APRIL 2023

Towards Ethical AI Homelessness Tools

*A report produced for LIFE*SPIN by Western University Philosophy students*

PHIL4901G // Honours Capstone Final Project

NATHALIE DIBERARDINO
FIRAS KAYSSI
LUCAS KOPP
MARTIN WHITTAKER

Table of Contents

Introduction

Artificial Intelligence (AI) is an umbrella term used to describe digital systems that can perform automated ‘intelligent’ tasks (Smuha, 2021). These technologies are increasingly being developed and deployed in our daily lives. They do everything from generating our social media feeds to matching us to our Uber rides, and even sorting through online job applications. They’re also increasingly being used in the public sector, and by public institutions and service agencies, to help them predict and manage the needs of the communities they serve.

One use of AI that is gaining particular attention is that of homelessness management and prediction. AI is being used to collect and categorize data to predict individual risk for homelessness, and to match and triage individuals with appropriate need-based housing support (Eubanks, 2018). These systems are becoming increasingly popular, with one model being used in over 1000 communities across the US, Canada, and Australia (Kithulgoda et al., 2022).

AI technology is what’s behind London’s Chronic Homelessness Artificial Intelligence (CHAI) tool. CHAI’s developers use a public dataset from the Canadian homelessness management information system, and uses AI and machine learning technology to generate predictions about the risk of individuals becoming chronically homeless. The tool makes predictions based on factors like monthly income, age, medical diagnoses, and number of public shelter visits (VanBerlo et al., 2020). Systems like CHAI are designed to identify individuals at risk of homelessness, which, according to London's manager of homeless prevention Jonathan Rivard, enables the public and private sectors to “provide [those at risk] with more support, and possibly reduce strain on the shelter system” (Lamberink, 2020).

However, the uptake of systems like CHAI comes with serious ethical questions. This report focuses on four such questions: (1) Are the tools informed by an understanding of homelessness that is inclusive enough? (2) Might they be biased against homeless people who are especially marginalized because of factors such as their gender or race—in other words, might the tools exhibit what’s called “algorithmic bias”? (3) Are the tools transparent enough to make their decisions understandable to all stakeholders? (4) How ought public and private institutions handle the collection and storage of homeless clients’ data/personal information?

The report proceeds as follows: *section 1* addresses question 1 above about **understanding homelessness**; *section 2* addresses question 2 about **algorithmic bias**; *section 3* addresses question 3 about **explainable AI**, and *section 4* addresses question 4 about **data privacy** and AI tools.

The goal of this report is to provide a conceptual background and analysis of the current problems around these four topics, which we see as central to the debate around AI's use in the public sector. We focus our examination on London's (CHAI) tool to better understand how these technologies interact with these issues and their potential for causing harm, but our analysis applies more broadly to any homelessness-management tools using machine learning AI technology. This report aims to raise awareness about the issues surrounding these tools amid their increasing popularity, and to serve as an important resource for community service organizations whose work might be impacted by the use of AI in the public sector. Ultimately, this report hopes at least to function as a source of caution for governments and regulatory officials considering using AI tools that threaten to hold such a stake in people's lives.

Defining Homelessness

Firas Kayssi

4th Year Philosophy & Economics Student

Are AI Homelessness tools informed by an understanding of homelessness that is inclusive enough?

PHIL4901G

HONOURS CAPSTONE FINAL PROJECT

1.1 The Importance of a Definition

Whenever a government, agency, or organization confronts the problem of homelessness, it is important that they first properly define what they mean by 'homelessness'. Why? Because the definition they use is a description of who gets access to help, and who does not. Therefore, choosing a bad definition might exclude some people who really do need help as a result of homelessness, leading to an unjust and inefficient distribution of resources. The CHAI program was an example of choosing a bad definition for homelessness. Thus, in an effort to better understand the problem, we suggest that a better definition can and should be adopted.

In a paper that documented the development of London's CHAI tool, they provided their own working definition of homelessness: "London's Homeless Prevention division identifies an individual as chronically homeless if they have spent 6 or more months (180 days or more) of the last year in a shelter. This definition was adopted from that of chronic homelessness outlined by the federal Canadian government's homelessness strategy directives" (VanBerlo et al., 2020). Although this description defines 'chronic' homelessness, we are concerned with homelessness, in general. The goal of this section is to show how this definition is problematic, and to recommend next steps in light of this.

Put simply, this definition of homelessness used by CHAI is insufficient, in the sense that it does not properly represent the population of people who are homeless. In their paper, "What A Home Does", David Jenkins and Kimberlee Brownlee distinguish between four standards of living. Within this theory, they provide a distinct and robust definition of the term 'home'. If the government can endorse this conception of a 'home' as defined by Jenkins and Brownlee, instead of the definition they currently use, they can then better identify who to consider 'homeless' (Brownlee & Jenkins, 2022). I will describe these four classifications in this section, and we will come to see that the government has only been concerned with one of the four standards of living.

