AI

Blindspots

healthcare

Knowledge Centre
Data & Society

vives
hogeschool

This AI Blindspots card set is inspired by [AI Blindspot](#) of Ania Calderon, Dan Taber, Hong Qu, and Jeff Wen, developed during the Berkman Klein Center and MIT Media Lab's 2019 Assembly program. It is available under a Creative Commons Attribution 4.0 International License.

The imec-SMIT-VUB team of the Knowledge Centre Data & Society adapted the original card set to a new version suitable to the Flemish context in order to support the development of trustworthy AI in Flanders, Belgium.

In a later phase, the Knowledge Centre Data & Society teamed up with [VIVES Hogeschool](#) to create a version of the AI Blindspots card set specifically for the healthcare context. VIVES Hogeschool worked on this AI Blindspots card set within the project 'AI in hospitals: ethical and juridical aspects', funded by ESF Flanders.

This adaptation is licensed under a [CC BY 4.0 License](#).

What are AI blindspots? Why is it important to detect them?

This card set is created to help you **avoid replicating societal biases, prejudices and structural disparities in your AI healthcare system**, which is imperative because of the sensitive character of the data and the vulnerable status of patients. Every person who is involved in the AI healthcare chain plays an important role in dealing with these blindspots, and thus this card set will help you as **doctors, nursing staff, managers*, IT staff and ethical committee members** to detect the blindspots proactively and provide you with strategies to cope with them.

By using the card set, you are able to **create more trustworthy and inclusive AI systems**. They allow you to tackle ethical issues, discuss these within your team, and communicate your actions in this regard with target groups and patients.

*The function profile 'managers' refers to various types of managing staff, operating at diverse levels and in different domains (e.g. legal, financial, ...) of the hospital.

What does the card set contain?

The card set consists of two card types: the AI Blindspots cards and the Ethical Dilemma cards.

The **AI Blindspots cards** are divided into two categories, i.e. **general blindspots** (rose cards; useful for all healthcare providers) and **specific blindspots** (blue cards; more relevant for a number of job profiles than for others). Each card presents one blindspot and contains a set of questions to help you uncover and discuss this blindspot. Next, a use case under the section 'How not to?' illustrates the importance of considering the blindspot, and lastly, a number of tools and tricks are listed to help you detect and mitigate the blindspot.

The **Ethical Dilemma cards** (green cards) are examples of healthcare situations in which an ethical issue arises, such as the need to access personal data or incompatible wishes of a patient.

The card set also includes 'Joker cards' that

allow you to include other AI Blindspots or ethical dilemmas, you and your team have identified and were not yet part of the card set.

The many ways in how you can make use of the card set

The card set **can be used in multiple ways**:

- A conversation starter for you and your health organisation to discuss the ethical, legal and societal aspects of AI technologies.

- A workshop with your team wherein you focus on one Ethical Dilemma card of your choosing and try to come up with one or several solutions for that particular dilemma. In a later stage, you look at the AI Blindspots cards and see if you can think of other solutions or make one of the already identified solutions more concrete with the help of the cards.

- A workshop with your team wherein you first divide the group according to job profile, and wherein the different groups look at the Blindspots cards specifically for their profile. You pick one Ethical Dilemma card and each group can think about how to solve the dilemma with the help of their AI Blindspots cards. Second,

you discuss these different perspectives together with the general Blindspots cards to come to an integrated solution.

- A workshop with your team wherein you identify one of your own ethical dilemmas and write it on the Ethical Dilemma joker card. Next, you can jointly think of one or several solutions with the help of the AI Blindspots cards.

- A workshop with your team wherein you create requirements for your (future) innovation using a reverse brainstorm. A reverse brainstorm does not aim to solve a problem but rather to make the problem worse. Use the AI Blindspots cards and ask yourself: "How can we ensure to make this blindspot a reality for our (future) innovation?". After the reversed brainstorm, reverse these unethical measures into requirements for your innovation.

- A double check for you and your team when you are designing, developing, implementing or thinking about using a healthcare AI system. Look at the AI

Blindspots cards and see if you are aware of and took into account the possible blindspots. If not, discuss how you and your team will deal with the blindspots.

