)[69]
National Institute for Health
and Care Excellence (NICE)
United Kingdom

Assessment approach and implementation model

NICE is a public body and an agency of the National Health Service (NHS) charged with promoting clinical excellence in NHS service providers in England and Wales.

NICE hosts a library of services that have been assessed for inclusion in their IAPT programme and integration to NHS services. The coverage is for clinical associated apps covering specific therapeutic issues. For inclusion, services undergo an intensive evaluation process that involves detailed reporting and peer review with external clinical assessors through the IAPT assessment briefings (IABs). This results in NICE either supporting, partially supporting or not supporting the service's case for adoption within the NHS.

The review processes cover:

– Content

– Technical standards

– Clinical effectiveness

– Cost and resource impact

Service testing is facilitated with services offered free to consumers during the evaluation process and funding is available to the providers to support the process. There is a facility for periodic review and assurance.

Discussion

– The process provides a very high assurance that programmes being evaluated meet the stringent standards for efficacy

– Assessing the services for cost vs. benefit is a core assessment to ensure improving service value over time

– There is also a high time and effort burden associated with meeting the requirements, and the process and requirements are not clearly laid out

– The process actively supports development of integration with the NHS and clinical models of care

– The evaluation covers a small subset of services that are highly clinical and does not cover the many thousands of applications that exist

– The evaluation process has been adapted from health application assessment and covers many of the clinical issues of health app assessment

Organization for the Review of Care and Health Apps (ORCHA)[70]

United Kingdom, Ireland and the Netherlands

Assessment approach and implementation model

ORCHA is a private company that specializes in health app evaluation and advisory services. It aims to help governments, health and social care organizations to deliver health apps that will safely make the biggest impact in terms of improving outcomes. Revenue and funding are raised through app recommendation and matching for clinicians, app evaluation and accreditation and app development and integration advisory services.

For application reviews, ORCHA uses a seven-stage process:

1. Exclusion filters and categorization of publicly available apps

2. App level classification

3. App functionality and features

4. App compliance with standards including Data and Security, Clinical Assurance and User Experience, using a yes/no questionnaire with approximately 350 questions based on publicly available app information

5. App scoring

6. Developer notification and scoring

7. Post-publication review monitoring and management

ORCHA then offer a pre-assessed library of applications for healthcare professionals to integrate into their practice or systems. ORHCA also offer a range of advisory services to vendors and healthcare professionals.

Discussion

– The yes/no questions offer an evaluation system with less overhead and that can be partially automated, greatly increasing the number of apps assessed

– The advisory and support services provided greatly encourage growth in the ecosystem, scaling and proper integration for digital/human combined models of care.

– As ORCHA is a private company, this can be seen to aid in its agility and focus on meeting market needs; however, there is less transparency about the evaluation process, which has been clinically validated but not publicly published, and the results, which are accessible only through a subscription

– The evaluation is initially, and largely, based on publicly available information on the apps being assessed. This encourages apps to be publicly transparent about all important aspects of their apps. However, without strict auditing of the evaluation or information provided, there is a higher risk of misrepresentation in outcomes

– The model has been assessed against user-based scores

OECD

OCD Recommendation on Responsible Innovation in Neurotechnology[71]
Organization for Economic Co-operation and Development (OECD)

Assessment approach and implementation model

In 2019 the OECD, the international organization that works to build better policies for better lives, developed its Recommendation on Responsible Innovation in Neurotechnology. The recommendation has 49 statements embodied in nine principles, which focus on:

1. Promoting responsible innovation

2. Prioritizing safety assessment

3. Promoting inclusivity

4. Fostering scientific collaboration

5. Enabling societal deliberation

6. Enabling capacity of oversight and advisory bodies

7. Safeguarding personal brain data and other information

8. Promoting cultures of stewardship and trust across the public and private sectors

9. Anticipating and monitoring potential unintended use and/or misuse

Discussion

– Like many statements of principles for ethical AI, such as the Google Objectives for AI Applications or the Future of Life Institute's Asilomar AI Principles, the OECD recommendation provides aspirational goals for all stakeholders to achieve

– The recommendation specifically calls out the need for oversight and advisory bodies

– The recommendation represents a broad set of stakeholder views.

– The recommendation does not have an implementation model (assessment or scoring framework) and lacks specific recommendations on how to follow each recommendation

Assessment approach and implementation model

One Mind is a not-for-profit project that aims to help people use technology to live a mentally healthier life.

