Legal & Ethics Report

Index

Introduction about the progress with ethics and legal requirements

In recent years, the city of Breda has seen a growing concern regarding public safety and nuisance. The increasing complexities of urban environments, coupled with various socio-economic factors, have led to a rise in crime rates and nuisance incidents within the city. It's important to adjust the data-driven strategy based on the public data of Breda.

The development of a DEDA framework specifically tailored to public safety and nuisance data in Breda is crucial. This framework will help with the establish guidelines, principles and best practices to get the responsible collection, analysis and utilization of data. By prioritizing ethical and legal considerations, stakeholders can maintain public trust, addressing the complexities of a safer and more livable environment for all residents of Breda.

Legal and Ethical aspects in the project

There are several factors that are included in the datasets of the project. Some of them have more potential than the other ones. To make it clearer and more relatable we need to figure out what can be missed or overlooked. Some of the commonly overlooked factors include socioeconomic factors, demographic variables, community characteristics, environmental factors, policing strategies, seasonal or temporal variations, technical advancements and policy and legislative changes.

Socioeconomic factors

Factors such as poverty levels, income inequality, unemployment rates, and educational attainment can significantly influence crime rates. Be aware that there can be a failure to these factors. It can lead to incomplete or biased analysis.

Demographic factors

Demographic variables, such as age, gender, ethnicity, and population density, can play a role in crime rates. If you fail for these factors, it can result in limited understanding of the underlying dynamics to crime rates.

Community characteristics

The characteristics of neighborhoods or communities can include cultural diversity, sense of belonging and social cohesion.

Environmental factors

Physical and environmental features, such as lighting, infrastructure, and presence of surveillance systems can influence crime rates.

Policing strategies

The effectiveness and efficiency of policing strategies, including community policing initiatives, can affect crime rates. Ignoring the impact of these strategies may lead to incomplete analysis and ineffective policy recommendations.

Seasonal patterns in the datasets

The crime rates in Breda can have different variations based on the seasons in the year. Also, some temporal factors, such as holidays, time of day, or specific events can be overlooked to an inaccurate understanding of crime trends.

Understanding the seasonal or temporal variations, we will analyze crime rates and nuisance incidents in relation to seasonal patterns per month that can gain valuable insights into the dynamics of public safety throughout the year.

Technical advancements

Staying updated on technological advancements is crucial for a comprehensive understanding of public safety. There is one factor included and that is cybercrime or changes in criminal methodologies.

Policy and legislative changes

Criminal justice can have an impact on the crime rates of the dataset. It can be changed by policies and legislation.

TLDR of ATLAI (Assessment list for Trustworthy Artificial Intelligence)

ATLAI was developed by the High-Level Expert Group on Artificial Intelligence (A group of experts to provide advice to the European Commission on its artificial intelligence strategy.) set up by the European Commission to help assess whether the AI system that is being developed deployed procured or used complies with the seven requirements of Trustworthy AI, as specified in these 7 key Ethics Guidelines for Trustworthy AI.:

1. Human Agency and Oversight
2. Technical Robustness and Safety.
3. Privacy and Data Governance.
4. Transparency.
5. Diversity, Non-Discrimination, and Fairness.
6. Societal and Environmental well-being.
7. Accountability.

Goal and Purpose of ALTAI.

ALTAI aims to provide a basis of evaluation process for Trustworthy AI self-evaluation.
Meaning that Organizations can draw elements relevant to the AI system from ALTAI or add elements to it as they see fir taking into consideration the sector, they operate in. It helps organizations to understand what trustworthy AI is, particularly what risks and AI system might generate. It raises awareness of the Impact of AI on society, the environment, consumers, workers, and citizens (particularly children and people belonging to marginalized groups.) It promotes involvement of all relevant stakeholders (within and outside of organizations). It helps gain insight on whether meaningful and appropriates solutions or processes to accomplish adherence to the requirements are already set in place (through internal guidelines, governance processes etc.) or need to be put in place.

A trustworthy approach is to enable “responsible competitiveness”, by providing the foundation upon which all those using or affected by AI systems can trust that their design, development, and

use are lawful, ethical, and robust. **ATLAI helps create responsible and sustainable AI innovation in Europe**. It seeks to make this a core pillar for developing a unique approach to AI, **one that aims to benefit, empower, and protect individuals and the common good of society**. ATLAI believes that this will enable Europe and European organizations to position themselves as global leaders in cutting-edge AI worthy of our individual and collective trust.

