



**Politecnico
di Torino**

Politecnico di Torino

Master's Degree in Computer Engineering

A.y. 2025/2026

Graduation Session March 2026

**Unhooked From the Smartphone:
Leveraging Large Language Models
to Detect Attention Traps**

Supervisors:

Alberto Monge Roffarello

Andrea Sillano

Candidate:

Matteo Vincenzo Petrera

Abstract

The rapid diffusion of short-form video platforms has intensified highly immersive interaction patterns, often leading to excessive and unintentional consumption behaviors commonly referred to as *doomscrolling*. Despite growing concern about its psychological and behavioral impact, limited research has focused on its real-time detection through objective behavioral data.

This thesis presents the design and experimental validation of a mobile application for the detection of doomscrolling behavior on the most widely used short-video platform, namely TikTok. The system uses Android Accessibility Services to collect processed interaction data, including scroll velocity, content skipping rate, interaction frequency, watch time, and session duration. A set of behavioral metrics was defined to model user activity and compute personalized detection thresholds.

A longitudinal study was conducted in two phases: a warm-up phase for baseline profiling and threshold computation, followed by a detection phase in which thresholds defined by the LLM are used to capture doomscrolling behavior, progressively adapting to new user habits. Objective behavioral indicators were integrated with subjective self-reports, including user-labeled doomscrolling episodes and perceived session duration.

The results suggest that specific behavioral features can effectively characterize problematic scrolling patterns. Finally, the thesis discusses the potential integration of Artificial Intelligence techniques, using Large Language Models (LLMs), to enable adaptive feedback generation and personalized digital well-being interventions.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisors, Alberto Monge Roffarello and Andrea Sillano, for their guidance throughout the development of this thesis. Their availability, valuable feedback, and constant support have been fundamental during every stage of this work. The discussions and insights they shared helped me approach the project with greater clarity and confidence, ultimately allowing me to complete this thesis in a constructive and stimulating environment.

A special thank you goes to my family, for their unwavering support throughout all these years. Knowing that I could always rely on your encouragement, patience, and trust has been an essential source of strength, especially during the most demanding moments of this journey. This achievement is not only mine, but also a reflection of everything you have given me along the way.

Finally, I would like to thank my friends, both the ones who have been by my side for many years and those I met during my university experience. Sharing this path with you has made these years far more meaningful. Between long study sessions, challenges overcome together, and countless memorable moments, your presence has played an important role in making this journey not only academically rewarding, but also personally unforgettable.

Table of Contents

List of Tables	VI
List of Figures	VII
Glossary	IX
1 Introduction	1
1.1 Objectives	4
1.2 Structure of the Thesis	4
2 Background and State of the Art	7
2.1 Digital Well-Being in the Age of the Attention Economy	7
2.1.1 The Attention Economy as a Business Model	8
2.1.2 The Emergence of the Digital Well-Being Concept	9
2.2 Attention-Capture Damaging Patterns	11
2.2.1 Definition and Taxonomy of ACDPs	11
2.2.2 Focus on Infinite Scroll and Autoplay	14
2.2.3 Impact on User Experience	16
2.3 Normative Dissociation in Social Media Use	18
2.4 Sense of Agency	20
2.4.1 Definitions and Theoretical Models	21
2.4.2 Sense of Agency in HCI	22
2.5 Short-Video Platforms as a Case Study	24
2.5.1 The case of TikTok	24
2.5.2 Interaction Patterns and Algorithmic Feeds	26
2.5.3 Evidence of Impact on User Behavior	28
3 Study Design	29
3.1 Research Objectives and Research Questions	30
3.2 Study Design	31
3.3 Variables and Operational Definitions	33

3.3.1	Independent Variables: Behavioral Predictors	33
3.3.2	Dependent Variables: Subjective and Analytical Outcomes	34
3.3.3	AI-Supported Threshold Variables	35
4	Design and Implementation of the Application	37
4.1	Design Objectives and System Requirements	37
4.2	Overall System Architecture	39
4.2.1	Warm-up Phase: Baseline Behavioral Observation	40
4.2.2	Detection Phase: Threshold-Based Evaluation	41
4.3	Mobile Application Design	42
4.3.1	Technology Stack	42
4.3.2	Application Modules	44
4.4	Session Management and Behavioral Metrics Computation Engine	46
4.4.1	The Central Role of <code>TikTokEventCaptureService.kt</code>	46
4.4.2	Logical Definition of a Session	47
4.5	AI-Supported Threshold Definition	48
4.5.1	Prompt Engineering Strategy	48
4.5.2	Threshold Application Logic	49
4.6	User Interface and Interaction Design	50
4.6.1	Main Application Screens	51
4.6.2	Overlays	55
5	Study Methodology	59
5.1	Participant Recruitment and Sample	59
5.1.1	Inclusion Criteria	59
5.1.2	Recruitment Procedure	60
5.1.3	Sample Characteristics	60
5.2	Experimental Procedure and Data Collection	60
5.2.1	Experimental Procedure	61
5.2.2	Data Collection Infrastructure	62
6	Results	65
6.1	Data Preparation	65
6.2	Research Questions Analysis	66
6.2.1	Characterizing Doomscrolling (RQ1)	66
6.2.2	Behavioral Changes Across Phases (RQ2)	68
6.2.3	Evaluation of LLM-Based Thresholding (RQ3)	76
6.3	Subjective Feedback	79
7	Discussion	83

8	Limitations and Future Work	88
8.1	Limitations	88
8.2	Future Work	89
9	Conclusions	91
A	LLM Prompts for Threshold Computation	93
A.1	File: thresholdDefinerPrompt.txt	93
A.2	File: thresholdUpdaterPrompt.txt	95
B	Participant Questionnaires	97
	Bibliography	107

List of Tables

2.1	A typology of 11 attention capture damaging patterns. [12]	14
2.2	Key dimensions of normative dissociation in social media use and their conceptual grounding.	20
4.1	Functional requirements of the mobile application.	38
4.2	Non-functional requirements of the mobile application.	39
4.3	Core application modules and their architectural responsibilities.	46
5.1	Session variables collected and stored in the Firestore database.	64

List of Figures

2.1	What happens online in 60 seconds	9
2.2	Social Media Timeline: Adoption	24
2.3	Time Spent Using Social Media Apps	26
2.4	The <i>For You</i> page interface of TikTok, characterized by full-screen video, vertical infinite scroll, and automatic content transitions. . .	27
3.1	Phases of the research process.	33
3.2	Session Example	35
4.1	Overall Application Architecture	40
4.2	Project Structure	43
4.3	Dashboard overview displaying average daily usage and doomscrolling indicators.	51
4.4	First part of Insights screen	52
4.5	Second part of Insights screen	53
4.6	Notifications history showing detected high-risk sessions.	54
4.7	Settings screen allowing adjustment of daily limits and doomscrolling thresholds.	55
4.8	Session labeling overlay shown at the end of selected sessions during the Warm-up Phase.	56
4.9	Real-time doomscrolling detection overlay displayed during high-risk sessions.	57
5.1	Structure of the user study, including baseline observation and threshold-based detection.	61
5.2	User Session saved on Firestore	63
6.1	Results derived from the Feature Activation Analysis done for RQ1	67
6.2	Behavioral Intensity Change	70
6.3	Doom Rate trend for participant 1	71
6.4	Session Duration trend for participant 1	71
6.5	Impulsive Scroll Rate trend for participant 1	72

6.6	Interaction Rate trend for participant 1	72
6.7	Content Skipped rate trend for participant 1	73
6.8	Doom Rate trend for participant 2	74
6.9	Session Duration trend for participant 2	74
6.10	Impulsive Scroll Rate trend for participant 2	75
6.11	Interaction Rate trend for participant 2	75
6.12	Content Skipped rate trend for participant 2	76
6.13	Participant responses on whether the application encouraged reflection on their social media usage.	79
6.14	Participant responses on whether the application helped them better recognize when they were doomscrolling.	80
6.15	Distribution of participant responses regarding whether the notifications were delivered at appropriate moments during usage.	81
6.16	Overall user rating of the application on a 5-point scale, showing an average evaluation of 4.2.	82

Glossary

HCI

Human-Computer Interaction

ACDP

Attention-Capture Damaging Pattern

DSCT

Digital Self-Control Tool

SoA

Sense of Agency

LLM

Large Language Model

RR

Risk Ratio

Chapter 1

Introduction

In the contemporary digital landscape, mobile devices and social media platforms have become deeply embedded in everyday life, shaping how people communicate, access information, and spend their free time. The widespread adoption of smartphones has enabled constant access to digital services that encourages patterns of interaction characterized by frequent, brief, and toxic engagement throughout the day [1]. In recent years, short-form content platforms have further amplified these dynamics by introducing interaction models based on rapid consumption, continuous scrolling, and algorithmically curated feeds.

Even if such systems offer clear benefits in terms of accessibility, their widespread use has also raised increasing concerns about their impact on user behavior and well-being [2, 3]. A recurring issue reported by users is the tendency to remain engaged with social media applications longer than originally intended. This prolonged usage emerges gradually during interaction rather than as a result of deliberate planning, causing users to lose track of time [4]. Experiences like this are often associated with feelings of dissatisfaction, frustration, or reduced self-control, suggesting a misalignment between user intentions and their actual behavior during interaction.

These concerns have contributed to the growth of the concept of digital well-being, which broadly refers to an individual's ability to maintain a healthy, balanced and intentional relationship with digital technologies [5]. Within the field of Human-Computer Interaction (HCI), a substantial body of research has focused on understanding how interface design choices and interaction mechanisms can support or undermine users' autonomy, awareness, and self-regulation [6, 7].

A key factor underlying many of these phenomena is the broader economic and technological context in which modern digital platforms operate. Most large-scale social media systems are developed within the framework of the attention economy, a paradigm in which user attention and engagement are the primary contended resources. [8]. In the attention economy paradigm, digital platforms

and social media companies prioritize metrics related to user engagement, such as session duration, interaction frequency, and user retention, as primary indicators of platform success. These metrics, while useful for optimizing advertising revenue and engagement, do not directly capture user satisfaction or well-being. Research has shown that attention-driven design and engagement optimization can even have negative consequences for user autonomy and well-being, highlighting the limitations of traditional success metrics in digital platforms. [9, 10]

This incentive structure has a direct impact on the design of the interface. To maximize engagement, platforms rely on interaction mechanisms that encourage continuous consumption. In the field of HCI, these mechanisms have been defined as Attention-Capture Damaging Patterns (ACDPs), that is, recurring design patterns that leverage cognitive biases and behavioral tendencies to sustain user attention over time [11, 12]. Examples of such patterns include infinite scrolling feeds, automatic content playback, and algorithmic recommendation systems that continuously adapt to user behavior.

Although these patterns are effective from an engagement perspective, they may also contribute to interaction styles that are less intentional and less reflective. In particular, they can facilitate prolonged usage sessions characterized by repetitive actions, minimal deliberation, and reduced awareness of time passing. Prior work has shown that habitual and low-effort interaction patterns can make smartphone use more pervasive and harder to regulate, even in the absence of explicit user intent [1].

These dynamics are especially present on platforms centered around short and fast consumable content. Among them, TikTok represents a relevant case study. Its interaction model is built almost entirely around the integration of infinite scrolling, automatic video playback, and a highly optimized recommendation algorithm. TikTok became popular for its interface, centered on a full-screen, vertically scrollable feed of algorithmically curated content, a paradigm that has subsequently been adopted by other platforms such as Instagram (Reels) and YouTube (Shorts). [13].

This design makes TikTok a suitable environment for observing behavioral patterns associated with prolonged unintentional use [13, 14]. Empirical studies and recent reports suggest that users spend a significant amount of time on the platform, often exceeding the average of other major social media services [15]. More importantly, the nature of this usage, characterized by rapid content consumption and low interaction cost, raises questions about users' awareness and well-being during consumption [16].

The term doomscrolling has been used in recent literature to describe a pattern of compulsive and repetitive consumption of online content. In this context, doomscrolling does not simply refer to a prolonged usage time, but to a specific mode of engagement characterized by continued content consumption despite reduced

interest or satisfaction and emotional fatigue [13]. Similarly to other problematic dynamics of social media use, doomscrolling emphasizes the interaction between user behavior and frictionless feed architectures that encourage uninterrupted scrolling [17]. Understanding this phenomenon therefore requires moving beyond purely temporal measures and considering how users interact with content at a behavioral level.

In response to concerns about excessive or problematic usage, both researchers and industry professionals have proposed a variety of Digital Self-Control Tools (DSCTs), such as screen-time dashboards, usage statistics, reminders, and hard usage limits. While these tools can increase user awareness, they often rely on static thresholds and require explicit user configuration. Moreover, they tend to focus primarily on how long an application is used, rather than on how it is used [6].

As a consequence, such approaches may fail to capture qualitative aspects of interaction that characterize doomscrolling behavior, including impulsive scrolling, low engagement with individual content items, and reduced intentionality [18]. In addition, rigid enforcement mechanisms, such as application blocking, may be perceived as intrusive, leading users to disable them or to circumvent their effects over time.

These limitations suggest the need for alternative approaches that emphasize behavioral understanding over strict control. In particular, there is increasing interest in systems capable of detecting potentially problematic usage patterns, and supporting user awareness through adaptive feedback [19].

Despite the growing body of literature on digital well-being and problematic technology use, a gap remains regarding the detection of doomscrolling behavior in real-world usage contexts using behavioral data collected directly on users' devices. While previous work has explored the psychological correlates of social media use and the effects of specific design patterns, fewer studies have attempted to treat doomscrolling as a measurable phenomenon grounded in interaction metrics [4, 7].

This thesis aims to contribute to this research area by investigating whether interaction data and metrics, such as scrolling behavior, interaction frequency, and session dynamics, can be used to identify patterns associated with doomscrolling during everyday usage. In addition, the proposed approach integrates Large Language Models (LLMs) to interpret behavioral metrics at a higher level. Specifically, LLMs are employed to support personalized threshold generation helpful for doomscrolling detection. Rather than focusing exclusively on restriction or prevention, the proposed approach explores interventions useful to help users reflect on their ongoing behavior.

1.1 Objectives

The primary objective of this thesis is to investigate whether doomscrolling behavior can be identified through the analysis of users' interaction patterns on mobile devices, going beyond traditional measures of application usage. Rather than assuming that prolonged usage alone is indicative of problematic behavior, this work explores doomscrolling as a qualitative phenomenon characterized by specific behavioral dynamics, such as repetitive scrolling, low engagement with individual content items, reduced intentionality and specific content category consumption.

To address this objective, the thesis adopts a user-centered and behavior-oriented perspective, aligned with recent research in the field of digital well-being and digital self-control tools [6, 7]. In particular, the goal is not to enforce usage reduction through rigid restrictions, but to support user awareness by providing timely and context-sensitive feedback based on observed interaction patterns.