PsyberGuide assesses mental health apps and publishes the results in its publicly available pre-assessed library. The assessment is scored based on four core tenets:

– Credibility – nine questions

– Transparency – three-option scale

– User experience – using the Mobile Application Rating Scale (MARS)

– Professional reviews – narrative review

PsyberGuide provides both the assessment and the pre-approved library. It also provides resources on mental health and digital mental health for consumers.within the NHS.

Discussion

– The assessment process is transparent and publicly available

– The process is based on publicly available information about the apps, encouraging transparency. However, it lacks independent audit verification

– The standards create an easily digestible score for the user

– The not-for-profit business model provides both a level of credibility to the process and the flexibility to work in a previously unchecked market, bringing a much-needed service to life. This should be balanced with long-term sustainability and scaling of the modelcare

– The evaluation covers a small subset of services that are highly clinical and does not cover the many thousands of applications that exist

– The evaluation process has been adapted from health application assessment and covers many of the clinical issues of health app assessment

Other examples

e-compared
European Union

The European project E-Compared conducts comparative effectiveness research in different mental care settings on the cost effectiveness of blended internet-based treatment for depression in comparison to standard care

ETICA Project
European Union

The ETICA Project (Ethical Issues of Emerging ICT Applications) is a research and consulting organization that collaborates with other organizations to identify black box algorithmic vulnerabilities and retrains AI-powered technology with better source data and content

Healthcare Information and Management Systems Society (HIMSS)
Unites States of America

HIMSS is a non-profit organization that acts as a global advisory and thought leader to support the transformation of the health ecosystem through information and technology. In particular, HIMSS has initiatives to advance the safety and security of devices and systems through the Health Technology Alliance - with shared knowledge and idea exchange - and public policy interpretation, advisory and advocacy projects

ImpleMentAll
Unites States of America

ImpleMentAll (IMA) is a European collaboration, funded by grants from the EU, aiming for faster and more effective implementation of eHealth interventions. IMA provides implementation support and research for mental health app assessment through frameworks such as the Normalization MeAsure Development questionnaire (NoMAD), based on Normalization Process Theory (NPT), and Model for Assessment of Telemedicine (MAST)-based mental health app assessment

NODE.Health
Unites States of America

NODE.Health is a not-for-profit organization in the US that brings together a large network of stakeholders to advise, advocate for, research and implement protocols for validation, clinical trials, publication and integration for the digital innovation ecosystem

Mobile App Rating Scale (MARS)[73]
Unites States of America

The MARS is a multidimensional instrument for classifying and assessing the quality of mobile health apps. It focuses on engagement, functionality, aesthetics and information quality. It can also be used to provide a checklist for the design and development of new high-quality health apps

Bibliography

1. Acumen Research and Consulting, "Behavioral Health Market Size to Hit US$ 240 Bn by 2026", GlobeNewswire News Room, 26 July 2019: https://www.globenewswire.com/news-release/2019/07/26/1892319/0/en/Behavioural-Health-Market-Size-to-Hit-US-240-Bn-by-2026.html (link as of 2/3/21).

2. American Psychiatric Association, "APA App Advisor": https://www.psychiatry.org/psychiatrists/practice/mental-health-apps (link as of 2/3/21).

3. American Psychological Association, "Ethical Principles of Psychologists and Code of Conduct", effective date 1 June 2003 with amendments effective 1 June 2010 and 1 January 2017: https://www.apa.org; https://www.apa.org/ethics/code (links as of 2/3/21).

4. Australian Commission for Safety and Quality in Health Care (ACSQHC), "National Safety and Quality Digital Mental Health Standards": https://www.safetyandquality.gov.au/standards/national-safety-and-quality-digital-mental-health-standards (link as of 2/3/21).

5. Brown University, *A Framework for Making Ethical Decisions,* May 2013: https://www.brown.edu/academics/science-and-technology-studies/framework-making-ethical-decisions (link as of 2/3/21).

6. Chisholm, Dan; Kim Sweeny, Peter Sheehan, Bruce Rasmussen, Filip Smit, Pim Cuijpers and Shekhar Saxena, "Scaling-Up Treatment of Depression and Anxiety: A Global Return on Investment Analysis", *The Lancet Psychiatry* 3, no. 5 (May 2016): 415–424: https://doi.org/10.1016/S2215-0366(16)30024-4 (link as of 2/3/21).

