Course Project for
Human Factors in AI

Managing Colony Collapse Disorder in Commercial Beekeeping Businesses

Ruth Shacterman
Oct 15, 2023

Task Analysis

- Product description
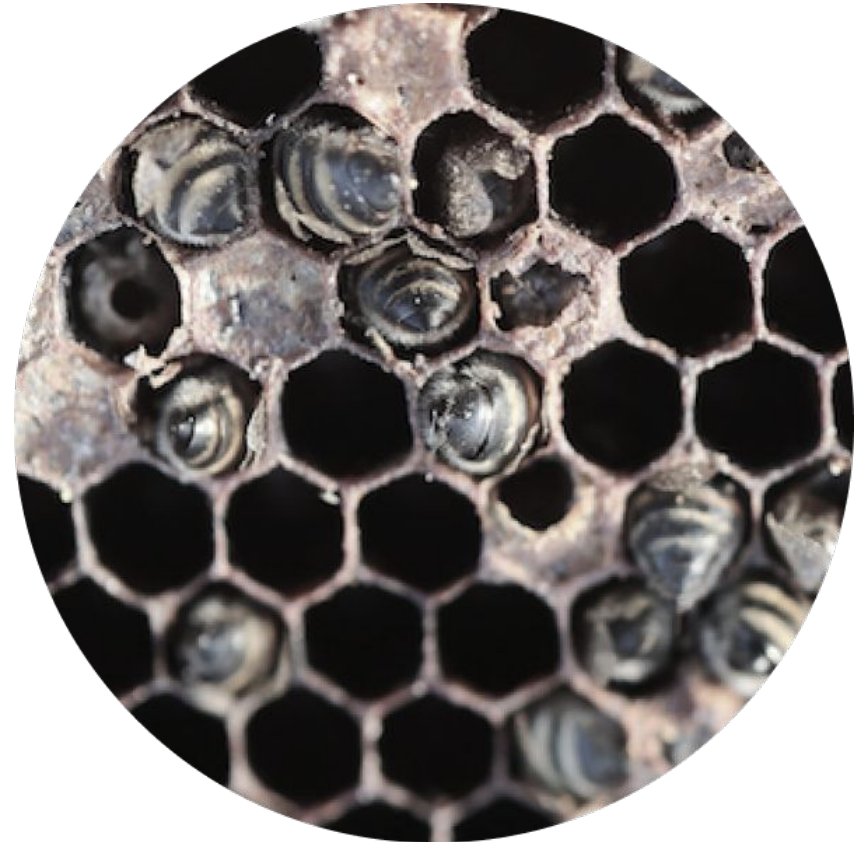- Problem description
- Task Analysis
- Insights

What is Colony Collapse Disorder?

Healthy Hive

Colony Collapse Disorder (CCD)

A Commercial Beekeeping Business

Commercial Beekeeping business (appx)

Hives: ~5000 hives
Beekeepers: ~10 full time

User
Beekeeper, Beekeeper business owner

Tasks
- Early detection of CCD
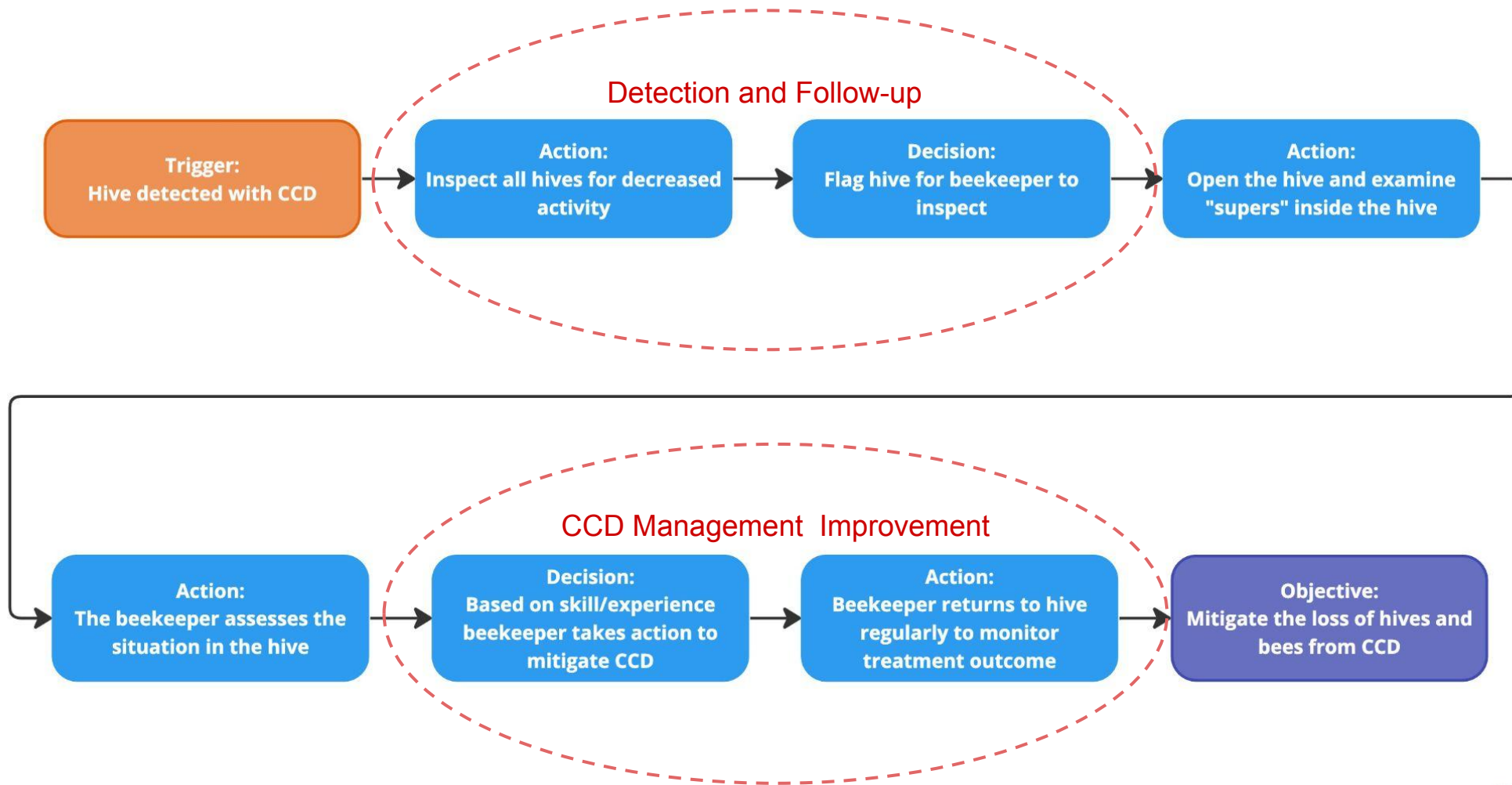- Better management of CCD outbreaks

