

## Concept design summaries

These summaries were auto-generated using ChatGPT (May 24 Version for workshops 1–3, and July 20 version for workshops 4–5)<sup>1</sup> with the following prompt:

“Can you summarize the following description of a design concept for a contestable algorithmic system in 150 words or less, using plain and straightforward language?” [Raw transcript of design concept verbal description pasted on subsequent line.]

We then checked the summaries against the original transcripts and lightly edited them for correctness.

---

<sup>1</sup><https://help.openai.com/en/articles/6825453-chatgpt-release-notes>

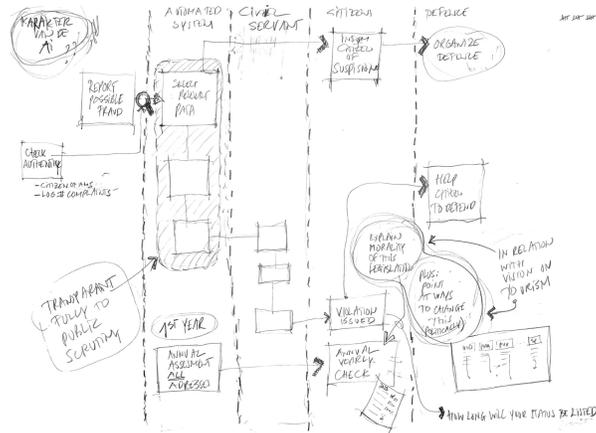


Figure 1: Concept design 1 from workshop 1 (C1.1). Suggests implementing a transparent and fair approach to assess the behavior of citizens in Amsterdam. Initially, everyone starts with a clean status, and annual assessments are conducted for all citizens, regardless of whether they have been reported or not. The system aims to detect illegal renting practices, and individuals who report potential violations will be checked anonymously. The number of complaints filed by a person will be logged and considered in the selection process, as it indicates their behavior patterns. Immediate reports will be provided to individuals under investigation, allowing them to organize their defense early on. The algorithm's selection mechanism will be open to inspection by organizations like Bellingcat and Follow the Money. The annual assessment will provide individuals with information on their fraud listing status and an explanation of why the city of Amsterdam has chosen this approach. Additionally, the system will offer guidance on how individuals can challenge the rules and seek support from political parties if they disagree with the city's vision on tourism. The AI character, representing the system, should be carefully designed to ensure it portrays a fair and informative entity rather than an intimidating presence. (Summarized from 771 words.)

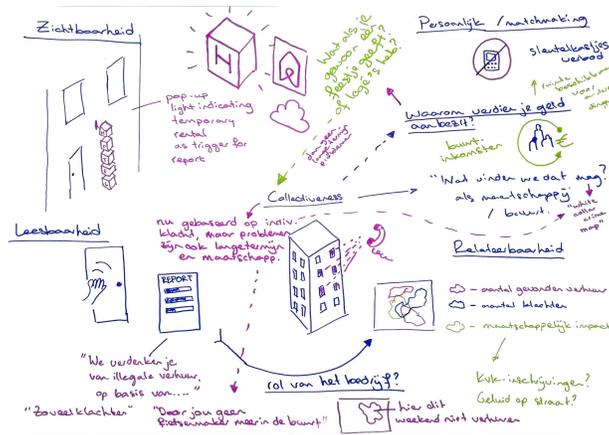


Figure 2: Concept design 2 from workshop 1 (C1.2). Addresses issues related to complaints and the impact of platforms like Airbnb on neighborhoods. The idea involves creating a visible indicator outside homes, similar to hotel signs, to indicate if a property is being rented out. This would help people understand the situation before filing a complaint. Additionally, when authorities investigate a complaint, they would bring a report indicating the specific indicators that raised suspicion, such as previous incidents or neighborhood characteristics. The concept also emphasizes the need to consider collective effects beyond individual complaints, such as changes in neighborhoods due to short-term rentals. It proposes using data from sources like the Chamber of Commerce to understand broader societal trends rather than solely relying on individual statistics. The role of companies is also highlighted, suggesting they could discourage renting in already saturated neighborhoods. Personal matchmaking and the redistribution of profits from rentals are mentioned as additional considerations. Overall, the concept aims to balance individual concerns with collective impacts, enhance transparency, and encourage a more comprehensive approach to address issues related to short-term rentals. (Summarized from 672 words.)

