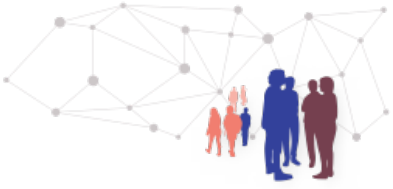
University of Zurich[UZH]

Digital Society Initiative



Artificial intelligence in medicine - Case Study «Allocation»

DSI Strategy Lab 2022

Editorial team: Nikola Biller-Andorno, Markus Christen, Jeffrey David Iqbal, Christian Kauth, Viktor Kölzer, Tania Krones, Michael Krauthammer 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.

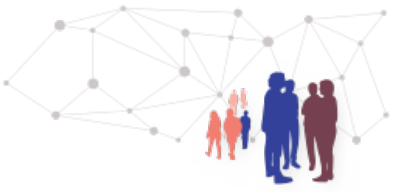
Participants in the expert workshop, in addition to the editorial team, were: Abraham Bernstein, Daniel Eberli, Philipp Fürnstahl, Sven Hirsch, Christian Kauth, Emanuela Keller, Birgit Kleim, Tanja Krones, Titus Neupert, Cristina Rossi, Bernd Stadlinger, Florent Thouvenin and Andreas Wicki. Participants 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: Depression, Health Data, Social Media Data, Wearables

1 Introductory remarks

Even in rich countries, resources in the health care sector are fundamentally scarce. Even where there is no capped budget for health care spending, the willingness and ability of citizens to pay eventually reaches a limit. Switzerland currently spends 11.8% of its gross domestic product on health care [1]. This expenditure competes with investments in education, infrastructure, social security, etc. More and more people have difficulties paying their health insurance premiums and receive tax-funded subsidies or forgo benefits. At the same time, it is well known that our health care system is inefficient – the potential savings are estimated at 20 to 25%.

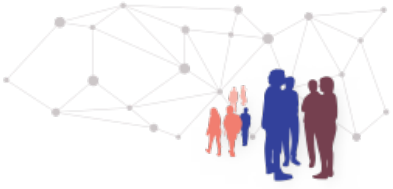
An important part of the waste is duplication and misuse of care. Patients do not receive the care that best meets their needs. Freedom of choice as well as the lack of successful treatment leads to clinic shopping, which in turn drives up costs. This also applies to the field of psychiatry. The clinical picture of depression in particular, which often accompanies sufferers for many years and which has become even more important since the pandemic, causes economic losses in the billions [2]. At the same time, it is known that not all people with mental illnesses in Switzerland have access to the diagnostic, therapeutic or preventive care they need [3].

In order to spare patients frustration and suffering and to use scarce public resources responsibly and economically, the best possible fit between patient and treatment is needed. In this hypothetical case study, we will use the example of depression to show how AI can improve access to medical interventions. This will not only optimize the fit between patient and treatment offer; at the same time, the targeted collection of large, relevant amounts of data, including patient-relevant treatment results, will allow continuous improvement of care pathways, closing gaps and eliminating waste. This will also make it possible to predict which interventions should take place for which type of patient, which combination, where (inpatient, outpatient or at home) and by whom (e.g. healthcare professionals, layperson, AI). The medical cases described do not exist in this form, all names of persons and companies are fictitious.

2 Case study level «Now»

2.1 Situating the scenario

Digitalization has diversified the forms of treatment: Patients no longer only have the choice between different therapeutic approaches and therapists with different training and approaches. There is also a range of telemedicine services and so-called «mental health apps» that are designed to support people with depression or anxiety disorders in everyday life, usually in combination with a human therapist. In



this way, inhibition thresholds can be lowered and barriers to care (e.g. waiting lists) can be reduced. This diversification of therapeutic offerings is also accompanied by a change in the processes surrounding depression treatment (treatment pathways) and a shift in treatment locations: Increasingly, care can be shifted from the outpatient or even inpatient setting to the home. The new forms of «home care» are often accompanied by monitoring, e.g., in case of acute deterioration (through an app such as «Depro-Detect» or DeDe, see case study «Diagnosis»).

