EXAMPLE: DETECTING ROAD ISSUES AND DEBRIS IN SAN JOSÉ

GovAI Use Case: Object Detection

<<example template, responses in green>>

Template Overview

This document serves as a template for others to create their own tailored use cases. It is meant to illustrate a strategic method for using AI technology in solving challenges in state and local governments.

All sections do not need to be completed when first filling this out. Sections can start as considerations (or blanks), and more information can be added in depth as the use case is continued.

Examples can be found in the AI Use Cases Working Group folder.

Proposed Use Case

In a sentence, what is your use case? What is the question you are trying to answer or problem to solve?

How can objects in the road, like trash and potholes, be better identified so that City staff can focus on addressing and mitigating these issues? These focus on 6 object categories: illegal dumping/trash, graffiti, potholes, tent encampments, lived-in vehicles/RVs, and parking violations.

Proposed AI

What kind of AI tool is being used and how is it being used? Is it a conversational chatbot or language model? Does the AI use computer vision or audio identification? Is the AI meant for prediction or for studying causal impact?

This project aims to use AI to identify the issues in City roads. This AI uses computer vision. It is meant to reduce the amount of time City staff need to be manually identifying road issues.

Proposed Project Phases

Consider what phases your project might have. Can start with your proposed phases, and update as you progress in your use case. Potential phases could include:

1. **Pilot Program Development:** Assess readiness and explore feasibility.

Complete a feasibility study to assess the effectiveness of utilizing AI to proactively identify objects and address issues that can improve the quality of life in San José. The study will include the following:

1. Object detection accuracy
2. Vendor evaluation and comparison for each type of object detection, technology, and used method (video streaming vs. image capture)
3. Implementation, analysis process, and support cost
4. Frequency of desired detection
5. Identify privacy concerns

The pilot is located in District 10 and will run for 6-10 weeks starting December 2023.

2. **Training Phase:** Focus on understanding AI training techniques and methodologies.

   The feasibility study has concluded that while object detection is possible, more training data is necessary to provide the model with a larger variety of issues to identify from different parts of the city. Training can include previous images of issues logged by the City. Additionally, training data should be curated in such a way to preserve as much privacy for residents as possible.

3. **Capacity Building:** Assemble a skilled team for project development and conduct initial assessments.

   The object detection team will work closely with San José's 311 contact center team. The pilot team, consisting of members from the Privacy team and project management team can be expanded to include members of the 311 contact center team.

4. **Testing and Validation:** Gather data on the performance of the system in the field. Once complete, add that performance report here for other agencies to learn.

   The object detection system can be tested on metrics such as accuracy in identifying common road issues in the field. We will add the performance report so other agencies who are also interested in object detection can see the performance of the system the City has decided to use.

Potential Benefits

When possible, try to quantify the benefits, like cost savings or added value. Even if we can't assign a number, showing that there are quantifiable benefits can often help.

We anticipate that there will be **x amount of time** and **y amount of money** saved. Additionally, this project provides a template for departments who would also like to incorporate AI systems into their decision making.

Immediate benefits

1. **Reduce cost and time:** Current method involves city staff and inspectors manually going through City to identify issues. Residents also report issues. AI can help city staff focus on addressing these issues.

2. **Identify patterns of need:** Cameras could proactively detect issues in real-time, in an automated way on a recurring basis, for City staff to identify areas that need more support. This helps the City become more effective at helping residents.

3. **Continuous Learning:** The AI system will evolve and improve its recommendations over time, based on the data fed into it.

Wider Applications:

1. **Statewide Application:** The solutions developed can inform policy and decision-making across the nation.

2. **System Integration:** Develop a system to track and analyze road issues, aiding in problem identification and mitigation.

System and Project Challenges

What kind of challenges do you anticipate in implementation? Can be very general in the planning phase as just considerations. As you implement your use case, add the major hurdles you experienced.