SUITABILITY

An AI system may provide support for the care for a patient, but is it the best solution for your specific goal? By examining other means, you will get a better idea of the possible solutions, your preferences and which one is worth investing in.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Will the **quality of care** for the patient improve by using the AI system?

B. Will the development and implementation of an AI system result in the **greatest benefits** in comparison to other solutions?

C. Do these **benefits outweigh the risks and changes** for your organisation and for the patient?

D. Will the **system (not) interfere** with (1) the norms and standards of your organisation, (2) the services you offer, and (3) the intentions you have as a care provider?

HOW NOT TO

Deep learning is very en vogue these days, therefore the board wants to invest in a deep learning application. This way the hospital will receive strongly awaited attention and recognition.

TOOLS & TRICKS

EXPLAINABILITY & TRANSPARENCY

When using an AI system in a healthcare context, the system will likely collect health data, which are personal and sensitive data. Clear communication about the benefits of the system, the data collection process, and the decision making process are crucial to gain trust from the envisaged users or others who will encounter the system.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. How will you communicate **the benefits of the AI system** to the envisaged users and others who will encounter the AI system?

B. How will you explain the **data collection process** to them?

C. How will you inform them about **the decision-making process** (incl. the underlying logic)?

D. What **difficulties** can you encounter when communicating about the AI system to your target groups, and how will you cope with them?

HOW NOT TO

A sales representative cannot give details about the datasets used other than them being 'massive' and 'free of all sorts of biases'. The model they use is proprietary and completely black box. Some measures of uncertainty are reported along with the output, though.

TOOLS & TRICKS

TRUST

The implementation of an AI system might lead to resistance as not everyone is acquainted with (digital) technologies.

For some, the implementation of an AI system might feel exciting and new, while others might feel stress and distrust.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. How will you gain **trust from your envisaged users** and others who may encounter the AI system?

B. Have you thought of **a sensibilisation campaign** for affected stakeholders?

C. How will you guarantee to not only reach the 'AI believers' but also **persons who are not yet convinced of the advantages of AI**?

HOW NOT TO

Only after the AI system (making automatic reports from consultations) was deployed, the management cared to brief the staff and communicate to the patients about the system and its promises.

TOOLS & TRICKS
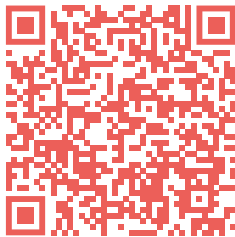
WILLINGNESS

The success of an AI application is partly dependent on the effort and willingness of the users (i.e. doctors, nursing staff, patients and other possible users) to learn, adopt and use the application for their (daily) needs and in their medical practice. The willingness to use an AI system goes hand in hand with trust in the AI system. When a person does not trust the system, why would he/she use it? (see also card on trust)

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Does the AI system address **a true need** according to the stakeholders?

B. Are the stakeholders **willing to adopt and use** the AI system or do they feel reluctant?

C. How will you **reduce the reluctance** of your stakeholders towards the AI system?

HOW NOT TO

It seems the new AI-based feature of the electronic health record software has barely been used by the staff. It turns out people don't like to change the way they work, as they tend to be less productive in the process of getting to know the new possibilities.

TOOLS & TRICKS

DIVERSITY

The persons (i.e. patients, nursing staff,...) confronted with the AI technology are a diverse group. They all have a different social and demographic background, and have a different level of digital maturity (see also card on (digital) inequality). This is not only the case on a personal level, but also on an organisational level. Between care institutions there are many differences in terms of digital maturity and experience with IT.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Can the AI system be **used and interpreted by a diverse set of possible users**?
B. Has your institution the **necessary experience and capabilities** to make use of an AI system in its daily practice?

HOW NOT TO

The company promised the AI system had endless possibilities. What they did not say is the user guide consists of 400 pages of technical writing full of equations and code.

TOOLS & TRICKS
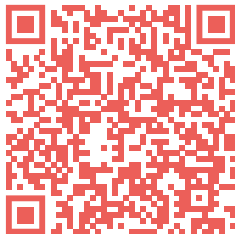
(DIGITAL) INEQUALITY

The implementation of an AI system might have a negative or unintended impact on the envisaged group of users, such as increasing existing inequalities.