ATLAI was developed over two years from June 2018 to June 2020.

What is ATLAI?

ALTAI aims to provide a basis of evaluation process for Trustworthy AI self-evaluation.

Why do we need ATLAI?

ATLAI helps create responsible and sustainable AI innovations in Europe and provides the guidelines to create safe and trustworthy AI.

Checklist for the DEDA and ATLAI

Data Collection

- Is the data collected from reliable and trustworthy sources?

Data collected should come from reliable and trustworthy sources to ensure accuracy and credibility.

The data provided is reliable and accurate cause it makes use of data time features and a constant value to the date and time.

- Are data collection methods transparent and well-documented?

Transparent and well-documented data collection methods are essential for replication, verification, and establishing the reliability of the data.

The data collected are transparent giving a clear understanding for what it's used for and they are well documented cause they're being updated monthly and are constantly being checked.

Data Quality

- Is the data accurate, complete, and up to date?

Data should be accurate, complete, and up to date to ensure the validity of the analysis and conclusions drawn.

The data is accurate, complete and up to date. The data is constantly being updated on a monthly basis by the provider of said data.

- Is there a clear understanding of the limitation's potential biases of the data?

Understanding the limitations and potential biases of the data is crucial for interpreting the results accurately and avoiding misrepresentation.

The data may develop some biases in terms of what neighborhood has the most crimes etc. But we are not making use of individual parties or ethnic groups that can be exploited with this data since this data only works with map and neighborhood data.

Confounding variables

- Have you taken steps to control these confounding variables that may influence both the cause and effect?

Controlling confounding variables is necessary to establish a clear cause-and-effect relationship.

Yes, we have found confounding variables that would assist us in defining the cause and effect.

- Is there transparency in reporting the inclusion and impact of confounding variables in crime rate analysis?

Transparency in reporting the inclusion and impact of confounding variables in crime rate analysis is important for the credibility and reproducibility of the research.

The confounding variables are transparent and clear to whoever look at it.

Ethical consideration

- Are there any expected benefits to participants?

Assessing the expected benefits to participants is important to ensure that their involvement in the project is justified and ethically sound.

- Will any reports/feedback include recommendations for improvement?

Providing feedback to participants and including recommendations for improvement in reports/feedback helps maintain transparency and accountability.

Interpretation and Policy Implementation

- Do the data provide insight into the personal lives of citizens?

It is essential to consider whether the data provide insight into the personal lives of citizens to avoid privacy violations or unethical use of personal information.

Our data is only based on neighborhoods and index numbers. There are no personal details such as last names, exact addresses, etc.

- Does this project make use of personal data?

If the project involves the use of personal data, additional safeguards may be necessary to ensure compliance with data protection regulations.

No, there is no personal data used. Only the total amount of the crime rates and public nuisance.

Bias and Fairness

- Is there a risk that your project could contribute to discrimination against certain people or groups? (Think about some regions and neighborhoods that are involved in Breda)

Assessing the risk of project contribution to discrimination against certain people or groups is important, especially considering specific regions and neighborhoods in Breda.

No, the data does not specify what part or region of the neighborhoud has the highes crime rate. So there is no group to single out.

- Are there measures in place to ensure fairness in decision-making and outcomes?

Implementing measures to ensure fairness in decision-making and outcomes helps prevent bias and promotes equitable treatment.

The measures that are in place, is that we do not show what specific neighborhood has the most crime rate, nuisance or what specific neighborhood brings down the livability score.

Causality Assumptions

- Have you considered the assumptions and limitations of the chosen algorithm for casual inference?

Considering the assumptions and limitations of the chosen algorithm for causal inference is important for accurate and reliable causal analysis.

Yes, we have considered the assumptions and limitation for casual inference since we have identified multiple causes and effects that can impact the crime rate and the livability score of a neighborhood.

- Have you clearly defined the casual direction between variables of interest?

Clearly defining the causal direction between variables of interest and stating the underlying assumptions in the causal analysis helps establish a solid foundation for the research.

