Personalized Medicine

Alexandru Bârsan, Matrikelnummer: 7396261
Konrad Wartke, Matrikelnummer: 6481209

Boundary and context

For our mid-term report we have chosen BenevolentAI Platform, a deep learning AI system which focuses “to get the right treatment to the right patients at the right time by using artificial intelligence to produce a better target selection and provide previously undiscovered insights through deep learning.” **[Daley 2020] [Benevolent]**
Benevolent is a private company established in 2013 in London, England which focuses on the development of artificial intelligence and computational medicine technologies which are intended to change the way new medicine is created and provided to the users. With the help of AI systems the company seeks to improve the patient’s life by generating more efficient decisions during the treatment while also lowering the costs when it comes to drug development **[Benevolent 2020a]**.

In our report we will assess the *trustworthiness* of the AI as based on the EU Ethics Guidelines for Trustworthy AI adapted to the healthcare domain. In the year of 2019 as indicated in **[ALTAI 2019]** the High-Level Expert Group of Artificial Intelligence (AI HLEG) published the ethical guidelines for a trustworthy AI. As specified before we will assess a few of the guidelines presented in the ALTAI which are: Privacy and Data Governance, Transparency, Diversity, Non-discrimination and Fairness

When developing an artificial intelligence system the developers have to take into consideration the above presented guidelines. In our case the platform aids scientists in the development of more efficient drugs and new ways to treat diseases, while it also seeks to improve patient’s lives with the appliance of technology when it comes to decision making as stated on **[Crunchbase].** Due to the fact that it interacts with the health of human beings when it comes to the prescription of personalised medicine or a treatment, from our point of view for the system to be *trustworthy* the compliance of the following guidelines provided by the AI HLEG are important:

- the users should be aware that the predictions and advised medicine are the result of an algorithmic decision and this should be overseen by the scientists developing the new drugs*(Human Agency and Oversight)*
- the vulnerabilities and accuracy of the system when prescribing a treatment should be taken in consideration *(Technical Robustness and Safety)*
- the data required from a patient, if personal it must respect the General Data Protection Regulation (GDPR), as well as the privacy of the individual *(Privacy and Data Governance)*

- the system should be traceable, which means that it could trace back the data which was used to take certain decision when prescribing a personalized medicine, as well as being able to explain those *(Transparency)*
- the data with which the system is trained should be diverse and non-discriminative, it should contain data from different races and groups of people, children and people with disabilities should be included in the data *(Diversity, Non-discrimination and Fairness)*

In this report we will not assess Societal and Environmental Well-being, and the Accountability guidelines of the ALTAI.

Socio-technical scenarios

A good example of the accuracy and importance of the AI system developed at Benevolent is stated in an article by dailymail **[Adams 2020]** . This is exemplified with a prediction done by the system regarding a medicine developed three years ago called baricitinib, which proved to also be helpful in the treatment of the COVID-19 virus **[Adams 2020]**. The drug was initially developed for those suffering from rheumatoid arthritis, a chronic inflammatory disorder. Benevolents AI anticipated that baricitinib, marketed as Olumiant would reduce “the ability of the virus to infect lungs cells” and so it has been. Pan-European researches which are led by the Sweden’s Karolinska Institute have validated the idea that the drug reduces the rate of deaths of the hospitalized patients by two-thirds and it also functions for the elderly persons, a vulnerable group which people which usually tend to be left a side as indicated by **[Adams 2020]**
This example highlights the importance of precision medicine and the introduction of AI technologies in healthcare, while also referencing this example to the 5th guideline(*Diversity, Non-discrimination and Fairness*) of the **[ALTAI 2019]** due to the fact that the drug helps in the treatment of those who are the most vulnerable to the covid-19 disease.

A blog post made on their website **[Benevolent 2020b]** states that there are “moral, scientific and economic reasons for having a more inclusive patient demographic represented in precision medicine.“ Having a more diverse gathering of data also means that the outcome of the precision medicine which is being or has been developed is more effective. In the article it is said that from a scientific point of view drug treatments can be influenced by the patient’s ethnicity, age, sex or even other demographic factors. **[Benevolent 2020b]** indicates that it is always good to improve “the diversity of patients represented in patient data”. For this they developed The Diversity Analysis Tool which they made public on GitHub **[GitHub 2020]**, so that other organizations can make use of it. The tool makes use of all datasets it is provided with, supposing those contain information about the age, sex, ethnicity and sex of different people, then combines them into a single dataset which then is able to provide clean reports reports about the data. Benevolent hopes that this tool will help other organizations which are also preoccupied with the development of precision medicine. When it comes to the guideline of *Diversity, Non-discrimination and Fairness* **[ALTAI 2019],** although sometimes it is not possible to gather information from all minorities, the company is trying its best to do so, since they know this is an important factor when it comes to precision medicine.

