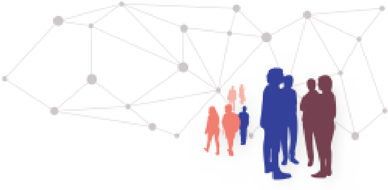
University of Zurich^UZH
Digital Society Initiative



Artificial intelligence in medicine - Case Study «Diagnosis»

DSI Strategy Lab 2022

Editorial team: Markus Christen, Nikola Biller-Andorno, Janna Hastings, Jeffrey David Iqbal, Viktor Kölzer, Michael Krauthammer, Viktor von Wyl and Claudia Witt

This case study was created in a systematic and participatory process as part of the «Strategy Lab Artificial Intelligence of Medicine» of the Digital Society Initiative (DSI) at the University of Zurich. The process, which involved various stakeholders, extended from the beginning of 2022 to mid-2023. In particular, a workshop with experts in June 2022, a workshop with doctoral students of the DSI Excellence Program in August 2022, and the testing of excerpts of the case study in various events with stakeholders were decisive for the creation of the case study.

Participants in the expert workshop, in addition to the editorial team, were: Abraham Bernstein, Daniel Eberli, Philipp Fürnstahl, Sven Hirsch, Christian Kauth, Emanuela Keller, Birgit Kleim, Tanja Krones, Titus Neupert, Cristina Rossi, Bernd Stadlinger, Florent Thouvenin and Andreas Wicki. Participants of the PhD workshop were Anais Aeschlimann, Ibrahim Al Hazwani, Joe Baumann, Giulia Frascaria, Marius Furter, Alexandra Ioana Georgescu, Maël Kubli, Alexander Lill, Eanuele Martinelli, Judit Martínez Moreno, Matteo, Micol, Markos Mpadanes, Kimon Papadopoulos, Amina Saleh, Jana Sedlakova, Kateryna Shapovalova, Lukas Tribelhorn, Morley James Weston, Basak Yalman, Federica Zavattaro and Donatella Zingaro. We thank all of these individuals is for contributing to this process; in particular, Jeffrey David Iqbal for providing operational guidance throughout the Strategy Lab process.

Further information: dsi.uzh.ch/strategy-lab.

Structure and use of the case studies

The case studies evolve along a temporal progression with a parallel increase in the level of autonomy of the AI application, as follows: the case study starts in the (fictional) «now» using AI with a very limited level of autonomy. It progresses to the «near future» with greater involvement of AI in the decision-making process. Finally, in a «far future scenario», it will be played out that an AI makes relevant decisions largely autonomously (as an «avatar» or «digital twin»). A precise temporal allocation of the scenarios is not attempted. The scenarios are fictitious in nature and were framed rather as «positive utopias» than «negative dystopias», although several aspects of the case studies are intended to be provocative and they should trigger ethical debates regarding the desirability of the scenarios as outlined here.

Comparatively simple language has been chosen; technical terminology has been avoided as far as possible so that lay people can understand the case studies. Factual statements are referenced only minimally (the scenarios are, after all, fictional in nature), supported by few references that are as

accessible to a general audience as possible. When creating the scenarios, care was taken to ensure that there were connections in terms of content between the fictional case stories.

The case studies will be available «open access» with the intention that they can be used in educational settings and/or workshops discussing AI applications in medicine. The keywords indicate topics that characterize the case studies and should help to decide upon their use in educational settings. Suggested questions after each «step» further guide their use in workshops and discussions.

Keywords: AI, Depression, Diagnosis, Digital Twin, Health Data, Mobile Coach, Social Media Data, Wearables

1 Introductory remarks

Depression is one of the greatest and often underestimated health challenges of our time. Up to one fifth of the population will suffer from depression in the course of their lives [1]. The consequences of depression include negative effects on quality of life as well as functioning in everyday life and work [2]. 10 to 15% of all individuals with recurrent major depressive episodes die by suicide [3]. At the same time, it is not easy to diagnose depression and to distinguish it from psychological states such as low mood or sadness. In this hypothetical case study, the aim will therefore be to use novel forms of behavioral observation to identify whether a person is suffering from major depression or is at high risk of developing such a condition in the near future. The aim is to minimize the risk of suicide in particular. Affected people should be made aware of support services in good time. The case study simplifies factors that in reality play a role in the diagnosis of depression (e.g., subjective moments and sociocultural factors).