More specifically, the objectives of this work can be summarized as follows:

1. To analyze and characterize interaction patterns associated with doomscrolling behavior, with a focus on scrolling dynamics and user engagement during content consumption.
2. To design and implement a mobile system capable of collecting interaction data directly on the device, without requiring modifications to the monitored applications.
3. To define a set of behavioral metrics aimed at capturing qualitative aspects of mobile interaction beyond session duration.
4. To explore awareness-oriented interventions, such as adaptive notifications, designed to prompt user reflection without imposing hard constraints on usage.
5. To evaluate the proposed approach through a longitudinal user study conducted in real-world conditions, assessing both detection effectiveness and user perception.

By pursuing these objectives, this thesis aims to contribute to ongoing research on digital well-being by providing empirical insights into the feasibility of detecting doomscrolling behavior through interaction data and by exploring design strategies that prioritize user agency over control.

1.2 Structure of the Thesis

This thesis is organized into eight chapters, guiding the reader through a structured path that begins with the analysis of the scientific literature, proceeds with the

design of the application and execution of an empirical study, and concludes with the interpretation and discussion of the experiment results.

Starting from Chapter 2, the thesis is structured as follows.

- **Chapter 2 – Background and State of the Art** presents a review of the scientific literature related to the theoretical foundations of this work. The chapter introduces key concepts in the field of Human-Computer Interaction (HCI), focusing on digital well-being, problematic technology use, and attention-capturing design patterns. Particular attention is devoted to the taxonomy of Attention-Capture Damaging Patterns (ACDPs), focusing on doomscroll, infinite scroll, and auto-play mechanisms.
- **Chapter 3 – Study Design and Methodology** describes in detail the empirical approach adopted in this thesis. The chapter outlines the research questions and hypotheses, the overall study design, the definition of variables and metrics, the participant recruitment strategy, and the data collection procedure. It also presents the instruments and methodologies employed to capture user interaction data and subjective feedback during real-world usage.
- **Chapter 4 – Design and Implementation of the Application** illustrates the architecture and functioning of the mobile application developed for the study. This chapter motivates the main design and technological choices underlying the implementation, describing how the system collects interaction data, processes behavioral metrics, and delivers awareness-oriented feedback to users.
- **Chapter 5 – Results** presents the data collected during the empirical study in a structured and systematic manner. The chapter reports the results of the analyses conducted to address the research questions, including descriptive statistics and comparisons across different usage conditions.
- **Chapter 6 – Discussion** interprets the obtained results in light of the theoretical framework introduced in the earlier chapters. The implications of the findings are discussed with respect to the relationship between interface design, user behavior, and digital well-being, as well as their relevance for the design of more mindful and user-respecting technologies.
- **Chapter 7 – Limitations and Future Work** provides a critical reflection on the methodological and practical limitations of the study, while also outlining promising directions for future research and potential extensions of the proposed approach.

- **Chapter 8 – Conclusions** summarizes the main contributions and final outcomes of the thesis, highlighting the significance of the findings within the broader context of research on digital well-being and interaction design.

Chapter 2

Background and State of the Art

This chapter aims to establish the theoretical framework upon which the entire thesis is grounded. The analysis begins with an examination of the economic model that governs the contemporary digital ecosystem, namely the Attention Economy, and then explores its direct consequences on individuals' well-being through the introduction of the concept of Digital Well-Being. The chapter subsequently focuses on the specific mechanisms through which user attention is captured and sustained, referred to as Attention-Capture Damaging Patterns (ACDPs). A taxonomy of these patterns is presented, with a targeted focus on Infinite Scroll and Autoplay, the two interaction mechanisms investigated in the experimental study. The relationship between these design elements and the central psychological constructs underlying this research, Normative Dissociation and Sense of Agency, is then examined. Finally, the chapter concludes by narrowing the discussion to the specific case study of TikTok, considered as a paradigmatic example of short-video platforms characterized by attention-intensive interaction models.

2.1 Digital Well-Being in the Age of the Attention Economy

To fully understand the dynamics governing the interaction between users and social media platforms, it is essential to examine the socio-economic context in which these technologies have emerged and continue to evolve. This context is shaped by the rise of a new economic paradigm, commonly referred to as the Attention Economy, which has redefined both the value and the nature of resources within the digital ecosystem.

The intense competition for this scarce resource, human attention, has led to a range of critical issues, ultimately contributing to the emergence of a new field of inquiry and design. This field, known as Digital Well-Being, seeks to address the challenges introduced by attention-driven technologies and to promote more balanced and intentional forms of interaction between individuals and digital systems.

2.1.1 The Attention Economy as a Business Model

The foundational principle underlying the Attention Economy was first articulated by Nobel laureate Herbert A. Simon. As early as 1971, in a pre-digital context, Simon observed that an abundance of information inevitably produces a scarcity of another critical resource: human attention. He famously argued that, in an information-rich world, information consumes the attention of its recipients, making attention the true limiting factor in cognitive and economic terms [8].

This concept was later analyzed by several scholars, most notably Michael Goldhaber, who in the 1990s explicitly connected the concept of attention shortage to the rise of the Internet. However, it was Davenport and Beck who formalized this idea into a specific economic framework with their innovative work *The Attention Economy* [20]. They described a critic inversion of the traditional economic paradigm: in a context where information is abundant and easily accessible, the scarce and valuable resource is no longer information itself, but rather the human attention required to process it.

Within this framework, attention becomes a form of currency, finite, valuable, and highly contested, around which digital markets are organized [21]. Technology companies, and social media platforms in particular, increasingly compete for users' attentional resources, investing heavily in interface designs and interaction mechanisms aimed at maximizing engagement.

Digital platforms have built their business models almost entirely on this principle. Rather than generating profit through the direct sale of products or services to users, many platforms monetize the time, focus, and behavioral data produced during user interactions. Engagement metrics such as session duration, frequency of access, and interaction volume are translated into economic value through targeted advertising and data-driven personalization strategies [22]. As highlighted by usability research, in an environment saturated with content, "attention, not information, has become the primary bottleneck" [23].

This economic logic imposes a clear direction on design practices. Interfaces are not primarily optimized for efficiency, utility, or user well-being, but for the continuous maximization of engagement-related metrics [11]. Design decisions such as frictionless navigation, continuous content streams, and personalized recommendations are therefore not incidental, but rather systematic responses to the

incentives imposed by the Attention Economy.

As a consequence, a structural conflict of interest emerges between platforms and users. While platforms are incentivized to prolong usage and intensify engagement, users may seek to maintain a controlled, intentional, and goal-aligned relationship with technology. This tension lies at the core of many contemporary concerns surrounding excessive social media use and represents a central motivation for the investigation of digital well-being and attention-aware design strategies [24, 25].

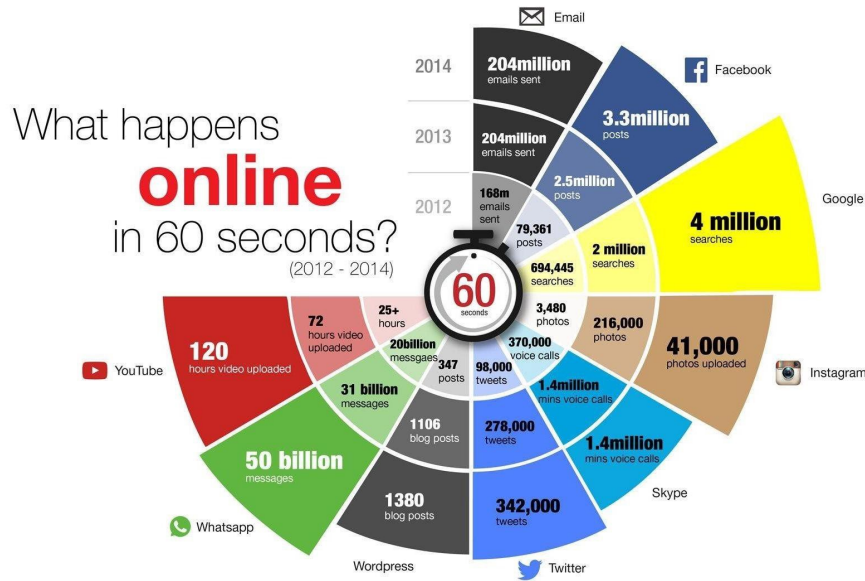


Figure 2.1: Graphic representation of everything that happens on average on the most used social media platforms in 60 seconds

2.1.2 The Emergence of the Digital Well-Being Concept

The pervasive exposure to digital environments explicitly designed to capture and retain user attention has been shown to produce significant effects on individuals’ psychological well-being and everyday functioning. A growing body of empirical research highlights how the constant availability of communication, entertainment, and personalized content fosters patterns of frequent and often automatic device checking, which may occur with limited awareness or conscious intention [24, 26]. Users are frequently motivated to engage in such behaviors by boredom, the desire to fill idle moments, or as a strategy to cope with socially uncomfortable situations [27, 23].

Over time, these repeated interactions can transform an initially goal-oriented or instrumental use of technology, such as replying to a message or checking a

notification, into habitual and prolonged engagement. This shift is often characterized by a progressive loss of temporal awareness, whereby users underestimate the duration of their interactions and remain engaged for longer than intended [1, 25]. Research in media psychology further distinguishes between active forms of use, such as content creation or direct social interaction, and passive forms of consumption, including continuous feed scrolling and algorithmically driven content viewing. Notably, it is passive usage that has been most consistently associated with declines in affective well-being and user satisfaction [28, 29].

Platforms centered around continuous content consumption, such as social media applications, video streaming platforms, and digital games, are frequently identified as contexts in which long and unintentional sessions are likely to occur [30, 31]. These sessions often end with negative outcomes, including feelings of frustration, guilt, or regret, with a clear distinction between users' initial intentions and their actual behavior [24, 25]. Also, problematic usage patterns have been linked to broader consequences, such as decreased sleep quality and difficulty with social interactions, increasing concerns about the impact of attention-intensive technologies on everyday life [32, 33].

In response to these new concerns, a new research area has taken shape with the name of Digital Well-Being. Rather than considering excessive use of technology exclusively in pathological terms, this field investigates how digital technologies can coexist in every day life in order for individuals to live a balanced, meaningful, and goal-aligned life [5]. From this perspective, digital well-being is not defined by the absence of technology use, but by the alignment between users' intentions, values, and actual interaction outcomes.

Several studies have contributed to explore this concept by emphasizing the subjective and experiential dimensions of technology use. For example, research has explored the distinction between meaningful and meaningless smartphone interactions, linking the latter to reduced concentration and lower satisfaction [6]. This line of work highlights that users may remain highly engaged with a system while simultaneously experiencing a loss of control or purpose, challenging traditional assumptions that equate engagement with positive user experience.

As awareness of these issues has increased, many users have attempted to adopt strategies to regain control over their digital habits, such as turning off notifications, uninstalling applications, or setting personal rules. However, study evidence suggests that the effectiveness of these strategies is quite limited, particularly in the context of applications optimized to engage users through persuasive design mechanisms [34, 35]. This gap between user intentions and actual behavior has contributed to the development of Digital Self-Control Tools (DSCTs), which aim to support self-regulation through features such as usage monitoring, reminders, and time limits [6].

From 2018 onward, major technology companies have incorporated native digital

well-being functionalities into their operating systems, publicly framing these tools as efforts to promote healthier usage patterns. Despite their widespread adoption, systematic evaluations indicate that many DSCTs struggle to produce lasting behavioral change [7]. These tools often operate as external constraints rather than as mechanisms that foster internal awareness or habit formation, leading users to bypass, disable, or ignore them over time [18].

This evidence has created a significant shift in the digital well-being world, considering not only users' self control as a primary way of taking action, but also single platform responsibilities in terms of persuasive designs. Research in Human-Computer Interaction has played a crucial role in this transition by demonstrating how interface designs actively shape user behavior and make disengagement very difficult [11, 12]. From this point of view, problematic usage patterns are not only the result of addiction and weakness of the user, but the result of specific design choices ruled within attention-driven business models.

Consequently, the study of digital well-being increasingly emphasizes the need to critically examine the relationship between economic incentives, interface design, and psychological experience. Understanding how these elements interact provides a necessary foundation for the analysis of specific attention-capturing mechanisms, which are examined in detail in the following section through the lens of Attention-Capture Damaging Patterns.

2.2 Attention-Capture Damaging Patterns

The analysis conducted so far has highlighted how engagement-oriented business models translate abstract economic incentives into concrete interface design choices. Within this context, Attention-Capture Damaging Patterns (ACDPs) represent a key analytical construct for understanding how digital interfaces systematically shape user behavior in attention-intensive environments.

This section introduces ACDPs as a distinct class of design patterns that operate aiming at economic interests, user vulnerabilities, and interaction design. After defining the concept and its theoretical foundations, the section presents the criteria used to identify and classify ACDPs, as proposed in recent HCI literature. These criteria provide a structured framework for analyzing how specific design patterns contribute to prolonged, unintentional, and potentially harmful engagement, setting a start point for the focused examination of Infinite Scroll and Autoplay in the following sections.

2.2.1 Definition and Taxonomy of ACDPs

Recent work in Human-Computer Interaction has emphasized the need to distinguish between interface designs that merely encourage engagement and those

that actively undermine users' ability to regulate their attention. In this vein, Attention-Capture Damaging Patterns have been introduced to describe recurring design strategies that exploit psychological and cognitive mechanisms in order to sustain user attention, often leading to misalignment between users' intentions and their actual behavior [12, 11]. Unlike traditional dark patterns, which are often associated with deception, coercion, or economic manipulation, ACDPs are primarily characterized by their impact on attention and agency. They operate by subtly reshaping interaction dynamics, making continued engagement easier than disengagement, and thereby fostering prolonged, habitual, and frequently unreflective use [21, 22]. This distinction is particularly relevant in the context of digital well-being, where harm often emerges not from isolated decisions, but from cumulative interaction patterns over time [5, 24]. To support the systematic identification of ACDPs, Monge Roffarello et al. propose a set of five criteria (C1–C5) that together capture both the mechanisms through which attention is captured and the experiential consequences for users. These criteria are not intended as strict binary conditions, but rather as analytical dimensions along which interface designs can be evaluated [12].

- **C1 – Exploit Psychological Vulnerabilities** Attention-Capture Damaging Patterns exploit predictable psychological heuristics and cognitive biases to induce behaviors that conflict with users' long-term goals. Unlike nudges designed to promote beneficial actions, ACDPs leverage mechanisms such as novelty seeking, short-term reward preference, and emotional memory bias to sustain engagement beyond users' intentions. Variable reward schedules, in particular, reinforce repeated interaction by making content discovery itself rewarding, often leading to continued use even when perceived value diminishes.
- **C2 - Automation of the User Experience and Normative Dissociation** ACDPs often automate interaction flows in ways that reduce opportunities for intentional decision-making and reflection. By minimizing explicit choices and promoting seamless, uninterrupted sessions, these patterns facilitate states of normative dissociation, characterized by reduced self-awareness and diminished goal monitoring. While such states can occasionally be beneficial, their deliberate induction in low-meaning activities may result in prolonged, unreflective engagement that is only recognized retrospectively.
- **C3 – Loss of Awareness of Usage Goals** A central impact of ACDPs is the erosion of users' awareness of their original goals for technology use. Through continuous attention demands and frequent distractions, these patterns divert users from goal-directed interaction toward behaviors aligned with platform objectives. As a result, users may engage in actions they would not have

consciously chosen, often at the expense of productivity and intentionality.