Benevolent aims to aid scientists in drug development by developing tool kits for personalized medicine. The personalization tries to improve treatment success by finding the drug components most effective for each person. Benevolent does this by analyzing existing research for relationships between the desired effect and the genome of a patient or a class of patients **[Neil et. al. 2020] [Sellwood 2018] [Myszczynska et al. 2020]**.

The patients themselves are not directly involved with Benevolent. Since the treatment costs of personalized drugs are high (up to 100.000 dollars / year and patient **[Iriart 2019]**), they expect better than normal treatment **[Miller 2014]**. On the other hand these patients often suffer from rare diseases, so they have no comparison to the normal treatment method. The patients have to trust the researchers developing the drug. These scientists and the organisations they work at are the main customers of Benevolent.

In regards to technology, Benevolent uses mostly natural language processing to parse research articles for interactions between genes and combine it with patient data obtained from the treating doctor **[Myszczynska et al. 2020] [Neil et. al. 2020]**. Together they form a knowledge graph containing recommendations which drugs to develop **[Benevolent 2019]** and how to chemically synthesize them. Researchers then take the suggested drug components to synthesize them **[Sellwood 2018]**.

Ethical tensions

Since Benevolent mostly uses existing research to recommend drugs and the existing genome data the recommender has a trade-off between accuracy and fairness. Genome-wide association studies (GWAS) are used to find relationships between genomes and potential drug effects **[Watcharapichat 2020]**. In these studies minorities are often misrepresented. In 2009 96% of the analyzed DNA was from western Europeans **[Popejoy and Fullerton 2016]** and sex-specific chromosomes are often ignored **[Khramtsova 2019]**. Minorities are therefore underrepresented in the research data and also in the knowledge graph produced by Benevolent. This leads in practice to a dilemma when having to decide between better treatment options for western males and worse treatment options for everyone else.

There is also a true dilemma between accuracy and explainability. Even if Benevolent finds an effective drug for a patient, the mechanism the drug uses is often still causally unexplained. Even if understood, the scientist developing the drug also has to explain the mechanism to the doctor treating the patient. The patient may trust the doctor prescribing the drug, but the doctor may not understand the drug mechanism himself. Because every developed drug is personalized, the doctor has no experience prescribing the drug.

The high costs of up to 100.000 dollars / year and patient **[Iriart 2019]** can also lead to a dilemma in practise when having to choose between equality of preferences and equality. The high cost compared to standard, non-personalized medicine may lead to only wealthy people having access to personalized medicine. Even health insurance providers may decide the costs are too high and other treatment options are more cost-effective **[Brothers 2015].**

Another dilemma in practice is the conflict between the quality of the drugs developed and privacy of the patient. For maximum efficiency the complete DNA of the patient should be searched for drug targets. This also includes highly personal information such as sex and

ethnicty. Because personalized medicine is often the last resort of the patients, they may be pressured by researchers or by themselves to offer everything in advance well beyond usefulness. This can be resolved by making sure to only use anonymized data. The treating doctor should be the only person interacting with the patient and therefore should not be included when researching drugs.

In conclusion, personalized medicine is a highly effective way to treat rare diseases. The ethical tensions mostly rest in the unequal distribution of the technology because of the high costs and unequal distribution of research data. The experimental nature of the drugs also leads to patients having to trust the researchers more than their treating doctor.

GRADE: 5 RVZ (Dec. 15, 2020)

Claims about the developed system

Benevolent claims to develop technology benefiting all of society **[Benevolent 2020b]**. One of their stated goals is for "no disease to go untreated" **[Benevolent 2020f]**.
The developed technology should be able to understand underlying causes for every disease and use that knowledge to treat them **[Benevolent]**.
Benevolent claims to develop better ("highest quality" **[Benevolent]**) drugs in fewer cycles, and for a lower price, due to lower development costs **[Benevolent 2020a]**.
The quality is also supposed to increase due to the data-driven methodology generating more hypotheses for further testing **[Benevolent 2020a]**.
Benevolent intends for more representative research data **[Benevolent 2020i, Benevolent 2020h]**. According to Benevolent, this will lead to better research **[Benevolent 2020g]**.
Benevolent claims the user data is secure according to the European General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act of 1996 (HIPAA) **[Benevolent 2020h]**.