Not all envisaged users will have the appropriate level of digital skills and digital health literacy to use and interpret the AI system.

Knowledge Centre
Data & Society

vives

(DIGITAL) INEQUALITY

HAVE YOU CONSIDERED...

A. Are there **alternatives** in place so persons who are not able to make use of or interpret the AI system are not left out?

B. Has it been ruled out that the implementation of the AI system will lead to more **stigmatisation and discrimination** of certain groups?

C. When implementing the AI system, will it enlarge or decrease **existing inequalities** because of...?
   i. The availability of patient information;
   ii. The differences in treatment between hospitals;
   iii. The level of competition between care institutions;
   iv. The effect of digitalisation on nursing staff; etc.

HOW NOT TO

The information desk is completely replaced by virtual assistants. People only have to scan the QR code they received by email after making their appointment online. From then on the assistant will talk them through everything they need to do.

TOOLS & TRICKS

TRADE-OFF

(Personal) data must be processed with respect for the privacy of the data subjects (see also card on data governance & privacy). But sometimes, a trade-off between personal privacy and public interest must be made: think for example of the contact tracing apps for the COVID-19 crisis. Even though the amount of personal data that needed to be shared was strictly limited, people still shared some personal data in order to trace corona infections. Furthermore, a weighing exercise sometimes needs to be made between the social interests/costs of implementing an AI system and the economic interests of commercial partners involved in this process.

GENERAL BLINDSPOT

HAVE YOU CONSIDERED...

A. What are **the benefits and the risks** of the AI system on **an individual** and on **a collective level**?

B. Does the AI system really **require personal data** in order to serve the public interest?

C. How long will the **personal data be stored**?

D. Are the **data subjects adequately informed** about the purposes of processing their personal data?

E. Which **interests** do the commercial partners have in developing/implementing this AI system? How do these **relate to the social interest** of providing good health care for all people?

HOW NOT TO

The pattern of a new disease is different for children with Down syndrome. An AI system can detect which of their characteristics strongly correlate with this divergent pattern. Their data can, however, only be pseudonymised and there is no time to ask for consent.

TOOLS & TRICKS
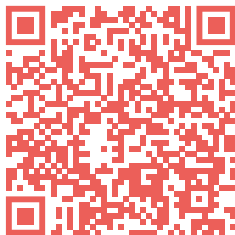
CONSENT

A patient in need of care may not be fully convinced of the advantages of the AI system that will be used during his/her care process. Be aware that a person (patient, healthcare staff, ...) might feel to have no other option than to give his/her consent in order to receive care or to continue his/her work. The power structure between patient and caregiver or between management and staff can play a role as well in the freedom of consent. And what with patients who are physically or mentally unable to give their consent?

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. What does **'giving consent'** entail?
   i. Is a patient giving his/her consent for the purpose of (1) optimising his/her diagnosis or treatment, and/or (2) further processing his/her personal data (e.g. training other AI systems)?
B. How will you **explain** to the affected stakeholders what 'giving consent' entails? Are they sufficiently **informed** about what they are consenting to?
C. Can stakeholders **withdraw their consent**?
D. Will you **review** the way(s) you are asking for consent?
E. When collecting data on other grounds than consent, the collection process may be **legal but not necessarily ethical**. Did you consider to not collect, store and analyse data, although you may be legally allowed to do so?

HOW NOT TO

As the performance of AI systems is in everyone's interest, the hospital assumes consent from the patients for the use of their data to train these AI systems.

TOOLS & TRICKS

DATAFICATION OF HEALTH DATA

Not all information can be quantified. One can for example think of mental health information or information regarding a person's perception of pain. This information is extremely interesting but harder to measure, to datafy and to standardise. When trying to datafy this type of information, a choice has to be made between the parameters that will/can be measured. Some parameters may be more easily measured than others because they are easier to datafy.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Are there **parameters** on which the AI system is based that are more difficult to measure and to datafy than others? If so, how will you deal with this?

B. Which **proxies** can you use to gather data about the parameters that cannot be identified? What are the limits/disadvantages of these proxies?

HOW NOT TO

In order to assess the risk of burn-out, a company's HR department decided to analyse facial micro-expressions when entering the building. The system did, however, not consider differences in facial structure. As a result, it arbitrarily estimated certain facial structures as a high risk of burn-out, while they were actually not.

