

TOPIC DETECTION in App Reviews

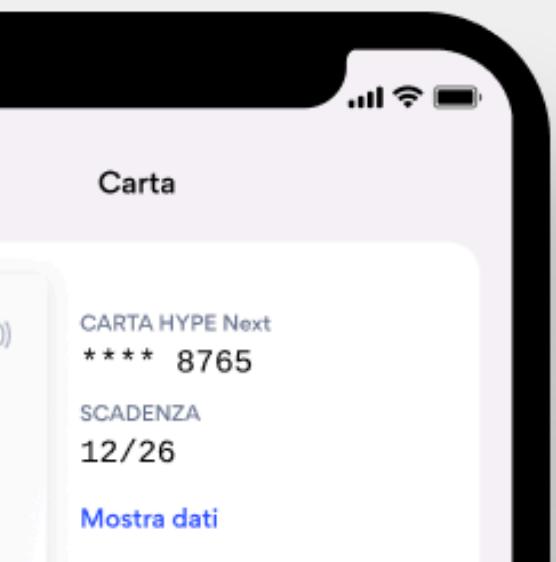
A Data-Driven Solution for HYPE



Alessia Ciccaglione - s344499
Amirhossein Lotf Ranaei - s323336
Sonia Foco - s343043

Identifying the **PROBLEM...**

Hundreds of HYPE app reviews are collected every month, mixing feedback with **complaints**, **bug reports** and **feature requests**.



OBJECTIVES of Our Solution:

KEY FEATURES

Automatically analyzes every HYPE app review with EDA to discover patterns

Detects **negative** topics (bugs, complaints, usability issues)

Identifies **recurring problems** and emerging trends

Tags each review **consistently** and at scale

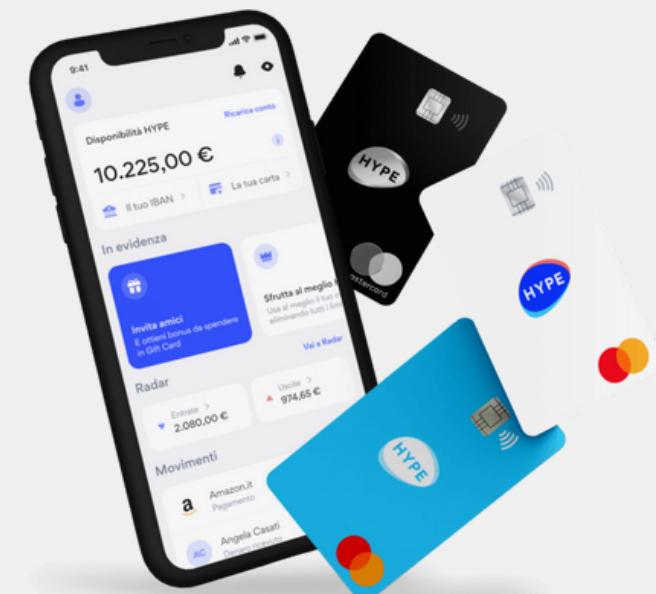
BENEFITS

Faster insights for the Research & Customer Insight team

Reduced manual reading and tagging **effort**

Summarizes large volumes of feedback in seconds
Improved visibility on customer **pain points**

AI TOOL FOR FEEDBACK AUTOMATION





Defining the VALUE PROPOSITION

8 DECENT WORK AND
ECONOMIC GROWTH



By using automated NLP, we enhance feedback analysis **efficiency**, reduce **manual effort**, and **accelerate** decision-making. Real-time insights from user reviews enable data-driven product improvements, **driving growth** and fostering a **more sustainable** business environment.

9 INDUSTRY, INNOVATION
AND INFRASTRUCTURE



By utilizing cutting-edge **AI technology**, we automate topic detection, improving **operational efficiency** and modernizing product development processes. This supports the growth of a more **data-driven** and **technologically advanced infrastructure**.

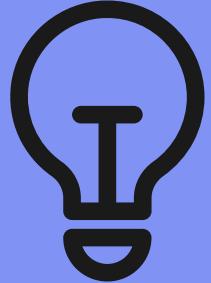


The Main TASK



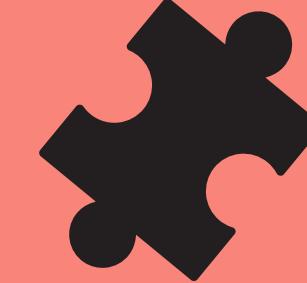
1. EDA & Pre-processing

Exploratory data analysis and preprocessing of text and features.