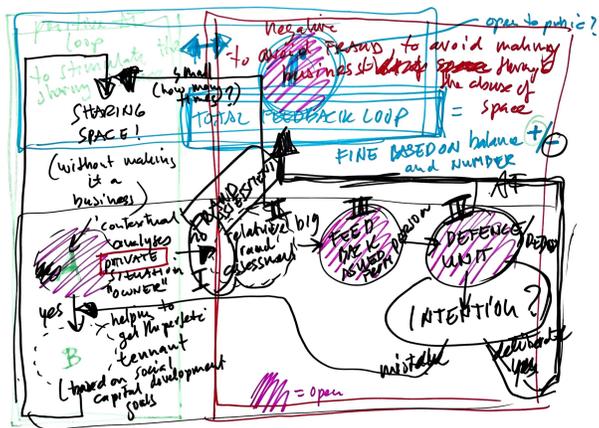


Figure 3: Concept design 3 from workshop 1 (C1.3). Proposes a system for sharing valuable space in a positive and fair way. Instead of focusing on fraud detection, the idea is to encourage individuals to share their empty space with others. A contextual analysis is performed to determine if a person meets the conditions for sharing their space. If they do, the system helps them find a suitable tenant, aiming to bring together diverse individuals who wouldn't have otherwise met. The feedback loop suggests that any financial gains from this sharing could be shared among participants in some manner. To prevent misuse of the sharing space as a business opportunity, a fraud assessment is conducted. However, it's important to assess the severity of the offense, differentiating between minor and major infractions. Small mistakes allow individuals to restart the process, whereas deliberate or significant offenses result in immediate action. Rather than imposing immediate fines, the system considers an individual's past behavior and weighs their positive contributions against negative actions. This approach values and rewards those who consistently contribute positively over time. The concept also suggests the possibility of redistributing capital as a form of justice within the system. (Summarized from 691 words.)



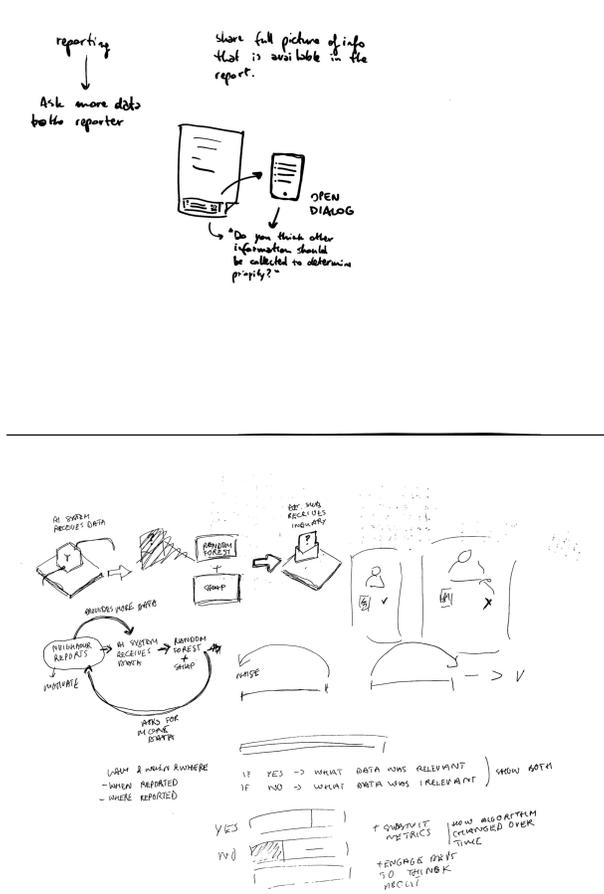


Figure 5: Concept design from workshop 3 (C3). Focuses on transparency, dialogue, and feedback. The system aims to provide a full picture of information, including metrics and sources, to understand why a report is generated and how decisions are made. It encourages dialogue and feedback from both decision subjects and users to improve the system and make it more transparent. The concept involves involving developers and the public in system development and sharing success rates to engage them. It also addresses the communication and impact on individuals being investigated, emphasizing a human approach and minimizing negative effects. The design includes monitoring and collecting feedback, considering both the human and technical aspects. It also highlights the importance of results and ethical discussions while improving the system's fairness and effectiveness. Additionally, the concept suggests involving law enforcement for valuable insights and patterns, and exploring ways to account for errors and improve accuracy. The aim is to create an open, collaborative system that continuously improves with public input and helps achieve desired societal goals. (Summarized from 3090 words.)