- **C4 – Distorted Sense of Time and Diminished Control** Exposure to ACDPs is commonly associated with a distorted perception of time and a reduced sense of control over interaction. Users frequently underestimate the duration of their engagement and report feeling “carried along” by the interface. This experience is closely linked to a diminished sense of agency, as continued interaction is framed as the default and disengagement becomes cognitively or perceptually difficult.
- **C5 – Post-Use Regret** Finally, ACDPs are often associated with a retrospective sense of regret following interaction. Users may experience dissatisfaction when reflecting on time spent in low-value activities, particularly when short-term gratification fades. This regret arises from the perceived discrepancy between actual behavior and alternative, more meaningful uses of time, reinforcing the tension between engagement maximization and digital well-being.

Based on this analysis, Monge Roffarello et al. proposed a precise taxonomy that identifies 11 ACDPs summarized in the following table (2.1) [12]

Table 2.1: A typology of 11 attention capture damaging patterns. [12]

Pattern Name	Description	Main Context(s) of Use
<i>P1 - Infinite Scroll</i>	As the user scrolls down a page, more content automatically and continuously loads at the bottom.	Social media (e.g., Facebook, Instagram, and Twitter).
<i>P2 - Casino Pull-to-refresh</i>	When the user swipes down on their smartphone, there is an animated reload of the page that may or may not reveal new appealing content.	Social media on smartphones.
<i>P3 - Neverending Autoplay</i>	A new video is automatically played when the current one finishes. There is never a point for the user to stop and reflect, and the option to turn off autoplay is hidden or non-existent.	Social media and video streaming platforms, e.g., YouTube.
<i>P4 - Guilty Pleasure Recommendations</i>	Personalized suggestions that prey on individual consumer frailty to target user’s guilty pleasures and increase use time.	Social media and video streaming platforms, e.g., YouTube.
<i>P5 - Disguised Ads and Recommendations</i>	Advertisements and recommendations, e.g., posts and sponsored pages, that are disguised as normal content into social networks’ newsfeeds.	Social media.
<i>P6 - Recapture Notifications</i>	Notifications that are deliberately sent to recapture users’ attention and have them start a new usage session, e.g., notifications with recommended content or notifications about content the user has never interacted with.	Social media, video streaming platforms, and messaging applications.
<i>P7 - Playing by Appointment</i>	Users are forced to use a digital service at specific times as defined by the service, otherwise the user may lose points and achievements.	Video games (mostly on social networks) and social media in general.
<i>P8 - Grinding</i>	Users are forced to repeat the same process several times to unlock an achievement, e.g., a new level in a video game or a badge on a social network.	Video games and social media.
<i>P9 - Attentional Roach Motel</i>	Registering to and accessing attention-capture digital services is easy, while operations like logout or canceling an account are painfully difficult.	Social media, e.g., Facebook.
<i>P10 - Time Fog</i>	A pattern through which designers reduce users’ awareness of time spent, e.g., by hiding the smartphone’s clock.	Video streaming platforms, e.g., Netflix.
<i>P11 - Fake Social Notifications</i>	The platform sends messages pretending to be another user or social notifications about some content the user has never interacted with.	Video games (mostly on social networks) and social media in general.

2.2.2 Focus on Infinite Scroll and Autoplay

Although the taxonomy of Attention-Capture Damaging Patterns is broad, two patterns emerge as particularly pervasive and influential in contemporary social media platforms, especially in short-video applications such as TikTok. Infinite Scroll and Never-Ending Autoplay play a central role in shaping the interaction

dynamics of these systems and are therefore the focus of the experimental investigation presented in this thesis. Both patterns closely align with the objectives of the Attention Economy, as they are designed to deliver a seamless, automated, and low-friction consumption experience aimed at maximizing user engagement and time spent on the platform [20, 21].

- ***Infinite Scroll*** Infinite Scroll is an Attention-Capture Damaging Pattern that eliminates natural stopping points in an interface by continuously loading new content as the user scrolls downward. Unlike paginated layouts, which implicitly signal the completion of a content unit, infinite scrolling creates an uninterrupted stream in which the end of available content is never reached [31]. This design choice has been widely discussed in the literature as problematic, as it systematically discourages users from reassessing whether to continue or terminate an interaction session [11, 12].

From a behavioral perspective, infinite scroll exploits users' sensitivity to novelty and intermittent rewards. As users scroll, each new content item represents an uncertain outcome, which may or may not be rewarding. This uncertainty, combined with occasional highly engaging content, reinforces continued interaction and promotes automatic scrolling behavior [36, 35]. Over time, such behavior can become habitual, reducing the role of deliberate decision-making and increasing the likelihood of prolonged, unintentional usage [1].

Empirical studies suggest that infinite scroll contributes to both within-session and cross-session engagement loops. Rixen et al. describe a dual mechanism in which prolonged scrolling during a single session coexists with habitual patterns of frequent application checking across the day [31]. These dynamics are closely associated with a distorted perception of time and diminished awareness of usage goals, two core impacts of Attention-Capture Damaging Patterns [24, 25].

Importantly, the effects of infinite scroll are not immutable. Design interventions that reintroduce stopping cues or visual boundaries have been shown to reduce passive consumption and overall time spent on a platform, highlighting the role of deliberate design choices in shaping user behavior [19].

- ***Never-Ending Autoplay*** Never-Ending Autoplay is an Attention-Capture Damaging Pattern that automates the transition from one content item to the next, removing the need for explicit user action to continue consumption. By design, autoplay minimizes decision friction and motor effort, lowering the cognitive cost associated with remaining engaged [12]. As a result, continued consumption becomes the default behavior, while disengagement requires an explicit and effortful decision.

This automation has been shown to directly affect users' sense of agency, understood as the perception of being in control of one's actions and their outcomes [37, 38]. When content playback proceeds automatically, users may experience reduced autonomy and a diminished ability to reflect on whether continued consumption aligns with their intentions [39]. Such effects are particularly relevant in the context of digital well-being, where agency and intentionality are central components [5].

The behavioral impact of autoplay has been empirically demonstrated in multiple contexts. A controlled experimental study on Netflix usage found that disabling autoplay led to statistically significant reductions in both average daily viewing time and the duration of individual viewing sessions [30]. Similar dynamics have been observed on social media platforms, where automated recommendations and playback mechanisms contribute to extended usage sessions and increased difficulty disengaging [40, 13].

In short-video platforms such as TikTok, autoplay is tightly integrated with vertical infinite scrolling. The immediate playback of each new video, combined with the absence of loading delays, creates a highly immersive environment in which users can remain engaged with minimal cognitive effort. This coupling amplifies the attention-capturing effects of both patterns and further reduces opportunities for reflective choice [41].

Taken together, Infinite Scroll and Never-Ending Autoplay do more than simply increase the quantity of content presented to users. They fundamentally alter the nature of interaction by removing both explicit decision points and intentional motor actions, two essential components of goal-directed behavior [6]. As a consequence, users are gradually transformed from active explorers of content into passive recipients of a continuous stream of stimuli.

2.2.3 Impact on User Experience

Prolonged interaction with interfaces dominated by Attention-Capture Damaging Patterns does not merely affect users' observable behavior, but also has profound consequences on their subjective experience and underlying mental states. Beyond generating feelings of wasted time and post-use regret, which have been widely reported in the literature [28, 29], sustained exposure to such patterns has been associated with a broad range of negative psychological outcomes.

Among these effects, research has highlighted cognitive depletion, defined as a reduction in available mental resources that may persist beyond the interaction itself. Excessive engagement with digital platforms—particularly during evening hours—has been shown to impair subsequent cognitive performance and self-regulation, with effects extending into the following day [23, 25]. Additionally, the

constant availability and demand for attention characteristic of attention-driven platforms has been linked to technostress, a modern form of stress arising from the perceived inability to disconnect and from the pressure to remain continuously responsive and online [24].

Furthermore, the fragmented and interruption-rich digital environments promoted by ACDPs have been associated with diminished attentional capacity and increased stress related to multitasking and frequent context switching [35, 2]. While these impacts are significant, the literature suggests that two psychological constructs are particularly central for understanding how ACDPs alter the user experience at a deeper level: Normative Dissociation and the Sense of Agency (SoA). These constructs form the theoretical core of the present thesis.

- ***Normative Dissociation*** Exposure to low-friction, continuous content streams can induce a psychological state commonly referred to as normative dissociation [42]. In this state, users experience a temporary reduction in self-awareness and reflective thought, interacting with digital content in a largely automatic and unintentional manner. Rather than actively choosing each action, users become absorbed in the interaction flow, often realizing the extent of their engagement only in hindsight.

Normative dissociation has been empirically observed in the context of social media and short-video platforms, where interface designs intentionally minimize decision points and encourage uninterrupted consumption [12]. Such states are frequently described by users in colloquial terms as being “on autopilot” or “zoned out,” reflecting a subjective experience of diminished control and awareness. While dissociative states are not inherently harmful and may occur in everyday activities, their induction through persuasive and attention-oriented design raises concerns when they emerge during interactions that lack intrinsic value or personal meaning [42, 19].

- ***Sense of Agency*** Normative dissociation is closely intertwined with the erosion of the Sense of Agency, defined as the subjective perception of being the initiator and controller of one’s actions and their consequences [37, 38]. A strong sense of agency is a fundamental component of intentional behavior and psychological well-being, as it supports autonomy, self-efficacy, and reflective decision-making [39].

Attention-Capture Damaging Patterns undermine the sense of agency by automating actions and reducing the predictability of their outcomes. Design mechanisms such as autoplay and algorithmically curated feeds act “on behalf of the user,” continuing interactions without explicit consent and presenting content that may not clearly reflect prior choices [12, 41]. As a result, users may feel less in control of their interaction trajectory, perceiving their behavior as driven by the system rather than by their own intentions.

Empirical evidence suggests that diminished sense of agency is associated with more negative experiences of technology use and lower satisfaction with social media engagement [5]. When users perceive that disengagement requires effort while continuation is automatic, the balance between autonomy and system control shifts in favor of the platform, compromising a core pillar of digital well-being.

Taken together, these findings support a coherent causal chain that underpins the theoretical framework of this thesis. Attention-Capture Damaging Patterns promote interaction styles characterized by low friction and automation, which facilitate states of normative dissociation. In turn, this dissociative experience erodes users' sense of agency, reducing their capacity to act intentionally and to regulate their behavior in accordance with personal goals [6, 7].

Within this framework, doomscrolling can be understood not simply as prolonged usage, but as a qualitatively distinct mode of interaction marked by passivity, reduced awareness, and diminished control. The centrality of normative dissociation and sense of agency in this process motivates their explicit consideration in the design, implementation, and evaluation of the system proposed in this thesis.

2.3 Normative Dissociation in Social Media Use

In this thesis, the term normative dissociation is used to refer to a non-clinical, everyday state characterized by reduced self-reflection and attenuated awareness of the ongoing interaction, which can emerge during prolonged and repetitive engagement with digital content streams. Importantly, the relevance of dissociation in this context does not depend on pathological interpretations; rather, it highlights a shift in the quality of interaction, away from intentional and goal, directed behavior and toward more automatic and less reflective engagement [42, 25]. This perspective is particularly useful when analyzing social media platforms that heavily rely on continuous feeds, algorithmic personalization, and low-effort content transitions [12, 41]. ACDPs such as infinite scrolling and autoplay are not only effective at increasing time-on-platform; they also transform the interaction into an experience that can become cognitively “lightweight” and repetitive [31, 30]. By removing explicit decision points (e.g., “Should I continue?”) and by making continuation the default option, these patterns facilitate states in which users keep consuming content with limited conscious deliberation [12]. In such conditions, the interaction may gradually detach from the user’s initial intention, turning a brief, purposeful check into an extended session that is only fully recognized after it ends [24]. This mechanism is consistent with broader evidence on smartphone habits and checking behavior. Oulasvirta et al. showed that smartphone use often becomes pervasive through habit formation, enabling frequent and sometimes automatic engagement

across the day [1]. When habitual checking is combined with interface architectures designed to sustain intra-session engagement, the conditions for prolonged, passive, and less self-aware interaction become especially strong [31, 12]. While normative dissociation is a multifaceted construct, research related to problematic and passive social media use suggests a set of recurring experiential dimensions that help characterize how dissociation manifests during interaction. In the context of this thesis, four dimensions are particularly relevant because they connect interface mechanisms to observable behavioral traces and post-use reports:

- **Diminished self-awareness.** During dissociative interaction, users may experience reduced monitoring of their internal state, intentions, and ongoing goals. This reduction in self-awareness is compatible with evidence linking intensive technology exposure to decreased cognitive resources and attentional availability [23, 25]. In ACDP-driven environments, where the interface continuously presents new stimuli and requires minimal deliberation, self-monitoring may progressively weaken [12].
- **Distorted sense of time.** A common feature of prolonged social media sessions is the subjective underestimation of time spent, often followed by retrospective realization of having exceeded intended use [24]. This dimension is particularly relevant for doomscrolling experiences, where the user’s perception of time passing may diverge from actual session duration.
- **Reduced meaningful recall of consumed content.** Passive consumption can produce an experience that feels “empty” or low-value, where users struggle to recall specific content items afterward. This effect is consistent with the idea that dissociation involves attenuated reflective processing during interaction [42, 35]. From a behavioral standpoint, rapid content switching and minimal engagement signals can be viewed as correlates of such reduced encoding.
- **Automatism and mindless continuation.** Dissociation is often associated with behavior that feels automatic—users continue scrolling or watching without an explicit decision to do so. This dimension resonates with evidence on habit-driven smartphone use [1] and with ACDP criteria emphasizing automation and reduced friction [12]. It is also the most directly linked to measurable interaction patterns (e.g., high scrolling frequency, low interaction diversity, sustained session continuation).

Together, these dimensions provide a structured lens to interpret dissociation not as an abstract construct, but as an experiential outcome that can plausibly be related to behavioral signals captured during real-world use [7]. This connection is crucial for the empirical approach adopted in this thesis, which aims to operationalize

doomscrolling through interaction metrics rather than relying only on time-based measures.

Table 2.2: Key dimensions of normative dissociation in social media use and their conceptual grounding.