Yes, we have defined the casual direction between variables of interests by analyzing the data we found the cause for low livability score ties with high crime rate and nuissances and the neighbourhood who had less nuisances also did have crimes but on a less frequent rate.

ATLAI Checklist

1. Human Agency and Oversight:

Could the AI system generate confusion for some or all end-users or subjects regarding the origin of algorithmic decisions?

Are end-users or subjects adequately informed that a decision, content, advice, or outcome is the result of an algorithmic decision?

Could the AI system lead to over-reliance by end-users, potentially affecting human autonomy?

Have procedures been established to prevent unintended interference with human autonomy?

Are human stakeholders involved in the decision-making process of the project?

Is there a clear understanding of the roles and responsibilities of human stakeholders?

Are there mechanisms in place for human oversight and intervention in the prediction process?

Is there a process to review and override automated decisions if necessary?

Are there provisions for accountability in case of errors or negative impacts?

2. Technical Robustness and Safety:

Could the AI system pose risks or threats due to design or technical faults, defects, outages, attacks, misuse, inappropriate or malicious use?

Have potential forms of attacks and vulnerabilities been assessed?

Are measures in place to ensure the integrity, robustness, and overall security of the AI system against potential attacks throughout its lifecycle?

Is there a mechanism to evaluate when changes to the AI system warrant a new review of its technical robustness and safety?

Is the prediction model robust and reliable, producing accurate and consistent results?

Are there measures in place to ensure the safety and integrity of the data used for predictions?

Has a risk assessment been conducted to identify and mitigate potential risks and hazards associated with the prediction model?

3. Privacy and Data Governance:

Have you considered the impact of the AI system on the right to privacy?

Have mechanisms been established to flag privacy-related issues related to the AI system?

Have you implemented measures mandated by data protection regulations?

Have you implemented the right to withdraw consent, the right to object, and the right to be forgotten into the development of the AI system?

Have you considered the privacy and data protection implications of data collected, generated, or processed throughout the AI system's life cycle?

Are there measures in place to protect personal data and comply with relevant data protection regulations?

Has a data governance framework been established to ensure responsible data handling practices?

4. Transparency:

Have you implemented measures to address the traceability of the AI system throughout its entire lifecycle?

Can you trace back which data was used by the AI system to make specific decisions or recommendations?

Can you trace back which AI model or rules led to the decisions or recommendations of the AI system?

Have you put in place measures to continuously assess the quality of the AI system's outputs?

Have you explained the decisions of the AI system to the users?

Have you provided appropriate training materials and disclaimers to users on how to adequately use the AI system?

Is the prediction model transparent and explainable?

Are explanations provided to stakeholders on how their data is used and how predictions are made?

5. Diversity, Non-discrimination, and Fairness:

Have you established a strategy or procedures to avoid creating or reinforcing unfair bias in the AI system, both regarding the use of input data and algorithm design?

Have you considered diversity and representativeness of end-users and/or subjects in the data?

Have you tested for specific target groups or problematic use cases?

Have you assessed and implemented processes to test and monitor for potential biases throughout the AI system's lifecycle?

Have you implemented educational and awareness initiatives to help AI designers and developers be more aware of possible biases in the AI system?

Have you ensured a mechanism for flagging issues related to bias, discrimination, or poor performance of the AI system?

Have you established clear steps and communication channels for raising such issues?

Have you assessed whether the AI system's user interface is usable by those with special needs or disabilities or those at risk of exclusion?

Have you ensured that information about the AI system and its user interface is accessible and usable to users of assistive technologies?

Have you involved or consulted with end-users or subjects in need of assistive technology during the planning and development phase of the AI system?

Have you considered the impact of the AI system on potential end-users and/or subjects?

Have you assessed whether the team involved in building the AI system engaged with possible target end-users and/or subjects?

Have you assessed whether there could be groups who might be disproportionately affected by the outcomes of the AI system?

Have you assessed the risk of possible unfairness of the system on the end-users' or subjects' communities?

Have you considered a mechanism to include the participation of a wide range of stakeholders in the AI system's design and development process?

6. Societal and Environmental Well-being:

Are there potential negative impacts of the AI system on the environment?

Which potential impacts do you identify?

Where possible, have you established mechanisms to evaluate the environmental impact of the AI system's development, deployment, and/or use (e.g., energy consumption, carbon emissions)?