This case study will focus on an area of life that increasingly influences the everyday life of many people: Interactions on social networks. It is therefore about the «digital behavior» of people within such networks: What information do people look at, what do they «like», what do they share and with whom, etc.? Four different groups are thus affected by the case study: the users of social networks (especially those at risk of depression), their relatives, medical professionals with a connection to the affected person (e.g., primary care physician, psychiatrist), and the providers of social networks. In the further development of the scenario, technologies that do not yet exist then also appear, with the aim of improving awareness itself [4]. The case study does not exist in this form, all names of persons and companies are fictitious.

2 Case study level «Now»

2.1 Situating the scenario

Let us imagine that an interdisciplinary consortium of researchers, citizens and designers has developed a groundbreaking analysis tool: Depro-Detect; or DeDe for short. DeDe is an app that can determine with a high degree of reliability, based on the behavior of users on social media such as Facebook or LinkedIn, whether a person is at risk of developing, or is already experiencing, severe depression. DeDe is operated on servers of the consortium with high security standards. We can therefore assume that the data and analyses of DeDe are protected and they serve the sole purpose of preventing suicide by depression. The cantonal data protection authority has reviewed and approved the app. Regular audits ensure compliance with data protection standards.

Social media providers can enter into an agreement with the DeDe operator: they can offer DeDe as an option in their social network. In concrete terms, this means the following: Users can decide for themselves whether DeDe should be activated or not. If they give their consent, they also provide contact details of close relatives and medical professionals who can be informed in an emergency. All these people, in turn, receive a notice that they have been specified as a contact – and only when they have given their consent does the contact become active, i.e. they would be contacted in an emergency. In this case, the DeDe app gains access to all interaction data that the users generate through their use of the social network. The providers of the social networks themselves do not have access to the evaluations of DeDe. However, they are compensated from health insurances if they offer DeDe as an option; especially since early diagnosis of depression has been shown to reduce health care costs for society as a whole; DeDe is thus financed by insurance premiums). DeDe is designed in such a way that persons who hardly use social media are informed that their online behavior generates too little data for a correct analysis and that they should therefore uninstall the app.

2.2 How it could work

Our story begins with Alice. She is prone to depression and her primary care physician made her aware of DeDe. Since she spends quite a lot of time on the social network Visionbook and because the company Giga offers DeDe as an option, she downloaded the app. Sometimes Alice feels bad and during these periods, she watches more news about the current uncertain world situation, which depresses her even more. DeDe notices this changing pattern and, in a first stage, provides recommendations to look at other news to lighten the mood. Sometimes this works, sometimes it does not.

When it does not work and Alice is looking at more and more depressing messages while sharing them less and less with her «friends» on Visionbook, DeDe notices that things are getting more serious. The app informs Alice that she is stuck in a depressive phase and recommends her to talk to one of the specified contacts after all, for example her brother Peter, who knows about Alice's depressive phases and has also agreed to be listed as a contact in DeDe. Unfortunately, even this does not always help, as she is often not following the recommendations of the App.

Sometimes Alice really sinks into the «dark corners» of the Internet and the smartphone's motion detector (DeDe can also read this information) shows that she hardly moves physically. In this case, Alice's psychiatrist, Sabine (also part of Alice's contact network stored in DeDe) receives a warning. Sabine now knows that the situation may be serious. She calls Alice; Alice picks up only after the third attempt. Sabine convinces Alice to come to her immediately, which Alice reluctantly accepts. At the doctor's office, the two discuss the situation. Sabine convinces Alice to go to an outpatient psychiatric clinic to deal with the acute depressive episode. Talks and drug interventions bring Alice out of the blues and prevent a possible suicide.