Dimension	Description
Diminished self-awareness	Reduced monitoring of intentions, goals, and internal state during interaction; facilitated by continuous attention demands and low-friction consumption
Distorted sense of time	Altered perception of time spent, with realization often occurring retrospectively after the session ends
Reduced content recall	Limited ability to recall specific items consumed during the session, consistent with reduced reflective processing and passive engagement
Automatism (mindless continuation)	Scrolling/watching continues with minimal conscious deliberation; linked to habit-driven smartphone behavior and automation-oriented interface design

2.4 Sense of Agency

The causal chain outlined so far, where attention-capture damaging patterns foster low-friction, automated consumption and facilitate normative dissociation, reaches a critical point in its impact on a fundamental psychological construct: the Sense of Agency (SoA). In broad terms, SoA refers to the subjective experience of being the initiator of one’s actions and of exerting control over their outcomes in the world. This experience is not an abstract philosophical notion; it is a pervasive component of everyday cognition that supports autonomy, intentional action, and self-regulation [37, 39].

Within the domain of digital well-being, agency is particularly relevant because many negative user experiences, such as feeling “pulled” into a platform, losing track of time, or continuing use beyond one’s intentions, can be interpreted as situations in which perceived control is weakened [5, 24]. In this sense, SoA offers a conceptual bridge between interface design and user well-being: when platforms systematically automate continuation and reduce decision points, users may remain engaged while experiencing a gradual shift from deliberate interaction to passive consumption [12, 7].

Understanding how SoA is defined, theorized, and measured is therefore essential for interpreting the psychological consequences of ACDPs and for grounding the

empirical approach adopted in this thesis. In particular, the study of doomscrolling behavior requires not only identifying what users do (e.g., scroll frequency or session duration), but also clarifying how they experience what they do—especially in terms of control and intentionality [6, 25].

2.4.1 Definitions and Theoretical Models

In its broadest sense, the Sense of Agency (SoA) refers to the subjective experience of being the originator of one’s own actions and of exerting control over their consequences in the external world [37, 38, 39]. Rather than being an abstract construct, agency represents a fundamental component of conscious experience, allowing individuals to perceive themselves as causal actors and to differentiate self-generated actions from externally driven events [39, 37].

As articulated by Haggard and Tsakiris, agency can be succinctly described as the experience of “being in control of one’s actions and, through them, events in the outside world” [37]. This experience underpins everyday interaction with both physical and digital environments, supporting autonomy, intentionality, and self-regulation. Importantly, however, the sense of agency is not a unitary or static phenomenon.

Neurocognitive research has shown that agency emerges from multiple interacting processes operating at different levels of awareness. In particular, the influential framework proposed by Synofzik et al. distinguishes between two complementary components of agency that are especially relevant for the analysis conducted in this thesis: the **feeling of agency** and the **judgment of agency** [43, 38].

- The *feeling of agency* refers to a low-level, pre-reflective experience that arises automatically during voluntary action. It is a non-conceptual and largely embodied sensation generated by rapid sensorimotor processes, reflecting the immediate coupling between action and perceived outcome [38, 37]. For example, when a user scrolls through a feed and the content moves smoothly in response to the finger’s motion, this fluent contingency produces an immediate sensation of control that characterizes the feeling of agency.
- In contrast, the *judgment of agency* operates at a higher, reflective level. It corresponds to an explicit, post-hoc attribution of authorship, expressed through conscious beliefs such as “I caused this” or “this happened because of my action” [38, 37]. This evaluative process integrates not only motor signals, but also contextual information, prior beliefs, expectations, and narrative interpretations of the interaction [43]. As such, it is more susceptible to cognitive biases and retrospective reinterpretation.

This layered nature of the Sense of Agency is crucial for understanding the impact of attention-capture damaging patterns. Design mechanisms such as infinite scroll

and never-ending autoplay may preserve, or even enhance, the feeling of agency at the level of micro-interaction by providing immediate, responsive feedback to user input. At the same time, their cumulative effect often manifests at the level of judgment of agency. After prolonged passive consumption, users may retrospectively perceive that the time spent did not align with their original intentions, leading them to attribute control to the platform rather than to themselves [44, 30]. This discrepancy between momentary control and reflective evaluation constitutes a key psychological basis for experiences of regret and dissatisfaction.

To explain how these experiences of agency are generated, the literature has proposed several theoretical models, now generally regarded as complementary rather than mutually exclusive. Predictive or comparator models posit that agency arises when the brain’s predictions about the sensory consequences of an action match the actual feedback received [38]. Conversely, postdictive accounts emphasize that agency judgments are constructed retrospectively through cognitive inference, based on how events are interpreted after they occur [43].

The most widely accepted contemporary perspective integrates these views within an optimal cue integration framework. According to this model, the brain dynamically combines multiple sources of information, such as sensorimotor predictions and contextual cues, to generate an overall experience of agency [38, 37].

2.4.2 Sense of Agency in HCI

Long before the emergence of the digital well-being discourse, the *Sense of Agency* was already recognized as a core principle in the design of effective interactive systems. Classical Human–Computer Interaction (HCI) literature emphasizes the importance of supporting an *internal locus of control*, grounded in the observation that users value the perception of being in command of the interface and of seeing the system respond coherently to their actions. As articulated by Shneiderman and Plaisant, successful interfaces are those that foster “the feeling that the user is in control of the interface and that the system responds to their actions” [45]. Within this perspective, a high Sense of Agency has traditionally been considered a hallmark of good interaction design, as it promotes feelings of autonomy, competence, and effectiveness.

With the rise of the *Attention Economy*, however, this principle has increasingly been put under strain. As digital platforms have become optimized for maximizing engagement, the relationship between Sense of Agency and user experience has shifted from a design objective to a critical point of tension. A growing body of research indicates that a weakened Sense of Agency is associated with passive interaction styles, excessive use, and negative post-use emotions such as frustration and regret [4, 28]. Consequently, the Sense of Agency has become a central construct

in the study of digital well-being.

In response to these concerns, HCI research has begun to systematically investigate how specific interface mechanisms, particularly *Attention-Capture Damaging Patterns* (ACDPs), contribute to the erosion of perceived control [12]. Empirical evidence illustrates this conflict clearly. For example, an experimental study by Schaffner et al. on Netflix revealed that features such as autoplay often place users in situations where their perceived intentions are misaligned with their actual behavior. Participants reported watching more content, and different content, than they had originally planned, highlighting a direct tension between engagement-oriented design and user agency [30].

Most work in this area has relied on **explicit measures** of agency to quantify these effects. A seminal contribution is the study by Lukoff et al. on YouTube, which combined surveys and co-design sessions to examine how different interaction mechanisms shape user experience [44]. Their findings show that features supporting deliberate planning and explicit choice, such as search and playlists, tend to strengthen the Sense of Agency, whereas automated mechanisms like recommendations and autoplay systematically undermine it. Similar conclusions have been reached in experimental studies on streaming platforms, where disabling autoplay led to statistically significant reductions in both viewing time and session length [30].

More recently, HCI researchers have started to explore the use of **implicit measures** of agency in order to capture more subtle aspects of user experience. For example, Didion et al. considered temporal binding paradigms to assess the *feeling of agency* in interactions with generative AI systems, demonstrating that agency can be studied beyond self-report measures even in complex and automated environments [46]. These approaches provide promising tools for analyzing agency as it unfolds during interaction, rather than only through retrospective evaluation.

Across different studies, a consistent pattern always emerges. Algorithmically curated feeds tend to minimize the Sense of Agency in a gradual and imperceptible manner. While users perform actions such as scrolling, the results of those actions are basically determined by the platform’s optimization objectives rather than by the user’s intentions. This introduces a form of *causal ambiguity*: the user acts, but the system decides [44, 41].

As concluded by Lukoff et al., contemporary platforms often provide strong support for ritualized use while offering limited support for instrumental, goal-oriented interaction. Users frequently report that their intentions are diverted by automated playback and endless recommendations, resulting in prolonged engagement that feels misaligned with their original goals [44]. This dynamic, in which interface design privileges passive and habitual consumption over intentional use, constitutes the core mechanism investigated in this thesis and will be analyzed in depth in the specific context of TikTok.

2.5 Short-Video Platforms as a Case Study

The rapid diffusion of social media platforms over the last two decades provides the structural background against which contemporary short-video applications have emerged. As shown in Figure 2.2, as stated by the Digital 2025 report of Simon Kemp [15], the adoption of social networks has followed a steady and sustained growth trajectory, reaching a majority of the world’s population in recent years. What initially started as niche online communities in the early 2000s, such as Facebook or Myspace in 2004 [15], has progressively evolved into a dominant mode of digital communication and content consumption.

This widespread adoption has created the conditions for increasingly competitive attention markets, in which platforms differentiate themselves not merely by functionality, but by their ability to maximize user engagement and time spent. Within this context, short-video platforms represent the latest evolutionary stage of social media, characterized by highly optimized interaction loops and content delivery mechanisms explicitly designed for rapid, continuous consumption.

Among these platforms, TikTok stands out as a particularly relevant case study due to both its scale and the extremity of its engagement-oriented design.

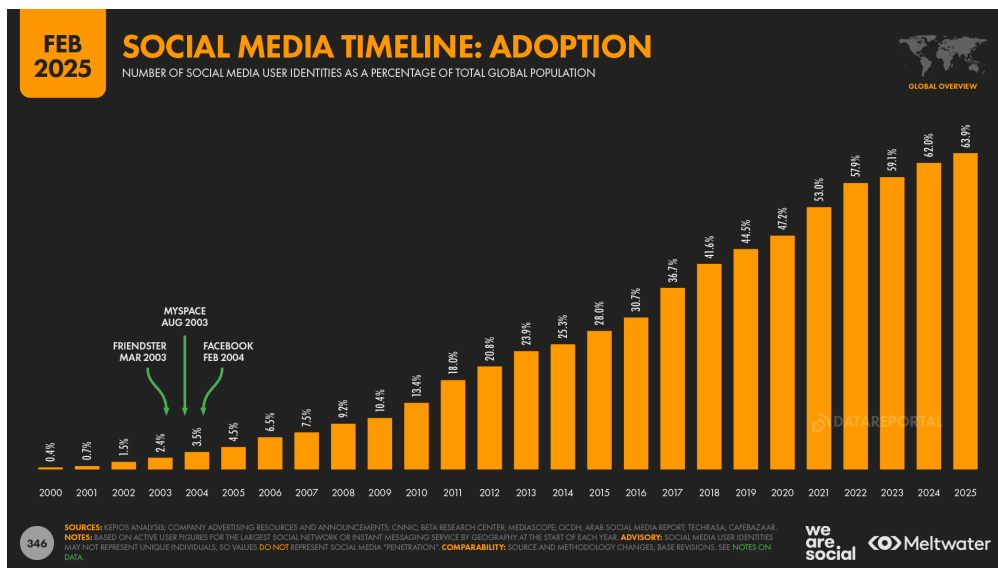


Figure 2.2: Number of social media user identities (single individuals with multiple social media accounts) as a percentage of total global population [15]

2.5.1 The case of TikTok

The social media application known as TikTok represents the most emblematic case study for analyzing the effects of attention-capture mechanisms discussed in

the previous sections. Its rapid global adoption and its exceptional ability to retain users for extended periods make TikTok a paradigmatic example of a platform designed according to the principles of the attention economy.

According to recent global reports, TikTok has surpassed 1.5 billion monthly active users worldwide, positioning itself among the most widely used social media platforms [15]. However, what truly differentiates TikTok from other social networks is not merely the size of its user base, but the intensity of user engagement. As illustrated in Figure 2.3, TikTok exhibits the highest average monthly time spent per user among major social media applications, exceeding long-established platforms such as YouTube, Facebook, and Instagram [47].

This result is particularly noteworthy given TikTok’s relatively recent emergence compared to its competitors. Prior research suggests that such high levels of engagement are not incidental, but rather the outcome of deliberate design choices aimed at minimizing decision friction and maximizing interaction continuity [31, 12]. TikTok systematically combines two central Attention-Capture Damaging Patterns (ACDPs): *Never-ending Autoplay* and an extreme form of *Infinite Scroll*, implemented through vertical swiping in a full-screen video interface.

Unlike platforms such as YouTube, which still support more goal-oriented modes of use (e.g., search, playlists, manual selection), TikTok is structured almost entirely around a single dominant interaction mode: the *For You page*. This algorithmically curated feed delivers a potentially infinite sequence of short videos, dynamically selected to maximize immediate engagement [41, 48, 40]. Users are not required to formulate explicit goals; instead, they are gradually absorbed into a continuous stream of stimuli, where each item represents a new potential reward.

From this perspective, TikTok should be regarded not only as a commercially successful platform, but also as a natural laboratory for studying the psychological and behavioral consequences of attention-optimizing interface design. Its global diffusion, extreme engagement levels, and structural reliance on ACDPs make TikTok an ideal context for investigating the relationships between interaction design, normative dissociation, and SoA core themes.

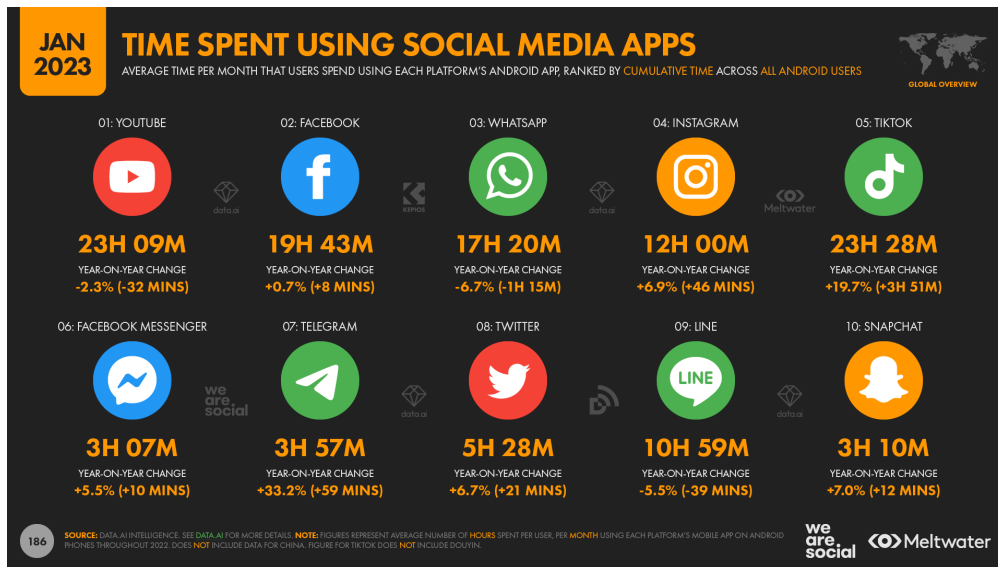


Figure 2.3: Average time per month that users spend using each platforms. Data collected from Android devices, ranked by cumulative time.