1. Global vs local approach:
    a. Global approach: The City drives a street and uploads the footage to its open data portal for anyone to build an identification algorithm for it.
    b. Local approach: A vendor can provide the City with their own cameras and instruct City employees how to install the devices on City vehicles. City employees drive down the street and identification is made.
2. How well does the AI model work:
    a. Metrics used: Includes false positive, false negative, true positive, and true negative
    b. Combatting bias: which fairness metric is used? How are the tradeoffs between fairness and performance determined?
3. Privacy concerns:
    a. No audio will be recorded, no video footage will be stored, no people will be tracked

People/Domains to Involve

Who or what groups need to be involved in this project? Where possible, consider when they should be involved as well. Some general groups to consider:

1. Agency or department leadership
2. People impacted by system (e.g., general public, applicants for a service)

3. Technical experts, such as the Information Technology Department

1. Department of Transportation (DOT):
    a. Fixing issues on the road is under DOT's purview. Additionally, the cameras used for the AI object detection system are affixed to parking compliance vehicles, which requires DOT's support.
2. Information Technology:
    a. SJ311 is the non-emergency services line that takes in many of the requests for road issues. Integrating the object detection AI with SJ311 streamlines SJ311 to ensure that the same issues aren't being addressed multiple times
    b. IT spearheads the project effort. Additionally, IT also has the technological expertise necessary to assess the usefulness of the object detection AI.
3. Legal:
    a. Legal professionals are often needed to navigate regulatory and code challenges.

Risks & Mitigation Strategies

1. **AI Risk**

    What would happen if the AI system were inaccurate or makes a wrong decision? How can you mitigate the risk of the system being wrong, and how can you fix the harm done when the system is wrong? For more detail on AI risk considerations, see BDO for an easy-to-understand resource[1] for an easy-to-understand resource and the EU for original source[2].

    Since the AI system will determine which issues are resolved, inaccurately identifying issues leads to the wrong issues or less urgent issues resolved. This will negatively impact residents' access to public services, since potholes or trash dumping may remain for longer than necessary.

    When the system is wrong, this may require additional manual inspection of identifying issues on the road. Additional fixes include updating the training data to include more examples of issues the AI system is likely to miss, such as illegal parking examples, and fixing missed issues.

2. **Privacy Risk**

[1] https://www.bdo.co.uk/en-gb/insights/advisory/risk-and-advisory-services/navigating-the-eu-artificial-intelligence-(ai)-act-implications-and-strategies-for-uk-businesses#:~:text=The%20EU%20Commission%20designed%20AI,and%20low%20or%20minimal%20risk.&text=AI%20applications%20would%20be%20regulated,address%20specific%20levels%20of%20risk

[2] https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf

Consider the risk to data privacy that affects residents. What would happen if the data collected was stolen? How can you mitigate this risk, and what plan is in place if data is compromised?

Resident faces and identifying features or possessions may be captured when taking video of City streets. If the data collected was stolen, this poses a risk for residents' whose faces or features are used without their consent. Mitigating this risk includes storing only photos of issues, not the entire video of the streets, and blurring out resident faces if they are in the photo.

3. **Other Risk**

   What are other kinds of risk your organization has to consider?

   Other kinds of risk include risk of new AI guidelines or laws put in place that may necessitate a redesign of the AI system itself due to the training data. Other risk includes the AI system developing a bias that leads to the AI system being unable to identify certain issues, which leads to the ineffective use of AI and inefficient use of resources.

Project Resources Needed

What costs do you anticipate or experience in this use case? Some common cost considerations include:

1. Upfront costs
    a. The money needed to run a pilot of the object detection AI
    b. Mounting cameras onto vehicles for the AI to run
    c. Collecting the initial training dataset
2. Ongoing costs
    a. Maintaining the contract with the vendor who provides the AI system
    b. Retaining new information from the streets to ensure that the AI system is retrained as necessary to be kept updated
3. Staffing needs, from procuring to managing the system. Where possible, include estimated capacity needed (e.g., number of staff, hours per month for staff)
    a. Require individuals proficient in navigating various sectors and data sources and understanding the nuances of different industry languages.
    b. Estimated capacity: 10 staff, each 40hrs a week, total: 1600 hrs/month
4. Cost for ending the system or project
    a. Retraining staff in original object detection method – manual detection
    b. Migrating the data stored by AI be compatible with the manual detection method

Data Sources

What kind of data does the AI system(s) need? Is this visual, audio, text, etc.? If possible, please include data classification (e.g., person identifiable, HIPAA, etc.), data quality, and data ownership to keep track of the data and when it needs to be updated.