Evidence for the claims

For better diversity, Benevolent released a tool to analyze research data **[Benevolent 2020b]**. This tool analyzes the research data according to several dimensions and creates charts showing research the diversity of the data set.
This tool to increase the diversity in research data is part of their „Data Diversity Initiative“ **[Benevolent 2020d]**. Benevolent tries to convince medical study organizers to get a more representative sample of the population. According to Benevolent, this will also increase the data quality and lead to more hypotheses generated and more patients treated **[Benevolent 2020b]**. When successful, this may aid in their efforts to develop drugs for rare diseases, since data for these diseases are distributed over the population.
Benevolents is aiding in the development of several drugs **[Benevolent 2020j]**, none of which have exited the trial phase of development. So Benevolent cannot yet provide evidence of better/”highest quality” **[Benevolent]** drugs.
Benevolent also does not provide evidence for the claim of generating more and better hypotheses for later drug development.
Benevolent provides no evidence for their claim of securing patient data, but notes the importance of these data protections **[Benevolent 2020h]**.

Map Ethical issues to Trustworthy AI Areas of Investigation

For an AI system to be considered trustworthy by us it must respect the principles outlined in Chapter 1 of the “Ethics Guidelines for Trustworthy AI” paper **[AIHLEG, 2019]**. As quoted “these requirements are applicable to different stakeholders partaking in AI systems’ life cycle: developers, deployers and end-users, as well as the broader society” **[AIHLEG, 2019].** By developers the authors are referring to those who are designing and developing the AI system, as for us BenevolentAI being the developer, while the deployers are the companies or people

which are using the AI system for their business or which offer services with the help of it. The end-users are those which engage with the application of the AI in a direct or indirect manner, in our case these are the patients.

Each stakeholder group plays a different role when it comes to ensuring that the principles of a trustworthy AI are met. The developers should respect and apply the requirements when developing the product, while those who are using and deploying the system should ensure the fact that the requirements are met. The end-users must request that those are being upheld **[AIHLEG 2019]**.

The ethical issues which we will evaluate and map are for the following requirements for trustworthy AI **[AIHLEG 2019]:**

- Privacy and data governance
- Transparency
- Diversity, non-discrimination and fairness

Privacy and data governance is a hotly debated topic in the 21st century and since the system interacts with human beings in the development of precision medicine and then with the patients which are getting the drugs prescribed, it is an area which has to be evaluated. The ethical issue that arises regarding this topic is the concern of those who are providing their own personal data. What happens with this data? Is the organisation respecting the ethical, as well as those imposed by the law rules?

Transparency is another factor which has to be taken in consideration, because the organization should be able to explain the origin of the drug and the decisions which lead to its creation/development. Also the end-user should be made aware that the medicine he is prescribed was discovered by an AI system or that the treatment he is advised is a result of an algorithmic decision **[ALTAI 2019]**. This is another ethical issue which has to be addressed, are those benefiting from the AI system aware of the fact that the treatment they are prescribed is the decision of a computer? Is the organisation able to explain the process behind the development of the new drugs?

Diversity, non-discrimination and fairness is a principle of great importance for the organisation. Because of the fact that Benevolent develops personalized medical treatments and precision medicine they need a diverse amount of that from people with different race, ethnicity, gender and age, so that the newly developed drugs are effective on a wide range of patients. The ethical issue that emerges regarding this principle is: Are minorities, the vulnerable or people at the risk of exclusion misrepresented?

ALTAI web tool to answer the mapped questions and evaluation of the results

When using the ALTAI web tool to evaluate the area of:

- **Privacy and data governance** of trustworthy AI, the tool suggests that there are no recommendations for this requirement.

- **Transparency** of trustworthy AI, the tool suggests to consider the continuously surveying of users**,** asking them whether they understand the decision(s) of the AI system
- **Diversity, non-discrimination and fairness** of trustworthy AI, the tool suggests to ensure a mechanism that allows for flagging of issues related to bias, discrimination or poor performance of the AI system. It also suggests that it "should assess whether there could be groups who might be disproportionately affected by the outcomes of the system."