2.5.2 Interaction Patterns and Algorithmic Feeds

What distinguishes TikTok is not merely the presence of individual Attention-Capture Damaging Patterns (ACDPs), but their tight integration and structural centrality within the overall user experience. Unlike other platforms, where mechanisms such as Infinite Scroll or Autoplay coexist with more articulated interfaces, in TikTok these mechanisms effectively *constitute the interface itself* [49, 31, 12]. The core interaction experience (illustrated in Figure 2.4) emerges from the convergence of three key design elements:

- **Immersive full-screen video:** Each piece of content occupies the entire visual space, maximizing perceptual impact, minimizing external distractions, and facilitating deep attentional absorption [49].
- **Never-ending autoplay:** The transition between videos is automatic and immediate, removing natural stopping points and eliminating explicit decision moments for the user [12].
- **Vertical infinite scroll:** Progression to the next item is reduced to a minimal motor action, a simple vertical swipe, revealing an apparently limitless stream of new stimuli [31].

Together, these elements produce an almost entirely frictionless consumption

experience, explicitly designed to reduce cognitive effort and promote passive, continuous engagement [40]. The resulting interaction model minimizes opportunities for reflection and disengagement, encouraging users to remain within the content flow for extended periods.

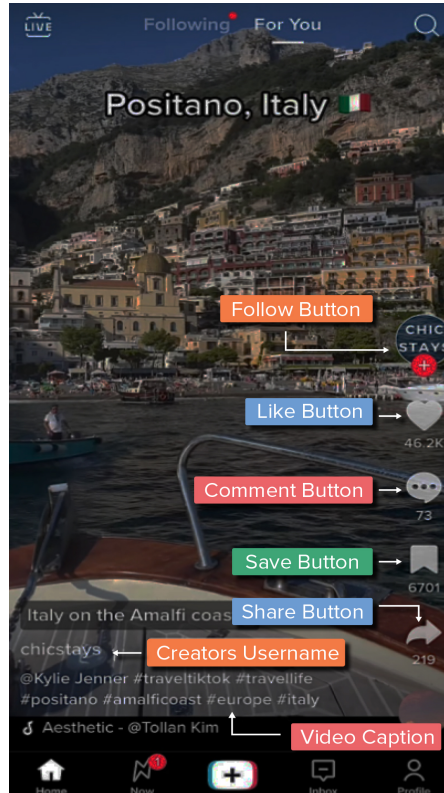


Figure 2.4: The *For You* page interface of TikTok, characterized by full-screen video, vertical infinite scroll, and automatic content transitions.

The orchestration of this experience is driven by TikTok’s powerful recommendation algorithm operating within the *For You* page. As highlighted by Feng et al., TikTok represents the most advanced example of a *content-centric* recommendation system, marking a clear departure from earlier *people-centric* social network models such as Facebook [49]. The primary objective is not to surface content from accounts explicitly followed by the user, but to continuously deliver videos that maximize engagement potential, regardless of their source.

The behavioral consequences of this architecture are effectively captured by the *double-loop* model proposed by Rixen et al., previously introduced in Section 2.2.2 [31]. TikTok excels both at triggering the *outer loop* of habitual app checking and, most notably, at sustaining the *inner loop* of session scrolling for unprecedented durations. As a result, brief and seemingly intentional interactions are frequently

transformed into prolonged sessions of passive consumption.

2.5.3 Evidence of Impact on User Behavior

An increasing body of empirical research has begun to document the specific behavioral and psychological effects associated with TikTok’s distinctive interaction architecture. One of the most robust quantitative findings is provided by the field study conducted by Rixen et al., which reports that usage sessions on TikTok are significantly longer than those observed on other major social media platforms such as Facebook, Instagram, and Twitter [31]. This objective evidence supports the hypothesis that TikTok’s design is particularly effective at sustaining user engagement, often extending usage well beyond users’ initial intentions.

Beyond effects on self-awareness and time perception, emerging research suggests that intensive use of TikTok can also influence higher-level cognitive processes. Experimental evidence by Jiang and Ma indicates that frictionless interaction paradigms typical of short-video platforms, characterized by minimal effort and rapid content transitions, can reduce analytical thinking and reflective processing, a phenomenon summarized as *“swiping more, thinking less”* [50]. Although such findings remain exploratory, they provide initial support for the claim that prolonged exposure to highly optimized attention-capture interfaces may have measurable cognitive consequences.

Recent conference-based studies further reinforce these observations. Zhang and Liu, for instance, report that frequent TikTok users exhibit patterns of repetitive engagement and reduced perceived control over content consumption, particularly during extended sessions [51]. Their findings highlight how algorithmically curated feeds and continuous playback contribute to habitual use cycles that are difficult for users to interrupt voluntarily.

Taken together, these converging lines of evidence indicate that TikTok’s combination of market dominance, unique interaction design, and documented behavioral effects makes it a particularly relevant case study for investigating attention capture, loss of intentionality, and digital well-being. This convergence provides a strong empirical foundation for the experimental investigation conducted in the remainder of this thesis.

Chapter 3

Study Design

This chapter presents the study design and the methodology adopted to explore doomscrolling as a specific form of problematic interaction with social media platforms. After discussing the theoretical background and related work in Chapter 2, this section shifts the focus to the practical side of the research, explaining how the phenomenon was observed and analyzed in real-world usage scenarios.

The main goal of the study is to analyze the patterns of user interaction that are associated with doomscrolling during daily use of TikTok, combining objective metrics collected with subjective user feedback. For this reason, a longitudinal in-the-wild study was conducted, allowing participants to use a custom mobile application in their natural environment over an extended period of time. This approach was chosen to ensure greater validity and to capture authentic interaction behaviors that may not emerge in controlled laboratory settings.

The methodology adopts multiple perspectives, integrating quantitative objective data derived from logged interaction, with qualitative subjective measures collected through questionnaires and user labeling (interactive pop-up). Behavioral metrics were defined to capture key aspects of user interaction, such as scrolling dynamics, engagement intensity, and temporal perception, which have been identified in the literature as relevant indicators of attention capture and potentially harmful usage patterns [52]. At the same time, we analyzed the perceived experience, awareness, and self-evaluation of the user's behavior.

This chapter is structured as follows. Section 3.1 introduces the research objectives and formalizes the research questions guiding the study. Section 3.2 describes the overall study design, including its longitudinal nature and experimental phases. Finally, section 3.3 defines the variables of interest and presents the operationalization of behavioral metrics used in the analysis.

3.1 Research Objectives and Research Questions

Given the exploratory nature of the study and the absence of experimentally manipulated interface conditions, the research adopts an observational approach in which research questions are formulated around naturally occurring usage sessions. From these research questions, a set of hypotheses is derived to guide the empirical analysis and to structure the interpretation of the collected data.

RQ1: Can doomscrolling be characterized through recurring behavioral patterns observable during real-world social media usage sessions?

This first research question addresses whether sessions perceived by users as doomscrolling present distinguishable interaction characteristics when compared to non-problematic sessions. Drawing on prior work on attention capture, automatic interaction, and loss of control, the study investigates whether specific behavioral metrics can serve as indicators of such sessions.

Based on the literature and on the conceptualization of doomscrolling as a state of prolonged, automatic, and passive consumption, the following hypotheses are formulated:

- **H1:** Sessions labeled by users as doomscrolling are associated with longer session durations compared to non-doomscrolling sessions.
- **H2:** Doomscrolling sessions exhibit higher scrolling intensity and more impulsive scrolling patterns than non-doomscrolling sessions.
- **H3:** Doomscrolling sessions are characterized by lower levels of active interaction compared to non-doomscrolling sessions.

RQ2: Does awareness-oriented feedback influence users' reflection on their social media usage during a longitudinal study in-the-wild?

Grounded in principles of reflective design and digital well-being interventions, the following hypothesis is formulated:

- **H4:** Repeated exposure to awareness-oriented feedback increases users' reflection on their social media usage over the course of the longitudinal study.

RQ3: Can a Large Language Model support the definition of threshold values for the detection of doomscrolling based on behavioral metrics?

Rather than evaluating the model's ability to produce objectively optimal thresholds, this research question focuses on the plausibility, consistency, and

empirical alignment of AI-generated thresholds within the context of an exploratory study. Accordingly, the following hypotheses are formulated:

- **H5:** Threshold values generated with the support of a Large Language Model are coherent with the distribution of observed behavioral metrics.
- **H6:** AI-supported threshold definitions exhibit meaningful alignment with users' session-level doomscrolling labels.
- **H7:** The use of an AI-based approach facilitates the adaptability of threshold-based doomscrolling detection in an exploratory setting.

3.2 Study Design

The study was structured into five sequential phases, spanning from the preliminary conceptual investigation to the final analysis of the collected data. This phased organization ensured methodological coherence and allowed the gradual refinement of the research instrument before its deployment in a real-world longitudinal setting.

Unlike controlled laboratory experiments based on manipulated interface conditions, the present research adopted a longitudinal in-the-wild approach. This choice was motivated by the need to observe doomscrolling in everyday contexts, without altering the platform interface or constraining user behavior, as doomscrolling usually appears spontaneously, capturing more authentic data. Participants interacted with social media platforms in their natural environment for **14 consecutive days**, allowing the observation of authentic behavioral patterns associated with doomscrolling. The overall study design is articulated as follows.

- **Phase 1 – Preliminary Study** The first phase consisted of a preliminary conceptual investigation aimed at conceptualizing the definition of doomscrolling and identifying relevant behavioral indicators of this phenomenon.

Based on the theoretical framework discussed in Chapter 2, this phase involved the identification of core indicators related to attention capture, automatic scrolling behavior, passive consumption, and temporal distortion. Then, based on these constructs, a set of potential behavioral metrics was defined to frame doomscrolling in a measurable way.

- **Phase 2 – Development of the Research Instrument** The second phase involved the design and development of a custom mobile application, later used as the main research instrument. The application was designed to collect interaction data during real-world social media usage (TikTok specifically) to provide meaningful suggestions (e.g., pop-ups, personalized feedback) to the user. The tool was designed to:

- compute session-level behavioral metrics (e.g., interaction count, average watch time, etc.).
- collect session-level subjective labels (e.g., perceived session time, subjective doom label).
- deliver awareness-oriented feedback.

The development phase enabled the operationalization of the constructs identified in the preliminary study into measurable variables suitable for empirical analysis.

- **Phase 3 - Participant Recruitment** Following the development of the research instrument, participants were recruited through a screening process. A preliminary questionnaire was distributed to identify eligible participants based on predefined inclusion criteria, such as regular social media usage (specifically TikTok), ownership of a compatible mobile device and willingness to participate in a 14-day longitudinal study.
- **Phase 4 - Longitudinal Experimental Phase** The fourth phase represented the core empirical component of the study. Participants used their smartphones as usual for 14 consecutive days, while the application collected interaction data during social media sessions. This phase can be further divided into two sub-phases: during the *Warmup phase* (7 days) the application passively collects data regarding user behavior; while during the *Detection phase* (7 days) the data collected is used to define the usage thresholds in order to activate the detection algorithm.
- **Phase 5 - Data Analysis** The final phase consisted of aggregating, cleaning, and analyzing the collected data. The main goal was to investigate the ability of the application to influence user behavior and mitigate toxic usage of TikTok. The data collected will display slight changes in behavioral metrics, towards a more conscious use of social media.

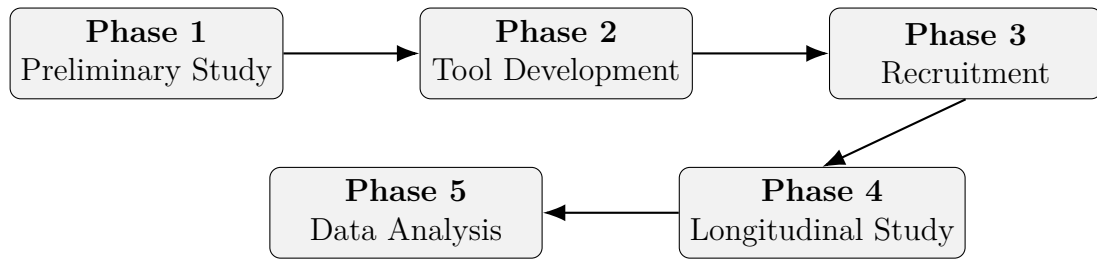


Figure 3.1: Phases of the research process.

3.3 Variables and Operational Definitions

In order to empirically investigate the research questions introduced in Section 3.1, it was necessary to define the variables involved in the study and clarify their respective roles. We evaluated independent and dependent variables, allowing the investigation to be structured in analytical terms. This distinction is essential to ensure that the collected data can be systematically analyzed and interpreted. Given the observational nature of the research and the absence of experimentally manipulated conditions, variables were defined in terms of behavioral predictors and outcome measures derived from real-world interaction data and subjective user input.

The study integrates objective behavioral metrics with subjective self-reported measures, enabling a quantitative analytical framework.

3.3.1 Independent Variables: Behavioral Predictors

In the present study, independent variables correspond to behavioral metrics extracted from interaction data collected during social media (TikTok) usage sessions. These variables were not experimentally manipulated but were observed as naturally occurring predictors of users' subjective perception of doomscrolling and temporal distortion. The following behavioral predictors were embedded at the session level:

- **Actual Session Duration** The total time elapsed between the beginning and the end of a social media session.
- **Scrolling Intensity and Velocity** Quantitative measures capturing the frequency and speed of scrolling gestures within a session, intended to reflect automatic or impulsive interaction patterns.
- **Impulsive Scroll Rate** The proportion of rapid, consecutive scrolling events occurring within short temporal intervals.

- **Interaction Rate** The frequency of deliberate user interactions (e.g., likes, comments, shares) relative to session duration (less than 2 seconds).
- **Content Skipping Rate** The proportion of content items viewed for very short durations.
- **Depth of Engagement** A composite metric that captures the extent to which users meaningfully watch the content encountered during a session. It is computed as the complement of the content skip rate, where a video is considered “skipped” if its watch time is shorter than a predefined threshold of 2 seconds.

Formally:

$$\text{DepthOfEngagement} = 1 - \frac{N_{\text{skipped}}}{N_{\text{total}}} \quad (3.1)$$

where N_{skipped} represents the number of videos watched for less than 2 seconds, and N_{total} is the total number of videos observed during the session.

- **Average Watch Time** The mean duration spent viewing individual content items within a session.
- **Content Distribution** A list of categories relative to videos watched during a scroll session

3.3.2 Dependent Variables: Subjective and Analytical Outcomes

The dependent variables in this study correspond to outcome measures reflecting users’ subjective perception and analytical classification results, such as perceived scrolling time, and a boolean information on doomscrolling perception.