For those affected, however, it is often a challenge to find the most suitable form of treatment for them. To find out which form of treatment is right for whom, a university team has developed an AI-based app called «deGPS» that helps navigate the jungle of treatment forms and generates a recommendation – in other words, they wanted to develop a GPS for depression treatment paths, hence the name. When you first use the app, you feed it with information about your own depressive illness and other conditions (comorbidities), as well as treatment goals, your own values, preferences (e.g., about the treatment setting), and priorities (e.g., rapid treatment success or low private costs). In addition, one can evaluate past interventions. All this can be done by the user or with the help of a health professional.

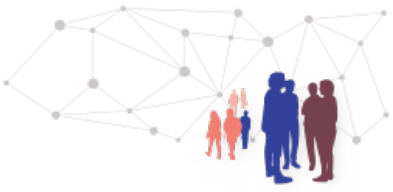
Based on this, the system makes suggestions for treatment options that can then be discussed with the primary care physician. No recommendations are made about specific providers (this is seen as a possible future add-on option), but rather options for treatment forms are presented, which are assigned a rating. For this rating, patient data are matched with evidence available on the Internet regarding (subgroup-specific) efficacy, risks, and side effects of various interventions. In addition, the experience with the recommended or perceived treatment options is rated by the user (possibly also in consultation with a medical expert), so that the system can learn what works for the individual and make increasingly targeted suggestions.

The data can be shared by users to improve the app and for healthcare research. In this way, the app's triage system and those who design the care system for psychiatric patients in Switzerland can learn which type of patient benefits most from which treatment at which point in time. If the app is widely used, it will also be possible to make statements about which treatment capacities should be expanded and which can be reduced. This is directly relevant for hospital planning, for example. Thus, in the sense of a learning system, treatment paths can be adapted and offers can be improved according to need. This possibility is particularly valuable in psychiatry, where treatment pathways are complex due to additional diseases or socioeconomic factors.

2.2 How it could work

Florian has had recurring depressive phases for many years now. As a teenager, he thought these were the usual torments of puberty, but the situation never really improved, even later. He hardly has any friends, let alone relationships. He had to abandon an apprenticeship, and often just sleeps away the days in a darkened room. The only thing he enjoys is eating, but this affects his weight and does not exactly raise his self-esteem. After much coaxing by his parents, because even at 25 he has not yet moved out, Florian contacts a psychiatrist. The discussions and subsequent clarifications make it clear that Florian is suffering from severe depression, which is also accompanied by concomitant illnesses. His unhealthy eating habits and lack of exercise have resulted in damage to his musculoskeletal system, and there is even the threat of diabetes if action is not taken soon.

Given the multitude of health challenges Florian faces due to his depression, the psychiatrist informs him about the new deGPS app. She still sees a certain will in Florian to get out of the situation through his own efforts. By initializing the app together with Florian, she wants to strengthen this will. Florian –



who sometimes spends hours on mobile game apps – is taken with the idea that an app could help him get out of his situation. For the first time, he feels some initiative of his own as they enter the relevant health information into the app together. For the first time, Florian also becomes aware of how comprehensively depression is determining his life and damaging his body.

Florian cannot enter all the information at the first meeting with the psychiatrist. He needs to gain an understanding about what he can and wants to realistically achieve. His psychiatrist has also told him that he is now facing a long process and that he cannot expect deGPS to simply spit out the «solution» for him. Once home, Florian immediately falls into a hole again – but a few days later he starts up deGPS again and he tries to enter the missing information himself. Shortly thereafter, the app comes up with a result. Unsurprisingly, it advises him to contact a psychiatric specialist as soon as possible. The app also indicates that in his case, inpatient treatment using medication combined with behavioral therapy that also focuses on eating behavior could probably be a promising approach. Florian is initially somewhat shocked. Nevertheless, he contacts his psychiatrist again to discuss this recommendation. He realizes that he has to do something - otherwise he will rot in his room.

2.3 Possible questions for discussions

What do you think about the treatment navigation app deGPS? What do you like, what do you dislike? Would you use deGPS yourself? Which aspects of the scenario description do you consider realistic, which not?

