AI Recognition Tool Design | Mindpetal

verify.io

Info Challenge '25

IC25076

Hansika Murugu and Nynika Badam

The challenge

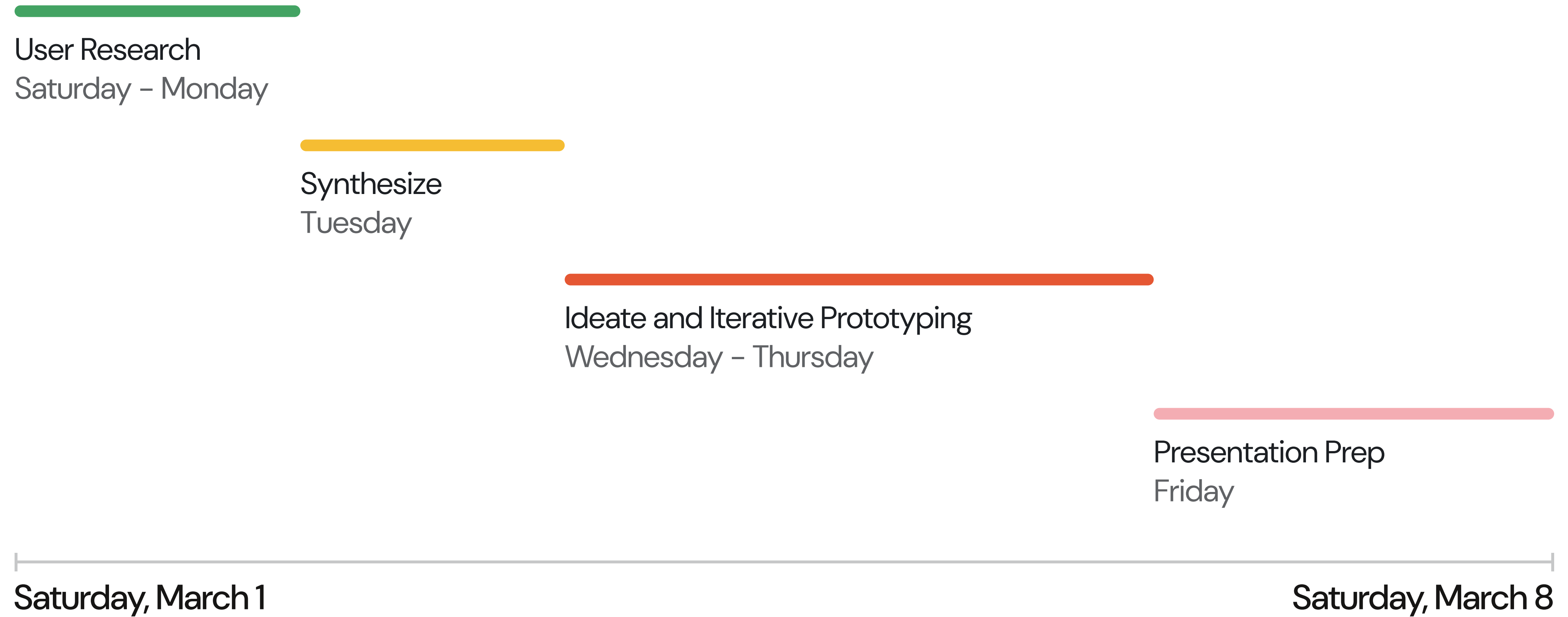# Design an AI-generated content checker tool

Specific goals

An AI-generated content checker that allows users to upload or input various media to determine AI involvement, providing insightful metadata

Design Priorities

- Intuitive user experience
- Comprehensive content verification
- Real-time and efficient feedback

Project timeline

User Research
Saturday – Monday

Synthesize
Tuesday

Ideate and Iterative Prototyping
Wednesday – Thursday

Presentation Prep
Friday

Saturday, March 1   **Saturday, March 8**

Agenda

1 ## Understand
Primary Research, Secondary Research

2 ## Synthesize
Competitor Analysis, User Personas

3 ## Define
Features, Interfaces

4 ## Implementation
Interactive prototypes

1

Understand

Primary Research, Secondary Research

We dove into research articles about how AI generated content is currently evaluated

49.9%
accuracy when people tried to differentiate between GPT-3 and human text

54%
of the time, GPT-4 text was identified as human

60%
average accuracy in the top 10 AI detector tools

We also looked into research on challenges in identifying AI-generated content

Do You Trust ChatGPT?

Huschens, M., Briesch, M., Sobania, D., & Rothlauf, F. (2023)

Enhance User Awareness of AI-Generated Content

An effective interface should clearly indicate the origin of the content, helping users make informed judgments.

Can you spot the bot? Identifying AI-generated writing in college essays

Waltzer, T., Pilegard, C., & Heyman, G. D. (2024)

Improve Detection of AI-Generated Writing

The interface should provide clear indicators or alerts when AI-generated content is detected, assisting users in distinguishing between human and AI authorship.

Enhance User Awareness of AI-Generated Content

The interface should offer explanations or evidence supporting its determinations to enhance user understanding and trust in the tool's assessments.