2. Topic Detection

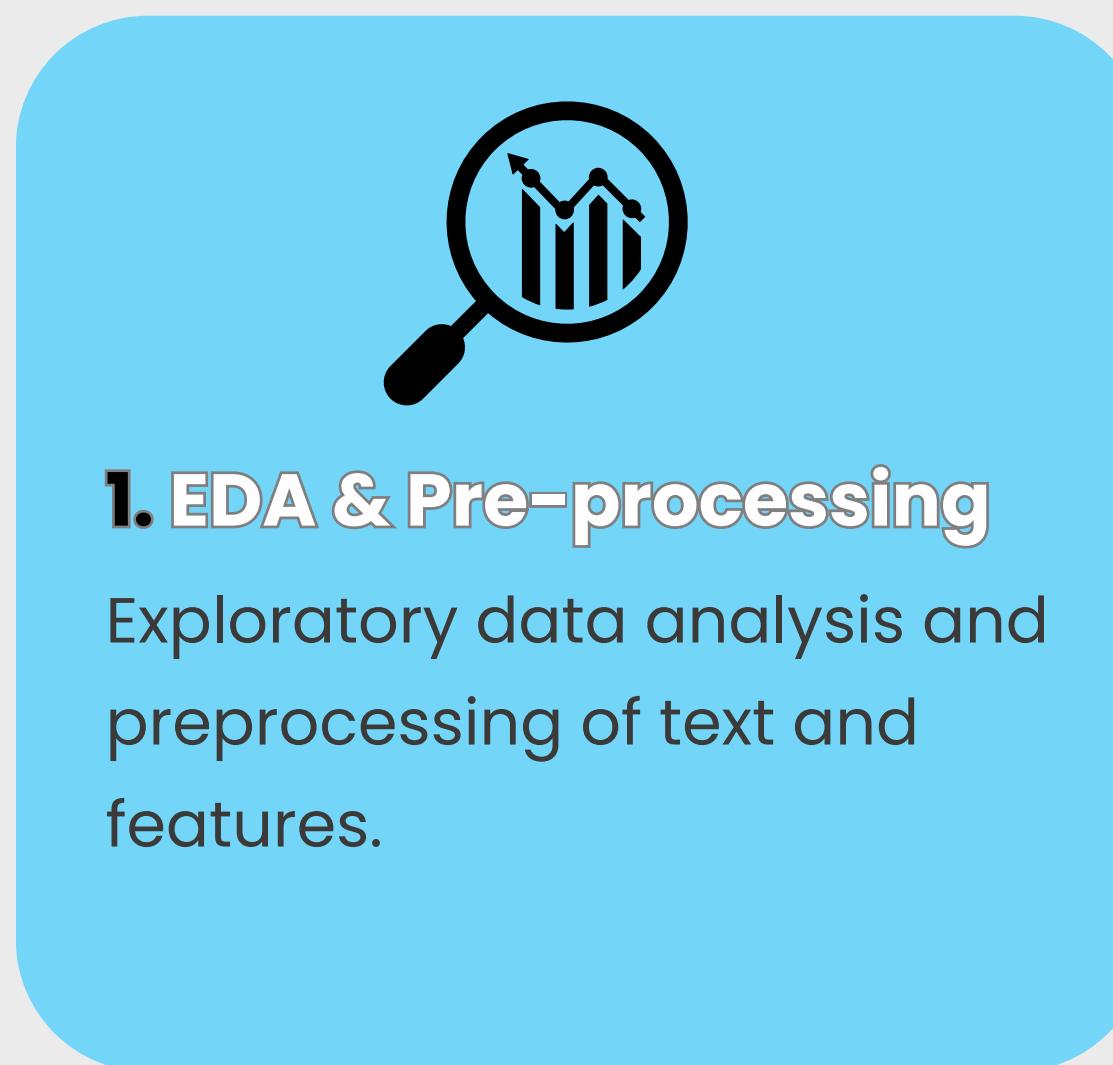
Detecting topics and comparing them with the HYPE taxonomy.



3. Extra Tasks

Future work: multi-label classification and hierarchical topic detection.