7. Elliott, Robert; Arthur C. Bohart, Jeanne C. Watson and Leslie S. Greenberg, "Empathy", *Psychotherapy 48*, no. 1 (2011): 43–49: https://doi.org/10.1037/a0022187 (link as of 2/3/21).

8. Evans-Greenwood, Peter; Rob Hanson, Sophie Goodman and Dennis Gentilin, *A Moral License for AI*, Deloitte Insights, 7 August 2020: https://www2.deloitte.com/content/dam/insights/us/articles/APAC43039_Ethics-and-AI/DI_APAC43039_Ethics-and-AI.pdf (link as of 2/3/21).

9. Farmer, P. and D. Stevenson, "Thriving at Work: The Independent Review of Mental Health and Employers 2017", 2017: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/658145/thriving-at-work-stevenson-farmer-review.pdf (link as of 2/3/21).

10. Fedder, Curt; Barb Renner and Jagadish Upadhyaya, "The Adoption of Disruptive Technologies in the Consumer Products Industry", Deloitte Insights, 11 December 2018: https://www2.deloitte.com/us/en/insights/industry/retail-distribution/disruptive-technologies-consumer-products/disruptive-digital-technologies-blockchain-potential.html (link as of 2/3/21).

11. Fiske, Amelia; Peter Henningsen and Alena Buyx, "Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy", *Journal of Medical Internet Research 21*, no. 5 (2019): e13216: https://doi.org/10.2196/13216 (link as of 2/3/21).

12. Furnham, Adrian, and Viren Swami, "Mental Health Literacy: A Review of What It Is and Why It Matters", *International Perspectives in Psychology* 7, no. 4 (October 2018): 240–257: https://doi.org/10.1037/ipp0000094 (link as of 2/3/21).

13. Gilbert, Barnabas J.; Vikram Patel, Paul E. Farmer and Chunling Lu, "Assessing Development Assistance for Mental Health in Developing Countries: 2007–2013", *PLOS Medicine* 12, no. 6 (2 June 2015): e1001834: https://doi.org/10.1371/journal.pmed.1001834 (link as of 2/3/21).

14. Global Wellness Institute, "GWI Finds Mental Wellness Is a $121 Billion Market", Global Wellness Institute: https://globalwellnessinstitute.org/press-room/press-releases/gwi-finds-mental-wellness-is-a-121-billion-market/ (link as of 2/3/21).

15. Graham, Sarah; Colin Depp, Ellen E. Lee, Camille Nebeker, Xin Tu, Ho-Cheol Kim and Dilip V. Jeste, "Artificial Intelligence for Mental Health and Mental Illnesses: An Overview", *Current Psychiatry Reports* 21, no. 11 (7 November 2019): 116: https://doi.org/10.1007/s11920-019-1094-0 (link as of 2/3/21).

16. Health & Consumer Protection Directorate-General, Green Paper: Improving the Mental Health of the Population: *Towards a Strategy on Mental Health for the European Union*. European Commission, COM(2005)484, 14 October 2005: https://ec.europa.eu/health/ph_determinants/life_style/mental/green_paper/mental_gp_en.pdf (link as of 2/3/21).

17. Heid, Markham, "New Mental Health Data Offers First Glimpse of COVID-19's Psychological Toll", *Time*, 7 May 2020: https://time.com/5833619/mental-health-coronavirus/ (link as of 2/3/21).

18. Hickie, Ian B., "The 'Uberisation' of Mental Health Care: A Welcome Global Phenomenon?" *Medical Journal of Australia* 211, no. 7 (October 2019): 315–316: https://doi.org/10.5694/mja2.50342 (link as of 2/3/21).

19. Hofmann, Stefan G.; Anu Asnaani, Imke J. J. Vonk, Alice T. Sawyer and Angela Fang, "The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-Analyses", *Cognitive Therapy and Research* 36, no. 5 (1 October 2012): 427–440: https://doi.org/10.1007/s10608-012-9476-1 (link as of 2/3/21).

20. IASC Inter-Agency Standing Committee and Reference Group on Mental Health and Psychosocial Support in Emergency Settings, *IASC Common Monitoring and Evaluation Framework for Mental Health and Psychosocial Support Programmes in Emergency Settings, 2017*, IASC, 2017: https://interagencystandingcommittee.org/iasc-reference-group-mental-health-and-psychosocial-support-emergency-settings/iasc-common (link as of 2/3/21).