2.3 Possible questions for discussions

What do you like about DeDe, what do you dislike? Would you use DeDe yourself? Which aspects of the scenario do you consider realistic, which not? What responsibilities lie with the people receiving data from DeDe? What are the responsibilities of the DeDe user?

3 Case Study Level «Near Future»

3.1 Situating the scenario

The DeDe app has now been in use for several years and has been consistently improved. Thanks to DeDe, many people could be diagnosed with a depressive phase and help could be provided at an early stage. The overall social burden of depression has thus been reduced – but difficult challenges remain. Although the number of people suffering from depression has decreased, those with diagnosed depression remain a medical challenge.

This challenge must be considered in the context of the following social developments. The digital transformation of society has led to some effects that still seem paradoxical today. On one hand, many everyday ojects are now part of a large «Internet of Things» (IoT). It has become normal for numerous objects equipped with sensors to optimize everyday challenges; for example, to adapt electricity consumption to fluctuating production (wind and solar power plants led to an increasingly volatile power supply). On the other hand, the «Back to Analog» (BATA) movement has led some people to develop an increasing aversion to being constantly in front of a screen. Social networking has become less important for those people. Especially for people diagnosed with depression, DeDe's approach – such as showing «positive» messages – no longer works so well. The DeDe consortium has therefore asked itself the following question: In a world shaped by the IoT and the BATA movement, how can we still diagnose people with depression?

3.2 How it could work

After years of research, the DeDe consortium has finally developed an innovative AI system that helps people with depression live normal lives. The Digital Depression Therapy system – DDTS for short – is based on two basic principles: First, DDTS can access a much broader database than the outworn DeDe app. After many attempts, Switzerland has also finally succeeded in establishing a dynamic electronic patient record in which a person’s health information is collected on an ongoing basis, protected from unauthorized access. Only the person concerned and the team of health professionals that each person has (in this era, the primary care physician no longer exists, but each person is assigned a «health cluster» of 3-5 health professionals who know the person personally and manage all interactions with other specialists) can access this data; researchers can use the data of the entire population in an anonymized form. Some health data is updated in real time – for example, if a person with diabetes has an implant that releases insulin when needed, that data is also available. The patient record also stores physiological data that correlate with depression if persons agree to use such sensors.

DDTS is now able to generate a comprehensive, personalized picture of the person's depression based on this broad database. This personalized depression model generated by DDTS is also important for the wider healthcare system. It is used, for example, to assess a person's ability to work and thus also to calculate the amount of unemployment compensation. After all, the goal of DDTS is to help people lead normal lives. Accordingly, services provided by the state or the health care system, as well as obligations of those affected, are to a certain extent subject to monitoring by DDTS. At this time, this is considered fairer than when health officials decide more or less arbitrarily, who should or should not receive a disability pension, for example.

The second basic principle is that DDTS is able to interact with the numerous devices of the «Internet of Things» with the aim of supporting the therapy of depression. For example, DDTS can access the home's lighting system and change the color of the light to best help remedy the current depressive episode. It is also known that depression is associated with a decreased desire to move. DDT can therefore act on the everyday digital helpers in such a way as to promote movement in depressed people. However, the following should also be noted: as described, the BTA movement has gained popularity; thus, there is also skepticism about digital paternalism. DDTS therefore never operates completely autonomously. Each person's health cluster regularly looks at how DDTS is functioning and intervenes when necessary. The affected person also receives regular reporting on how DDTS is helping to combat the depression. The same applies to the person's close caregivers – they also receive regular reports (as specified by the person) on how the depression therapy is progressing. As with the DeDe app, they are also encouraged to contribute to the therapy, for example by talking to the affected person on a regular basis.

3.3 Possible questions for discussions

What do you like about DDTS, what do you dislike? How wide should the «decision space» for DDTS be? Who decides on the autonomy level of DDTS (the patient, the care team, the society)? Would you use DDTS yourself? Which aspects of the scenario do you consider realistic, which not?