Task 1: EDA & PRE-PROCESSING



Handle **junk reviews** and **outliers**

Manage reviews written in **different languages**

Understand how to treat **neutral and positive** reviews compared to **negative** ones

Perform initial **analysis** on already assigned labels

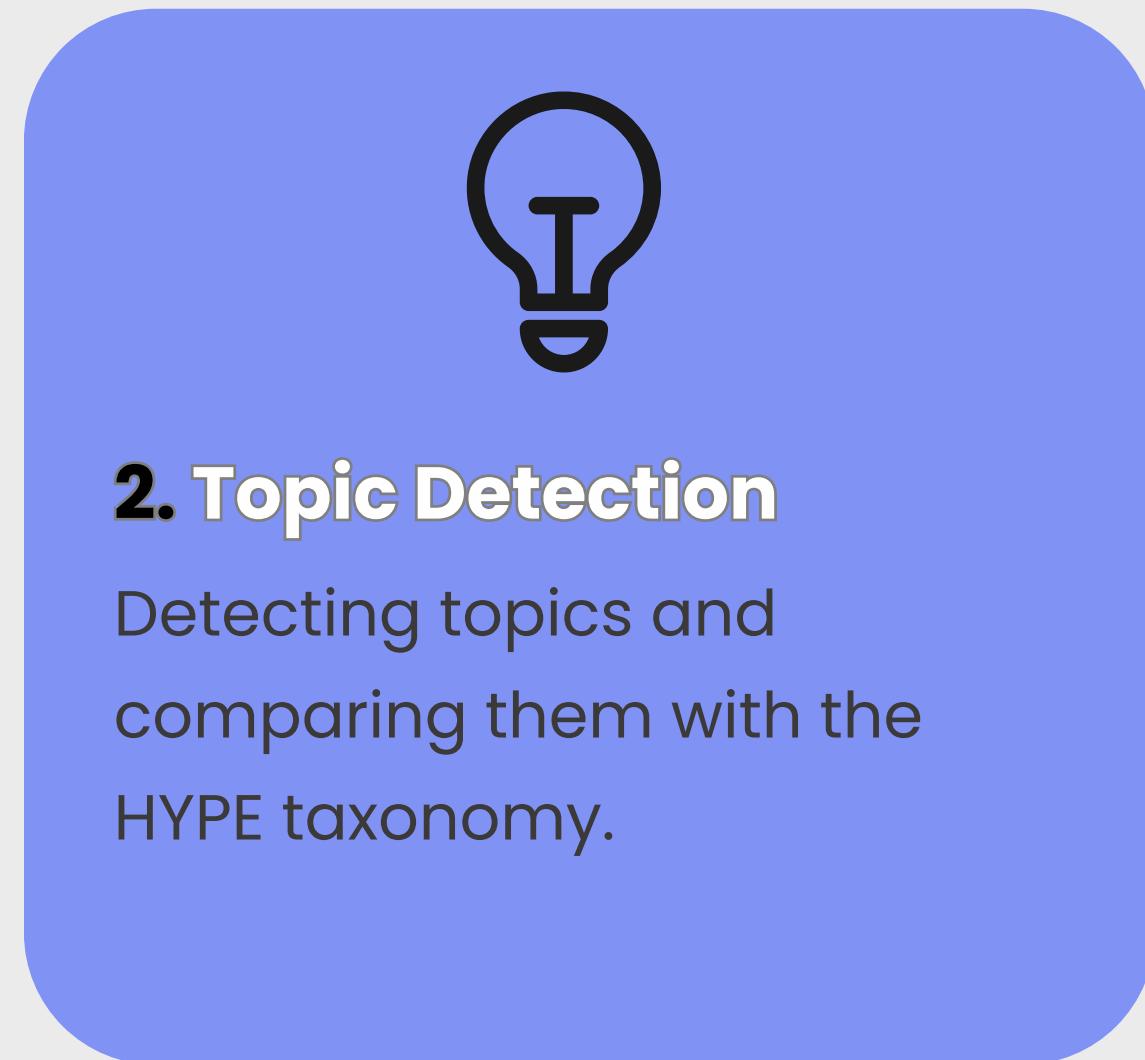
Understanding which parameters are **critical** (e.g., the number of ratings)



Task 2: TOPIC DETECTION

UNSUPERVISED APPROACH

TAXONOMY COMPARISON



WHY?