21. | Jorm, Anthony F.; Scott B. Patten, Traolach S. Brugha and Ramin Mojtabai, "Has Increased Provision of Treatment Reduced the Prevalence of Common Mental Disorders? Review of the Evidence from Four Countries", *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)* 16, no. 1 (February 2017): 90–99: https://doi.org/10.1002/wps.20388 (link as of 2/3/21).

22. | Karyotaki, Eirini; Orestis Efthimiou, Clara Miguel, Frederic Maas genannt Bermpohl, Toshi A. Furukawa, Pim Cuijpers and the Individual Patient Data Meta-Analyses for Depression (IPDMA-DE) Collaboration, "Internet-Based Cognitive Behavioral Therapy for Depression: A Systematic Review and Individual Patient Data Network Meta-Analysis", *JAMA Psychiatry*, 20 January 2021: https://doi.org/10.1001/jamapsychiatry.2020.4364 (link as of 2/3/21).

23. | Kay Lambkin, Frances J.; Amanda L. Baker, Terry J. Lewin and Vaughan J. Carr, "Computer based Psychological Treatment for Comorbid Depression and Problematic Alcohol and/or Cannabis Use: A Randomized Controlled Trial of Clinical Efficacy", *Addiction* 104, no. 3 (March 2009): 378–388: https://doi.org/10.1111/j.1360-0443.2008.02444.x (link as of 2/3/21).

24. | KFF, "Mental Health Care Health Professional Shortage Areas (HPSAs)", 5 November 2020: https://www.kff.org/other/state-indicator/mental-health-care-health-professional-shortage-areas-hpsas/ (link as of 2/3/21).

25. | Larsen, Mark Erik; Jennifer Nicholas and Helen Christensen, "Quantifying App Store Dynamics: Longitudinal Tracking of Mental Health Apps." *JMIR MHealth and UHealth 4*, no. 3 (9 August 2016): e96: https://doi.org/10.2196/mhealth.6020 (link as of 2/3/21).

26. | Lui, Joyce H. L.; David K. Marcus and Christopher T. Barry, "Evidence-Based Apps? A Review of Mental Health Mobile Applications in a Psychotherapy Context", *Professional Psychology: Research and Practice* 48, no. 3 (June 2017): 199–210: https://doi.org/10.1037/pro0000122 (link as of 2/3/21).

27. | Martinengo, Laura; Louise Van Galen, Elaine Lum, Martin Kowalski, Mythily Subramaniam and Josip Car, "Suicide Prevention and Depression Apps' Suicide Risk Assessment and Management: A Systematic Assessment of Adherence to Clinical Guidelines", *BMC Medicine 17*, no. 1 (December 2019): 231: https://doi.org/10.1186/s12916-019-1461-z (link as of 2/3/21).

28. | Mental Health Commission of Canada, "Newfoundland and Labrador, Stepped Care 2.0 E-Mental Health Demonstration Project, Final Report", 2019: https://www.mentalhealthcommission.ca/sites/default/files/2019-09/emental_health_report_eng_0.pdf (link as of 2/3/21).

29. | Naslund, John A., and Kelly A. Aschbrenner, "Risks to Privacy with Use of Social Media: Understanding the Views of Social Media Users with Serious Mental Illness", Psychiatric Services (Washington, DC) 70, no. 7 (2019): 561–568: https://doi.org/10.1176/appi.ps.201800520 (link as of 2/3/21).

30. | Naslund, John A.; Kelly A. Aschbrenner, Ricardo Araya, Lisa A. Marsch, Jürgen Unützer, Vikram Patel and Stephen J. Bartels, "Digital Technology for Treating and Preventing Mental Disorders in Low-Income and Middle-Income Countries: A Narrative Review of the Literature", *The Lancet Psychiatry* 4, no. 6 (June 2017): 486–500: https://doi.org/10.1016/S2215-0366(17)30096-2 (link as of 2/3/21).

31. | Naslund, John A.; Pattie P. Gonsalves, Oliver Gruebner, Sachin R. Pendse, Stephanie L. Smith, Amit Sharma and Giuseppe Raviola, "Digital Innovations for Global Mental Health: Opportunities for Data Science, Task Sharing, and Early Intervention", *Current Treatment Options in Psychiatry* 6, no. 4 (1 December 2019): 337–351: https://doi.org/10.1007/s40501-019-00186-8 (link as of 2/3/21).