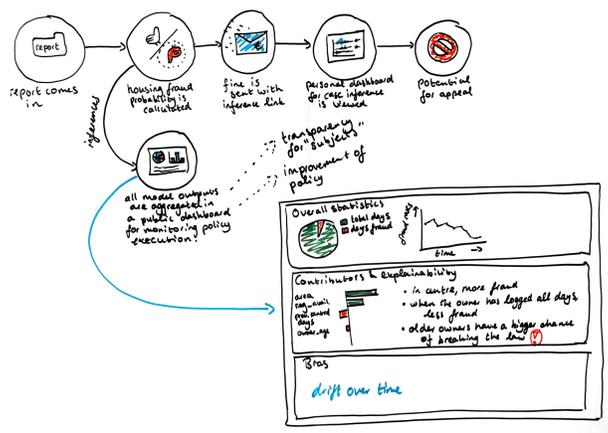


Figure 6: Concept design 1 from workshop 4 (C4.1). Aims to address fraud cases and provide transparency and accountability. When a report indicates potential fraud, a fine is issued, and the person can access an “inference dashboard” showing the factors influencing the decision. The key addition is a “monitoring dashboard” that operates on an aggregated scale, visible to both policymakers and the public. This dashboard has three parts: (1) Overall statistics: Showing the proportion of normal days versus fraud cases over time, giving context and aiding policy adjustments. (2) Contributors: Highlighting features and their impact on the system (e.g., location or age) with explanatory statements for better policymaking. (3) Bias overview: Monitoring model drift and bias evolution over time to identify potential issues. By implementing this system, it becomes possible to steer policy decisions based on data, promote fairness, and build public trust in algorithmic processes. (Summarized from 388 words.)

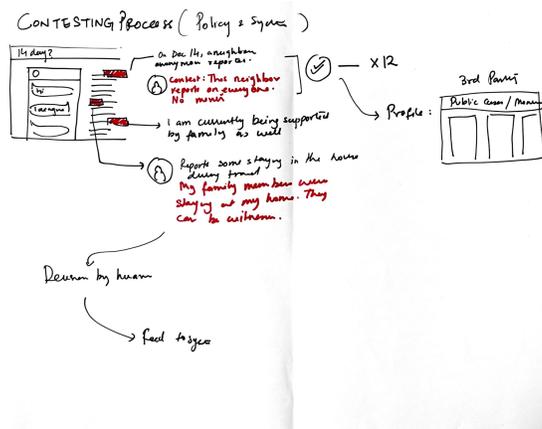
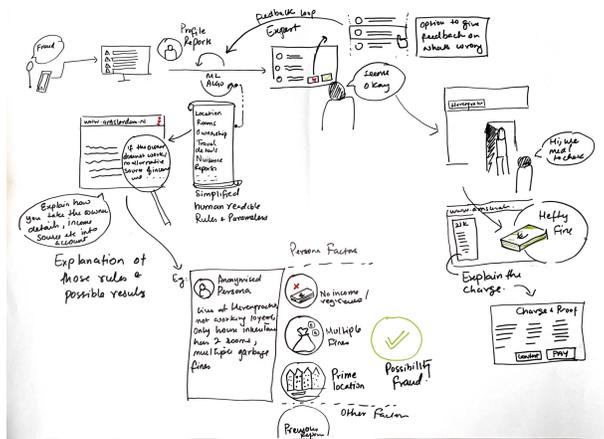


Figure 7: Concept design 2 from workshop 4 (C4.2). Focuses on creating an explainable and transparent process for handling fraud reports. The system aims to allow citizens and experts to understand how the algorithms work and how decisions are made. To achieve this, the system would route algorithm parameters and results to a website where they can be presented in a more human-readable manner, using anonymized personas to illustrate examples. The process involves human intervention, where experts can review algorithm-generated results and provide qualitative feedback, which is then recorded in the system. If a charge is levied based on the algorithm's assessment, the system ensures that the charge is clearly explained, and the individual has the option to contest it before paying. The contesting process is designed to allow citizens to challenge decisions with relevant evidence. However, some challenges remain, such as determining time limits for contestation and addressing potential biases in the system. Implementing the system requires considering policy implications and finding ways to aggregate and present bias-related data to address recurring patterns. Overall, the design concept aims to empower citizens by giving them agency in influencing decisions made by the algorithmic system while ensuring transparency, fairness, and accountability throughout the process. (Summarized from 1197 words.)

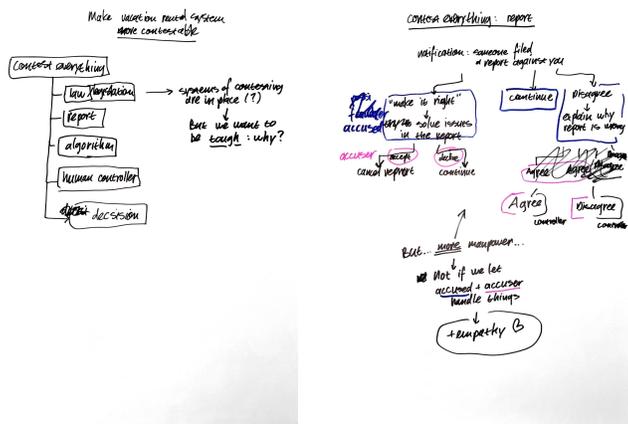


Figure 8: Concept design 3 from workshop 4 (C4.3). Involves contesting various aspects of the process. It begins with legislation, suggesting the need for more empathy and understanding rather than just imposing fines. The second step is contesting the report, where the accused is notified and given options to make things right before the process starts. The accuser can also receive feedback and decide whether to proceed or not. Instead of relying solely on a human controller, the idea is to let the accuser and accused work things out together, fostering empathy and understanding during the process. The final aspect is contesting the algorithm, allowing individuals to challenge the analysis provided by the algorithm along with the report. This approach aims to improve the system's fairness and effectiveness while promoting collaboration and empathy between parties involved. (Summarized from 643 words.)