Useful for identifying topics that **change** or **emerge** over time, adapting to new patterns in user feedback.

Supports the **dynamic adjustment** and enhancement of the current topic taxonomy.

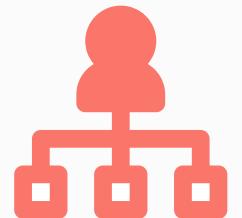
Ensures consistent and **efficient classification** by automatically tagging reviews.

Task 3: EXTRA TASKS



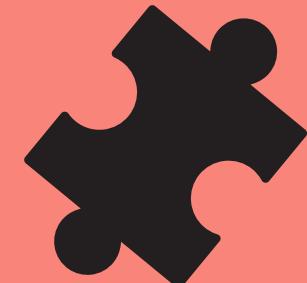
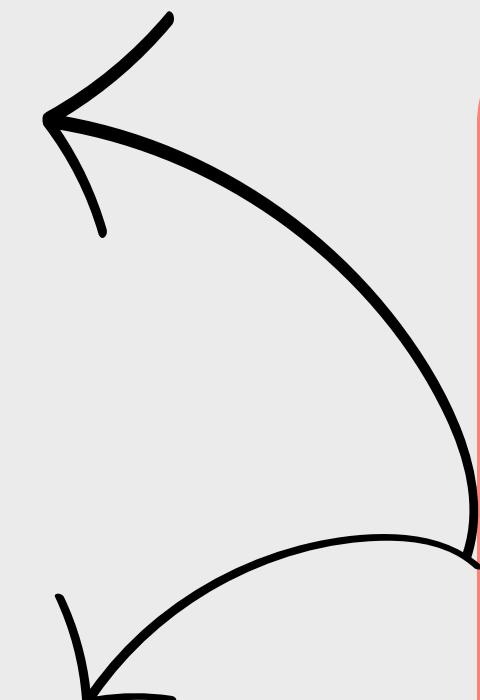
Multi-label Classification

Reviews often cover multiple topics, requiring classification into several categories at once.



Hierarchical Detection

Allows analysis of topics at different levels to better understand their relationships and importance.



3. Extra Tasks

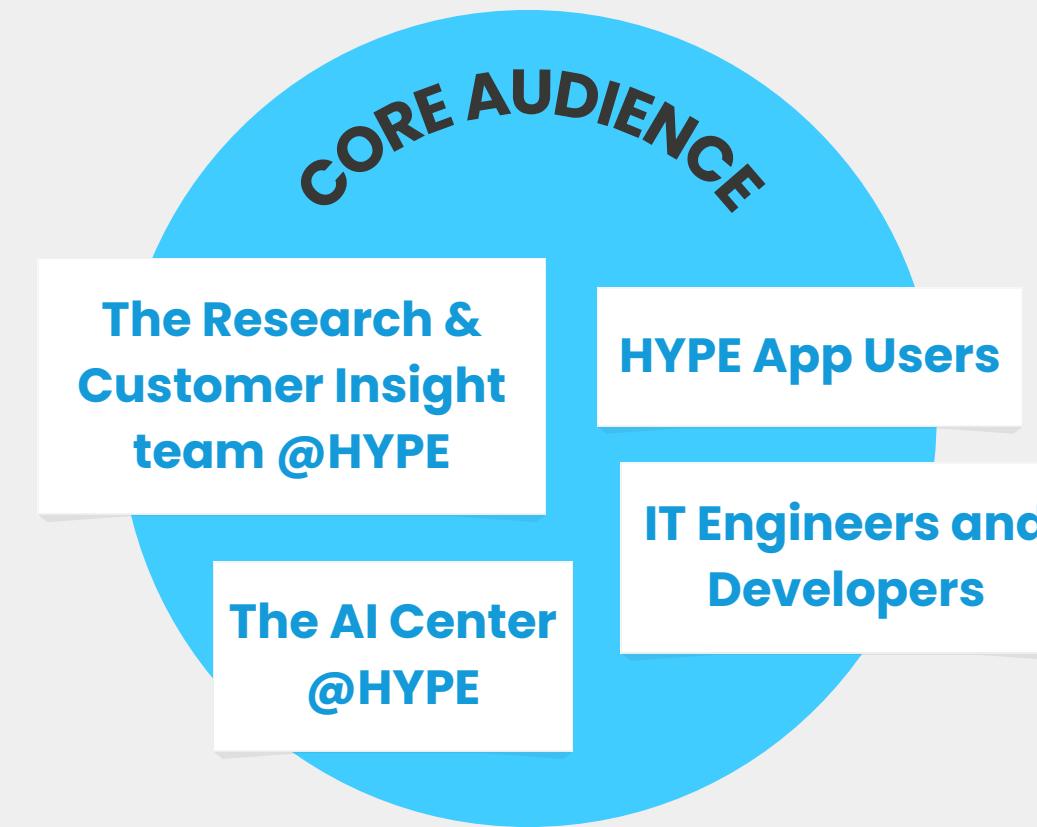
Future work: multi-label classification and hierarchical topic detection.