TOOLS & TRICKS

ACCURACY & QUALITY

It is important that an AI system is fed with accurate and high-quality data to ensure that the results and outcomes can be interpreted in a correct and adequate way. This is especially the case when the system is for example used to improve care or well-being, such as examining which type of treatment will be the most appropriate for a patient.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Are the data on which the AI system is based **accurate** and **representative** for the current time, space and population?

B. Are the data **carefully entered/integrated**?

C. Is there **a test phase** to examine whether the AI system predicts the right symptoms/treatment?
   i. Are the appropriate performance estimates used for this test phase? Are these estimates unbiased?

D. Are there **control mechanisms** to monitor the analysis and the outcomes of the AI system?

E. Is there a **strategy to detect and counteract** users who give false information on purpose?

HOW NOT TO

An AI system analyses the course of pain complaints of patients with fibromyalgia. They must enter their level of pain into an app three times a day. These estimations, however, vary per patient, and patients often forget to submit the data.

TOOLS & TRICKS

IMPACT ON WORKFLOOR

When implementing an AI system, the impact of the system on the daily practices and work

of e.g. healthcare professionals must be assessed and measured.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. How will the AI system influence the **current workflow** and more generally **the workfloor**?
   i. Will it not create additional workload?
   ii. Will it not cause feelings of oppression (e.g. medical staff may feel controlled by the system as it collects metrics about patients and their treatments)?
   iii. Will it not affect and compromise the human aspect of providing care?
B. Are stakeholders able to give **feedback** with regard to the implementation of the AI system?

HOW NOT TO

When measuring blood pressure, an app automatically enters the data on the blood into the patient's electronic health record. The app has some bugs and, therefore, gives warnings of unidentifiable values. Health workers need to verify all the data and, if necessary, enter the correct values themselves.

TOOLS & TRICKS

UNDESIRABLE IMPACT

Which unintentional consequences and perverse effects could be caused by the AI system? One can for example think of (1) the stress and worries that users may feel because of false positives, and the unwarranted feeling of joy and reassurance due to false negatives, or (2) a health application that constantly reminds its users that they are not healthy.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. What is the **perceived negative impact** of the AI system on (the experiences of) patients?

B. What are **possible unintentional negative consequences** caused by the AI system (e.g. unnecessarily alarming patients because they are able to monitor their data)?

HOW NOT TO

An AI system that delineates the zones in the brain that need to be radiated in order to treat brain tumours is widely valued for its accuracy. Hospitals using the system are seen as pioneers in the treatment of that type of tumour. All this positive attention led to patients believing that nothing can go wrong, while the system cannot always delineate the zone and treat the tumour.

TOOLS & TRICKS

OWNERSHIP

Who should be qualified as the owner of health data? This question is the subject of a vivid discussion on whether patients can remain the owners of their data and decide who they grant access to the data.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Who is the **current owner** of a patients' health data from a legal point of view?

B. Is **ownership being affected** when part of these data is stored on a private or public health platform?

C. Is there a possibility, from a legal point of view, for **patients to become the owner** of their own data?

D. If so, what **actions** must be made to (1) guarantee secure ownership by patients and (2) inform patients about their role as owner and it's associated responsibilities?

HOW NOT TO

During a hospital's experiment heart patients were given a wearable that collected medical data. The patients themselves owned the data. The hospital also collected data about how patients interact with their data. These data were stored on and owned by a company's platform.

TOOLS & TRICKS

JOKER CARD

························· | ·························
························· | ·························
························· | ·························
_____ | _____
_____ | _____
························· | ·························
························· | ·························
························· | ·························

Knowledge Centre
Data & Society

vives

JOKER CARD

HAVE YOU CONSIDERED...

HOW NOT TO

TOOLS & TRICKS

BROAD DISCUSSION

The implementation of an AI system in your organisation will bring several opportunities but also risks or challenges. To minimise the risks and challenges for all, every relevant stakeholder (this includes patients) should be involved in the discussion regarding the purpose of the AI system and its functionalities. The more diverse these stakeholders are, the more you will be able to limit and counter these risks and challenges.