The primary dependent variable is:

- **User Doomscrolling Label (Subjective)** After selected sessions, participants were asked to indicate whether they perceived the session as an instance of doomscrolling. This binary self-report measure helps the LLM with the computation of the initial thresholds during the *detection phase*.
- **Temporal distortion index** Defined as the discrepancy between perceived session duration (captured at the end of some sessions through an interactive pop-up) and actual session duration, this variable captures potential time distortion effects.

```
dayId: "20260211"
labelType: "LABELED"
metrics
  avgScrollVelocity: 1.5724521279335022
  avgWatchTime: 230348
  contentDistribution: [{"category": "cooking", "su...}
  contentSkippedRate: 0
  depthOfEngagement: 1
  durationMinutes: 5.131
  endTime: 1770809869121
  impulsiveScrollRate: 0
  interactionRate: 2
  interactionsCount: 4
  startTime: 1770809561261
  totalScrollEvents: 2
sessionId: "177bc371-00a0-45ef-b50e-d0e2da0baf30"
timestamp: 1770809877329
userEstimatedMinutes: 8
userLabelDoom: false
```

Figure 3.2: Sample of session data, the metrics attribute shows each independent variable defined

3.3.3 AI-Supported Threshold Variables

In relation to RQ3, personalized threshold values were derived with the support of a Large Language Model (LLM). The motivation for using an LLM was not to “predict” doomscrolling directly, but to support the definition of user-specific habits in a scenario where fixed global thresholds may be inadequate due to strong variability across users (e.g., different baseline scrolling speed, interaction frequency, or typical session length). For this reason, the resulting thresholds were treated as analytical parameters: they define how raw behavioral signals are interpreted by the detection pipeline.

Threshold initialization occurs at the end of the *Warm-up phase*. During this stage, the application collects a first dataset of sessions and associated user-provided labels. This labeled dataset is then summarized and provided to the LLM, which returns threshold values grounded in the observed distribution of the user’s behavior.

In practice, the Warm-up phase serves to estimate an initial baseline and to avoid starting the detection phase with arbitrary values.

After initialization, the thresholds are periodically updated to account for behavioral drift and adaptation over time. Updates are performed in batches: at fixed intervals, a new set of recent sessions is aggregated and used to refine the previous thresholds. This step follows an *in-context learning* approach: instead of retraining a model, the LLM is prompted with the current threshold set together with a compact representation of newly collected sessions. The output is an updated set of thresholds, designed to preserve continuity with the previous baseline while gradually incorporating new evidence from ongoing use.

Chapter 4

Design and Implementation of the Application

This chapter presents the design and implementation of the mobile application developed to support the empirical study described in Chapter 3. This chapter describes the system architecture, the internal application structure, the behavioral metrics computation logic, the threshold-based detection algorithm, the AI integration workflow, and the backend data model. Through this detailed analysis, the chapter demonstrates how theoretical constructs related to doomscrolling were translated into a functioning, deployable software system.

4.1 Design Objectives and System Requirements

The development of the application was guided by a set of design objectives derived from the research questions defined in 3.1. Since the application served as the primary research instrument for the in-the-wild study, its design needed to balance data collection capabilities with usability, privacy protection, and system reliability.

The primary design objectives of the system were the following:

- **Behavioral Monitoring** The application was required to monitor user interaction patterns without altering the natural usage behavior of participants. Also, the system was designed to collect only interaction metadata (e.g., scroll events, timestamps, interaction counts), without accessing sensitive data.
- **Real-Time Session-Level Detection** The application needed to support session-level doomscrolling detection in near real-time. The detection logic was implemented locally on the device to minimize latency and reduce dependency on network availability.

- **Interpretability of Detection Logic** Rather than employing machine learning classification models directly on the device, the detection mechanism relied on thresholds applied to behavioral metrics to increase feasibility.

These design objectives were translated into a set of Functional Requirements (FR) and Non-Functional Requirements (NFR), reported in Tables 4.1 and 4.2.

Functional requirements define what a system must do (features, user actions, data processing), while non-functional requirements define how the system performs those functions (speed, security, usability, reliability).

Table 4.1: Functional requirements of the mobile application.

ID	Requirement Description
FR1	The system shall compute session behavioral metrics (e.g., depth of engagement, interaction count, content distribution, etc.).
FR2	The system shall store aggregated session data in a cloud database.
FR3	The system shall generate threshold parameters from a backend service supporting AI-based generation.
FR4	The system shall apply threshold-based detection logic to session metrics locally on the device.
FR5	The system shall deliver awareness notifications when detection criteria are met.
FR6	The system shall prompt users to provide session-level subjective labeling (doomscrolling yes/no and perceived duration).
FR7	The system shall allow user to manually set his preferences about session duration threshold and daily usage limit.

Table 4.2: Non-functional requirements of the mobile application.

ID	Requirement Description
NFR1	The system shall operate continuously in the background during social media usage sessions without noticeable impact on device performance or battery consumption.
NFR2	The system shall guarantee data privacy by collecting only interaction metadata.
NFR3	The system shall use anonymized user identifiers to prevent direct identification of participants.
NFR4	The system shall be compatible with most Android devices.

4.2 Overall System Architecture

The overall system architecture is structured around a client which acts as the primary data collection, processing, and intervention unit. The architecture is composed of three main layers:

- the client-side application running on the Android device;
- a cloud storage layer for persistence;
- an external server used for threshold generation and adaptation.

The application is responsible for monitoring real-time interaction inside TikTok through an Accessibility Service able to compute and map behavioral metric, and the triggering of awareness feedback (e.g., overlays and notifications).

The cloud storage layer, implemented using Firestore Database, is used for persistence of interaction data and synchronization. The stored data include aggregated behavioral metrics and lightweight content category summaries.

Finally, the external server is used to correctly prompt OpenAI LLM (GPT-4.1 and GPT-4.1 mini) to generate thresholds and adjust them during the last phase.

As illustrated in Figure 4.1, the application operates in two distinct but interconnected phases:

- **Warm-up Phase** (baseline observation)
- **Detection Phase** (active threshold-based detection)

This separation ensures a clear distinction between baseline behavioral observation and threshold-based doomscrolling detection and is managed by the `AppPhaseManager.kt` component. This component allows for the transition between phases and manages the life-cycle of the longitudinal study.

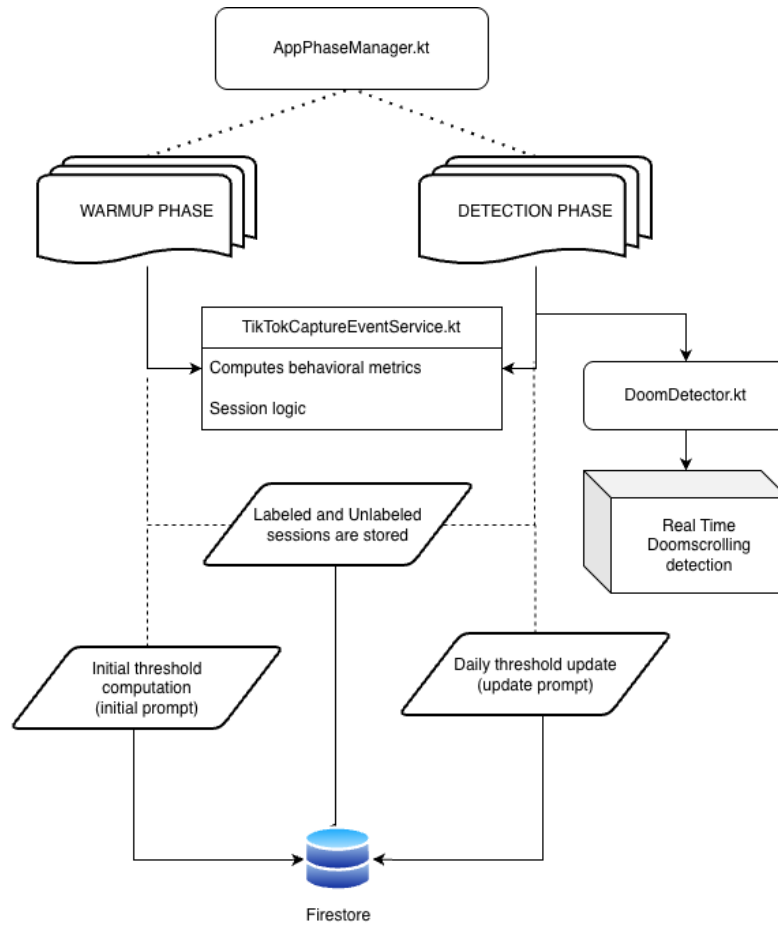


Figure 4.1: Overall System Architecture, displaying the main structural components of the application (e.g. `TikTokCaptureEventService.kt`)

4.2.1 Warm-up Phase: Baseline Behavioral Observation

During the Warm-up Phase, the system operates in passive monitoring mode, collecting data from everyday usage without providing feedback. The primary objective of this phase is to collect representative behavioral data without activating detection mechanisms. The procedural flow is as follows:

1. The `TikTokCaptureEventService` intercepts user interaction events through the Android Accessibility framework which logs general system events such as `TYPE_VIEW_SCROLLED` or `TYPE_VIEW_CLICKED`.
2. Raw interaction events are used to compute aggregated session-level structures and behavioral metrics.

3. Sessions may be labeled or unlabeled, depending on whether the user is asked for subjective input. The decision of showing the interactive pop-up is delegated to the `LabelingScheduler.kt` component. A maximum of 7 pop-ups can be shown per day, with a probability of 30% for each session.

At the end of the warm-up phase (**seven days**), during the first session of the detection phase, the application initiates a threshold initialization procedure. This procedure consists of sending to the backend server all labeled sessions collected during the warm-up phase, including both objective behavioral metrics and the subjective labels provided by the user ("doomscrolling" vs "not doomscrolling").

The backend service then builds a structured prompt (shown in Appendix A.1) and forwards it to the OpenAI API (GPT-4.1). The purpose of this request is to elaborate personalized numeric thresholds that best separate DOOM and NON-DOOM sessions based exclusively on the user's observed behavior.

A structured prompt refers to a carefully designed textual input to the Large Language Model (LLM), organized into semantically distinct sections that guide the model's reasoning process. Rather than sending raw data alone, the prompt embeds: role specification, task instructions, reasoning constraints, an in-context learning example, and the actual user dataset.

This structure serves three main purposes:

1. To constrain a response based on data contained in the dataset.
2. To enforce output format consistency.
3. To reduce generic assumptions.

The LLM is explicitly instructed to produce strictly formatted JSON containing only numeric threshold values, without commentary or explanatory text. This JSON output is then parsed by the backend service and stored as the personalized threshold set for the user.

4.2.2 Detection Phase: Threshold-Based Evaluation

In this phase, the system continues to capture interaction events through `TikTokCaptureEventService` and computes session-level metrics. However, unlike the Warm-up Phase, each session is now evaluated using the threshold-based detection algorithm implemented in `DoomDetector.kt`.

Thresholds are defined as either minimum (*min*) or maximum (*max*) numeric values associated with specific behavioral metrics. Each threshold represents a boundary condition that, when satisfied, contributes to the classification of the current session as a potential doomscrolling episode.

More precisely, a *min* threshold is triggered when the observed metric exceeds the defined value, while a *max* threshold is triggered when the observed metric falls below the specified value. Doomscrolling detection is not based on a single metric, but on the simultaneous activation of multiple threshold conditions.

Another key feature of this phase is the update of the thresholds. At the first session of each new day, the application triggers a backend request to recompute thresholds based on the sessions collected during the previous day. This process follows the same AI-supported workflow used at the end of the Warm-up Phase (as described in 4.2.1).

This recalibration mechanism was designed to allow the detection system to progressively adapt to evolving user behavior patterns.

4.3 Mobile Application Design

The mobile application was designed as a modular, event-driven Android system capable of operating continuously in the background. Since modern Android versions restrict background execution to preserve battery life and system performance, the application leverages an Accessibility Service declared as a foreground-bound system service. By registering as an accessibility listener, the application remains active even when not in the foreground, allowing it to capture interaction events in real time.

Particular attention was devoted to ensuring the separation of concerns between event capture, domain logic, data persistence, and user interface components.

The application structure reflects a layered architecture, organized into distinct packages corresponding to functional responsibilities as shown in Figure 4.2.

4.3.1 Technology Stack

The application was developed for Android devices using Kotlin as the main programming language. Kotlin was selected for its modern language features, strong type safety, concise syntax, and seamless interoperability with the Android SDK.

The core components of the technology stack include:

Android SDK and Accessibility Framework

TikTok events, such as scrolling, video interactions, or moving across tabs, are captured using the Android AccessibilityService APIs.

The `TikTokCaptureEventService` extends the `AccessibilityService` class, allowing the application to intercept basic interaction events and compute session-level behavioral metrics.

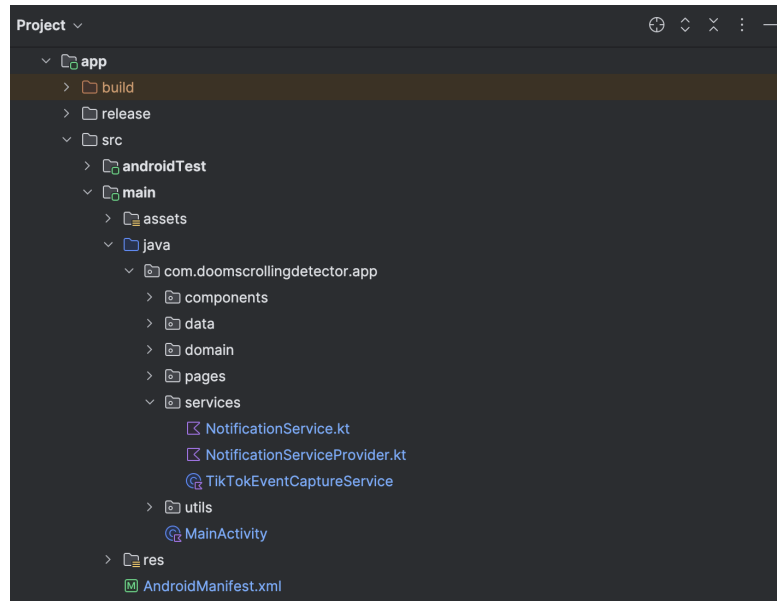


Figure 4.2: Package structure of the application in Android Studio

The main events processed by the `TikTokCaptureEventService` are:

- **`AccessibilityEvent.TYPE_VIEW_SCROLLED`** This event is handled by the corresponding function `handleScrollEvent(event, ts)`, which manages the computation of the scroll gesture in TikTok.
- **`AccessibilityEvent.TYPE_VIEW_CLICKED`** This event is handled by the corresponding function `handleClickEvent(event, ts)`, which correctly maps each click to a precise behavior (video interaction or tab switch).

In order to filter only events coming from TikTok, the `onAccessibilityEvent()` function applied a filter on the package-name of each events, which has to correspond to "com.zhiliaoapp.musically"

Firestore

Cloud persistence is implemented using Google Firebase Firestore, a NoSQL document-based database. Firestore was chosen for the following reasons:

1. its real-time synchronization capabilities,
2. structured hierarchical document storage,
3. seamless Android SDK integration,

Sessions and thresholds for each user are stored and uploaded in Firestore using the `FirestoreUploader` class, which includes different functions that serve this task.