Opportunities

- Monitor, Detect, and flag CCD
- Assist with assessment and management of CCD (Human Factors in AI)

Why is CCD a Problem worth solving?

- CCD is a significant concern for beekeepers and agricultural communities because honeybees play a crucial role in pollination. In the US, honeybee pollination accounts for over $15 billion in agricultural products each year.

- Early detection and monitoring can provide valuable data on the progression of CCD. This information can help Beekeepers with strategies to prevent and/or better manage the CCD.

Task Analysis and Insights

Detection and Follow-up

Trigger:
Hive detected with CCD

Action:
Inspect all hives for decreased activity

Decision:
Flag hive for beekeeper to inspect

Action:
Open the hive and examine "supers" inside the hive

CCD Management Improvement

Action:
The beekeeper assesses the situation in the hive

Decision:
Based on skill/experience beekeeper takes action to mitigate CCD

Action:
Beekeeper returns to hive regularly to monitor treatment outcome

Objective:
Mitigate the loss of hives and bees from CCD

miro

Opportunity Evaluation:

Sample Numbers:

Hives in medium size commercial business:
5,000 hives

Average inspection time per hive about:
6 min

Average inspection per month:
2

CCD mortality rates per season:
X < 5% considered low (100 hives)
X > 20% considered high (1000 hives)

Very rough estimates

Regular Inspection:

Expert Beekeeper hours per month:
6min x 5,000 hives)/60min x 2
= 1000 Hr/Mon

Note: Regardless of number of CCD cases, if number of hives remains more or less constant

With Detect and Alert App:

Given a "bad case" of 20% cases to inspect
(15min x 1000) /60
= 250 Hr/Mon

Notes:
- *On demand alerts, no need for recurring inspection*
- *Added 12 min for longer average inspection time*
- *Cases (< 20%) total hrs will decrease as well.*

Success Criteria

Outcomes	Metrics	Targets
Improve the beekeeper's CCD Detection Process	**Outcome**: Increased detection stats and reduced CCD related labor and cost.	**Outcome:** Reduction in CCD detection labor/cost >= 60%.
Maximize the effectiveness of expert beekeeping skills.	Improved assessment and treatment results	Colony mortality with ML <= mortality with the expert inspection.
Offer diagnostic options		
Offer treatment suggestions		
Improve diagnostics and suggestion	**Output**: The product uses a Binary Classification Tree model (0/1) for CCD detection.	**Output:** Detection probability approaches 100% (with a possible bias towards false positive)
	Output include prediction confidence level.	

Task Analysis - AI opportunity summary

The idea is to augment the Beekeeping workflow by automating tasks which require pattern detection from large data:

- Early detect of CCD outbreaks through prediction (using a binary classification tree model and sensor input)
- Assist the beekeeper with assessments and strategies to mitigate outbreaks (most likely using a regression model for suggestions)
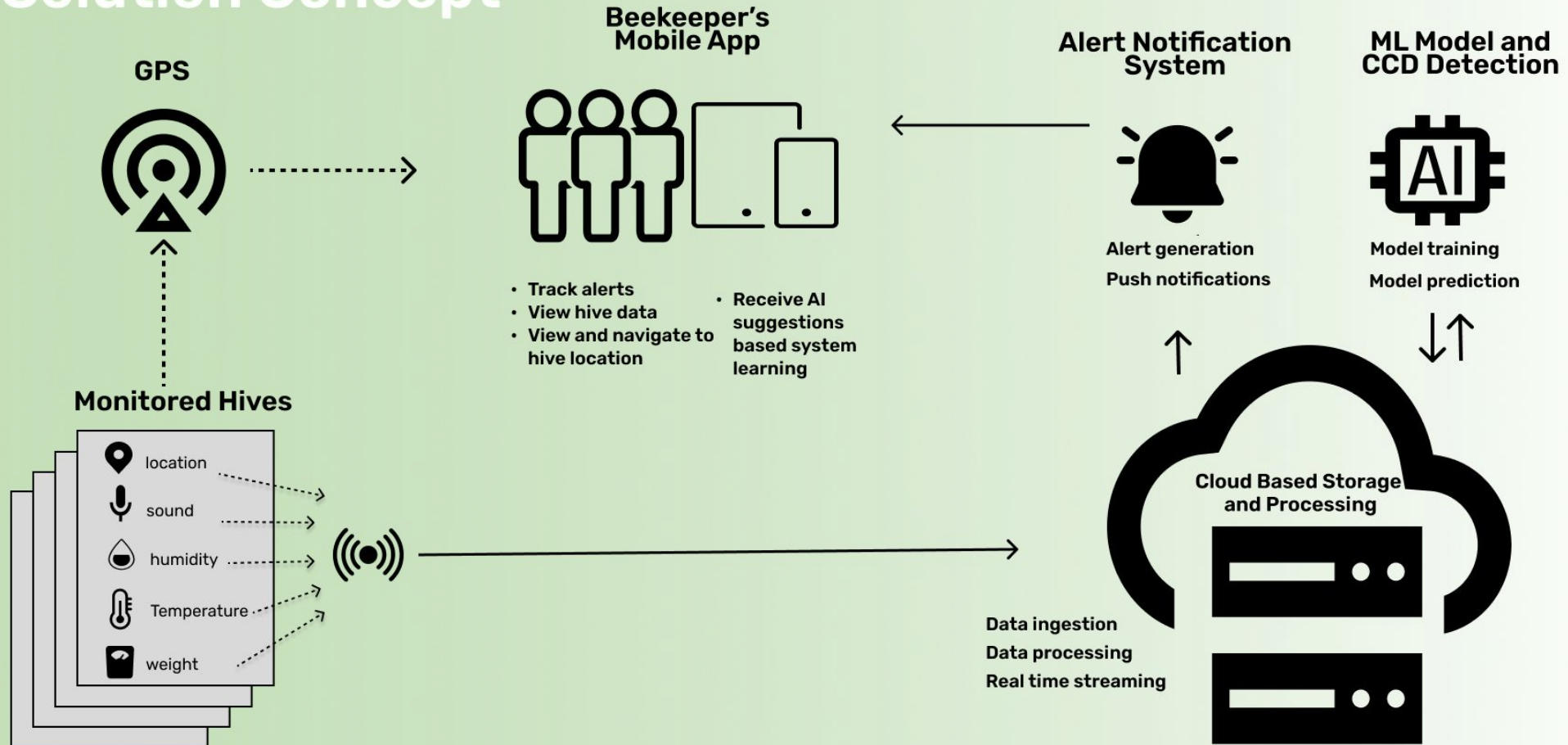
This product uses an ML- driven mobile app, and sensor data to

- Detect and alert Commercial Beekeepers to the onset of Colony Collapse Disorder (CCD)
- Assist the beekeepers with CCD management
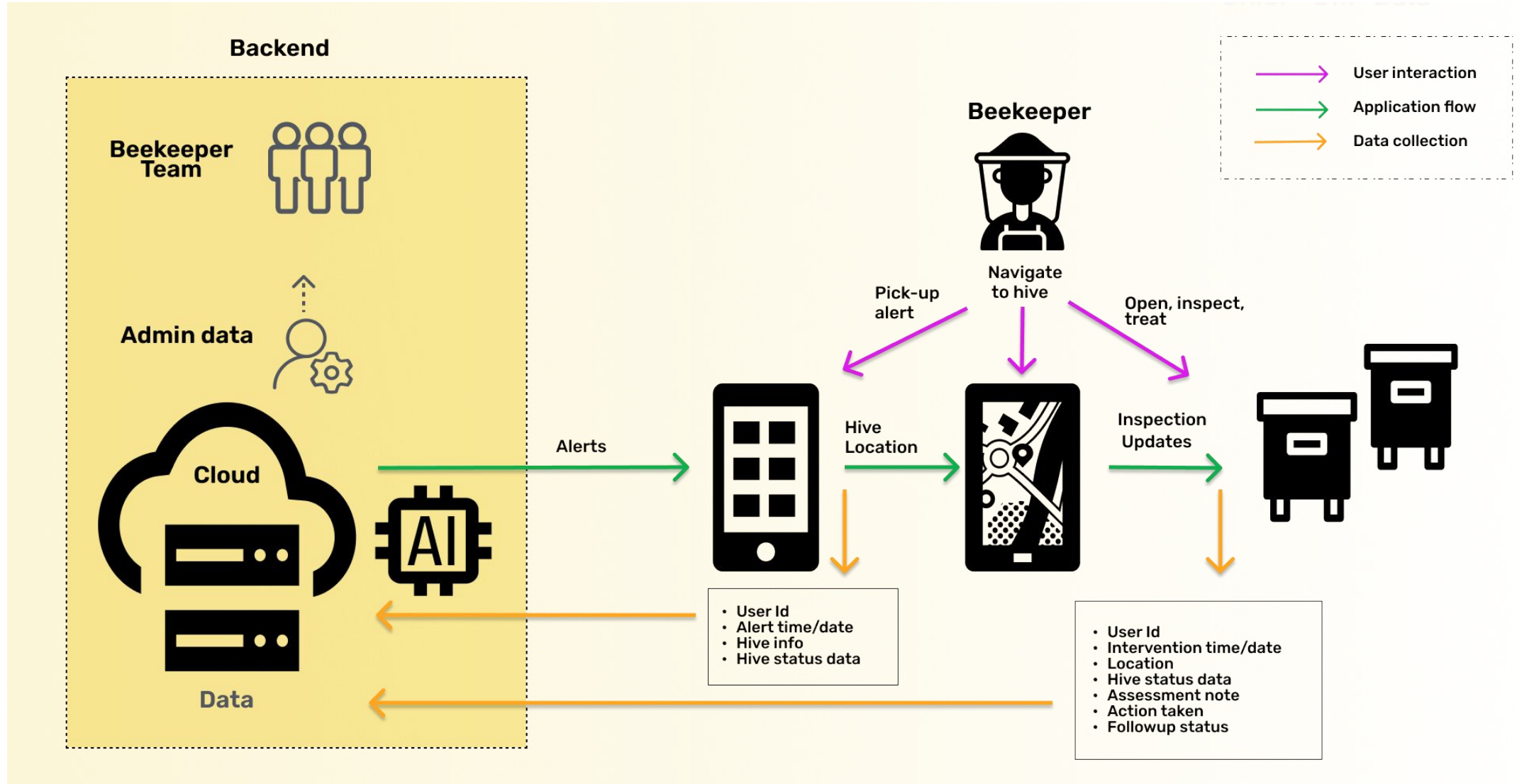
The Product - Solution Concept