1.2 A Good Way to Properly Define Homelessness

Brownlee and Jenkins start by distinguishing between two narrow notions of homelessness: *temporary shelter* and *persistent shelter*. The first notion refers to people who have unreliable access to a roof over their heads and a space to fulfill their "primal needs," such as sleeping and using a bathroom throughout the night. The second notion refers to people with reliable access to the same (Brownlee & Jenkins, 2022).

Their third classification takes a step up from these first two notions of shelter by distinguishing shelter from the notion of *housing*. A *house* has two important characteristics. Firstly, it is a place where our 'primal needs' are met, which makes it similar to shelter. Secondly, it is a place over which we have control (Brownlee & Jenkins, 2022). The idea of control can be understood more easily in terms of 'permission.' If we need permission to occupy a space, then we cannot say it is our house. This requisite control comes from legal ownership or a rental contract. Thus, being housed, on this understanding, involves having property rights.

Finally, the highest order of classification is the *home*. According to Jenkins and Brownlee, a home has more to do with how we feel and how we are treated by others in a space than what the space looks like, or what is in it. In their own words, a home is ‘‘a place of intimate belonging, in which our deepest social needs are met [and it is a] social space in which we are welcome, respected, and accepted” (Brownlee & Jenkins, 2022). If individuals were asked whether they have a space that they can call a home, according to this definition, it would be clear to answer. Further, this conception is much richer because of its ability to convey the feeling of security - something that the government’s current definition does not seem to do well.

It is important to note that, according to Jenkins and Brownlee’s theory, one can (for example) be sheltered and houseless at the same time - or inhabit some other combination of these situations. For example, let us consider a living situation where multiple families live in one housing unit. Only one family might have the right to live in this unit and, therefore, the other family does not have property rights over it. Despite having a space to sleep, having no permission to control that space means they are houseless according to Jenkins & Brownlee’s theory. This shows us how homelessness is more than just what we can see physically.

Therefore, Jenkins and Brownlee’s theory suggests that homelessness is the lack of having a place of intimate belonging, where one feels welcomed, respected, and has their social needs met. In comparison to how the government currently classifies homelessness, Jenkins and Brownlee’s theory also provides the terms ‘shelter’ and ‘house’ to better identify the living standard of an individual. Having more precise terms allows an entity like the government to distinguish between the living situation of different people. This is important, because a poor and unclear definition would end up leaving some people out of the picture, and helping others who might not need it.

Now that we understand Jenkins and Brownlee’s theory, it is clear that London has only been concerned with people who are considered to be temporarily sheltered. Further, it is important to recognize that the government does not appear to be concerned with homeless people, nor the houseless, nor the persistently sheltered. This is not a problem in-and-of-itself, but it can be if they decide to help only the temporarily sheltered and fail to identify them, or even end up identifying the wrong people.

1.3 Next Steps

Moving forward, we suggest that the City of London (as well as any other governments, agencies, and organizations that want to tackle the problem of homelessness) incorporate the terms defined by Jenkins and Brownlee into their project. This change will improve a few things. First, it will give them a better arsenal of language to use for future policy making, in the sense that they can better distinguish between one group of people and another. Second, and most relevant for this paper, is the point that governments can better inform their AI technologies with these more precise terms. As we will see in the next section, AI tools are biased, and this bias is informed by how we construct them. Incorporating Jenkins and Brownlee’s theory will essentially minimize the possible damages thereof.

Further, Jenkins and Brownlee’s definition of ‘home’ and ‘homelessness’ could help destigmatize the problem itself, since this definition allows us to extend our considerations of the problem to more than just visible housing insecurity, like sleeping on the street.

In light of accepting this definition, London might decide that their true focus is elsewhere. For instance, London might recognize that their interest is instead about the temporarily sheltered. Either way, becoming aware of Jenkins & Browlee’s theory of homelessness will help them clarify their goal. The important part, and what we hope to show is, that the definition of a targeted demographic matters, and that there are real consequences in doing this improperly - especially when we are in position to inform AI technology about it.

Algorithmic Bias

Nathalie DiBerardino

4th Year Philosophy Student

Could AI tools biased against homeless people who are especially marginalized because of factors such as their gender or race?

PHIL4901G

HONOURS CAPSTONE FINAL PROJECT

2.1 Understanding AI

“Algorithmic bias” is an issue that has been the subject of much discussion, both within the media and in academic literature. To understand the problem of algorithmic bias, and how it can impact AI homelessness tools, we first need to get into a little more detail about what AI actually is.