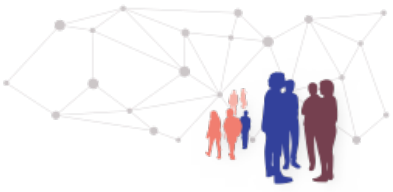
3 Case Study Level «Near Future»

3.1 Situating the scenario

When deGPS was launched by the research group, the app had relatively few features to offer. Users are reminded of the early days of Amazon's «You might also like this» recommendations. Or of navigation systems in cars, which sometimes take you on strange routes when you select the «shortest possible connection». This also led to some legal disputes that almost ruined the start-up that the research team had founded. However, after important liability issues were resolved, an investor stepped in to enable the professional development of the app. In addition, a license was sold to Cyriakus, a large private psychiatric hospital chain, which has already expanded its offerings into the outpatient and homecare/telehealth markets in recent years and recognized the app's value. Now that the tariff system in psychiatry has been adapted in such a way that the most efficient possible treatment of patients has become economically attractive, an app that allows the best possible fit between treatment needs and offers have become interesting for the clinic chain for quality and cost reasons. Insurers have also already expressed interest.

The now not-so-small startup hires a team of programming, design and marketing professionals and has the app reprogrammed to impress with its graphics and ease of use. The app is linked to the electronic patient record and also interfaces with the now widely used DeDe app (see case study diagnosis) – an app that gathers clues for depressive behavior patterns from the use of social network data and can thus also indirectly provide clues about the success of suggested treatment paths. «De GPS» is now called «Your Best De-Treatment» (YBDT).

However, the company's developers are now faced with three controversial questions: All tests have so far taken place with the sole licensee Cyriakus, the large hospital chain. After the successful pilot operation, Cyriakus would like to limit the recommendations that YBDT generates to those offers that



are available within the company's clinical ecosystem. In addition, the recommendations should also take into account the costs for the various treatment paths – including profits that the chain could generate. This is justified by the efficient use of resources.

Second, the developers have improved outcome data collection. YBDT now captures both patient-reported outcome measures (PROMS) and the impact of the care process on patient experience, such as communication (patient reported experience measures or PREMS). These additional data have the effect of making YBDT better and better at suggesting the optimal treatments – but at the cost of no longer being able to understand in detail how the AI arrives at its recommendations.

Thirdly and finally, one member of the team proposes to equip YBDT also with a chatbot that can have a diagnostic, therapeutic or preventive effect via interaction with the user himself. YBDT would thus not only become a tool that suggests suitable therapies, but would also have a therapeutic effect itself.

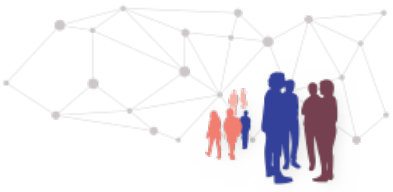
It becomes clear to the team: The combination of all three points results in a complex allocation problem: YBDT will be able to provide both evidence-based suggestions and (with the chat function) therapeutic effect. However, in order to be able to evaluate whether these goals are really achieved, one should ideally work with a partner company that allows a controlled application – for example, with regard to the actual success of a therapy evaluated by a third party. But what if the partner's resources are limited and YBDT makes recommendations that the affected person cannot take advantage of? Should one refer the affected person to offers that one can no longer control? Or should one offer them the «second best solution»? And how do you explain this to the affected person? The team concludes that it cannot really solve these problems and programs YBDT in such a way that it continues to make only recommendations and ultimately leaves it up to the affected person to decide whether he or she wants to follow the recommendation or not.

3.2 How it could work

Florian has now embarked on the long road to a cure for depression with the help of deGPS. Florian regularly had to seek inpatient psychiatric treatment when his episodes exacerbated. In most cases, he followed the recommendations of deGPS. He was able to move out of his home in the meantime, and he keeps trying new offers by adjusting his data in deGPS every now and then to see if a new offer pops up. However, it is clear to him that this trial and error is probably not really worth it.