User Experience Design

User interactions with my proposed solution

Key considerations of AI product user experience design

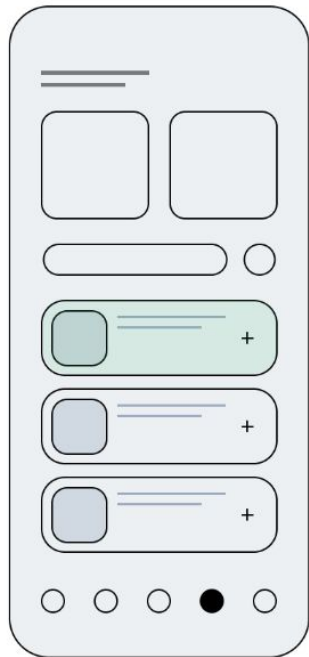# User Experience - The Human Factor

User Experience and Inputs

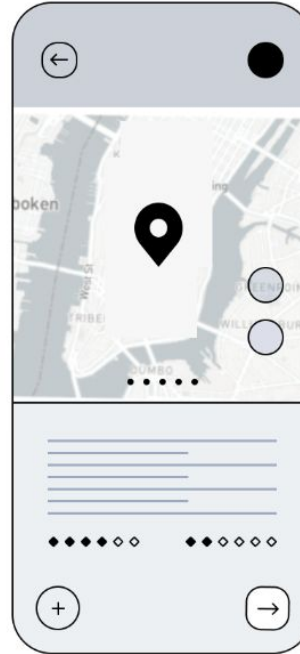**1. CCD Alerts**

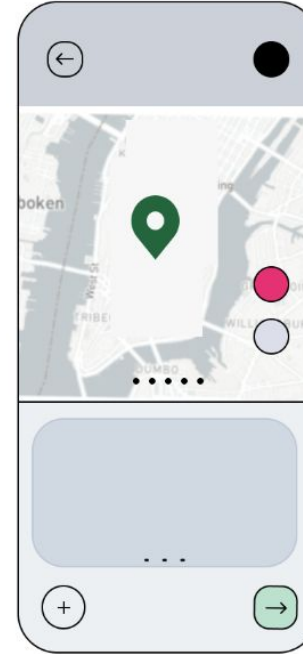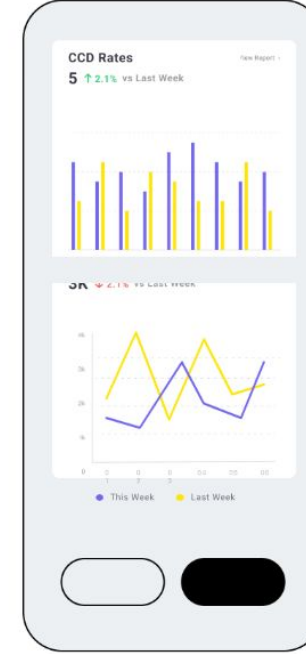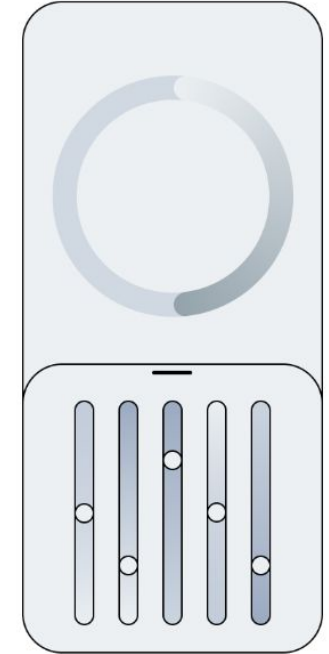**2. Hive Locate, Assess, Update**