This AI system uses video of City streets and photos of issues. Video of City streets will be taken by the City department. Photos of issues can be submitted by residents or by City staff.

Public Data. Include the specific dataset if data is from a particular department (more applicable to cities or counties).

1. Name of dataset: San Jose Districts Roads and Streets Map
    a. Owner/maintainer: City of San Jose, Department of Transportation
    b. How to access: Link to the dataset is found here:
    c. Details on dataset: This dataset contains street and road information to help the City staff determine which roads to drive down to gather video data.

Private Data

1. Name of dataset: Vendor Training Dataset
    a. Owner/maintainer: Vendor of AI system
    b. How to access or collect: Accessible only to the vendor.
    c. Details on dataset: Contains examples of objects in the road from all over California.

Combatting AI Bias

What biases did you consider and how are we mitigating them? How can you track if the system is unintentionally impacting different communities differently (e.g., by race, age, gender, skin tone, socioeconomic status, language, immigration status), and how would you go about fixing any unintended bias? See more detail on algorithmic (aka "AI") bias from the Greenlining Institute[3], with industry-specific examples starting on page 8[4]. You can also see a brief video on AI bias from PBS[5].

Biases considered:

1. AI violates definitions of fairness
    a. Group-based: statistical measures between different groups are similar
        i. Achieving group-based fairness means that different demographic groups should be treated similarly. For example, the AI system's

[3] https://greenlining.org/wp-content/uploads/2021/04/Greenlining-Institute-Algorithmic-Bias-Explained-Report-Feb-2021.pdf

[4] https://greenlining.org/wp-content/uploads/2021/04/Greenlining-Institute-Algorithmic-Bias-Explained-Report-Feb-2021.pdf#page=8

[5] https://www.youtube.com/watch?v=gV0_raKR2UQ

accuracy rate at identifying issues should be similar between different demographic groups based on where they live.

   b. Fairness through unawareness: remove the use of legally protected characteristics from the data
      i. The training data should not have legally protected characteristics such as race or age attached to it. The real-world scenarios in which the AI system is used also should not have legally protected characteristics attached to them.
   c. These can be tracked by comparing statistical measures (like accuracy or false positive rate) between communities (such as race, gender, socioeconomic status). If the measures are similar across all groups, we can conclude that group-based fairness has been achieved. However, this is not a perfect solution. For example, assume that Group A and Group B has the same percentage of issues. Although Group A and Group B have the same accuracy for the AI identifying issues, this may be because the AI consistently identifies most objects as issues for Group A and consistently identifies most objects as non-issues for Group B.

2. User values and AI values do not align
   a. Users may believe that issues pertaining to basic needs, like tent encampments, are more important than issues pertaining to quality of life, like graffiti. The AI may have been trained so that all issues are of equal importance.
3. Model is used in a different context than the one that it was trained in
   a. Vendor models may be trained on video data that was only taken at a certain time of day (ie morning) while the use of the model takes place at all times of the day. This change in time may reduce the performance of the model.
4. Biased Data:
   a. Non-representative input data
      i. Piles of trash may only appear near intersections in video data when trash piles are likely to appear near roads without intersections.
      ii. Training data only includes images of issues from wealthier areas that usually contain a particular road marker. When the AI system is deployed, low-income areas that do not have the particular road marker are less likely to have issues identified even though issues exist.
   b. Encoded bias from data sources
      i. Older cars are more likely to be marked as abandoned vehicles, but older cars may also be more used in lower-income areas. As a result, older cars in lower-income areas may be identified as abandoned vehicles when they are not.

Mitigating Bias:

1. Values encoded in the AI system are the same as the values intended by the decision-makers
2. Ensure the model is being used in the same context as the model was trained in by testing several cases from the use case

    a. Test the model on example problems with solutions to check model performance.
3. Ensure the input data is as unbiased as possible (correcting for historical biases, ensuring that examples from different groups are evenly represented in the data)
    a. Resample the training data so that all scenarios are evenly represented, or create training examples to correct for historical bias