4 Case Study Level «Far Future»

4.1 Situating the scenario

We are now in a distant future. The functionality of DeDe can now only be admired in an app museum, and DDTS has become obsolete in the form it was conceived at the time. The digital sphere has become as natural as the air we breathe. Just like clothes, digital devices are part of everyday human existence. Of course, there is still – to use the analogy of clothes – something like «digital nudists»; i.e., people that meet each other in protected zones without any digital interference; «naked», so to speak. However, it has become common for everyone to have «digital twin services» to whom people can assign different tasks (e.g., they pay the bills for the person, they get official paperwork done on digital platforms, etc.). The service is like a «second self», hosted by different service providers depending on the relevant sphere of life (e.g. for work, for health etc.). People interact with their digital twins in natural language.

Thanks to a sophisticated system, the personal data that forms the basis for the functioning of the digital twin services is protected in such a way that only the person concerned can access it. The individual differences regarding the nature of the digital twin services are great. While those who are committed to the ideals of the old BTA movement use this service only in a very rudimentary form only the most necessary functions, other people use very sophisticated digital twin services that sometimes almost life a «life of their own» Some narcissists even fall in love with their digital twin and enter into a form of marriage with it. Many discussions in this future time concern the question of whether certain forms of digital twins form a legal personality, what happens to them of the «original person» dies or whether they are even morally responsible for some of their actions in a certain sense.

4.2 How it could work

Of course, digital twin services also play an important role in health care. Depending on the preferences of the user, the digital twins contain sophisticated physiological and cognitive models. This can be used, for example, to simulate the effect of therapeutic interventions on both the physical functions and mood of the human «original». Of course, digital twins are never perfect – but many undesirable side effects of therapeutic interventions can be much better estimated and minimized. As a result, the notion of a «disease» has changed significantly. Previously, a disease was thought to be a comparatively stable condition, resulting from the malfunction of certain physiological processes caused by genetic predispositions, unfavorable socio-economic conditions or harmful environmental influences. The goal of medical intervention was to bring people out of these condition – once that was achieved, medicine had done its job. Today, people look at the problem from a completely different direction: there is a selection of positively occupied health states from which people can choose – examples include the state of «calm mindfulness» or the state or «human dynamo». The parameters of these states are well known and the digital sphere of each person is used to optimize the respective selected health state. In this sense, we no longer speak of diseases at all; the «International Classification of Diseases» (ICD) has been replaced by the «International Classification of Health States» (ICHS).

The former DeDe consortium, which also developed the DDTE system in the past, is in the distant future a network of international researchers whose own digital twins are highly advanced and who develop new upgrades for digital twins in the human-AI network. This network is dedicated to difficult sentiment modeling. It has ensured that an increasing number of people are giving their digital twin a «mood upgrade» – that is, a model that can also capture and influence people's moods. This upgrade is closely linked to each person's digital sphere and controls, for example, the delivery of active ingredients from implants to prevent unwanted states of consciousness. People also use these upgrades to test out in their digital twin new forms of mental health states with which the ICHS is increasingly supplemented. In other words, they first allow their digital twin to experience the state and then exchange information with it in order to decide whether they themselves also want to enter this state. Given that the digital twins are still not perfect simulations of the original, persons never can be sure whether the intervention on themselves will indeed generate the same mood effects as reported by their digital twins.

An interesting phenomenon has emerged in this regard: Some users test with their digital twins states that used to be called «depression», as medical historians have found out. Now there is a hot debate whether this condition should be included in the ICHS or not...

4.3 Possible questions for discussions

How would you design your digital twin? Could you describe the «health state» in which you would like to be? Would you give your digital twin a «mood upgrade»? Which aspects of the scenario description do you consider desirable, which not?

5 References / Links

[1] https://www.medix.ch/wissen/guidelines/psychische-krankheiten/depression/

[2] https://link.springer.com/article/10.1007/s40273-021-01040-7

[3] https://www.neurologen-und-psychiater-im-netz.org/psychiatrie-psychosomatik-psychotherapie/stoerungen-erkrankungen/depressionen/auswirkungen

[4] https://pubmed.ncbi.nlm.nih.gov/33958228/