Aligning **USER NEEDS,** **DESIGN** and **STAKEHOLDER** goals for impactful solutions

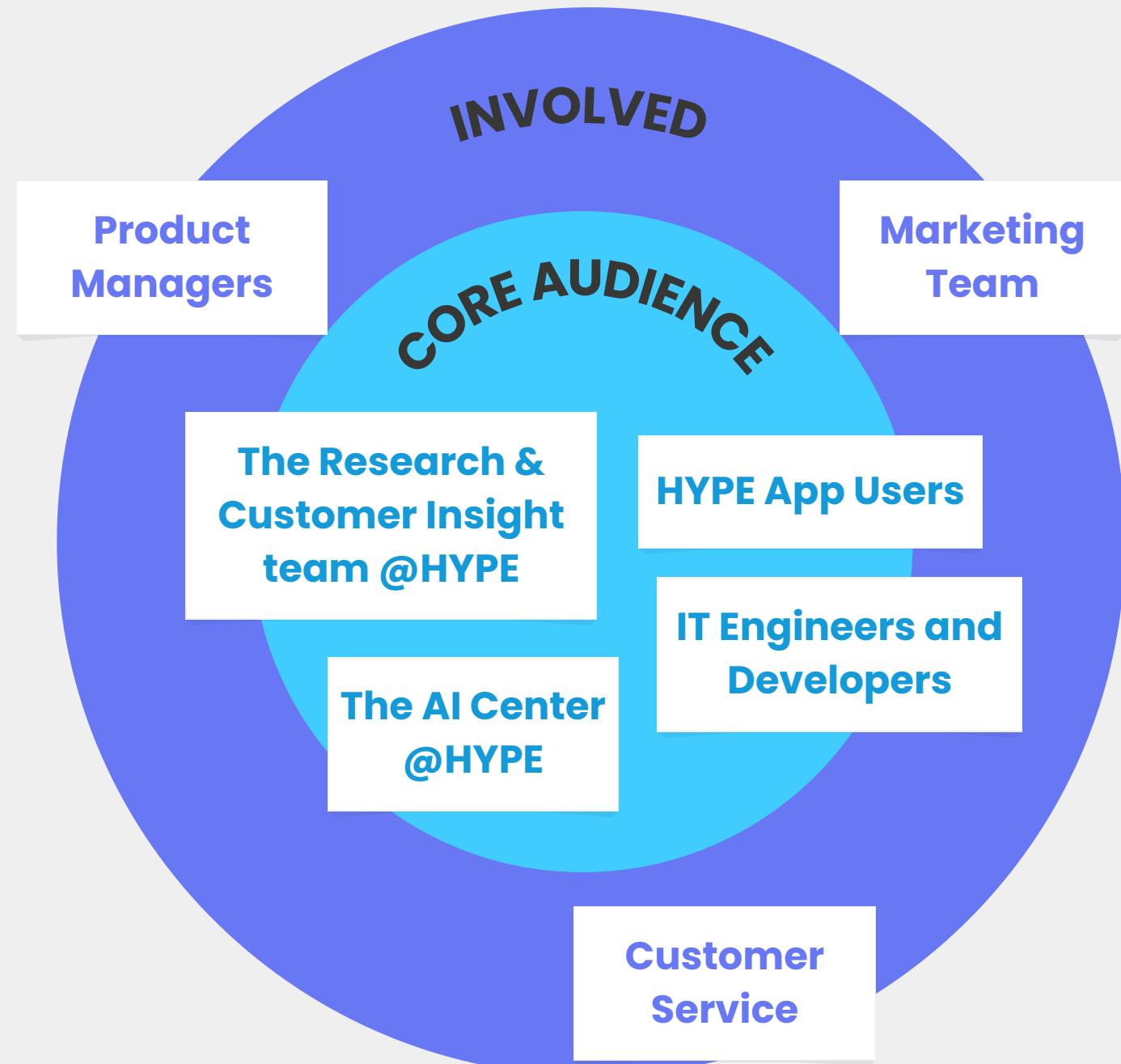
D
E
S
I
G
N