When it comes to privacy and data governance BenevolentAI seems to respect requirements suggested by the **[AIHLEG 2019]**. At the conference "Conversation on Intellectual Property and Artificial Intelligence" held in Geneva, Switzerland by the *World Intellectual Property Organization* the Benevolent team were asked questions about their system and its Intellectual Property, while data privacy also came into question. According **[Benevolent 2020c]**: " It would need to overcome licensing restrictions for data in-licensed from a 3rd party to the patent holder. It would also need to overcome privacy law restrictions for any personal data and it may need security etc. to maintain confidentiality of the data set until patent expiry. " This together with the ALTAI web tool results suggest the fact that the organization respects the laws and requirements of privacy and Data governance of trustworthy AI.

As mentioned before in the report there's a dilemma between accuracy and explainability. Although researchers find an effective drug for a patient, the mechanism through which the drug was advised is often just casually explained. The same thing applies for the new drugs being discovered. **[Benevolent 2020c]** states that "the explainability issue will be more problematic in certain areas than in others". This suggests the fact that in the field of medicine it is harder for the developers to explain the process of development to the end-users.
[Benevolent 2020c] also remarks the fact that "Requiring higher explainability may reduce accuracy", which could lead to a significant risk of harm to the end-user. This and the result of the evaluation being taken into consideration, the BenevolentAI team should try to explain the mechanism behind their AI system to the point that it doesn't reduce it's accuracy, while also surveying the customers about their thoughts on the process.

When it comes to Diversity, non-discrimination and fairness BenevolentAI is aware of the fact that it needs a lot of diverse data, which led to the "Data Diversity Initiative" **[Benevolent 2020d]** started in february of 2020 by their team. The point of these actions is to encourage the gathering of more diverse data from people with a different race,ethnicity, gender and age.
In an article, in which they present the initiative of data diversity **[Benevolent 2020e]** remarks the fact that "Diseases, treatments and drugs affect people differently according to factors such as race, ethnicity, gender or age, meaning these data gaps in research drastically limit our knowledge to effectively diagnose and treat disease." Analysing the ALTAI web tool results and the above statements, Benevolent is already trying to gather as much diversified data as possible, so that the drugs it develops and treatment it provides is effective on a wider range of people and that those who are vulnerable or at the risk of exclusion aren't left aside. As a conclusion although the medicine and treatments aren't perfect the organisation tried to be as diverse, non-discriminative and fair towards all end-users.

To conclude all the above presented ethical issues together with the approach of the organisation when assessing those, Benevolent is a company which respects the "ethical rules" and also the laws imposed by the state. It's work and development in the medical industry is

making the life of those who suffer from diseases better, while also sometimes being a life saver, which is of great importance for the future of humanity.

GRADE: 5 RVZ (Febr. 15, 2021)

FINAL GRADE: 10/10 -> 1.0 (DV)

References

[Adams 2020] Stephen Adams (2020, Nov 14). Arthritis drug 'cuts elderly Covid-19 deaths by two-thirds', say researchers - raising hopes that it will save the most vulnerable. Retrieved 25.11.2020 from https://www.dailymail.co.uk/news/article-8949605/Arthritis-drug-cuts-elderly-Covid-19-deaths-two-thirds-say-researchers.html

[AIHLEG 2019] Independent High-Level Export Group on Artificial Intelligence (2019), *Ethics guidelines for trustworthy AI. European Commission.* Retrieved 02.02.2021 from https://ai.bsa.org/wp-content/uploads/2019/09/AIHLEG_EthicsGuidelinesforTrustworthyAI-ENpdf.pdf

[ALTAI 2019] *The Assessment List for trustworthy Artificial Intelligence (ALTAI) for self assessment*, by Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission. (2019, Jun 26) Retrieved 25.11.2020 from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342

[Benevolent] BenevolentAI Official Website, https://www.benevolent.com/

[Benevolent 2019] Benevolent (2019, Mar 31). GuacaMol, a new framework to compare models for de novo molecular design. Retrieved from https://www.benevolent.com/news/guacamol-a-new-framework-to-compare-models-for-de-novo-molecular-design

[Benevolent 2020a] BenevolentAI |What We Do. Retrieved 02.12.2020 from https://www.benevolent.com/what-we-do

[Benevolent 2020b] The Diversity Analysis Tool: Towards Better Diversity In Data For Precision Medicine (2020, Sep 22). Retrieved 04.12.2020 from https://www.benevolent.com/news/the-diversity-analysis-tool-towards-better-diversity-in-data-for-precision-medicine