32. | National Institute for Health and Care Excellence (NICE), "Improving Access to Psychological Therapies (IAPT)": https://www.nice.org.uk/about/what-we-do/our-programmes/nice-advice/iapt (link as of 2/3/21).

33. | New Zealand and Mental Health and Addiction Inquiry, *He Ara Oranga: Report of the Government Inquiry into Mental Health and Addiction*, 2018: https://mentalhealth.inquiry.govt.nz/inquiry-report/he-ara-oranga/ (link as of 4/3/21)

34. | New Zealand Ministry of Health, *Framework for Psychosocial Support in Emergencies*, 2016: https://www.health.govt.nz/system/files/documents/publications/framework-psychosocial-support-emergencies-dec16-v2.pdf (link as of 2/3/21).

35. | One Mind, "One Mind PsyberGuide: The Mental Health App Guide Designed with You in Mind": https://onemindpsyberguide.org/ (link as of 2/3/21).

36. | ORCHA, "Organization for Review of Care and Health Apps": https://www.orcha.co.uk/ (link as of 2/3/21).

37. | Organisation for Economic Co-operation and Development (OECD), "OECD Principles on Artificial Intelligence", June 2019: https://www.oecd.org/going-digital/ai/principles/ (link as of 2/3/21).

38. | Organisation for Economic Co-operation and Development (OECD), "OECD Recommendation on Responsible Innovation in Neurotechnology", 11 December 2019: https://www.oecd.org/science/recommendation-on-responsible-innovation-in-neurotechnology.htm (link as of 2/3/21).

39. | Patel, Vikram; Shekhar Saxena, Crick Lund, Graham Thornicroft, Florence Baingana, Paul Bolton, Dan Chisholm, et al., "The Lancet Commission on Global Mental Health and Sustainable Development", *The Lancet* 392, no. 10157 (October 2018): 1553–1598: https://doi.org/10.1016/S0140-6736(18)31612-X (link as of 2/3/21).

40. | Psychiatry & Behavioral Health Learning Network, "Prevalence of Mental Illness Might Be Underreported": https://www.psychcongress.com/article/prevalence-mental-illness-might-be-underreported (link as of 2/3/21).

41. | Santa Clara University and Markkula Center for Applied Ethics, *A Framework for Ethical Decision Making*, May 2009: https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/a-framework-for-ethical-decision-making/ (link as of 2/3/21).

42. | Schwab, Klaus and Guido Vanham, "How to Prevent a Global COVID-19 Depression", *Time*: https://time.com/5817922/science-collaboration-global-covid-depression/ (link as of 2/3/21).

43. | Singh, Tanvir and Muhammad Rajput, "Misdiagnosis of Bipolar Disorder", *Psychiatry (Edgmont (Pa.: Township))* 3, no. 10 (October 2006): 57–63: http://europepmc.org/article/PMC/2945875 (link as of 2/3/21).

44. | Stiglic, Neza and Russell M. Viner, "Effects of Screentime on the Health and Well-Being of Children and Adolescents: A Systematic Review of Reviews", *BMJ Open 9*, no. 1 (1 January 2019): e023191: https://doi.org/10.1136/bmjopen-2018-023191 (link as of 2/3/21).

45. | Stoyanov, Stoyan R.; Leanne Hides, David J. Kavanagh, Oksana Zelenko, Dian Tjondronegoro and Madhavan Mani, "Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps", *JMIR MHealth and UHealth* 3, no. 1 (11 March 2015): e27: https://doi.org/10.2196/mhealth.3422 (link as of 2/3/21).

46. | Sukhera, Javeed, "From Recognition to Transformation: How Digital Technology Can Reduce Mental Illness Stigma", *The Conversation*, 30 January 2020: http://theconversation.com/from-recognition-to-transformation-how-digital-technology-can-reduce-mental-illness-stigma-130371 (link as of 2/3/21).

47. | Thorne, James, "Driven by Pandemic Demand, Mental Health Startups Surpass 2019 Funding", PitchBook, 14 October 2020: https://pitchbook.com/news/articles/mental-health-startups-venture-capital-outpace-2019 (link as of 2/3/21).