SPECIFIC BLINDSPOT: MANAGEMENT

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Are **all relevant stakeholders** involved in the discussion regarding (1) the purpose of the AI system, and (2) the necessary functionalities of the AI system?

HOW NOT TO

Because the hospital wants to invest in AI, an AI working group is assembled. The hospital director is proud that this working group reflects the different departments of the hospital. However, only managers of these different departments are involved in the working group.

TOOLS & TRICKS

EDUCATION & SENSIBILISATION

When implementing an AI system, the involved stakeholders may not be familiar with the technology. It is important to not only sensibilise the stakeholders about the goal, the benefits and the impact of the technology but also educate them and guide them in how the AI system can be used and/or interpreted.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Is there **a learning trajectory** for stakeholders/users on the topic of AI (**general information**)?
B. Are **stakeholders**/**users informed** about the goal, the benefits and the impact of the AI system?
C. Is there **a learning trajectory** for stakeholders/users that guides them in **using and interpreting an AI system**?

HOW NOT TO

An innovative AI system in radiology is able to optimize the quality of CT scans that were taken with a small radiation dose. However, it turns out that only those radiologists who are convinced of the benefits of innovations and AI are using the system. The other radiologists are not informed about the benefits of the system and how to work with it.

TOOLS & TRICKS

LARGE SELECTION

There are many AI systems on the market, many of which are not transparent about their gains and pains. This makes it harder (1) to examine which AI system will be the most effective for a specific situation and (2) to estimate which third parties (i.e. developer and supplier of the AI technology) are trustworthy.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Have you **compared** the different developers and suppliers of the AI system? If so, with whom did you have the **most comfortable feeling**?
B. Are there **specific risks** related to working with the developer/supplier you want to choose?
C. Did you estimate the **return on investment** working with the developer/supplier you want to choose?

HOW NOT TO

A hospital wants to invest in AI but has little expertise at hand. It decides to work with an American company that develops medical AI apps, but it turns out that the data collected in the hospital to feed the AI app is stored on the company's insufficiently secure platform.

TOOLS & TRICKS

CHANGING PURPOSE

The data controller, i.e. the person or company that determines the purpose and means of personal data processing, has the responsibility to set the correct boundaries for personal data processing. What if the processing of personal data was approved for a specific purpose, but this purpose has changed after a certain period of time?

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Do you regularly examine if the AI system is **still the best and most efficient solution** for the intended purpose?

B. If the purpose has changed, did you **reset the boundaries for personal data processing**?

HOW NOT TO

A hospital is using, with the consent of the patients, an AI system that can predict a patient's risk for several diseases, based on his/her medical data in the electronic health record. The system was trained by another company using a large dataset, but now the hospital wants to use its own electronic health record data to train the AI system. However, it does not inform the patients about this.

TOOLS & TRICKS

AVAILABILITY OF DATA

The data on which you want to build your AI system may not be available or easily accessible

to you, or may not be allowed to be shared with other actors.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Are the data you need for your AI system **available and accessible** to you?
B. Are the data you want already **digitised**?
C. If not, will the **effort and cost** to digitise the data outweigh the benefits of being able to access the data?

HOW NOT TO

Strong advances can be made in treating multiple sclerosis (MS) if an AI system could analyze the data from MS patients all over Europe. Unfortunately, the data is in many cases locked, unfindable or unreadable by the system...

TOOLS & TRICKS

CONTEXTUAL FACTORS

The implementation of an AI system is preceded by a great deal of research and development in order to maximise the functioning of the system. But what if the predetermined context in which the system is to function changes and no longer corresponds to what was analysed? As a result, the system might not work as accurately as estimated.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Do you regularly **compare the training and testing data** with the current situation?

B. Do the input data and predicted values align with the **expectations**?

C. If needed, do you have **a plan to remedy or to phase out** the use of the AI system?

HOW NOT TO

A smartphone app was developed which allows patients to test if they are infected with the coronavirus by coughing to their smartphone. The app was trained with a mass of recordings of coughs. The app is very functional and a lot of people use it. However, when the coronavirus mutates, the app is not retrained with new data, leading the app to miss a significant amount of positive cases.