Have you defined measures to reduce the environmental impact of the AI system throughout its lifecycle?

Does the AI system impact human work and work arrangements?

Did you inform and consult with impacted workers and their representatives (e.g., trade unions, work councils) in advance of introducing the AI system in your organization?

Have you adopted measures to ensure a clear understanding of the impacts of the AI system on human work?

Did you ensure that workers understand how the AI system operates and its capabilities?

Could the AI system create the risk of de-skilling the workforce?

Have you taken measures to counteract de-skilling risks?

Does the system promote or require new (digital) skills?

Have you provided training opportunities and materials for re- and up-skilling?

Could the AI system have a negative impact on society at large or democracy?

Have you assessed the societal impact of the AI system's use beyond the end-users and subjects, considering potentially indirectly affected stakeholders or society at large?

Have you taken action to minimize potential societal harm caused by the AI system?

Have you taken measures to ensure that the AI system does not negatively impact democracy?

Has the potential impact of the prediction on society and the environment been considered?

Are there measures in place to mitigate any negative consequences or externalities?

Does the project align with sustainable and ethical practices in urban planning and neighborhood development?

Has the project considered the potential social and environmental impacts of predicting public nuisance?

Are there mechanisms in place to address and mitigate any negative social or environmental consequences?

Has the project considered the broader well-being and quality of life of the neighborhoods and communities involved?

7. Accountability:

Did you establish mechanisms that facilitate the AI system's auditability (e.g., traceability of the development process, sourcing of training data, logging of processes, outcomes, positive and negative impact)?

Did you ensure that the AI system can be audited by independent third parties?

Did you foresee any external guidance or third-party auditing processes to oversee ethical concerns and accountability measures?

Does the involvement of these third parties go beyond the development phase?

Did you organize risk training, and if so, does it inform about the potential legal framework applicable to the AI system?

Did you consider establishing an AI ethics review board or a similar mechanism to discuss overall accountability and ethics practices, including potential unclear areas?

Did you establish a process to discuss, continuously monitor, and assess the AI system's adherence to this Assessment List for Trustworthy AI (ALTAI)?

Does this process include identification and documentation of conflicts between the requirements or between different ethical principles, and explanation of the "trade-off" decisions made?

Did you provide appropriate training to those involved in the process, including coverage of the legal framework applicable to the AI system?

Did you establish a process for third parties (e.g., suppliers, end-users, subjects, distributors/vendors, or workers) to report potential vulnerabilities, risks, or biases in the AI system?

Does this process foster revision of the risk management process?

For applications that can adversely affect individuals, have redress-by-design mechanisms been put in place?

Is there a clear framework for accountability and responsibility within the project?

Are stakeholders aware of their roles and responsibilities regarding the prediction model?

Are there mechanisms in place for addressing any concerns or issues that may arise?

Is there a process for addressing complaints, concerns, or disputes related to the predictions made?

Are there mechanisms in place to track and evaluate the impact and effectiveness of the prediction model over time?

Relevance ATLAI Checklist

It is important to notice that there are some aspects that are more necessary to look at than others. The decision should be based on careful evaluation of the project's scope, context and available resources.

Human agency and Oversight: These are the importance of human involvement and decision making in the project. It needs to be allowed to for responsible and context-specific interventions to address public nuisance and safety effectively.

Technical Robustness and Safety: It ensures the reliability, accuracy and stability of the data analysis and algorithms used in crime ratings.

Privacy and data governance: We need to protect individuals' privacy and ensuring proper data governance are essential for crime ratings.

Diversity, Non-discrimination, and Fairness: It is crucial to avoid biases and ensure fairness in crime ratings. Promoting diversity, non-discrimination, and fairness aligns with ethical principles of equal treatment in neighborhoods.

Accountability: This aspects in the checklist promotes accountability and transparency in the development and use of the AI system. This transparency builds trust among stakeholders and community members.

The relevance of legal and ethical aspects from the checklist

There are many aspects of the checklist that need to be understood in order to provide a relevant explanation of why it is so important. The inclusion of legal aspects involves compliance with laws and regulations, such as data privacy and protection, to ensure that the project operates with legal boundaries. This helps protect the rights and privacy of individuals involved in this project.