48. | Torous, John and Laura Weiss Roberts, "Needed Innovation in Digital Health and Smartphone Applications for Mental Health: Transparency and Trust", *JAMA Psychiatry* 74, no. 5 (01 2017): 437–438: https://doi.org/10.1001/jamapsychiatry.2017.0262 (link as of 2/3/21).

49. | Torous, John; Hannah Wisniewski, Gang Liu and Matcheri Keshavan, "Mental Health Mobile Phone App Usage, Concerns, and Benefits Among Psychiatric Outpatients: Comparative Survey Study." *JMIR Mental Health* 5, no. 4 (November 16, 2018): e11715: https://doi.org/10.2196/11715 (link as of 2/3/21).

50. | Trautmann, Sebastian; Jürgen Rehm and Hans-Ulrich Wittchen, "The Economic Costs of Mental Disorders: Do Our Societies React Appropriately to the Burden of Mental Disorders?" *EMBO Reports* 17, no. 9 (2016): 1245–1249: https://doi.org/10.15252/embr.201642951 (link as of 2/3/21).

51. | United Nations Human Rights Council, "Resolution 6/29: Right of Everyone to the Enjoyment of the Highest Attainable Standard of Physical and Mental Health", 14 December 2007: https://www.right-docs.org/doc/a-hrc-res-6-29/ (link as of 2/3/21).

52. | US Food & Drug Administration, *Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device*, 2 April 2019: https://www.fda.gov/media/122535/download (link as of 2/3/21).

53. | Uusitalo, Susanne; Jarno Tuominen and Valtteri Arstila, "Mapping Out the Philosophical Questions of AI and Clinical Practice in Diagnosing and Treating Mental Disorders", *Journal of Evaluation in Clinical Practice* (30 September 2020): https://doi.org/10.1111/jep.13485 (link as of 2/3/21).

54. | Victoria, Royal Commission into Victoria's Mental Health System, and Penny Armytage, *Interim Report*, 2019: https://rcvmhs.vic.gov.au/interim-report (link as of 2/3/21).

55. | Villongco, Christopher and Fazal Khan, "'Sorry I Didn't Hear You.' The Ethics of Voice Computing and AI in High Risk Mental Health Populations", *AJOB Neuroscience* 11, no. 2 (June 2020): 105–112: https://doi.org/10.1080/21507740.2020.1740355 (link as of 2/3/21).

56. | Wan, William, "The Coronavirus Pandemic Is Pushing America into a Mental Health Crisis," *Washington Post*, 4 May 2020: https://www.washingtonpost.com/health/2020/05/04/mental-health-coronavirus/ (link as of 2/3/21).

57. | World Economic Forum, *Empowering 8 Billion Minds: Enabling Better Mental Health for All via the Ethical Adoption of Technologies*, 2019: http://www3.weforum.org/docs/WEF_Future%20Council_Mental_Health_and_Tech_Report.pdf (link as of 2/3/21).

58. | World Health Organization, "eHealth at WHO", World Health Organization: http://www.who.int/ehealth/about/en/ (link as of 2/3/21).

59. | World Health Organization, *Mental Health Action Plan 2013–2020*: https://apps.who.int/iris/bitstream/handle/10665/89966/9789241506021_eng.pdf? sequence=1 (link as of 2/3/21).

60. | World Health Organization, *Mental Health Atlas 2017*, 2018: https://apps.who.int/iris/bitstream/handle/10665/272735/9789241514019-eng.pdf?ua=1 (link as of 3/4/21).

61. | World Health Organization, *Policies and Practices for Mental Health in Europe*, 2008: https://apps.who.int/iris/handle/10665/107366 (link as of 4/3/21).

62. | Wykes, Til and Stephen Schueller, "Why Reviewing Apps Is Not Enough: Transparency for Trust (T4T) Principles of Responsible Health App Marketplaces (Preprint)", Preprint, *Journal of Medical Internet Research*, 4 October 2018: https://doi.org/10.2196/preprints.12390 (link as of 2/3/21).