TOOLS & TRICKS

STANDARDISATION

Standardisation of data is important when the data are used for statistical purposes. An ontological framework in which the characteristics of the data are defined is essential to ensure unity in the standardisation process. Standardisation will also help the sharing of data between stakeholders.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Is there an **ontological framework** to ensure the data are interpreted in an identical way?
B. Is it possible to **share the data** with other stakeholders/other parties? What must be done to make this possible?

⚠ HOW NOT TO

All Flemish hospitals want to team up to train an AI predicting tool on the basis of all electronic health records in Flanders. However, because the electronic health record has many different formats depending on the hospital and region, this plan is abandoned.

TOOLS & TRICKS

DATA MINIMISATION

According to the General Data Protection Regulation (i.e. GDPR) you are not allowed to collect more (personal) data than needed for the functioning of your AI system. More data also means more data analysis, and more costs and effort to process and analyse the data.

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. What **data** should be collected for the **proper functioning of the AI system**?

B. How **long** should the data be stored for the **proper functioning of the AI system**?

C. Have data been collected that are **not strictly necessary for the functioning of the system**?

D. Can you comply with the **data subject's rights of the GDPR**?

HOW NOT TO

The hospital has put up an experiment for a group of patients with a specific disease and uses an app developed by a company that gathers data about a standard set of parameters. This set, unfortunately, includes data that are unnecessary for analysing the disease type of the group of patients.

TOOLS & TRICKS

SECURITY

Because of the sensitivity of healthcare data an adequate security policy for the collection, storage and sharing of data with other parties is required.

Knowledge Centre
Data & Society

vives

💡 HAVE YOU CONSIDERED...

A. Do you have a **data security policy**?
B. Did you assess **who** can access the data and who cannot, and more importantly **why** actors can or cannot access the data?
C. Do you **register** who has accessed the data, at what time and for what purpose?
D. Is there a procedure to **report violations**?
E. Did you **test** if the AI system **cannot be hacked**? Do you test the AI system periodically?

⚠️ HOW NOT TO

Two doctors are experimenting with a new AI system and store the medical images which the AI system analyses on a server that is not perfectly secure. By doing so, they potentially expose the diagnoses and the individuals to whom they are related to hackers.

✎ TOOLS & TRICKS

HUMAN VS MACHINE

An AI system can facilitate the human aspect of care, by giving care professionals more time to interact with patients. Yet the suggestions made by the AI system about the right diagnosis or treatments can deviate from the personal intuition, 'gut feeling', and evaluation of the care provider. It can be difficult for care professionals to deal with this balance between the human assessment and the machine.

SPECIFIC BLINDSPOT: DOCTORS, NURSING STAFF & MANAGEMENT

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. Are the results of the AI system **reviewed** by a human professional?

B. Is there a **policy** on how the assessment of health professionals relates to the one of the AI system?

C. Can health professionals **freely report** their concerns with regard to the decisions of an AI system?

D. Who is **responsible** for (1) the collection, the storage, the sharing, and the analysis of the data, and (2) for choosing the right treatment and monitoring the patient?

E. Who has the **final responsibility**?

HOW NOT TO

An AI system analyses a patient's tumour and makes a suggestion about the zone that needs radiation. One day, a radiologist overlooked a parameter that was not submitted to the AI system. This led to a wrong sugesstion and, eventually, the wrong zone was being radiated.

TOOLS & TRICKS

DATA GOVERNANCE & PRIVACY

Privacy with regard to the data subjects is crucial when making use of health data. However, privacy is not a clear-cut concept. Data can, for example, be made anonymous by removing the names of the patients. Yet, because of other parameters in the data set (which are often particularly interesting to feed AI models), it is sometimes still possible to identify patients. Furthermore, the difference between anonymisation and pseudonymisation is not always clear. And is it righteous to use the data you have collected?

SPECIFIC BLINDSPOT: ETHICAL COMMITTEE & MANAGEMENT

Knowledge Centre
Data & Society

vives

HAVE YOU CONSIDERED...