3. Analytics

List
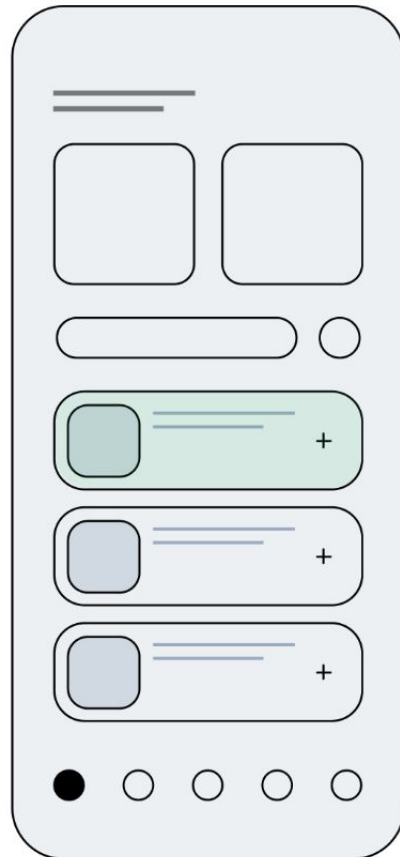
Details

Hive locate

Capture intervention

CCD trends

Alert workflow

User Experience - CCD Alerts

Alert List

Alert details

System Captures

- Beekeeper's Id

- Time/date of assignment

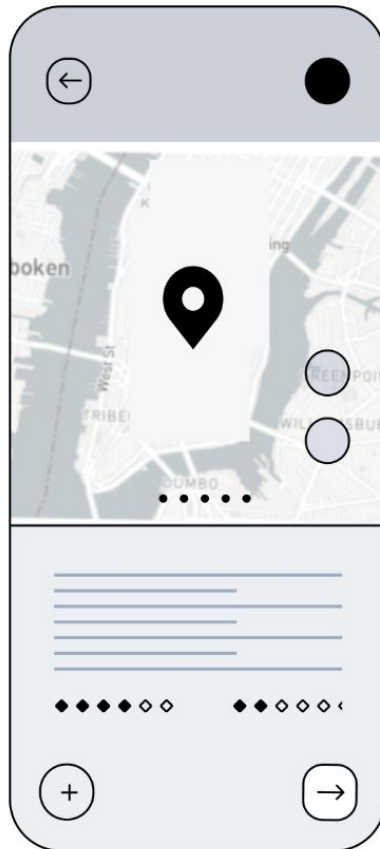- User's current location

System Displays

- The hive info

- Hive location

- CCD diagnosis with probability

- Hive status

- Intervention history

Locate

User Experience - Hive Visit

Hive locate
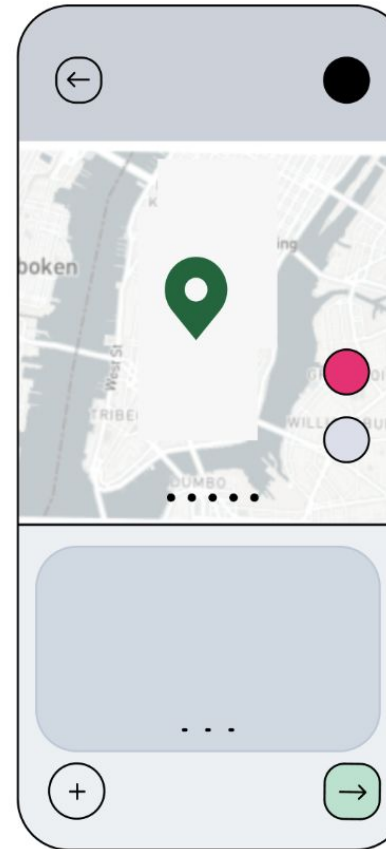
System Displays

- The hive's exact location (map)

- Current hive readings

- Suggested assessment options

- Treatment suggestions

Capture intervention

System Captures (learning)

- Beekeeper's Id

- Time/date at the hive

- Current hive readings

- Beekeepers selected assessment (from list)

- Selected treatment

- Beekeeper notes and uploaded photos

- Updated alert status

User Experience - Learning trends

Product uses the captured data and AI to generate:

- Alert status report

- An intervention process report based on user response timeline

- Intervention followups

- "Effectiveness of intervention" report

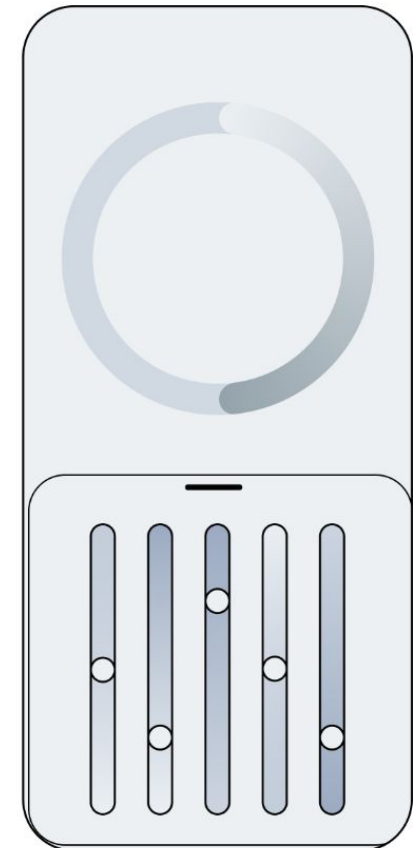**It enhances the algorithm to**

- Improve the assessment suggestions

- Improve the treatment suggestions

- Learn more about CCD

CCD trends

CCD Rates

5 ↑ 2.1% vs Last Week

3K ↓ 2.1% vs Last Week

- This Week
- Last Week

Alert Status / Process

User Experience - ML Feedback Loops

This CCD App uses **implicit feedback loops**, where the Beekeeper's interactions with the app, and the follow-up sensor data, influence the outputs (predictions, assessments, suggestions) that the Beekeeper will receive form the system over time.

With implicit feedback loops it is important to be on the lookout for biases that are introduced through unaccounted for elements in the UX and in the process flow for example:

- The order in which options are presented (i.e. first option is most often selected).
- The level of difficulty of a suggested treatment option (causing it not be selected)

Preventing bias from implicit feedback loops is an ongoing process that requires continuous monitoring, testing, and improvement. It would be good practice to inform the Beekeeper of the possible limitations of the model and his/her role and responsibility in the hive management process. This could be done by implemented reminders, and by using infographics on the capture the beekeeper's diagnostic input.

Privacy Considerations

Privacy-related concerns stemming from the data collected
The applicability of privacy laws and privacy obligation

PII - Why is the User Id needed?