Contributors

Lead Authors

Stephanie Allen
Global Health Care Sector Leader, Deloitte Global,
Australia

Rohan Hammett
Asia Pacific Health Care Sector Leader, Deloitte,
Australia

Marcus Schweizer
Fellow, Disruptive Technology in Mental Health,
World Economic Forum
Manager, Deloitte, Australia

Acknowledgements

Arnaud Bernaert
Head of Health and Healthcare, World Economic
Forum, Switzerland

Sir Philip Campbell
Former Editor-in-Chief, Nature, United Kingdom

Kenneth Carswell
Technical Officer, World Health Organization, Switzerland

Murali Doraiswamy
Professor of Psychiatry and Medicine, Duke University
Medical Centre, USA

Thorsten Engel
Partner, Deloitte, New Zealand

Cameron Fox
Project Specialist, World Economic Forum, USA

Alvaro Fernandez
Chief Executive Officer and Editor-in-Chief,
SharpBrains, USA

Thomas R. Insel
Co-Founder, Mindstrong, USA

Ralph Judah
Senior Managing Director, Monitor Deloitte, USA

Rotem Kopel
Chief Executive Officer, MindMatters Foundation,
Switzerland

Kelly McCain
Project Lead, World Economic Forum, Switzerland

David Rabinowitz
Senior Manager, Deloitte, USA

Karen Rommelfanger
Director, Emory University Centre for Ethics, USA

Shekhar Saxena
Professor, Harvard School of Public Health, USA

Danielle Schlosser
Behavioural Health Lead, Verily, USA

Richard Schweizer
Policy Advisor, Mental Health Commission of New
South Wales, Australia

Robyn Shearer
Deputy Director General of Mental Health and
Addiction, New Zealand Ministry of Health, New
Zealand

Peter Varnum
Associate Director, Orygen Global, Norway

Endnotes

1. World Health Organization, *Mental Health Action Plan 2013–2020.*

2. Ibid.

3. Trautmann, Rehm and Wittchen, "The Economic Costs of Mental Disorders".

4. Farmer and Stevenson, "Thriving at Work".

5. Schwab and Vanham, "How to Prevent a Global COVID-19 Depression".

6. Wan, "The Coronavirus Pandemic Is Pushing America into a Mental Health Crisis".

7. Fedder, Renner and Upadhyaya, "The Adoption of Disruptive Technologies in the Consumer Products Industry".

8. Hickie, "The 'Uberisation' of Mental Health Care".

9. Torous and Roberts, "Needed Innovation in Digital Health and Smartphone Applications for Mental Health".

10. Torous et al., "Mental Health Mobile Phone App Usage, Concerns, and Benefits Among Psychiatric Outpatients".

11. World Economic Forum, *Empowering 8 Billion Minds: Enabling Better Mental Health for All via the Ethical Adoption of Technologies*.

12. Farmer and Stevenson, "Thriving at Work".

13. Chisholm et al., "Scaling-Up Treatment of Depression and Anxiety".

14. Furnham and Swami, "Mental Health Literacy".

15. Patel et al., "The Lancet Commission on Global Mental Health and Sustainable Development".

16. United Nations Human Rights Council, "Resolution 6/29: Right of Everyone to the Enjoyment of the Highest Attainable Standard of Physical and Mental Health".