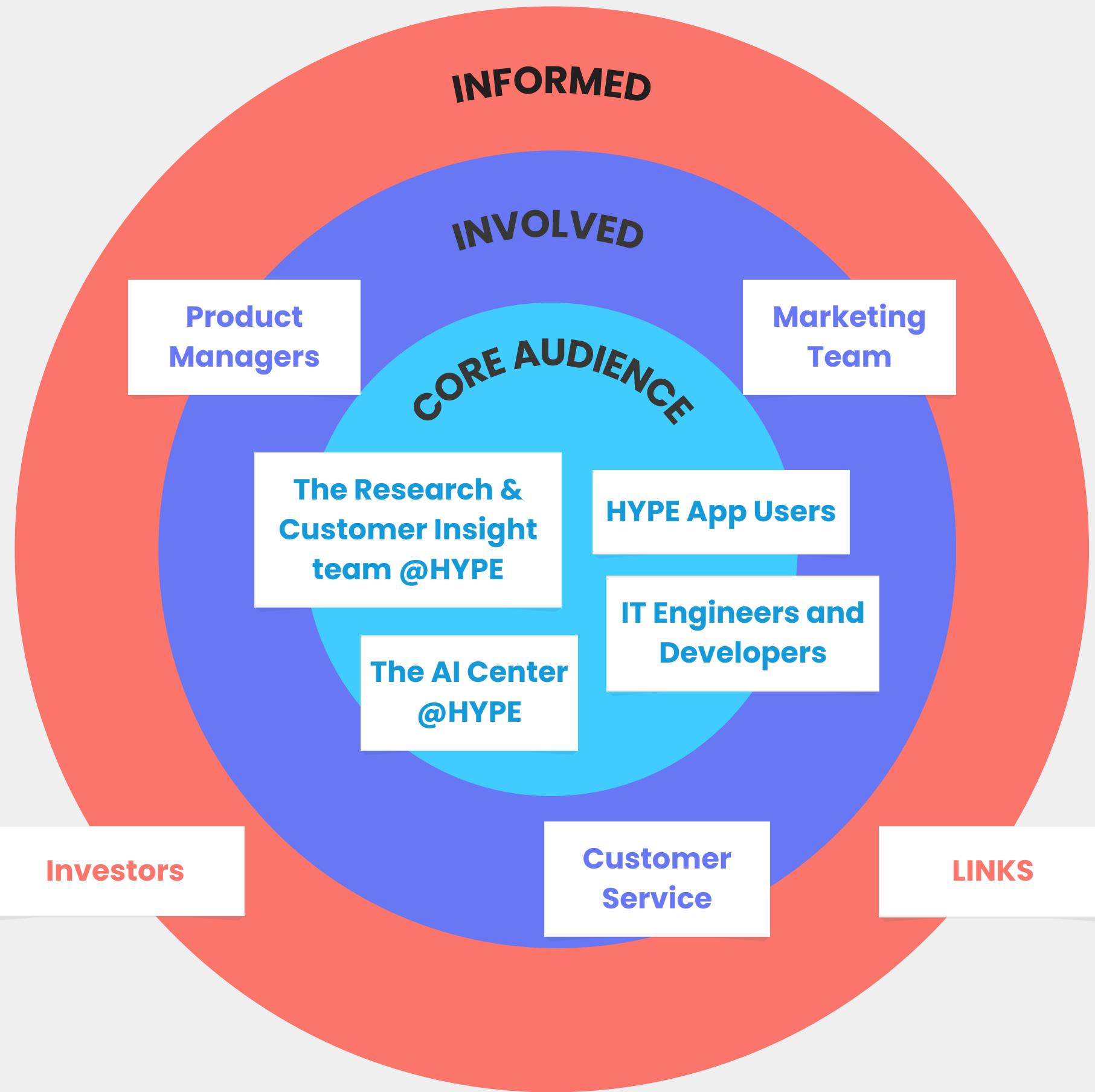
Stakeholders: CORE AUDIENCE



Stakeholders: INVOLVED



Stakeholders: INFORMED



Who We Designed For...