The database follows the hierarchical schema:

`users/{userId}/day/{day}/sessions/{sessionId}`

`users/{userId}/day/{day}/thresholds/current`

Backend Service and OpenAI API

The application utilizes a backend server used to communicate with OpenAI. The main tasks are to:

- Compute the initial thresholds for a specific user
- Update daily the previous thresholds for a specific user
- Retrieve the context and category of videos watched during a session, using screenshots

The choice of using GPT-4.1 and GPT-4.1 mini through the OpenAI API service is due to the simplicity of using such an accurate LLM instead of pretrained smaller models.

4.3.2 Application Modules

The application is organized into multiple packages that reflect a layered architectural model. The main modules are described below.

Service Layer

The `services` package contains components responsible for system operations and background execution.

- **TikTokCaptureEventService** The core component of the system. This Accessibility Service intercepts interaction events, computes session behavioral metrics, and sends raw interaction data to the domain layer.
- **NotificationService and NotificationServiceProvider** These services manages notification creation, configuration, and triggering based on detection outcomes.
- **Overlay Services** They are responsible of displaying a pop-up window while using TikTok in real-time, mainly in case of doomscrolling detection or user subjective labeling.

Domain Layer

The `domain` package contains the core behavioral logic of the application.

- **MetricsAggregator** Aggregates raw interaction events into session behavioral metrics.
- **DoomDetector** Implements the detection algorithm. It evaluates session metrics against threshold values (min and max) and returns a structured decision object containing: detection level (**WARNING** if time metric exceeds but no other thresholds are triggered or **DOOM** if time metric exceeds with additional triggers), boolean doom decision and a list of triggered signals.
- **LabelingScheduler** Manages the timing and frequency of user labeling prompts during sessions (maximum of 7 pop-ups per day showed with a probability of 30%) .
- **AppPhaseManager** Controls the transition between Warm-up and Detection phases.
- **TikTokUsageEventsPoller** Supports periodic checks to verify opening and closing of the TikTok application

Data Layer

The data package encapsulates persistence and data transfer logic. The most important packages includes:

- `model/` Defines structured data classes such as `SessionMetrics`, `Thresholds`, `AggregateMetrics`, `ScrollEvent` and more.
- `repository/` Implements data access abstraction and data operations.
- `remote/` Manages communication with Firestore, by correctly mapping and uploading data.

This separation ensures that domain logic remains independent from the persistence mechanism.

UI Layer

The user interface components are organized under `pages/`, `components/`, and the main activity entry point. The UI layer is intentionally lightweight, as the primary computational logic resides in background services and domain components.

Table 4.3: Core application modules and their architectural responsibilities.

Module	Layer	Core Responsibility
TikTokCaptureEventService	Services	Captures TikTok interaction events via Accessibility Service and detects session boundaries.
MetricsAggregator	Domain	Aggregates raw interaction events into session-level behavioral metrics.
DoomDetector	Domain	Applies threshold-based detection logic and produces structured session decisions.
AppPhaseManager	Domain	Controls the transition between Warm-up and Detection phases.
LabelingScheduler	Domain	Manages scheduling of user labeling prompts during sessions.
NotificationService	Services	Delivers awareness-oriented feedback based on detection outcomes.
Repository / Remote Layer	Data	Handles persistence, Firestore synchronization, and communication with the backend threshold service.
UI Layer	Presentation	Manages application pages, and visualization of pop-up windows.

4.4 Session Management and Behavioral Metrics Computation Engine

The core component of the application is the `TikTokEventCaptureService`, which acts as the central orchestrator of interaction capture, session life cycle management, and real-time behavioral aggregation. Unlike simple usage trackers (such as Screen Time for iOS or Digital Well-being for Android) based exclusively on the application foreground time, this service implements a fine-grained event-driven architecture capable of detecting scroll gestures, user interactions, session boundaries, and contextual transitions within the TikTok interface.

This section describes in detail the logical definition of a session, the mechanisms used to capture interaction events, and the procedural flow that transforms raw events into structured data structures.

4.4.1 The Central Role of `TikTokEventCaptureService.kt`

`TikTokEventCaptureService` extends the Android `AccessibilityService` class and operates continuously in the background. It intercepts `AccessibilityEvent` instances generated by the TikTok application (package "com.zhiliaoapp.musically") and filters them according to relevance.

The service performs three fundamental tasks:

1. Detection of session start and end.
2. Capture and classification of interaction events (scrolls, clicks, content transitions).
3. Upload of session-level metrics.

4.4.2 Logical Definition of a Session

A session is defined as a continuous period of active interaction within the TikTok “For You” page. The session life cycle is managed through a state-based mechanism governed by internal flags such as `isSessionActive` and contextual tab tracking. A session begins when the following conditions are all true: TikTok enters foreground (detected by `TikTokUsageEventsPoller`), the active tab corresponds to "home" and there are no active sessions.

The method `maybeStartSession(ts: Long)` performs these checks and initializes:

- `sessionStartTime`
- `lastScrollTime`

Instead a session is terminated when one of these two conditions occurs: TikTok moves to the background or the user navigates away from the Home tab. Termination is handled by `endScrollSession()`.

At session end:

- Real-time detection is stopped (in the detection phase)
- Screenshot routines are stopped
- All interaction and scroll events are retrieved from `EventRepository`
- The session is persisted locally and uploaded to Firestore

Scroll Gesture Detection Logic

Scroll gestures are not directly interpreted as single events. Instead, they are constructed from consecutive `TYPE_VIEW_SCROLLED` accessibility events, which are frequently emitted during a single scroll gesture. When a scroll event is received, the service first verifies that a session is currently active and that the user is inside the Home tab (“For You Page”). The first useful scroll event initiates a gesture. At this moment, the system records the initial timestamp and the vertical position.

Then subsequent scroll events contribute to build the ongoing gesture by updating the cumulative vertical delta and incrementing the number of new events.

The termination of the scroll gesture is handled by a time out mechanism. If no new scroll events are detected within `NO_SCROLL_TIMEOUT` milliseconds (0.25 seconds), the gesture is considered complete.

Interaction and Click Detection

In addition to scroll gestures, the system monitors user interactions through `TYPE_VIEW_CLICKED` accessibility events. These Android events are used to capture the engagement in TikTok, such as likes, content interactions, and navigation between different tabs of the application. When a click event is received, the service first extracts specific information from the event data, including `event.text` and `event.source.viewIdResourceName`. This parsing enables the system to differentiate between navigation actions (e.g., switching between Home, Inbox, or Profile tabs) and content interactions.

4.5 AI-Supported Threshold Definition

A central design requirement of the system is the use of an LLM to compute and update the detection thresholds used by the doomscrolling application. Instead of manually defining fixed thresholds for each behavioral metric, the application relies on a dedicated backend service that prompts a Large Language Model to derive personalized threshold values from the user’s own session dataset.

This approach was adopted because doomscrolling is highly individual: a session duration or scrolling intensity that is problematic for one user may be completely normal for another.

From an architectural perspective, all LLM interactions are performed server-side. This ensures secure handling of API credentials, centralizes prompt management, and allows server-side validation of the returned thresholds before they are stored and delivered to the mobile client.

4.5.1 Prompt Engineering Strategy

The backend service employs two distinct prompts, reflecting two different moments of the longitudinal workflow: (i) the initial threshold computation at the end of the Warm-up Phase, and (ii) the daily update during the Detection Phase.

Initial threshold computation (`thresholdDefinerPrompt.txt`). At the end of the Warm-up Phase, the system collects a dataset of user labeled sessions containing objective behavioral metrics (e.g., session duration, scroll velocity,

impulsivity rate, content skipping rate, watch times, interactions), and subjective user labels (*doomscrolling* vs *not doomscrolling*), as well as subjective estimates of perceived duration. This dataset is serialized and injected into the prompt placeholder `{{DATASET}}`.

The prompt, shown in Appendix A.1, explicitly constrains the model to derive thresholds *only* from patterns present in the provided dataset, discouraging generic assumptions about doomscrolling. In addition, the prompt provides an in-context reasoning template and forces the model to return *only* a strict JSON object with a fixed schema.

This is important for reliability: the application expects a complete threshold set, and the output must be readable without additional processing.

In the event of a failure in the LLM request (e.g., network error, timeout, or invalid response format), the system falls back to a deterministic local strategy. In this case, each threshold is computed individually as the arithmetic mean of the corresponding metric values observed during the warm-up phase. Since all session metrics are stored locally on the device, this fallback mechanism guarantees continuity of operation and prevents blocking the transition to the Detection Phase.

Daily threshold update (`thresholdUpdaterPrompt.txt`). During the Detection Phase, thresholds are updated regularly to account for the behavioral shifts of the user over days. The update prompt, shown in Appendix A.2, receives the current threshold configuration (`{{CURRENT_THRESHOLDS}}`) and a batch of new sessions from the previous day (`{{NEW_SESSIONS}}`), which may include labeled and unlabeled sessions.

The main design principle in this prompt is conservativeness: the prompt instructs the model to avoid drastic changes unless multiple sessions consistently contradict the current threshold model. Labeled sessions are treated as stronger evidence than unlabeled ones as stated in the prompt, but even in that case the update is gradual. As in the initial computation prompt, the output is constrained to the same strict JSON schema, and the prompt explicitly requires returning every threshold field, even if unchanged.

4.5.2 Threshold Application Logic

Once generated, thresholds are stored and later applied entirely on-device using a decision mechanism implemented in `DoomDetector.kt`. This separation is intentional: the mobile application remains in full control of the detection logic, while the AI component only provides the values. This is important to reduce OpenAI API calls and therefore costs.

The threshold application logic follows a two stage structure. First, the detector applies a *time gate*: if the session duration does not exceed `minSessionDuration`,

the session is not problematic. This design choice keeps the duration of the session as the primary gatekeeper.

If the duration gate is passed, the detector evaluates a set of secondary behavioral metrics by comparing them with the corresponding thresholds. In the current implementation, doomscrolling is detected if at least one secondary condition is triggered (e.g., impulsive scrolling rate above `minImpulsiveScrollRate`, content skipping above `minContentSkippedRate`, scroll velocity above `minAvgScrollVelocity`, or depth of engagement below `maxDepthOfEngagement`). The detector returns both the final decision and the list of trigger signals, making the outcome observable for the user study.

This design choice reflects the role of secondary thresholds as complementary behavioral signals that characterize the mode of consumption, rather than its simple duration. While the primary condition (session duration exceeding `maxSessionDuration`) captures extended exposure, the secondary thresholds serve as indicators of attention traps, such as rapid content rejection (high impulsive scroll rate and skipping), automatic interaction patterns (high scroll velocity), and reduced intentional engagement (low depth of engagement).

Requiring at least one secondary condition (in addition to the primary one) was chosen to higher the accuracy of the model of doomscrolling detection. In contrast, relying solely on the duration of the session would be insufficient, as long sessions can also represent intentional and engaged usage.

4.6 User Interface and Interaction Design

Although the main functionality of the application operates primarily in the background, the user interface plays a crucial role in providing structured feedback to guide the user to mitigate excessive doomscrolling. The application was intentionally designed to remain minimally intrusive during normal usage, while still providing clear feedback and meaningful interaction when required.

The UI design follows three main principles:

- **Non-intrusiveness:** monitoring occurs silently in the background without interfering with normal TikTok usage.
- **Transparency:** users can understand what is being measured and why feedback is provided.
- **User-Centered Feedback:** the application has to provide personalized feedback and meaningful information.

The following sections describe the main UI components and interaction flows.

4.6.1 Main Application Screens

The application is primarily made of 4 different sections to navigate: *Dashboard*, *Insights*, *Notifications*, and *Settings*. These sections are accessible through a bottom navigation and each of them serves a distinct purpose within the overall interaction model.

Dashboard

The Dashboard represents the main overview screen of the application (Figure 4.3). Its purpose is to provide a concise summary of the user's recent activity.

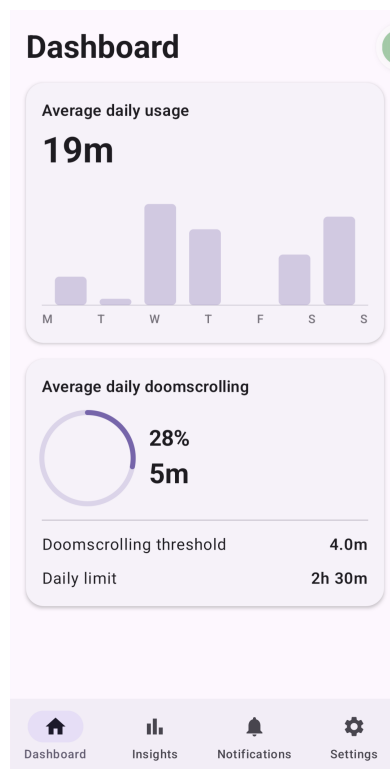


Figure 4.3: Dashboard overview displaying average daily usage and doomscrolling indicators.

The upper card presents the **average daily usage** throughout the week, visualized through a simple bar chart. This allows users to quickly identify patterns and fluctuations in their engagement over time.

Below, a circular indicator summarizes **average daily doomscrolling**, showing both the percentage of doomscrolling in relation to total usage time and the corresponding value in minutes. Importantly, the screen also displays the current

session duration threshold and daily limit. This reinforces transparency, so users can see the parameters against which their behavior is being evaluated.

Insights

The Insights section (Figure 4.4) provides a more detailed breakdown of the last doomscrolling session.

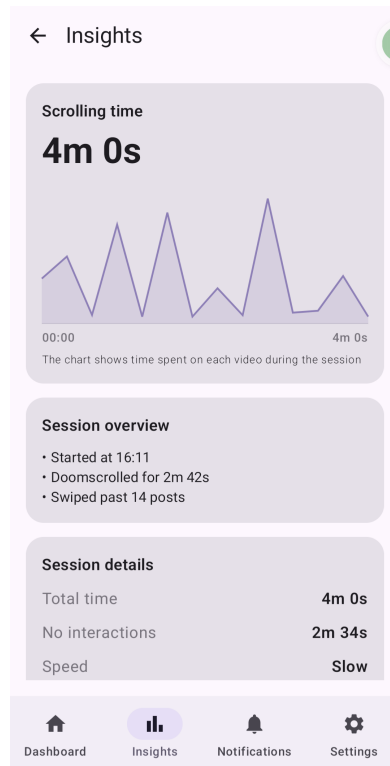


Figure 4.4: First part of Insights screen showing session-level metrics and scrolling patterns.

The screen includes a **time graph** that visualizes the temporal distribution of watch times during a session, giving a useful insight to the user.

Below the graph, a **session overview** summarizes key elements such as session start time, total doomscrolling duration, and number of posts swiped. An additional section **session details** provides further contextual information, including total time, interaction patterns, and scrolling speed.

The Insights screen presents a dedicated section to personalized and subjective suggestions, i.e. a set of informative instructions built around triggered thresholds (e.g., suggesting to diminish scroll speed if one of the triggers includes velocity).