For Workflow and Collaboration

- Assigning and tracking alerts
- Follow-up reminders
- Identifying and following up on issues
- Treatment inquiries and clarifications

For Product UX

- Filtering and sorting options:
  - i.e. filter by my open alerts
- Process reporting:
  - Show me all treated hives?
- Follow-up alert status
  - open, in progress, completed

Learning

- The user workflow binds the hive, the CCD predictions, the sensor reading, response times, the selected assessment and treatment to **outcomes.**
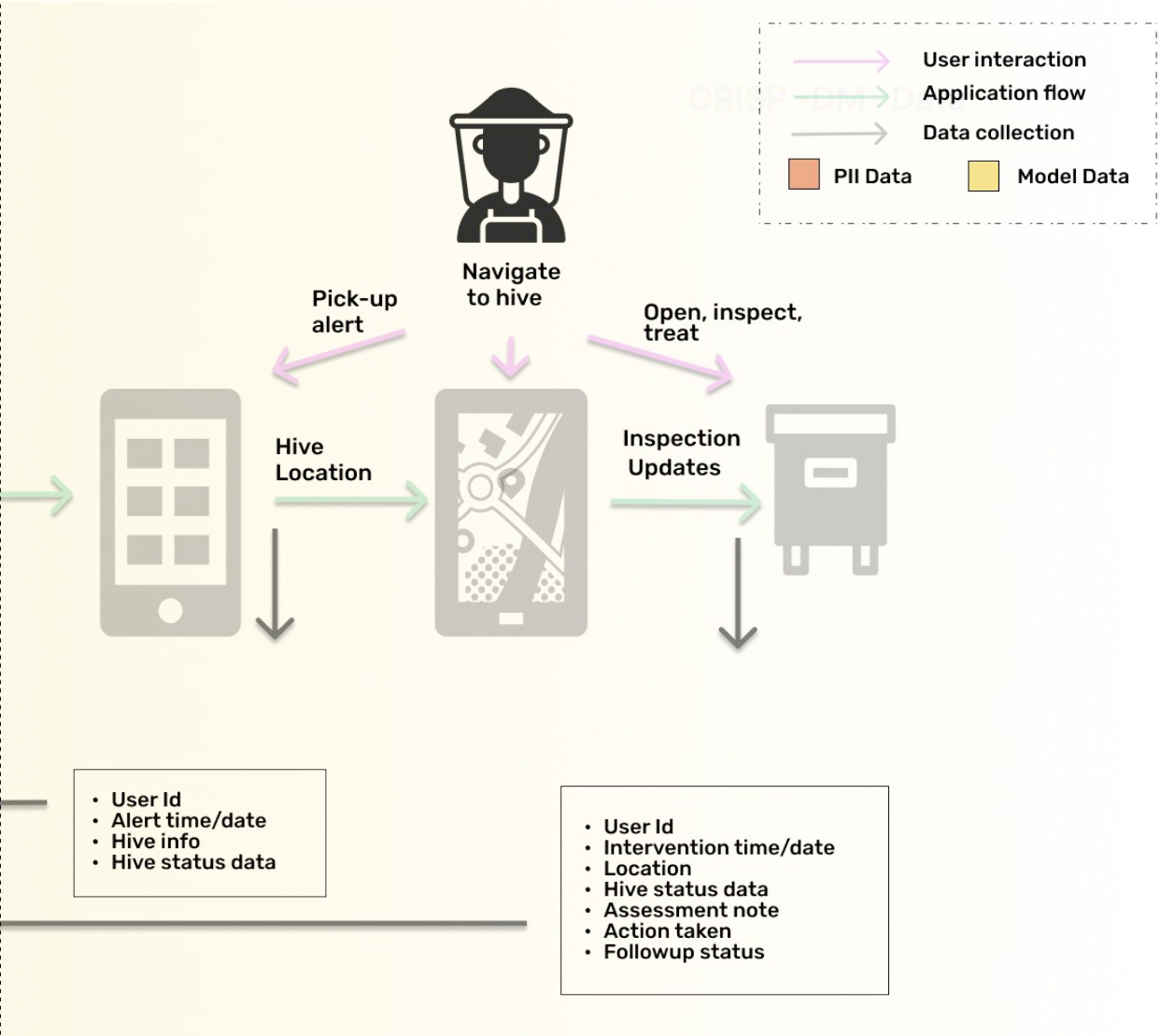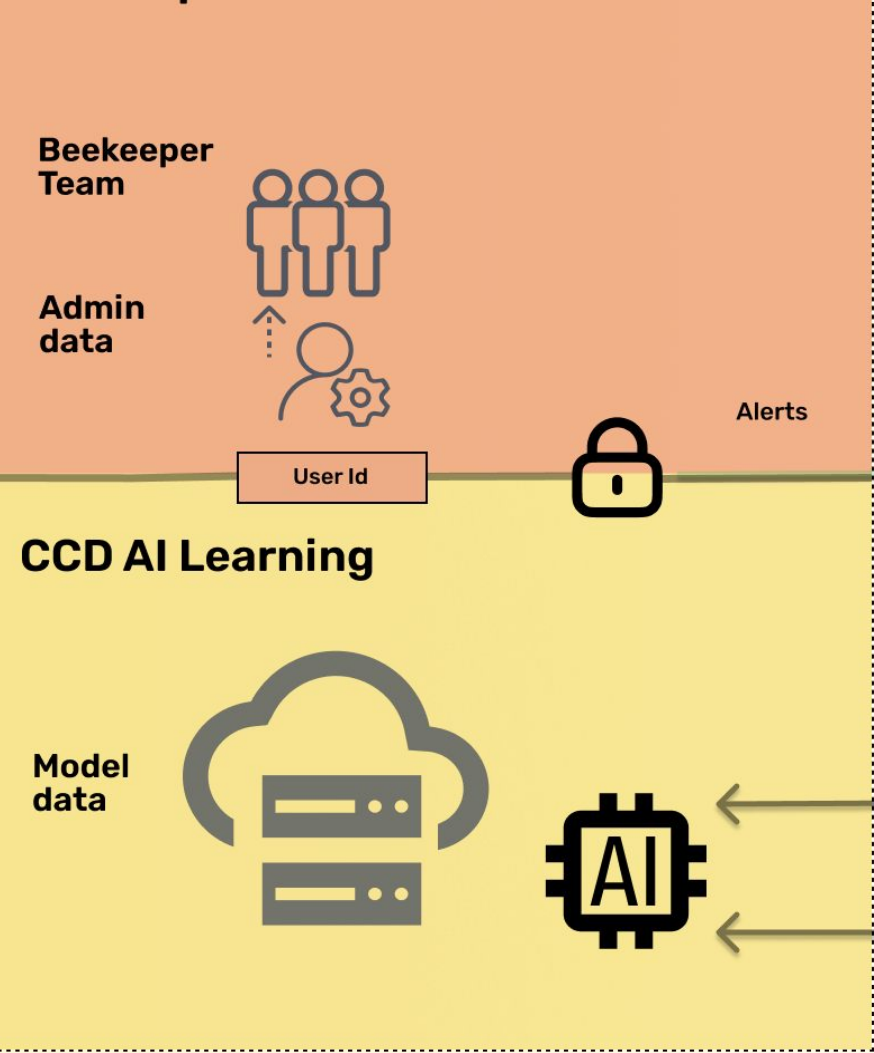
Approach for the learning Model