AI tools are made possible through algorithms: abstract sets of steps of tasks for computers to follow. One especially powerful type of algorithm is one that has machine learning (ML) capacities: those which replicate and mimic human thinking, decision making, and problem solving, as they build on their own prior decision-making (Noble 2019). To ‘teach’ an algorithm to make decisions, we input ‘training data’ into them that helps them sort through data in a certain way. Over time, machine learning algorithms become more efficient and accurate at doing this, and can eventually come to generate new predictions and decisions about a given set of data on their own. But, like any other AI tool, CHAI is vulnerable to *bias*.

2.2 Algorithmic Bias

AI or algorithmic bias happens when algorithms inherit social patterns that are reflected in their training data—patterns that developers did not intend to include (Johnson, 2020).

Given that we live in a society stratified by race, gender, class, and the like, these social patterns often take the form of discrimination and prejudice towards certain groups. For example, research has uncovered that AI is often biased against women and people of colour (Birhane, 2017). Importantly, algorithmic bias can have serious consequences. Take, for instance, the COMPAS recidivism tool, which was used in the United States to predict the risk of criminal defendants re-offending. The scores generated by the algorithm were used for parole assessments and even given to judges at the time of sentencing (Angwin et al., 2017). However, COMPAS was discovered to be biased against Black defendants: it incorrectly flagged them as future criminals almost twice as much as white defendants (Angwin et al., 2017). This bias then interfered with the freedom and future of people’s lives as it influenced court decision-making.

An AI tool dealing with homelessness would be no different. In predicting individual risk for homelessness for the purpose of matching individuals with appropriate and needed housing support, these tools have the power to supply and withhold access to the fundamental goods of shelter, privacy, and ideally a home. There is thus an urgent need to investigate the possibility of bias in these systems; we need to make sure that resources are allocated to the right people, and for the right reasons.

2.3 Bias in homelessness AI tools

There is evidence that homelessness AI tools are biased against women, and especially women of colour. One reason for this might be that when female victims of domestic abuse—a leading cause of homelessness in women—seek public support and housing, they are often categorized and counted as “women who are victims of domestic violence”, rather than *as* “homeless” (Bretherton, 2017, p.3).

There is also evidence that women, and particularly Black women, under-report their own experiences of criteria for homelessness services in self-assessment surveys, especially on questions on stigmatized topics like mental health history (Kithulgoda et al., 2022). This can prevent them from being triaged and matched with the support that they, in reality, did qualify for.

But perhaps the most serious evidence of AI bias has to do with homelessness *definitions*. As we learned in the previous section, CHAI developers follow the City of London in understanding chronic homelessness as over 180 shelter visits over 365 days (VanBerlo et al., 2020). However, evidence suggests that homeless women are less likely to live on streets and utilize public shelter systems, and instead tend to make private arrangements to couch surf, or temporarily reside with friends or acquaintances (Bretherton, 2017; Oudshoorn et al., 2021). If our working definition of homelessness in AI homelessness tools is one that only considers public shelter stays, and homeless women tend *not* to stay in public shelters, then women will be underrepresented or even absent in the datasets these tools rely on.

The result is that when these tools are used to algorithmically match and triage individuals who qualify as 'chronically homeless' with housing support, women can be *systematically excluded from these opportunities*. This comes with enormous consequences for these women's wellbeing and livelihoods. If our social biases creep into algorithmic tools being used to allocate fundamental resources—thanks to narrow definitions that fail to capture the reality of everyone living with housing insecurity —then the use of AI homelessness tools poses a serious risk of harm.

2.4 Next Steps

How can we contend with this problem of bias in homelessness AI tools? One obvious solution might be to aim to eliminate any and all algorithmic bias in these tools. Indeed, the goal of 'de-biasing' AI has been much discussed in computer science and in corporations adopting AI tools. But a major contribution from the philosophy literature to this problem is that we can't just 'de-bias' AI.

Philosophy professor Gabbrielle Johnson argues that algorithms can never be value-free, or completely objective and free from bias (2022). She argues that the design of algorithms themselves necessarily comes with choices about which values to include in those algorithms, such as accuracy or consistency (Kuhn, 1962). But values are often chosen on the grounds of social, ethical, and political considerations. For example, a corporation might prioritize the values in their AI tool that further their goal of increased political power. This can thereby inject bias into the tool as it works via those chosen values—at the exclusion of others.

Bias can even be imbued into values themselves. Johnson uses the example of clinical testing of the sleep aid Ambien, which prioritized the value of simplicity and thus tested the drug only on a homogenous group of participants: one entirely of men. Why? Because men have historically been seen as the 'typical' research subject. The value of simplicity imbuing that social pattern later had disastrous consequences for women taking the recommended dosage, since it was based on metabolisms significantly different from their own (Johnson, 2022, p.13). And since we've learned that AI tools necessarily come with decisions about values, they are vulnerable to this issue of chosen values being bias-laden.

