

Inspired by our own experiences, we want to create a filter between us and the kind of social media we have at our fingertips.

We believe that through allowing users to make conscious choices about the content they want to see, we can foster a more positive online space focused on connecting people instead of polarising them for engagement and profit.

We want to foster critical thinking by encouraging users to choose the way they engage with their social media feeds, through the lenses of a humancentric technological solution.

Anna

NLP Expert – with experience in data science, she is responsible for the selection and implementation of transparent algorithms for the project.

Jan

AI Engineer – with experience in endto-end AI applications building, he is responsible for transforming the idea into a reality with coding expertise.

Jesus Carlos

AI Ethics Expert – with a background in NLP and AI policies, he ensures alignment with the requirements of EU AI Act.

Maria

Design Expert – focused on user experience and problem-solving, she is responsible for the strategic and creative direction of the project.

What's the solution?

A middleware that returns power over social media platforms to the people.

Redefine algorithmic control: Empowers users to choose between algorithms that focus on social connection, values or positivity instead of solely engagement.

Regain trust: Notifies users about particularly engagement-sparking content that might foster an environment for hate speech or ragebait, as well as flags AI-generated content.

Rebuild focus. Equips users with tools that help increase mindful consumption of content, such as grayscale mode, ad blocking, and usage statistics.

Problem Understanding

Social media algorithms optimize for engagement, therefore the very <u>design of these platforms</u> creates a cycle of polarization, disinformation, and endless scrolling.

Current platform responses to misinformation – top-down fact-checking and content moderation – often backfire by triggering psychological reactance where people resist when their autonomy is threatened. This makes <u>direct fact-checking labels ineffective or even harmful</u>, possibly leading to a <u>Streisand effect</u> where flagged content receives increased engagement.

Instead of telling users what to think, our solution builds on three core principles:

Transparency builds trust - users can directly see and understand how their feeds are ranked.

Agency reduces resistance - self-selected exposure to diverse viewpoints is more effective than forced exposure.

Customization enables gradual exposure to diverse viewpoints - users can adjust their information diet at their own pace, <u>fostering critical thinking skills</u>.

Market Analysis & Target Group

Our approach aligns with successful products that prioritize user control, including Ground News & AllSides, Reddit Enhancement Suite, and Minimalist Phone.

These solutions show a strong user demand for customizable content experiences, need to reduce screen time and optimize user experience represents a demand for technological governance, as well as the preferance to control the information diet over being told what's "correct".

Considering the target group, although we aim the solution to be used for an intergenerational audience, our approach would be to consider early adopters as young adults that are conscious about the risks of using social media.

We have used many evaluation frameworks to analyze and craft our solution, including 5WHY, PESTLE, as well as SWOT analysis, which we include below.

Strenghts	Opportunities
usefulness backed by research focused on the user	built for scalability to other platforms, algorithms and features
created by a diverse team	rising interest in tech mindfulness
Weaknesses	Threats
accessibility to API	complex code requirements
harder to develop for mobile use	less appeal for younger audiences

Features & Technical Feasibility

This project empowers users to take control of their social media experience by giving them the ability to choose how their content is ranked and displayed. Instead of being passively fed content by opaque algorithms, users can select from a range of feed-sorting strategies. To illustrate this, we have already designed a <u>first version of our user interface</u>.

Simple UX for Custom Feed Logic: Users can toggle between different algorithms via a clean interface – e.g., "Chronological," "Critical Thinking," "Feel Good," or their own trained model (plug-and-play). Our client-side engine will re-order or filter the scraped posts based on these preferences.

Target the Browser Level with a Cross-Platform Extension: A plugin that is built as a WebExtension which is supported in standard browsers, e.g., Chrome, Firefox. This gives us control at the client level, regardless of what algorithms the platforms use.

DOM-Based Content Extraction: We will use JavaScript (injected by the extension) to observe and extract posts from the user's feed. We avoid official APIs to bypass rate limits and platform restrictions.

Account-Friendly, Not Platform-Hostile: We do not interfere with login/authentication or post content; we just reorganize what the user already sees. This makes us legally and technically non-invasive.

Built for Extensibility: Although Instagram is our first target, the underlying architecture supports scaling to other social media platforms.

Hackathon Scope & Next Steps

The project will be rapidly prototyped during the Hackathon using a modular, teamparallel approach. Early efforts focus on front-end mockups, algorithm selection, and a clear UX narrative.

During Hackathon

Frontend: Connect buttons to feed, polish UI, test interactions, prep demo

Backend: Implement ranking, link to UI, ensure variants work, add comparison

UX/Design: Add tooltips, finalize text and layout, ensure smooth user flow

Strategy: Frame ethics, refine pitch, align with impact, plan next steps

After Hackathon

Month 1: Extend to more realistic data inputs with DOM-based extraction

Month 2: Prototype browser extension for overlaying custom feed logic

Month 3: Pilot test with educators, civic tech orgs, or disinfo researchers

Ongoing: Publish results, open up ranking API, invite contributors