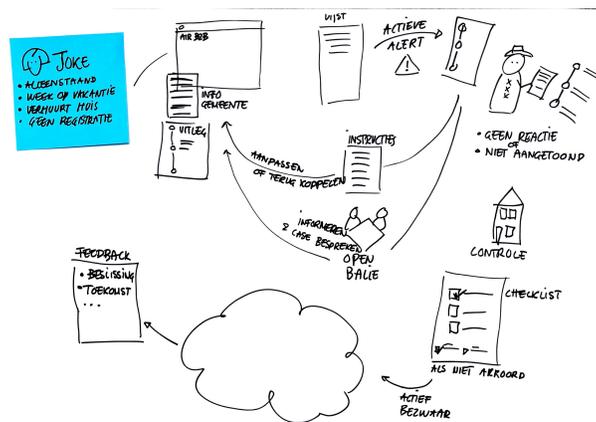


Figure 9: Concept design 1 from workshop 5 (C5.1). Involves identifying vulnerable individuals related to Airbnb rentals and using a step-by-step process to handle potential issues. The primary target is "Joke," an Airbnb host who unintentionally commits fraud by forgetting her registration number. The system triggers an alert and provides instructions for her to rectify the situation. If she fails to respond or disputes the decision, a handhaver (enforcement officer) intervenes. During this process, there are opportunities for her to provide feedback and challenge the algorithm's decisions. It's essential to give enough time between steps to accommodate adjustments and feedback. The system aims for transparent decision-making and swift resolution while considering the user's circumstances and ensuring fairness. (Summarized from 949 words.)

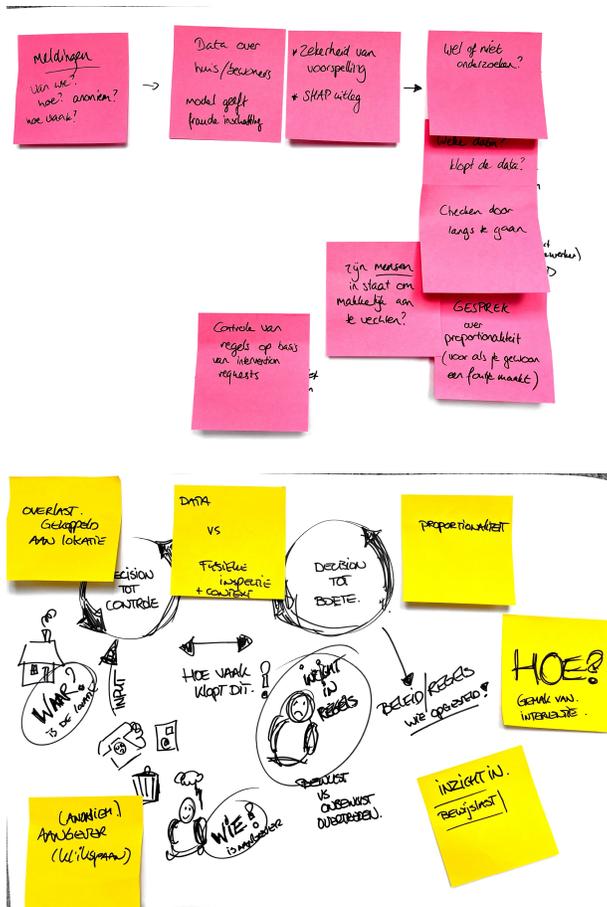


Figure 10: Concept design 2 from workshop 5 (C5.2). Involves a circular process. It begins with collecting reports of potential issues or disturbances. The system aims to make this information transparent to the affected individuals, allowing them to verify its accuracy. Using the collected data about the property and people involved, the model then assesses the likelihood of fraud. When a staff member reviews the case, they have access to the prediction's certainty level, aided by explanations from SHAP. The system facilitates communication with the affected person without immediately accusing them of fraud, allowing for validation and potential corrections. The decision-making process also considers human judgment and explores whether policies need adjustment. Two key stages are determining whether to initiate an investigation and deciding whether to impose a penalty, with a focus on ensuring proportionate consequences for intentional versus unintentional errors. The system aims to provide insight into the rules while minimizing administrative burdens for individuals and allowing intervention when needed. (Summarized from 863 words.)