USER PERSONAS

Chiara



- **Role:** Analyst, Research & Customer Insight Team @HYPE
- **Goals & Needs:**
 - Detect recurring negative topics to quickly identify user issues.
 - Analyze topic frequency to help prioritize what needs to be fixed first.
 - Get faster insights while significantly reducing manual reading and analysis effort.
- **Tasks & Responsibilities:**
 - Refine model outputs by merging topics or discarding irrelevant ones, leveraging her expertise to ensure the system's results are accurate and actionable.
- **Frustrations:**
 - "Manual analysis is too time-consuming and difficult to scale."
 - "I can't tell if an issue is a small bug or an emerging trend."

User requirements

Mo

FUNCTIONAL

Detect: the system **must** detect recurring topics in user reviews, particularly negative ones, to support the identification of issues

S

Adapt: the system **should** be able to be updated with newer reviews and adapt to evolving new topics

Analyze: the system **should** provide analysis on how often each detected topic appears in the reviews, to help prioritize the most common issues

Co

Classify: the system **could** perform multi-label classification to tag reviews with the most relevant topics

Organize: the system **could** identify hierarchical relationships among topics

W

Process Real-time Data: the system **won't** process reviews in real-time as they are submitted

NON-FUNCTIONAL

Accuracy: the system **must** ensure that reviews are correctly tagged with topics, minimizing false positives or missed topics and ensuring reliability

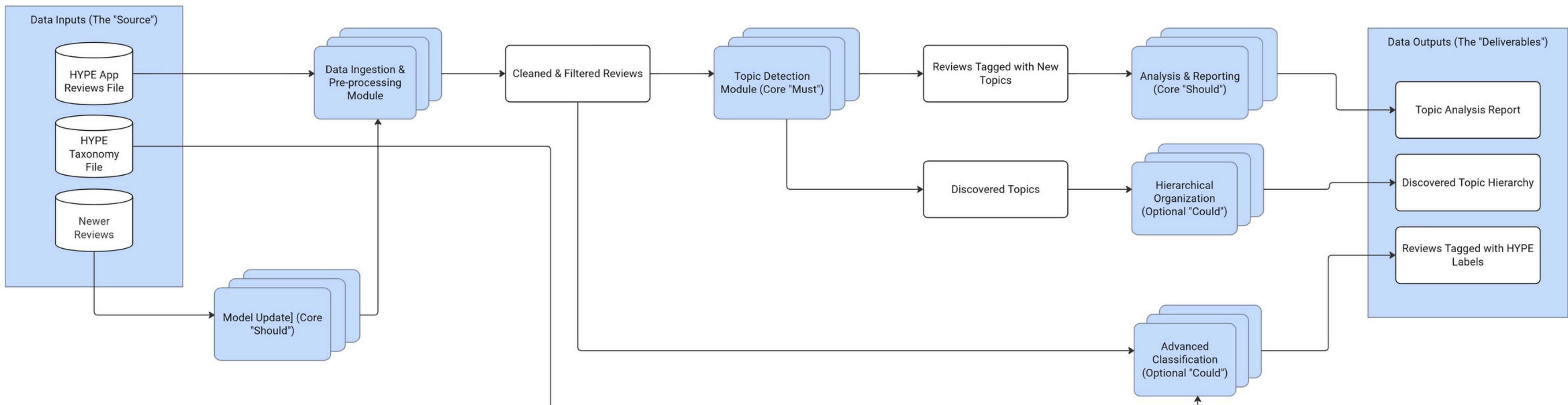
Robustness: the system **should** be robust by being able to process large volumes of reviews

Efficiency: the system **should** perform efficiently, providing fast results without requiring excessive computational resources or time

Modular Architecture: the system **could** adopt a modular structure, allowing components to be extended, replaced, or upgraded without affecting the overall pipeline.

User-Friendliness: the system **won't** provide a user interface, therefore it won't be designed to be user-friendly

Whole system FUNCTIONAL DIAGRAM



How we work as a TEAM



M
A
N
A
G
E





Project Value Proposition & Design Pillar

Defined the **value proposition** and structured the design pillars based on the **stakeholder map** and **user requirements**.



Project Objectives, Slide Refinement & Graphics

Outlined the project's **main tasks** and refined the **presentation slides**, including layout and visual design.