- Do not store an identifiable user Id in the AI learning model.
- Use an instance identifier that is not related to the UserId
- Encrypt User Id when passing data back to the workflow model

Data Privacy - CDD App Overview

CCD App - Privacy by Design?

1 - Beekeeper workflow

AlertID/UserID(?)
User location
Hive Location
Alert Status (open, close)
Date time of the alert
Date time of the assignment
Date time of the intervention

2 - CCD AI Learning

Hive temperature
Hive humidity
Hive sound
Hive Weight
Assessment
Treatment
Duration
Hive Location
Success status
Environmental conditions

miro

FIPs Considerations in CCD App

The product will use User Id only to drive the front-end application and user interface with the product.

Because off this, PII information lifecycle, control and management will not be center to this product's learning model.

| Rights of Individuals | ? Information Lifecycle | ? Control on Information | ? Management of PPI |

FIPs - Privacy Rights of Individuals

However, since user location, response time, and input will be captured, the user has the right (Rights of Individuals) be informed of why this information is collected and how it will be used.

The product will follow the FIPs recommendations

- The user will be notified that the data above is collected and how it will be used.
- That this information will be captured, used in calculation, but not stored as personal data.
- The data collected will only be used for the purposes presented above.
- The data will be visible to the user in the application interfaces.
- If needed, differential process methods will be used to prevent linking the UserId to user's PII.

Impacts on Data Privacy
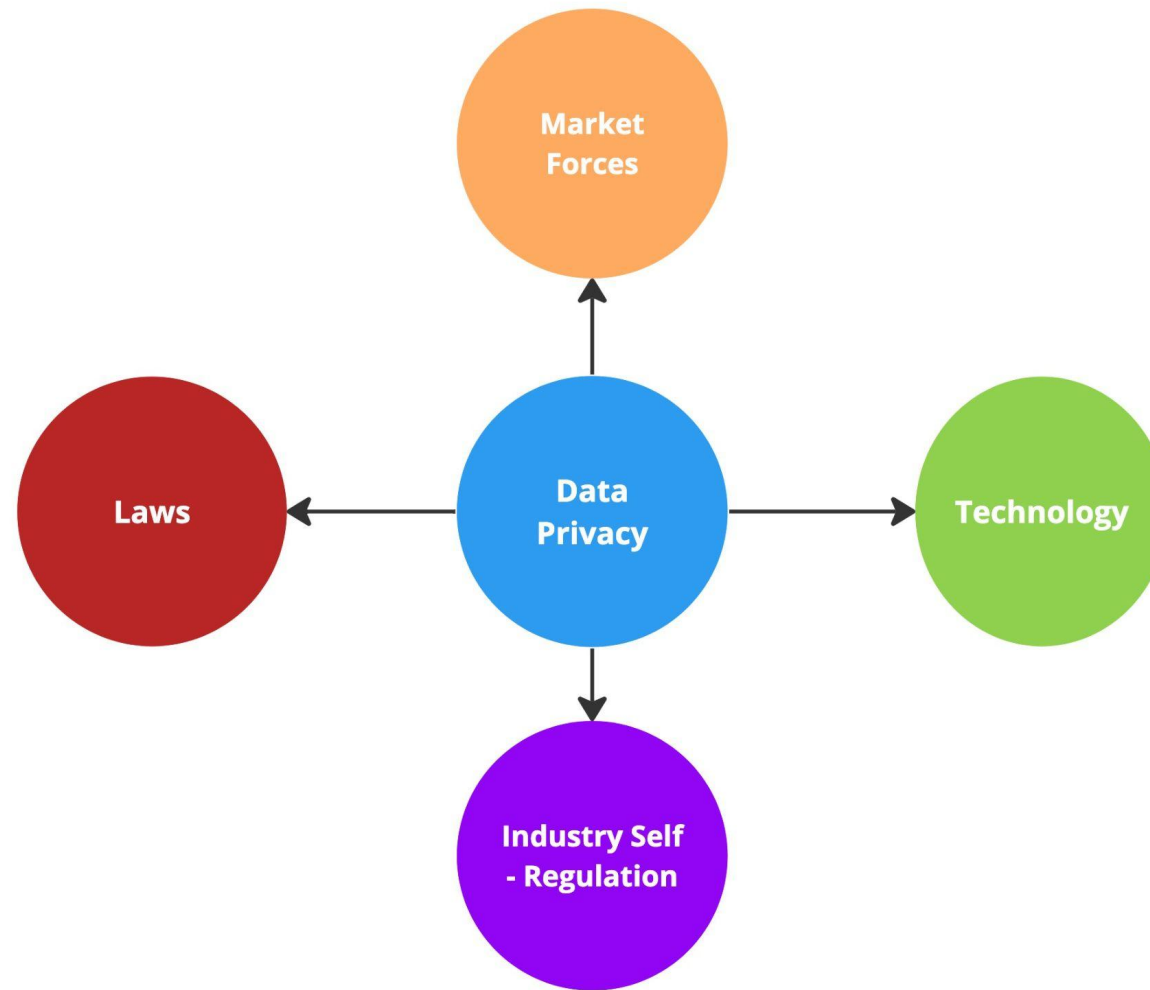
Privacy considerations

Technology

Privacy by design: Separating user workflow from learning data.