Legal aspects

This aspect includes legal requirements that govern the collection, use and dissemination of crime-related data.

Ethical aspects

Ethical considerations build public trust and confidence in this project. When stakeholders see that the project operates with transparency and respect guidelines, they are more likely to trust the findings to effort the public safety in Breda.

These considerations also promote a lot of fairness and equity in addressing public safety and nuisance issues. This means avoiding bias, discrimination or disproportionate impacts on specific groups or cultures.

Be aware that there are some aspects that can be very crucial to consider these legal and ethical aspects. From the DEDA checklist they may be valid justifications for not including certain elements. Due to the project's limited time, you need to prioritize what are the most necessary aspects to use.

Decision-making for the project

In the context of the project focusing on public safety and nuisance in Breda, ethical decision-making plays a critical role in ensuring responsible and transparent data analysis. There are some keywords that are highlighted to get a better overview of the aspects.

- Data Collection: Ethical decision-making begins with ensuring that data is collected from reliable and trustworthy sources. Transparent and well-documented data collection methods are employed to maintain transparency and accountability.

- Data Quality: Ethical considerations demand that the collected data is accurate, complete, and up to date. Recognizing the limitations and potential biases of the data is crucial to avoid misinterpretation or unfair representation.

- Ethical Consideration for Participants: The project takes into account the potential benefits to participants. Feedback and reports, including recommendations for improvement, are provided to ensure transparency and mutual benefit.

- Bias and Fairness: Efforts are made to identify and mitigate any risks of contributing to discrimination against certain regions, neighborhoods, or groups in Breda. Measures are in place to ensure fairness in decision-making and outcomes.

- Causality Assumptions: The project acknowledges the assumptions and limitations of the chosen algorithm for causal inference. Clear definitions of causal directions and the underlying assumptions are stated to avoid misinterpretation and ensure accurate analysis.

By adhering to these ethical considerations, the project promotes trust and confidence among stakeholders. It demonstrates a commitment to fairness, equity, and the protection of individuals' rights and privacy. Ethical decision-making is an integral part of the project's overall success in addressing public safety and nuisance issues in Breda while upholding the values of responsibility and transparency.

What is excluded?

There are some questions that are not going to be related to the project. Here are some explanations why and which ones we need to be aware why it is excluded.

Special needs or disabilities: While inclusivity is important, if your research specifically focuses on public nuisance in neighborhoods, the direct impact on users with special needs or disabilities may not be the primary objective. The focus of your study may be more centered on understanding the factors contributing to public nuisance rather than the specific needs of different user groups.

Environmental impact of the AI development and deployment: Our project is related to public safety and nuisance. These questions may be outside the scope of our research. However, if environmental concerns are a central aspect of public nuisance in neighborhoods, you may reconsider including them.

Diversity and representation about the end-users for the project: Our research does not specifically aim to analyze the impact of public nuisance on diverse populations or explore disparities in its occurrence across different groups, these questions may not be essential for your study.

Work related questions about human oversight: While it impacts of AI systems on human work is an important consideration, if your research primarily focuses on public nuisance in neighborhoods, the labor impact may not be the primary objective. However, if your research investigates the broader socio-economic implications of public nuisance, these questions could be relevant to include.

User's interface and usability with AI: If the user interface does not directly affect the prediction or prevention of public nuisance, these questions may not be directly relevant to your research. Usability concerns are more pertinent when the interface plays a significant role in the use and effectiveness of the AI system, which may not be the case for your study.

By excluding these aspects, the research can focus more specifically on the factors contributing to public nuisance in Breda, providing a targeted analysis and understanding of the problem at hand.

References

High-Level Expert Group on Artificial Intelligence. (2020). Ethics Guidelines for Trustworthy AI. Retrieved from https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

Data.breda.nl. (n.d.). Open Data Breda. Retrieved from https://data.breda.nl/

European Commission. (2020). The Assessment List for Trustworthy Artificial Intelligence (ALTAI) (ISBN: 978-92-76-20008-6). doi:10.2759/002360 KK-02-20-479-EN-N.

Politie Nederland. (n.d.). Politie Open Data. Retrieved from https://data.politie.nl/#/Politie/nl/

Statistics Netherlands. (n.d.). CBS - Statistics Netherlands. Retrieved from https://www.cbs.nl/