Since algorithms can never be value free, they can never be ‘de-biased’. But this isn’t necessarily bad news; algorithms can, and should, include *good values* that influence the decisions we make. If algorithms can’t be value free, then **the task that algorithmic developers must turn to now is making active decisions about *which values* ought to be included and prioritized in their AI tools.** In the case of CHAI, developers need to carefully consider how to design their algorithms to minimize anyone’s exclusion from access to housing support. This might look like prioritizing inclusivity or equity. Or they might work with a framework in their system design that appreciates how individuals can be oppressed or disadvantaged in multiple ways (that is “intersectional”), and that works to recognize—and not exclude—the unique needs people have in virtue of their social identities (Crenshaw, 1989).

It’s wrong to assume that algorithms can free us from these kinds of judgments. When this much is at stake, philosophers stress that we need to take on the role of actively stopping the perpetuation of oppression and discrimination in AI tools that we’re falsely told can be objective or unbiased. Put another way: if value commitments in algorithms are inevitable, then we need to choose the values that will best serve the ends of justice as AI in the public sector becomes an increasingly popular strategy for making decisions.

Explainable AI (XAI)

Lucas Kopp

4th Year Philosophy & Computer Science Student

Are AI Homelessness tools transparent enough to make their decisions understandable to all stakeholders?

PHIL4901G

HONOURS CAPSTONE FINAL PROJECT

3.1 What is XAI and why might we want it?

One of the problems with advanced AI and machine learning (ML) systems is that their operations are “opaque”, or inaccessible, to humans. These systems, such as CHAI, are often called “black boxes” because it is as though the internal workings of the AI system are hidden within a locked, black box. One solution to the opacity problem that has gained popularity in the past few years is the idea of explainable AI (XAI) (Zerilli, 2022), which is an algorithm that approximates, or simplifies, how an AI generated its output, in order to make it understandable to a human.

A primary goal of XAI is to give enough insight into how the AI “thinks” such that a human is able to decide on their own whether or not they can, or should, rely on the AI system (Turek, 2017). In other words, XAI translates how an AI system processes information into clear explanations, so that we have insights into why the input caused a particular output. This is relevant to tools like CHAI, because it is important that users can understand and rely on the decisions that will impact peoples’ lives, especially when the decision-maker is a non-human black box.

3.2 Problems with XAI

Although there is a general consensus about the need for XAI among philosophers, lawyers, machine learning specialists, and regulators, there is very little consensus on what these explanations need to contain (Lipton, 2016). Brent Mittelstadt, Sandra Wachter, and Chris Russell (2019) of the Alan Turing Institute note that the most popular XAI algorithms offer explanations that resemble scientific models, a mathematical or numerical representation of a system, rather than explanations and, as a result, what they provide is only comprehensible, and reliable to technical experts or specialists. Moreover, these scientific models might show *that* certain factors were relevant to the decision, but it does not say *why* the factors were important, which is vital for a good explanation.

Another issue for XAI is that advanced AI decision-makers find incredibly complex relationships between millions of data points that it uses to generate its decision. Peter Lipton, one of the leading philosophers of explanations, notes that, behind every event, is a long list of causes; however, not all of these causes are *explanatory*: that is, vital to understanding why an event occurred (2001). For example, if you are in a car accident, the fact that you were driving in the first place is one of the causes of the accident, but this cause does not explain *how* or *why* the accident happened. To connect this back to XAI, since millions of connections are being made, the factors that are truly explanatory are not always clear. Mittelstadt and colleagues (2019) acknowledge that this is one of the problems with XAI: it is difficult to select which causes actually explain the decision. If XAI needed to explain the car accident example, the explanation might say that the most important cause of the accident was that you were driving, and although it is not wrong - the accident would not have occurred if you were not driving; this does not really capture *why* the accident occurred.

This point, determining which causes you use to explain a decision, is important because the way information is presented in an explanation can alter or manipulate which factors the recipient of the explanation (explainee) perceives as important. For example, if you explain your car accident to your friend, and tell them you were due to get your brakes serviced, that may cause your friend to think the accident was caused by faulty brakes, when in reality you were on your phone and ran a red light, which caused the accident. This is important in XAI, because developers can alter which factors (causes) their XAI algorithm can use in its explanation, so there is a worry that explainers can present explanations that will discourage "explainees from critically questioning or contesting a decision" (Mittelstadt et al, 2019).

3.3 Explanations

Peter Lipton argues that the purpose of an explanation is to convey the appropriate causes of an event, such that the explainee understands *why* the event took place (2001). In the case of AI decision-making algorithms, a good explanation will explain which factors caused the AI system to make its particular decision, for example, the primary facts that contributed to why a certain individual was selected to be matched with housing support. For Lipton, a causal, *contrastive* explanation is, in general, the best form of explanation for doing this (2001). Mittelstadt, Wachter, and Russell agree that a contrastive explanation is the most effective way for an explainer to ensure an explainee understands why an event took place. A contrastive explanation explains by informing the explainee why another, different, result did not occur (Mittelstadt et al, 2019). These are valuable because they inform the user, simply and effectively, why the AI system did not give them an expected result, along with the smallest change in the input that would have to occur for the system to have arrived at a different conclusion (Wachter et al, 2017). For example, suppose a bank uses an AI system to approve mortgage loans. If someone does not receive an approval and they demand an explanation of the system, a contrastive explanation would be of the form:

> <Applicant> would have been approved for a loan if their annual salary was $80,000 per year, however <Applicant>'s application was rejected because their annual salary is $65,000.