He is therefore pleased when his psychiatrist, who had told him about deGPS and whom he still consults from time to time, tells him about a further development of the app. YBDT is supposed to generate significantly better suggestions. He said that the app was still in a test phase and that one would have to undergo treatments in an institution of the Cyriakus Group if one wanted to use the app – but his psychiatrist explained that this clinic company was one of the best in the world in the field of psychiatry and had a wide range of services. Accordingly, Florian is persuaded and he reports to a specialist at the institution, who instructs him on how to use YBDT. The specialist also makes it clear that he has the ultimate responsibility of whether or not to follow the treatment recommendations. Although YBDT also takes population-level factors into account in the calculation – for example, with regard to the presumed availability of a therapy – it cannot ultimately make a final judgment here. It remains his responsibility to choose; also taking into account that some treatment will require approval by medical experts, e.g., when medication is involved. Accordingly, the specialist suggests that Florian should sometimes also consider the second-best option, as it may only end up in second place because YBDT suspects, based on the current data, that this option is currently difficult to access.

Florian takes a liking to the app, which indeed generates significantly better suggestions for therapy options – adapted to his personal mood and situation. He occasionally meets with the specialist, who



asks him about his experience with the app for study purposes. Florian asks him if he knows why YBDT makes such good recommendations – but the specialist disappoints him: «It's impossible to know, the model is too complicated and is constantly learning. But the most important thing is that it works, right?»

Florian is not entirely satisfied with the answer – and the dissatisfaction increases when he reads something about YBDT in an article of an investigative media portal: the app would exclusively offer therapy offers of the Cyriakus Group and systematically ignore offers of alternative providers in particular. Florian is unsettled – although he himself did not have good experiences when he clicked through the somewhat confused offers of deGPS; and the current therapeutic approaches work well. Nevertheless, he wonders what actually influences the therapy paths suggested by YBDT...

3.3 Possible questions for discussions

What do you think about «Your Best De-Treatment» (YBDT)? What do you like, what do you dislike? Would you use YBDT yourself? Which aspects of the scenario description do you consider realistic, which not?

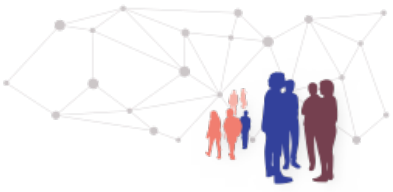
4 Case Study Level «Far Future»

4.1 Situating the scenario

After YBDT established itself and was widely used despite a few scandals at the beginning – such as the accusation that the app's suggestions were biased towards certain providers – the startup was acquired by tech giant Pineapple a few years ago. The YBDT app has been continuously improved and, as the Internet of Things has spread, has been able to tap into more and more data sources, which have been used to generate better-customized treatment paths. Even data that at first glance seemed far-fetched, such as online shopping, social interactions, or sleep, but also real-time physiological parameters such as blood pressure or oxygen saturation, contributed to this improvement. Soon, the first press articles appeared under titles such as «Pineapple Mental Health App Outperforms Psychiatrists – better, faster, and cheaper». Public discussions followed on the question of whether a human professional needed to be involved at all, as the data on treatment success showed no difference.

Consulting firms saw these further developments as a milestone that could eliminate access and equity problems for patients with depression and perhaps other mental illnesses in the future. Although the monthly subscription price for the enhanced app is quite considerable, it is still cheaper than the usual treatment costs to date, which is why the costs for insured persons are covered by many health insurance companies.

Since Pineapple's development department has almost unlimited resources, a next step has been initiated: instead of an external device that both determines the optimal treatment path for patients with depression and at the same time still has a therapeutic effect (through targeted electrostimulation of certain brain areas), an implant in the brain tissue is to take over this function. Technologically, a number of breakthroughs have been achieved: Sensors based on carbon nanotubes can record the neuronal activity of wide areas of the brain far better and less invasively than classical electrodes. New types of three-dimensional computer chips make it possible to run AI algorithms in the smallest possible space with low energy requirements. Photonics has found ways to modulate neuronal networks in the brain precisely. Finally, the problem of energy supply could also be solved: analogous to brain cells, the implants can obtain energy from blood oxygen and glucose. Thus, these implants can be transplanted



into the brain and function there for many years. The systems can also be controlled wirelessly so that they can be monitored and controlled – but only at a very short distance, and a device similar to headphones must be worn. Finally, the difficult problem of securing the system against «hacking the brain» was mainly solved – although not perfectly, as no technological systems is 100% secure.