A. What measurements did you take to **protect the data subjects** and their (personal) data?

B. Can you ensure the **anonymity** of the data subjects? If not, are they informed about this?

C. Have you considered how **proxy-data categories** (e.g. shoe size for gender) can also identify personal information?

D. Are you complying with the **General Data Protection Regulation** (i.e. GDPR) and other regulations?

E. Is it still **righteous** to use the data?

HOW NOT TO

A hospital wants to use its electronic health records to train an AI system. They replaced the patient names with a code so they are able to retrace patients that need to be informed about the risk of having a hearth attack. The hospital didn't ask patients for consent.

TOOLS & TRICKS

JOKER CARD

.... | ....
.... | ....
.... | ....
.... | ....
.... | ....
.... | ....
.... | ....
.... | ....

SPECIFIC BLINDSPOT: ....

JOKER CARD

HAVE YOU CONSIDERED...

....
....
....
....
....
....
....
....
....
....

HOW NOT TO

....
....
....
....
....
....
....
....
....
....
....

TOOLS & TRICKS

....
....
....
....
....
....
....
....
....
....
....
....

A hospital wants to put itself "in the picture" by investing in a new **Electronic Health Record system** that uses AI to make predictions about patients' health, based on the data saved in the EHR system. The data is **pseudonymized** before it is inputted in the AI system, yet a lot of other personal and identifying data is saved, such as address, gender, age,... The goal is to track and thus prevent different disease patterns. However, doctors signal that they do not know how they will have to deal with the predictions from the AI system because they already have not enough time to see all their patients and provide them with the care they need.

Knowledge Centre Data & Society

vives

A patient has a very rare type of throat cancer. The radiologist would like to use an AI system that can help to **analyse the complicated form of the tumour** and can make a suggestion about how to radiate this tumour. However, the patient is afraid of a computer "deciding about his life" and he declines the proposition of the radiologist to use the AI system.

Knowledge Centre
Data & Society

vives

Hospitals use **wearables** to collect data about patients' condition. This results in a vast amount of data, as data is collected about multiple parameters at different moments of the day. Before this data is sent to the hospital, it is stored and structured on a cloud from a data company. When the AI systems of the hospital have analysed these data, the results are again stored on the company's cloud, and relevant results are sent to the apps of the patients. The hospital tries to inform its patients about this **data storage and processing**, but they notice that patients don't understand what happens with their data, yet they do give their consent to process the data when they are asked.

Knowledge Centre Data & Society

vives

A new **wearable** was developed that monitors the heart rate, blood pressure, and the presence of a certain protein in the blood of the patient. The data collected by this wearable is sent in **real-time** to an AI system in the hospital which can give a warning when a patient is likely to have a heart attack. This AI application was specifically designed for older patients who have more risk of having a specific type of heart attack. However, these older patients often do not understand how to work with this wearable; they do not put on the sensors in the right way or panic when the application gives fault notifications and thus they decide to put the wearable away.

Knowledge Centre
Data & Society

vives

A young father of two kids finds out he has a brain tumour. He is treated by a neurologist who is experimenting with an **AI model that screens brain tumours** and suggests to either radiate or operate the patient. For this patient, the AI model suggests to operate. The patient is very afraid of having an operation and has a bad feeling about this. However, because the neurologist is so keen on using the AI model he doesn't dare to say anything about this.

A hospital struggles with **keeping the data of its patients safe** and a **few hacks** occurred. Therefore, the hospital collects data about the actions of the hospital personnel on the hospital's computers. An AI system is used to identify big deviations of the type of data transfers that are normal for a hospital's activity, which can show that a data leak has occurred. However, the hospital's personnel has the feeling that they are being controlled by the AI system and the hospital's management and they decided to protest against it.

Knowledge Centre
Data & Society

vives

A hospital is struggling because they have few nurses but a lot of patients who need to be looked after. It installs a camera in the rooms of the patients that records the **pain expressions on the patients' faces**. An AI application analyses the expressions and gives them a score to help prioritize the nurses' work. It turns out that the nurses can more efficiently use their time and they have the feeling that at the end of the day they have helped all patients in a better way than when they react to the calls of patients themselves. Patients on the other hand sometimes have the feeling of not being heard.

ETHICAL DILEMMA

Knowledge Centre
Data & Society

vives

JOKER CARD

....
....
....
....
....
....
....
....
....
....

Knowledge Centre
Data & Society

vives