This example shows both of the aforementioned benefits: the loan failed because the applicant does not make enough money, and it also vitally informs the applicant that the easiest way to be approved for a mortgage in the future is to increase their salary.

3.4 Solutions

Contrastive explanations are needed in XAI to explain to non-experts because, regardless of competency with AI, they offer insight into how the decision was made, how the system values different data points, and gives the explainee important information regarding how to modify their behaviour in order to receive a different decision in the future (Wachter et al, 2017). This idea is highlighted by Virginia Eubanks, in her book *Automating Inequality,* in which she discusses a man named "Uncle" Gary Boatwright in Los Angeles, who is homeless, and his experiences with VI-SPDAT (an AI system that stands for "Vulnerability Index – Service Prioritization Decision Assistance Tool") that was used to determine his eligibility for housing aid. He filled out the form three separate times with three different agencies and failed to receive aid each time.

He remains unsure why he was ineligible, whether it is his psychiatric record or his police record or some other factor. If he had received a contrastive explanation, he would understand why he did not receive aid.

The other problem: how to determine which causes are explanatory and relevant. One way to help overcome this, presented by Mittelstadt et al (2019), is to make the explanations *communicative*, a solution supported by Wolter Pieters in his work on building trust in AI systems through explanations (Pieters, 2011). This idea is supported by the way humans often explain to one another. When we do not understand why a friend did something, we often ask them questions until they have explained their action to the point that we understand their motivation. More generally, an explanation is communicative when the explainer and explainee can engage in a dialogue, so the explainee can raise some of their concerns and the explainer can justify their decision, through a series of explanations, that address the explainee's concerns.

3.5 Analysis of CHAI as an XAI

The Chronic Homelessness Artificial Intelligence (CHAI) tool developed by the City of London implements a neural network (which is a black box) to predict whether or not someone is likely to become chronically homeless in the next six months. Therefore, to make CHAI more transparent, the developers use an XAI algorithm to generate an explanation of any decision produced by CHAI. The particular XAI tool used by the developers is called LIME (Local Interpretable Model-Agnostic Explanations). Essentially, a neural network is too complex to explain as a whole, so LIME locates a small subsection of the neural network that *can* be explained. The idea is that you can get a good sense of how CHAI values various factors based on the small subsection of the network. From this smaller, simpler network, LIME can generate an explanation of CHAI.

The team who developed CHAI did an admirable job tuning LIME to produce the best explanation they possibly could. In testing, they used the explanations to remove any unintended biases (see section 2) and they also removed any features that proved to be uninfluential. Most importantly, they collaborated with "domain experts at Homeless Prevention" to ensure CHAI did not fixate "on bizarre or unrealistic correlations" (VanBerlo et al, 2020: 9).

There are two primary problems with LIME as a tool for XAI. First, because it bases its explanation on a small subsection of the entire algorithm, LIME sometimes gives too much, or too little weight to certain factors because the subsection that was used to generate the explanation did not include information that was reflective of the entire neural network. Consequently, the explanation may not *fully* reflect how CHAI made its decision. The other problem is stability (or consistency). Because it looks at a subsection, not the entire algorithm, it sometimes generates two different explanations based on the same input data. Although LIME is *typically* accurate, and the team that developed CHAI did a good job at minimizing these concerns, both of these problems weaken our ability to rely on CHAI's explanations.

In conclusion, XAI has to explain to many different stakeholders, and it is important that the explanations cater to all of these stakeholders, but especially those who are most directly impacted by the AI decision. Based on the philosophy literature, contrastive and communicative explanations are the best explanations for non-expert stakeholders, such as homeless people, because it will best allow them to understand the decision.

As it stands, even though CHAI offers an explanation, it does not produce the kind of explanation that is necessary for non-expert stakeholders to actually understand CHAI's decisions. LIME provides a good enough explanation for the needs of expert stakeholders; however, LIME simply does not explain to non-experts in a way that conveys genuine understanding of the decision. Looking forward, there are reasons to be optimistic, because AI researchers have begun to recognize some of the problems with XAI, which is the first step towards addressing them. Importantly, if governments want the public to be able to rely on their AI tools, they should support XAI research that will help create explanations that are both communicative and contrastive.

Data Privacy

Martin Whittaker

4th Year Philosophy Student

Is the private data/personal information of homeless clients being collected and handled responsibly?