17. World Health Organization, *Mental Health Action Plan 2013–2020.*

18. Gilbert et al., "Assessing Development Assistance for Mental Health in Developing Countries".

19. Patel et al., "The Lancet Commission on Global Mental Health and Sustainable Development".

20. Health & Consumer Protection Directorate-General, *Green Paper: Improving the Mental Health of the Population.*

21. World Health Organization, *Policies and Practices for Mental Health in Europe.*

22. World Health Organization, *Mental Health Atlas 2017.*

23. KFF, "Mental Health Care Health Professional Shortage Areas (HPSAs)".

24. Jorm et al., "Has Increased Provision of Treatment Reduced the Prevalence of Common Mental Disorders?".

25. Victoria, Royal Commission into Victoria's Mental Health System, and Armytage, *Interim Report.*

26. New Zealand and Mental Health and Addiction Inquiry, *He Ara Oranga.*

27. Psychiatry & Behavioral Health Learning Network, "Prevalence of Mental Illness Might Be Underreported".

28. Singh and Rajput, "Misdiagnosis of Bipolar Disorder".

29. Heid, "New Mental Health Data Offers First Glimpse of COVID-19's Psychological Toll".

30.	World Health Organization, "eHealth at WHO".
31.	Hickie, "The 'Uberisation' of Mental Health Care"; Naslund et al., "Digital Technology for Treating and Preventing Mental Disorders in Low-Income and Middle-Income Countries".
32.	Sukhera, "From Recognition to Transformation".
33.	Kay-Lambkin et al., "Computer-based Psychological Treatment for Comorbid Depression and Problematic Alcohol and/or Cannabis Use".
34.	Mental Health Commission of Canada, "Newfoundland and Labrador, Stepped Care 2.0 E-Mental Health Demonstration Project, Final Report".
35.	Karyotaki et al., "Internet-Based Cognitive Behavioral Therapy for Depression".
36.	Hickie, "The 'Uberisation' of Mental Health Care".
37.	Ibid.
38.	Elliott et al., "Empathy".
39.	Martinengo et al., "Suicide Prevention and Depression Apps' Suicide Risk Assessment and Management".
40.	Lui, Marcus and Barry, "Evidence-Based Apps?"; Larsen, Nicholas and Christensen, "Quantifying App Store Dynamics".
41.	Hofmann et al., "The Efficacy of Cognitive Behavioral Therapy".
42.	Fiske, Henningsen and Buyx, "Your Robot Therapist Will See You Now".
43.	Graham et al., "Artificial Intelligence for Mental Health and Mental Illnesses".
44.	Ibid.
45.	Naslund et al., "Digital Innovations for Global Mental Health".
46.	Graham et al., "Artificial Intelligence for Mental Health and Mental Illnesses".
47.	Fiske, Henningsen and Buyx, "Your Robot Therapist Will See You Now".
48.	Villongco and Khan, "'Sorry I Didn't Hear You.' The Ethics of Voice Computing and AI in High Risk Mental Health Populations".
49.	Stiglic and Viner, "Effects of Screentime on the Health and Well-Being of Children and Adolescents".
50.	Uusitalo, Tuominen and Arstila, "Mapping Out the Philosophical Questions of AI and Clinical Practice in Diagnosing and Treating Mental Disorders".
51.	Fiske, Henningsen and Buyx, "Your Robot Therapist Will See You Now".
52.	Naslund and Aschbrenner, "Risks to Privacy with Use of Social Media".
53.	American Psychological Association, "Ethical Principles of Psychologists and Code of Conduct".
54.	Brown University, *A Framework for Making Ethical Decisions.*
55.	University and Markkulla Center for Applied Ethics, *A Framework for Ethical Decision Making.*
56.	World Economic Forum, *Empowering 8 Billion Minds: Enabling Better Mental Health for All via the Ethical Adoption of Technologies.*
57.	Organisation for Economic Co-operation and Development, "OECD Principles on Artificial Intelligence".
58.	US Food & Drug Administration, *Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device.*

59. Evans-Greenwood et al., *A Moral License for AI.*

60. US Food & Drug Administration, *Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device.*

61. Wykes and Schueller, "Why Reviewing Apps Is Not Enough".

62. Inter-Agency Standing Committee and Reference Group on Mental Health and Psychosocial Support in Emergency Settings, *IASC Common Monitoring and Evaluation Framework for Mental Health and Psychosocial Support Programmes in Emergency Settings,2017.*

63. New Zealand Ministry of Health, *Framework for Psychosocial Support in Emergencies.*

64. Global Wellness Institute, "GWI Finds Mental Wellness Is a $121 Billion Market".

65. Acumen Research and Consulting, "Behavioral Health Market Size to Hit US$ 240 Bn by 2026".

66. Thorne, "Driven by Pandemic Demand, Mental Health Startups Surpass 2019 Funding".

67. American Psychiatric Association, "APA App Advisor".

68. Australian Commission for Safety and Quality in Health Care (ACSQHC), "National Safety and Quality Digital Mental Health Standards".

69. National Institute for Health and Care Excellence, "Improving Access to Psychological Therapies (IAPT)".

70. ORCHA, "Organization for Review of Care and Health Apps".

71. Organisation for Economic Co-operation and Development (OECD), OECD "Recommendation on Responsible Innovation in Neurotechnology".

72. One Mind, "One Mind PsyberGuide: The Mental Health App Guide Designed with You in Mind".

73. Stoyanov et al., "Mobile App Rating Scale".

WORLD ECONOMIC FORUM

The World Economic Forum,
committed to improving
the state of the world, is the
International Organization for
Public-Private Cooperation.

The Forum engages the
foremost political, business
and other leaders of society
to shape global, regional
and industry agendas.