Conducting our own user survey allowed us to ask questions tailored to the tool

17

Responses

18–59

Age Range

3

Key Takeaways

Familiarity with AI-Generated Content

88%

of the respondents have *doubted the trustworthiness* of the content they consume

So, what type of content do they mostly consume?
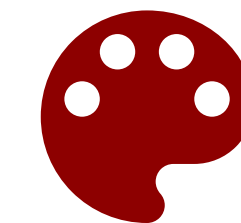
Social Media Posts
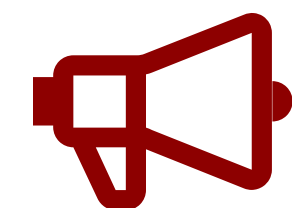
News Articles and Blogs

Emails and Messages

Learning Material

Along with the social media posts and news articles and blogs, surprisingly, respondents felt the need to verify the content from:

Creative Content

Marketing Content

Reaction to AI-Generated Content

majority of the respondents feel **neutral or somewhat confident** AI-generated information.

majority of the respondents would feel **more cautious** before sharing or using the content and **concerned** about its accuracy or reliability.

So, what factors did users identify as key to increasing their confidence in content credibility?

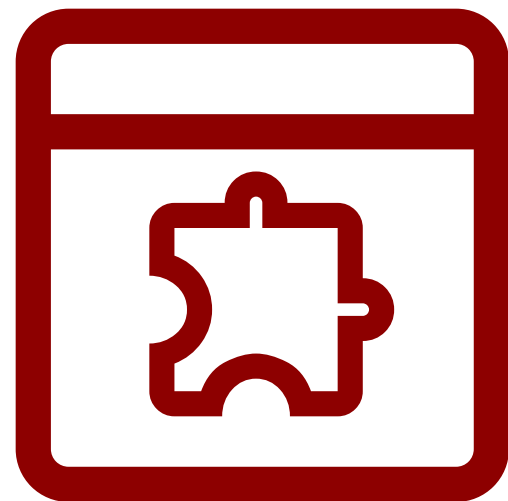**citations or links** to verified sources

transparency about AI's accuracy

clear explanations about AI's conclusions

Preferences of the Interface

88% of the respondents have would prefer to have some basic instructions or tooltips while using the tool.

majority of the respondents have believe that a website and/ or a browser extension would be the most effective for AI-generated content detection

70% of the respondents would look for additional verification or source if the content is AI-generated.

How might we...

 Design a non-disruptive AI-generated content checker for a wide range of users across different age groups who frequently consume digital content, so that they can seamlessly verify authenticity and gain clear, trustworthy insights without interrupting their workflow?

2

Synthesize

Competitor Analysis, User Personas

We tried to understand the current market to position ourselves accordingly.

	GPTZero	Quillbot	Scribbr	SightEngine	Hive	Deepware	Mindpetal
Instant Access Across Devices	✕	✕	✕	✕	✕	✕	✓
Multiple Media Type Support	✕	✕	✕	✕	✓	✕	✓
Convenient and Seamless	✓	✓	✓	✕	✓	✓	✓
Process Transparency	✓	✕	✓	✓	✕	✕	✓

We stepped into the shoes of our tool's prospective users to understand their pain points and needs

Woke Wanda, she/her
Age: 22
Occupation: University student

Goals

- Stay up to date on breaking news in politics, environment, and daily life
- Consume entertaining content like cat videos

Needs

- Accurate information from trusted news sources and influencers

Pain Points

- Has taken up to 30 seconds to realize that a video was a deep fake
- Thought that articles she was reading was written by reporters but upon checking saw it was AI generated
- No dependable AI detectors that work on her phone

Competency

AI content:

Not familiar ———●——— Very familiar

Not confident ———●——— Very confident
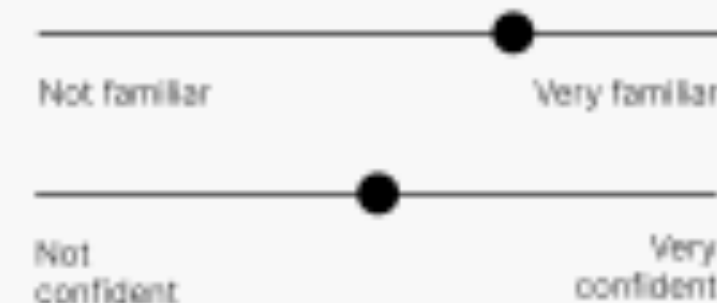
Vulnerable Vincent, he/him
Age: 55
Occupation: Software Engineer

Goals

- Decompress by using social media
- Stay informed about current events
- Pick up life hacks

Needs

- The ability to make sure that what he's viewing is authentic
- Easy to use tools that don't disrupt his media consumption

Pain Points

- Been tricked by AI generated social media videos encouraging inaccurate health content
- Misinformed by fake news
- Deep fakes are hard to spot

Competency

AI content:

Not familiar ——●———— Very familiar

Not confident ——●———— Very confident