PHIL4901G

HONOURS CAPSTONE FINAL PROJECT

4.1 Data Privacy and its Importance

The responsible collection and handling of personal data has proven to be a significant concern. Data privacy and the ethics of data collection is serious enough to warrant serious discussion on the topic (Véliz, 2020). Many people who have expressed that they care about privacy reported negative experiences related to breaches in their personal data, and hold a degree of mistrust or skepticism towards institutions in the handling of their data (Brooke & Véliz, 2019). The rise of artificial intelligence tools will likely compound existing privacy concerns. In order to “train” AI tools to perform their tasks, they require large datasets to work with, which can contain substantial amounts of personal data.

Data privacy is a unique and relatively recent kind of privacy in the wake of the information age and the phenomenon of “big data.” The concern over data privacy has roots in common intuitions about privacy, such as its connection to autonomy, identity, and the “right to be let alone.” If an individual generates enough personal data that can be traced back to them, then that data can potentially confer a level of power or influence over that individual to whoever comes into possession of it; the greater the amount of information, the greater the power. For instance, an institution may have enough information about you to make inferences about your character, aptitudes, circumstances, etc., which they can use to make decisions that will affect you. A credit bureau can determine how credit-worthy you are, which can have all sorts of cascading effects (Pasquale, 2016); an insurance company can know how risky you might be to insure (NAIC, 2022; Blake, 2022); and so on.

When thinking about the concept of privacy, it will be beneficial to have a pragmatic, bottom-up conception of privacy, rather than trying to locate some “essence” of privacy. A promising theory of privacy has been developed by Daniel Solove, a professor of law at George Washington University. In his theory, how we view privacy depends on the context in which it is deemed valuable; an invasion of privacy can manifest in unique ways in the context of different social practices (Solove, 2002). “Social practices” can refer to “various activities, customs, norms, and traditions,” which can include writing letters, talking to a therapist, and making certain decisions about yourself, and so on (2002). All these practices admit of a dimension of privacy that is more-or-less unique to them, and integral to their function. When this dimension of privacy is invaded, a practice is disrupted or destroyed.

Data privacy is important for preserving the dimensions of privacy of the social practices that people engage in. By engaging in certain social practices, people generate or leave behind enough data for parties to collect it, and potentially interpret it into identifying information, leaving those individuals vulnerable to privacy invasions. Depending on what sort of party has that data, and what their motivations are, an individual’s privacy can be invaded in different ways, leading to the disruption of one or more social practices. With the increasing ability of AI tools to detect patterns and draw inferences, it is important to consider how this technology might increase the potential for privacy violations. As organizations assemble massive datasets to train AI algorithms, it is imperative that they adopt responsible policies on the collection, handling, and storage of data.

4.2 Data Privacy and the homeless

It is exceedingly rare that the homeless are mentioned in discussions of privacy. The homeless have no private property of their own to retreat to, and thus are forced to occupy public space indefinitely, or temporarily stay in the private space of someone else. The average citizen, being integrated into society, is likely to think that the homeless have nothing to lose in terms of privacy, and thus have no real claim to it. However, homeless individuals still express the same desire for privacy as anyone else, as well as a level of privacy regarding personal information about them (Sparks, 2013). These individuals do not want to be 'seen as homeless,' nor would they want—much less deserve—to be haunted by their past in the form of digital records.

Recall earlier what has been said about the CHAI tool. One of the databases used to train the CHAI model is London's Homeless Individuals and Families Information System (HIFIS), which "joins the service usage information for over a dozen shelters and related homeless services," and contains "approximately 4 years of 6521 clients' records" (VanBerlo et al, 2020). The data includes a client's use of social services, such as number of shelter stays, number of days receiving a housing subsidy, times they were refused service from a shelter, and SPDAT assessment (2020, p. 4). It also includes "total monthly income, total monthly expenses, medical diagnoses, shelters they stayed at, as well as demographic information, such as age, citizenship and gender" (2020). While this is all sensitive information, it is worth mentioning that client anonymity was preserved, "as names and other identifiable information was not fetched by the query," and clients "were identified by a unique ClientID" (2020, p. 3).

Another tool used by organizations is the Service Prioritization Decision Assistance Tool (SPDAT). As the name suggests, it assists in deciding the priority of homeless clients for receiving homeless prevention services. Using the SPDAT system, a homeless client is asked for a substantial amount of personal information, of which a large portion can be very intimate. This includes a client's mental and physical health, medication use, substance abuse, experience of abuse and trauma, risk of harm to self and others, legal issues, history of homelessness, and so on (SPDAT manual, 2015, pp. 5-19). Since the purpose of the SPDAT is to determine the priority of assistance for homeless clients, these clients may likely feel compelled to provide as much information as they can, even if it is very intimate, so that they can receive the assistance they need.