The result is «Perfect Brain Health» (PBH) – a system that can recognize all known forms of psychiatric disorders and take appropriate action. Most interventions run automatically. For example, sleep-wake cycles can be controlled, with appropriate advance notice and confirmation by the user. External specialists can also be called in to suggest further appropriate measures. These suggestions can cover broad areas of life. Depending on the individual situation, a change of working conditions or a change of relationship may also be suggested. Both individual aspects and overall social considerations are taken into account in order to avoid «excessive therapy». The aim is to counteract the pressure that people feel to constantly optimize themselves and thus harm themselves and the community. There is also an ongoing debate claiming that socio-economic factors still play an important role developing for depression risks – factors that are beyond the control of the individual. Public health experts warn that a focus on individual behavior optimization may miss the point.

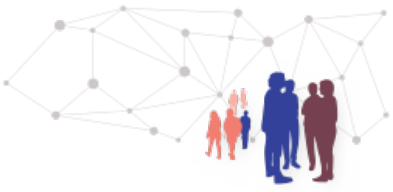
For their part, the experts who are constantly developing Perfect Brain Health are using sophisticated AI models that also take efficiency and cost aspects into account. Finally, it is known that psychiatric disorders can have a strong impact on the economic and social productivity of individuals. Due to low birth rates, labor is a scarce resource – and the societal need to secure this capacity for the benefit of all is widely recognized. Accordingly, treatment pathways are designed to optimally allocate medical resources and achieve the best possible effect – the PBH system developed by Pineapple Health is an important component in this effort.

4.2 How it could work

Florian looks back on his life so far on his 75th birthday, the date of his retirement: He has managed to escape the clutches of his depressive illness. It took him until well into midlife before he was finally able to start living a normal life. Rather late – quite common for that time – he became a father at the age of 60. His daughter Kora is now facing the third educational sequence; that time when young people are expected to develop their autonomy skills so that they can better understand and control numerous digital devices that have accompanied them in their lives so far.

Florian is somewhat concerned, because modern diagnostic systems have long since recognized that, due to his genetic predisposition and the epigenetic imprinting of his genetic makeup as a result of his long depressive phase, Kora also carries a significant risk of developing depression. Florian's personal digital coach has been pointing out options to reduce this risk since she was 10 years old. Pineapple Health's newly developed implant may be the best solution for this, the coach said. Florian wants to spare Kora what he himself had to go through. But because Kora is not legally allowed to make her own decision about such an intervention until she has completed the third educational sequence, i.e. the training of her autonomy skills, he has to make the formal decision.

Of course, he asked Kora – but she is unsure. For one thing, she has heard the strange stories of «brain hackers» – people who suddenly put headphones on people with the PBH system in crowds or something to try to gain control of the system. But what makes Kora even more insecure are the stories of her friends who have already had such a system transplanted for medical reasons. They are always uncertain about the target values – i.e., the mental and emotional state – that the system is supposed to achieve. Sometimes they test the whole spectrum from «cheerful» to «melancholic» and the coaches then warn against such arbitrariness – warnings that are regularly countered with the objection that



they should only be trimmed to a «cog in the system» anyway. «Should I really have the PHB system implanted before autonomy training?» she wonders....

4.3 Possible questions for discussions

What do you think about PBH? What do you like, what do you dislike Would you ever consider using PBH yourself, assuming it was safe and reversible? Which aspects of the scenario description do you think are realistic, and which do you think are not?

5 References / Links

[1] https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/kosten-finanzierung.html

[2] Swiss Health Observatory: Mental Health in Switzerland: Monitoring 2020. Report 15/2020. https://www.obsan.admin.ch/sites/default/files/obsan_15_2020_bericht_2.pdf

[3] https://www.bag.admin.ch/bag/de/home/strategie-und-politik/politische-auftraege-und-aktionsplaene/politische-auftraege-im-bereich-psychische-gesundheit.html