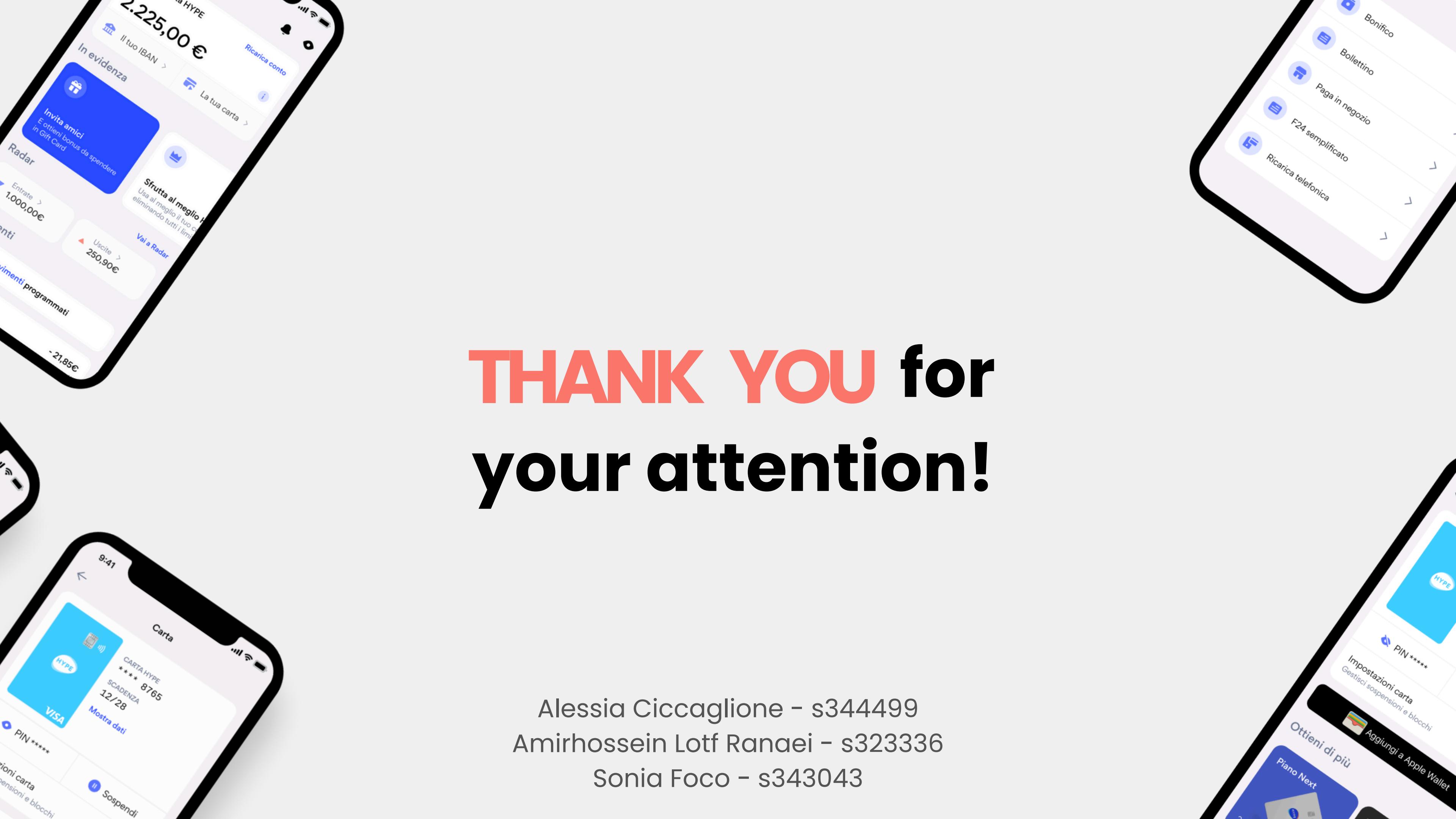
User Personas & Functional Diagram

Developed **user personas** and designed the system's **functional diagram**.

Stay tuned for
the next
MILESTONE...



Data preprocessing → EDA →
Research questions → Topic
Modeling overview



**THANK YOU for
your attention!**

Alessia Ciccaglione - s344499

Amirhossein Lotf Ranaei - s323336

Sonia Foco - s343043