Data encryption: Encrypting the exchange of user identifying information between the models.

Data minimization: Only the minimal required data is collected and stored.

User access and data portability: users can access and download all data they entered in the product.

Laws

- The CCD App falls outside the main regulated privacy areas: Health, Finance and Education.
- Therefore, the product is not subject to mandatory laws such as **HIPPA, FCRA & GLBA, or FERPA**
- **GDPR** should be considered

Market Forces

The App tracks user activity and treatment outcomes. Would market forces push for user performance tracking? If so what would be the implications? (see next section)

Industry Self Regulation

- General data Privacy and Security measures
- User Consent for data collection
- Data retention limit and deletion
- Guidelines for data sharing and collaboration

Ethical Consideration

Main potential sources of bias

How the solution might meet the goals of ethical AI (fairness, accountability and transparency)
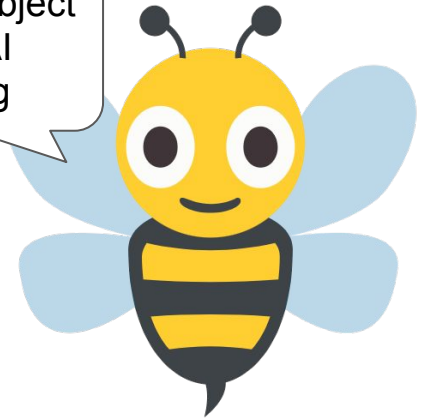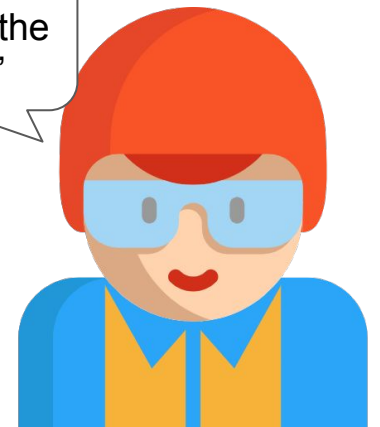
Ethical Considerations in CCD App

Risk of Allocative or Representational Harm

The "CCD Detect Alert and Manage" application is not optimal for discussing Ethical Biases in AI. The reason is that the main subject of learning in the honeybee, not the human. Human factors such as Race or Gender are not incorporated into the learning algorithm. The algorithm is at of very low risk of Allocative or Representational Harm.

I am the subject of the AI learning

Deployment Bias

For the reason above, I would like to start this section with a discussion of possible deployment bias risk for the CCD system, and use this in my following discussion of ethical consideration.

I am the instrument, the "sidekick"

Sources of Bias

Deployment Bias

Deployment bias refers to a mismatch between how a tool was intended to be used when it was developed and how it is used in practice.

The CCD application tracks Beekeeper (user) workflow activity and system outcomes. For the industry, it could be tempting to use this information for "tracking" Beekeeper performance, even though the original intent behind the tool.

If market forces push to evolve in this direction, we would need to look more closely at possible allocative and representation biases based on the distribution of gender and race in the Beekeeping profession.

CCD App - Deployment Bias

1 - Beekeeper workflow

AlertID/UserID(?)
User location
Hive Location
Alert Status (open, close)
Date time of the alert
Date time of the assignment
Date time of the intervention

2 - CCD AI Learning

Hive temperature
Hive humidity
Hive sound
Hive Weight
Assessment
Treatment
Duration
Hive Location
Success status
Environmental conditions

3 - Performance tracking

UserID
AlertID
User Location
Date/time of alert
Date time of assignment
Assessment selections
Treatment selection
Success status

miro

Ethical Considerations in CCD App

Fairness considerations

As the product is designed there is no issue of fairness. However, if the product evolves to incorporate performance tracking of Beekeepers, then questions of fairness as related to gender and race as these relate to performance and hiring practices of beekeepers will require consideration.

Accountability Measures
- **Preemptive:** Define an ethical checklist when assembling the model, and document with a datasheet for the dataset to be used
- **Values:** A commitment that any bias reported as caused by the algorithm/data should be addressed promptly, and the user affected should be updated.
- **Responsible:** The product should have a built-in performance check. In addition, the PM should assign a representative for the timely handling of bias issues.
- **Recourse:** A user should be able to reach this responsible person through a contact form or number that is clearly displayed in apps "Contact Us" section.

Ethical Considerations in CCD App

Transparency
Transparency is built into the CCD application, since all user data collected is displayed in the workflow and in the process reports.

In case that Beekeeper Performance Tracking is added to the application, we will take the following measure for transparency:

- Create an information section for performance scores and predictions
- Share the model information - most likely a simple linear regression model
- Share the features used and their relative importance. For example
  - Response time to alert
  - Number of hive visited during a period.
  - Treatment success rates
- We will also provide an input interface for users to enter more specific questions and assign a Role in the product or CS teams to address such inquiries (could eventually evovle into a FAQ page).

The End