The sheer volume and sensitive nature of data collected from homeless clients poses a risk in and of itself. It is important, therefore, to take steps to ensure that data collected from the homeless stays within the context of helping the homeless. Should this data escape that context, a homeless or formerly homeless individual can be made vulnerable to severe privacy violations, potentially impacting their ability to engage in social practices as they establish themselves in society.

4.3 Respecting the Data Privacy of the homeless

Given the risks involved in handling the sorts of data collected from the homeless, their data privacy concerns should be taken as seriously as anyone else's. The purpose of these services and algorithms is to assist homeless clients in escaping homelessness and to reintegrate into society; given this, the data privacy of homeless individuals ought to be treated with the same respect as any other citizen. I will propose a list of basic data privacy principles for organizations to keep in mind as they collect and store the personal information of clients:

[1] Ensure that the homeless clients whose data is being collected and analyzed are the main beneficiaries of this collection and analysis; [2] ensure that clients are aware, to the best of their abilities, of your organization's data collection practices, and that their informed consent is obtained; [3] only collect data that is necessary for the purpose of helping the client, or for addressing homelessness in a given population; [4] have a system in place for certain personal or identifying client data to be expunged after a period of time, or upon request; [5] implement, within reason, tools or methods of privacy protection like certain forms of encryption or anonymization; and [6] ensure that client data is only shared within a system of legitimate organizations that provide service to the homeless and will remain in that shared system.

This is a tentative list of general principles that can easily admit of refinement, amendment, or expansion as needed. This list would benefit greatly from expert knowledge in the fields of data management, data privacy, artificial intelligence, and social services related to homelessness. Nonetheless, it is important to have principles like these to inform the data policies of organizations, especially as AI tools become increasingly widespread and sophisticated.

Conclusion

This report has examined four key issues surrounding the use of AI in the context of homelessness management and prediction: homelessness definitions, algorithmic bias, explainable AI, and data privacy. We have analyzed these issues through the particular case of the City of London’s Chronic Homelessness Artificial Intelligence tool. In our investigation of CHAI, we have contended with, and provided answers to, the following questions:

(1) Are these tools informed by an understanding of homelessness that is inclusive enough? We suggest that they are not, and that the theory of homelessness depicted by Jenkins and Brownlee in their paper “What A Home Does” is a better way to understand the problem and its parts. **(2) Might they be biased against homeless people who are especially marginalized?** We have shown that there is serious evidence of bias against women, and especially women of colour in these tools. **(3) Are the tools transparent enough to make their decisions understandable to all stakeholders?** We have shown that XAI fails to provide explanations that allow all stakeholders to understand the AI-decision. **(4) How ought public and private institutions handle the collection and storage of homeless clients’ data/personal information?** Institutions that utilize or plan to utilize AI tools should implement a set of strict principles to guide their data collection practices and prioritize the privacy and security of client data.

In all, we have argued that CHAI and AI tools like it—and the organizations using them—are falling short of adequately responding to these questions. We suggest that more work needs to be done to ensure fairness and equity in each of these dimensions before AI homelessness tools are used in the public sector.

References

Angwin, J., Larson, J., & Mattu, S. (2016, May 23). Machine Bias. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. Patterns, 2(2), 100205. https://doi.org/10.1016/j.patter.2021.100205

Blake, E. (2022, June 21). Is your insurance company watching you online and is it legal? The University of Sydney. https://www.sydney.edu.au/news-opinion/news/2022/06/21/is-your-insurance-company-watching-you-online-and-is-it-legal-data-privacy.html

Bretherton, J. (2017). Reconsidering Gender in Homelessness. 11(1). European Observatory on Homelessness.

Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. University of Chicago Legal Forum, 1989(1).

Data Use, Privacy, and Technology. (2023, March 4). NAIC. https://content.naic.org/cipr-topics/data-use-privacy-and-technology

Essert, C. (2016). Property and Homelessness. Philosophy & Public Affairs, 44(4), 266–295. https://doi.org/10.1111/papa.12080

Eubanks, V. (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press, Inc.

Jenkins, D., & Brownlee, K. (2022). What a Home Does. Law and Philosophy, 41, 1–28. https://doi.org/10.1007/s10982-021-09414-w

Johnson, G. (2022). Are algorithms value-free? Feminist theoretical virtues in machine learning. Journal of Moral Philosophy, forthcoming.