[Benevolent 2020c] BenevolentAI (2020), *Conversation on Intellectual Property and Artificial Intelligence.* Retrieved 02.02.2021 from https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/conversation_ip_ai/pdf/corp_benevolentai.pdf

[Benevolent 2020d] BenevolentAI (2020), *Data Diversity Initiative.* Retrieved 02.02.2021 from https://www.benevolent.com/data-diversity-initiative

[Benevolent 2020e] BenevolentAI (2020), *Data Diversity Initiative: Tackling The Medical Research Data Gap.* Retrieved 02.02.2021 from https://www.benevolent.com/news/data-diversity-initiative-tackling-the-medical-research-data-gap

[Benevolent 2020f] BenevolentAI (2020), *Working together to meet a 'moonshot moment' in drug discovery and development.* Retrieved 02.02.2021 from https://www.benevolent.com/news/working-together-to-meet-a-moonshot-moment-in-drug-discovery-and-development

[Benevolent 2020g] BenevolentAI (2020), *Designing better medicines using AI and big data.* Retrieved 02.02.2021 from https://www.benevolent.com/news/designing-better-medicines-using-ai-and-big-data

[Benevolent 2020h] BenevolentAI (2020), *AI in Healthcare: Towards a purposeful future.* Retrieved 02.02.2021 from https://www.benevolent.com/news/ai-in-healthcare-towards-a-purposeful-future

[Benevolent 2020i] BenevolentAI (2020), *Mind the gap: the diversity issue in medical research.* Retrieved 02.02.2021 from https://www.benevolent.com/news/mind-the-gap-the-diversity-issue-in-medical-research

[Benevolent 2020j] BenevolentAI (2020), *BenevolentAI Drug Programmes.* Retrieved 02.02.2021 from https://www.benevolent.com/benevolentai-drug-programmes
[Brothers 2015] Brothers, K. B., & Rothstein, M. A. (2015). Ethical, legal and social implications of incorporating personalized medicine into healthcare. *Personalized medicine*, 12(1), 43–51
[Crunchbase] BenevolentAI. Retrieved 25.11.2020 from https://www.crunchbase.com/organization/benevolent-ai
[Daley 2020] Sam Daley (2020, Jul 29). 32 Examples of AI in Healthcare That Will Make You Feel Better About the Future. Retrieved from https://builtin.com/artificial-intelligence/artificial-intelligence-healthcare
[GitHub 2020] Diversity Analysis Tool. Retrieved from https://github.com/BenevolentAI/benevolentai-dat
[Iriart 2019] Iriart, J.A.B. (2019). Precision medicine/personalized medicine: a critical analysis of movements in the transformation of biomedicine in the early 21st century. *Cadernos de Saúde Pública 35*, e00153118. Epub March 25
[Khramtsova 2019] Khramtsova, E.A., Davis, L.K. & Stranger, B.E. The role of sex in the genomics of human complex traits. *Nature Reviews Genetics* 20, 173–190 (2019)
[Miller 2014] Miller, F., Hayeems, R., Bytautas, J. *et al.* Testing personalized medicine: patient and physician expectations of next-generation genomic sequencing in late-stage cancer care. *European Journal of Human Genetics* 22, 391–395 (2014).
[Myszczynska et al. 2020] Myszczynska, M. A., Ojamies, P. N., Lacoste, A. M., Neil, D., Saffari, A., Mead, R., ... & Ferraiuolo, L. (2020). Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nature Reviews Neurology* 16, 440-456.
[Neil et. al. 2020] Neil, D., Briody, J., Lacoste, A., Sim, A., Creed, P., Saffari, A. (in press) Interpretable Graph Convolutional Neural Networks for Inference on Noisy Knowledge Graphs. Retrieved from https://arxiv.org/abs/1812.00279
[Oshisanya 2019] Oshisanya, P. (2019, Oct 6). Mind the gap: the diversity issue in medical research. Retrieved from https://www.benevolent.com/news/mind-the-gap-the-diversity-issue-in-medical-research
[Popejoy and Fullerton 2016] Popejoy, A. B., & Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature News* 538, 161.
[Sellwood 2018] Sellwood, M. A., Ahmed, M., Segler, M. H., & Brown, N. (2018). Artificial intelligence in drug discovery. *Future Medicinal Chemistry 10*(17),
[Watcharapichat 2020] Watcharapichat, P. (2020, Mar 8). How historically neglecting sex differences has impacted women's health [Blog post]. Retrieved from https://www.benevolent.com/news/how-historically-neglecting-sex-differences-has-impacted-womens-health