Johnson, G. M. (2020). Algorithmic Bias: On the Implicit Biases of Social Technology. Synthese, 198(10), 9941–9961. https://doi.org/10.1007/s11229-020-02696-y

Khalifa, K. (2013). The Role of Explanation in Understanding. The British Journal for the Philosophy of Science, 64(1), 161–187. http://www.jstor.org/stable/23356461

Kithulgoda, C. I., Vaithianathan, R., & Parsell, C. (2022). Racial and gender bias in self-reported needs when using a homelessness triaging tool. Housing Studies, 0(0), 1–13. https://doi.org/10.1080/02673037.2022.2151986

Kuhn, T. S. (1962). The Structure of Scientific Revolutions. Chicago (University of Chicago Press) 1962.
Lamberink, Liny. (2020, August 17). A city plagued by homelessness builds AI tool to predict who's at risk. CBC News London. https://www.cbc.ca/news/canada/london/artificial-intelligence-london-1.5684788

Lipton, P. (2001). What Good is an Explanation. In Explanation (Vol. 302, pp. 43–59). Springer. https://doi.org/10.1007/978-94-015-9731-9_2

Lipton, Z. C. (2017). The Mythos of Model Interpretability (arXiv:1606.03490). arXiv. http://arxiv.org/abs/1606.03490

Miller, Tim. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1–38. https://doi.org/10.1016/j.artint.2018.07.007

Mittelstadt, B., Russell, C., & Wachter, S. (2019). Explaining Explanations in AI. Proceedings of the Conference on Fairness, Accountability, and Transparency, 279–288. https://doi.org/10.1145/3287560.3287574

Noble, S. U. (2018). Algorithms of oppression. In Algorithms of oppression. New York University Press.

Oudshoorn, A., May, K., Berkum, A. V., Schwan, K., Nelson, A., Eiboff, F., Nichols, N., & Parsons, C. (2021). Exploring the Presence of Gender-Based Approaches to Women's Homelessness in Canadian Communities.

Pasquale, F. (2016). "Chapter 2: Digital Reputation in an Era of Runaway Data." In The Black Box Society (pp. 19–58). Harvard University Press.

Pieters, W. (2011). Explanation and trust: What to tell the user in security and AI? Ethics and Information Technology, 13(1), 53–64. https://doi.org/10.1007/s10676-010-9253-3

Poulin, Brett, Eisner, Roman, Szafron, Duane, Lu, Paul, Greiner, Russ, Wishart, D.S., Fyshe, Aona, Pearcy, Brandon, MacDonell, Cam, & Anvik, John. (2006). Visual explanation of evidence in additive classifiers. IAAI'06: Proceedings of the 18th Conference on Innovative Applications of Artificial Intelligence, 2.

Service Prioritization Decision Assistance Tool (SPDAT) Manual. (2015). OrgCode Consulting, Inc. https://www.empowersimcoe.ca/wp-content/uploads/2017/12/single-spdat.pdf

Smuha, N. A. (2021). Beyond the individual: Governing AI's societal harm. Internet Policy Review, 10(3). https://policyreview.info/articles/analysis/beyond-individual-governing-ais-societal-harm

SNIA. (2022, October 25). What is Data Privacy? Storage Networking Industry Association. https://www.snia.org/education/what-is-data-privacy#_ftn1

Solove, D. J. (2002). Conceptualizing Privacy. California Law Review, 90(4), 1087–1155. https://doi.org/10.2307/3481326

Sparks, T. (2010). Broke Not Broken: Rights, Privacy, and Homelessness in Seattle. Urban Geography, 31(6), 842–862. https://doi.org/10.2747/0272-3638.31.6.842

Turek, Matt. (2017). Explainable Artificial Intelligence [..Mil]. DARPA. https://www.darpa.mil/program/explainable-artificial-intelligence

van den Hoven, J., Blaauw, M., Pieters, W., & Warnier, M. (2014). Privacy and Information Technology. https://plato.stanford.edu/archives/sum2020/entries/it-privacy/

VanBerlo, B., Ross, M. A. S., Rivard, J., & Booker, R. (2020). Interpretable Machine Learning Approaches to Prediction of Chronic Homelessness (arXiv:2009.09072). arXiv. http://arxiv.org/abs/2009.09072

Véliz, C. (2020). Data, Privacy, and the Individual. In Center for the Governance of Change. https://philpapers.org/rec/VLIPM

Vold, K., & Whittlestone, J. (2019). Privacy, Autonomy, and Personalised Targeting: Rethinking How Personal Data is Used. Report on Data, Privacy, and the Individual in the Digital Age, 5.

Wachter, S., Mittelstadt, B., & Russell, C. (2017). Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3063289

Zerilli, J. (2022). Explaining Machine Learning Decisions. Philosophy of Science, 89(1), 1–19. https://doi.org/10.1017/psa.2021.13

Acknowledgements

Thanks to LIFE*SPIN for their guidance and extremely helpful insights as we completed this project, and to the members of Western's Homelessness Research Team (including Luke Stark, Joanna Redden, and Janet Allen) for their help. Special thanks to Carolyn McLeod, our professor who facilitated the development of this project and was integral to its success.

Contact

ndibera@uwo.ca
lkayssi@uwo.ca
lkopp3@uwo.ca
mwhitta9@uwo.ca