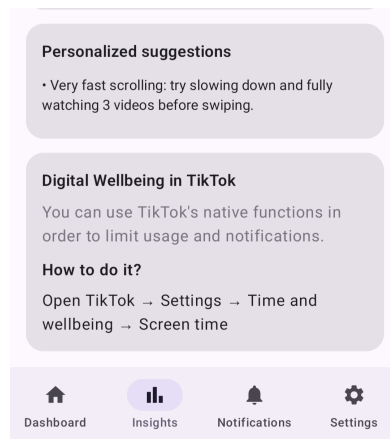


Figure 4.5: Second part of Insights screen showing personalized subjective suggestions and TikTok digital well-being tutorial.

Notifications

The Notifications section (Figure 4.6) contains notifications about previous doomscrolling sessions and awareness messages generated by the detection engine.

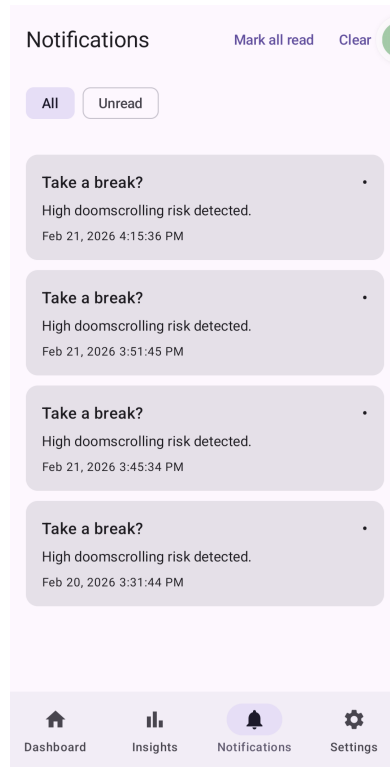


Figure 4.6: Notifications history showing detected high-risk sessions.

When a session is classified as *doom* during the Detection Phase, a notification is generated and stored in this section. The interface allows users to review past alerts, providing temporal context, and aiming to reinforce accountability.

Settings

The Settings section (Figure 4.7) allows users to set important parameters such as the daily usage limit and the doomscrolling time threshold.

Although threshold values are determined primarily through the AI mechanism described in Section 4.5, users still have the ability to adjust certain limits manually based on their preferences.

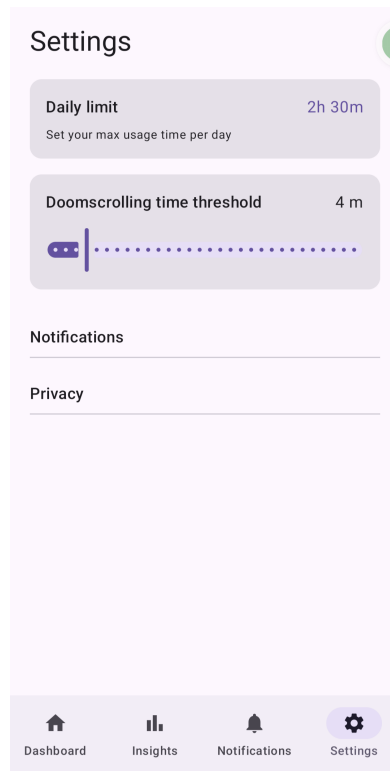


Figure 4.7: Settings screen allowing adjustment of daily limits and doomscrolling thresholds.

4.6.2 Overlays

Unlike the main application screens, which are accessed intentionally by the user, overlays represent visual interventions that appear directly on top of TikTok during usage. They are implemented as interactive banners/interfaces and are only triggered in specific moments.

Two distinct overlay types are used within the application: the **session labeling overlay** displayed at the end of selected sessions during the Warm-up phase, and the real time **doomscrolling detection overlay**.

Session Labeling Overlay

At the end of selected sessions during the Warm-up Phase, the application displays a short feedback overlay designed to collect subjective information from the user (Figure 4.8).

The purpose of this overlay is not to interrupt the user experience aggressively, but to ask sporadically for subjective information. The interface asks the user

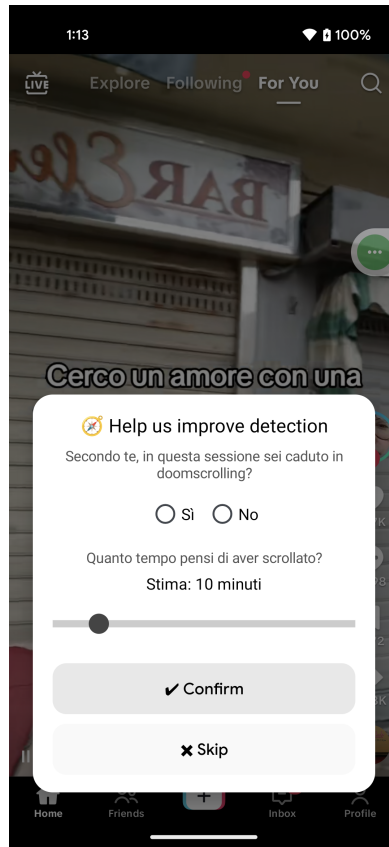


Figure 4.8: Session labeling overlay shown at the end of selected sessions during the Warm-up Phase.

whether the session felt like doomscrolling or not. In addition, users are asked to estimate how long they believe they have been scrolling.

From a technical standpoint, the overlay is triggered by the `LabelingScheduler` after `endScrollSession()` is executed and only during the Warm-up Phase. The overlay can appear at most 7 times during the day, and with a strict percentage of appearance of 30%, in order to limit excessive intrusiveness of the application in daily usage of TikTok.

Doomscrolling Detection Overlay

During the Detection Phase, a different overlay can appear while the session is still ongoing (Figure 6.1).

This overlay is triggered when the real-time detection engine classifies the current session as DOOM. The alert is based on the threshold application logic described in Section 4.5.2.

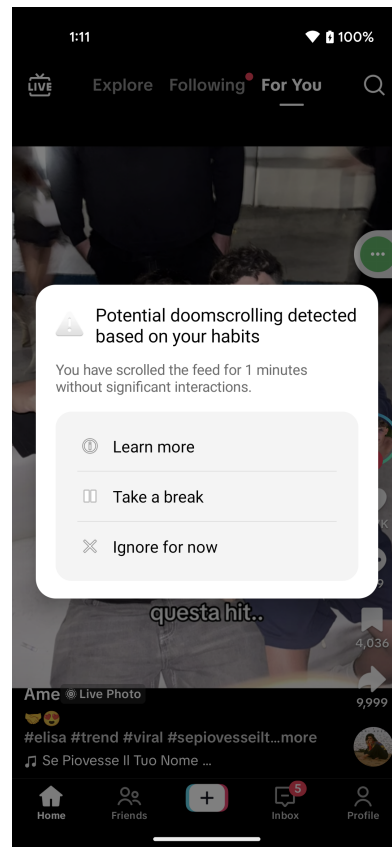


Figure 4.9: Real-time doomsscrolling detection overlay displayed during high-risk sessions.

Rather than blocking TikTok usage, the overlay acts as a reflective interruption, displaying a personalized message about excessive usage of TikTok.

Three possible actions are offered:

- *Learn more*, which redirects the user to the Insights page;
- *Take a break*, closing the application (TikTok);
- *Ignore for now*, allowing immediate continuation.

Importantly, this interaction design was directly informed by RQ2 and its corresponding hypothesis H4. Rather than enforcing restrictive measures, the system was designed to provide awareness feedback that encourages user reflection while preserving autonomy. The *Learn more* option supports reflective insights by going into detail about behavioral metrics; *Take a break* suggests voluntary self-regulation; and *Ignore for now* preserves user agency, avoiding forced interruption.

This design aligns with the research hypothesis that repeated exposure to awareness-oriented feedback increases reflection over time. By allowing users to repeatedly encounter, and optionally act on feedback, the UI becomes a core experimental instrument to evaluate whether it directly influences behavioral awareness during the longitudinal study.

Additionally, the system includes a "*cool down*" mechanism to prevent repeated alerts within a short time window. This ensures that interventions remain meaningful rather than disruptive.

Chapter 5

Study Methodology

5.1 Participant Recruitment and Sample

The selection of an appropriate sample was essential to ensure the validity and reliability of the study results. Since the research aims to analyze natural doom-scrolling behaviors in everyday contexts, it was important to recruit participants who actively use TikTok and exhibit heterogeneous usage patterns. This section describes the inclusion criteria, the recruitment procedure, and the characteristics of the final sample that participated in the longitudinal study.

5.1.1 Inclusion Criteria

Participants were recruited based on a set of predefined criteria, with the objective of ensuring that the sample consisted of individuals with relevant and regular experience in social media usage.

The inclusion criteria were defined as follows:

1. Being at least 18 years old.
2. Owning a personal Android smartphone compatible with the research application.
3. Being an active user on social media platforms.
4. Being an active user of TikTok specifically.
5. Being willing to install the research application and participate in a 14-day longitudinal in-the-wild study.

These criteria ensured that participants were familiar with the interaction patterns under investigation and that sufficient behavioral data could be collected during the study period.

5.1.2 Recruitment Procedure

The recruitment was conducted through the distribution of an online screening questionnaire (the full text of which is reported in Appendix B). The questionnaire was distributed through university-related messaging groups (e.g. Telegram and WhatsApp) and through a snowball sampling process.¹

Candidates who satisfied all inclusion criteria were subsequently contacted via email and provided with detailed instructions regarding the installation of the mobile application, the study procedure, and ethical considerations, including informed consent and the possibility to withdraw from the study at any time without consequences.

Recruitment continued until the desired number of participants was reached.

5.1.3 Sample Characteristics

The study involved a small sample of five participants. The sample consisted of 2 females and 3 males, with a mean age of 27,4 years (range 24-59). This choice was motivated by the exploratory nature of the research, which required continuous data collection and active user labeling over an extended period of time. Given the high granularity of the logged behavioral data, each participant generated a substantial number of sessions and interaction events.

Based on the information collected through the screening questionnaire, all participants reported regular use of at least one social media platform characterized by infinite scrolling mechanisms (40% of participants stated of using socials more than 2 hours, 40% stated of using socials between 1 and 2 hours and 20% less than 60 minutes, with TikTok and Instagram Reels as the most used platforms).

All participants completed the 14-day longitudinal study. The resulting dataset includes multiple session-level observations per participant, allowing for analysis across different usage contexts and days.

5.2 Experimental Procedure and Data Collection

This section describes the experimental procedure adopted during the longitudinal study and the mechanisms used to collect behavioral and subjective data. The procedure was designed to ensure methodological rigor and data consistency.

¹Snowball sampling is a recruitment technique in which already enrolled participants help identify and recruit additional eligible subjects, starting from a small initial group and progressively expanding the sample.

5.2.1 Experimental Procedure

After completing the screening questionnaire and satisfying all inclusion criteria, eligible participants were contacted via email and added to a dedicated mailing list of research participants. Through this mailing list, participants received detailed instructions regarding the installation procedure, study timeline, and general participation guidelines.

The participants were instructed to install the application on their personal Android devices. In order to enable the data collection required for the study, three specific system-level permissions were requested: Accessibility Service access, used to monitor scrolling events and interaction patterns within TikTok; Notification access, required to allow the application to provide notifications; and Overlay permission, allowing the system to display pop-ups for session labeling and detection feedback.

Participants were explicitly informed about the purpose of each authorization and about the type of data that would be recorded.

Each participant used TikTok as usual over a period of 14 consecutive days. No artificial tasks or constraints were imposed, in order to preserve naturalistic usage conditions.

The user study was structured into two sequential phases: a *warm-up phase* and a *detection phase*. This division allowed the system to initially observe a baseline behavior of the specific user in order to then activate detection mechanisms through computed thresholds.

Longitudinal Study

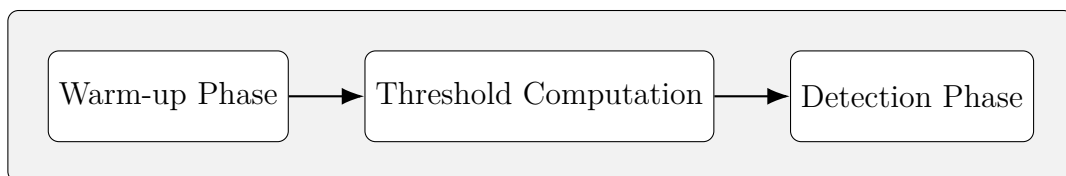


Figure 5.1: Structure of the user study, including baseline observation and threshold-based detection.

Warm-up Phase

The first phase of the study, called *warm-up phase*, was designed to collect interaction data without triggering any detection feedback. During this initial period, the application passively saved interaction events and computed session metrics, then uploaded online using Firestore.

The purpose of this phase is precise. It enabled the collection of initial behavioral metrics and sessions **subjectively labeled** by the user, which were later used to

determine more precisely the values of computed thresholds used for doomscrolling detection.

During this phase, no notifications and session feedback were provided, and participants were not explicitly informed about specific detection criteria to avoid priming participants towards particular usage patterns.

Detection Phase

Following the warm-up phase, the study entered the *detection phase*. During this phase, the application continued to collect interaction data while activating doomscrolling detection through a set of specific thresholds.

When session’s behavioral metrics exceeded predefined threshold values, the system generated awareness-oriented feedback. During this phase, user’s subjective evaluation of random sessions continued in order to always have a more accurate dataset.

The thresholds defined at the beginning of this phase are updated daily. The update process relies on in-context learning using the GPT-4.1 model [53]. At each iteration, a structured dataset of sessions is provided to the model through a carefully designed prompt.

The daily batch includes unlabeled sessions collected during the previous day. Together with the previously computed threshold values, these session-level metrics are passed as contextual input. The model is instructed to revise the thresholds while maintaining consistency with the previously established baseline and incorporating new observed patterns.

5.2.2 Data Collection Infrastructure

Behavioral and subjective data were collected through the mobile application and securely stored in a cloud-based database infrastructure.

The system was designed to record only interaction metadata and a brief summary of the content category watched. Content categorization is performed through periodic visual sampling: approximately every 30 seconds, a screenshot of the currently displayed video is captured. The image is then processed using the GPT-4.1-mini model, which provides a high level categorization of the content (e.g., entertainment, news, fitness, etc.).

Only the computed category label is stored for analysis, while the raw image is not permanently captured. This approach enables contextual understanding of viewed content, providing more accuracy to doomscrolling evaluation.

For data persistence and synchronization, the study used a *Firestore* database, part of the Firebase platform provided by Google. Firebase is a cloud backend-as-a-service (BaaS) solution that offers authentication, storage, and real-time data

management functionalities for mobile and web applications.

Firestore Database is a NoSQL document-oriented database designed to support scalability. In the context of this study, it was used to securely store session metrics and user labels for each participant. Each participant was assigned an **anonymous identifier**, ensuring that collected data could be linked correctly while preserving user privacy.

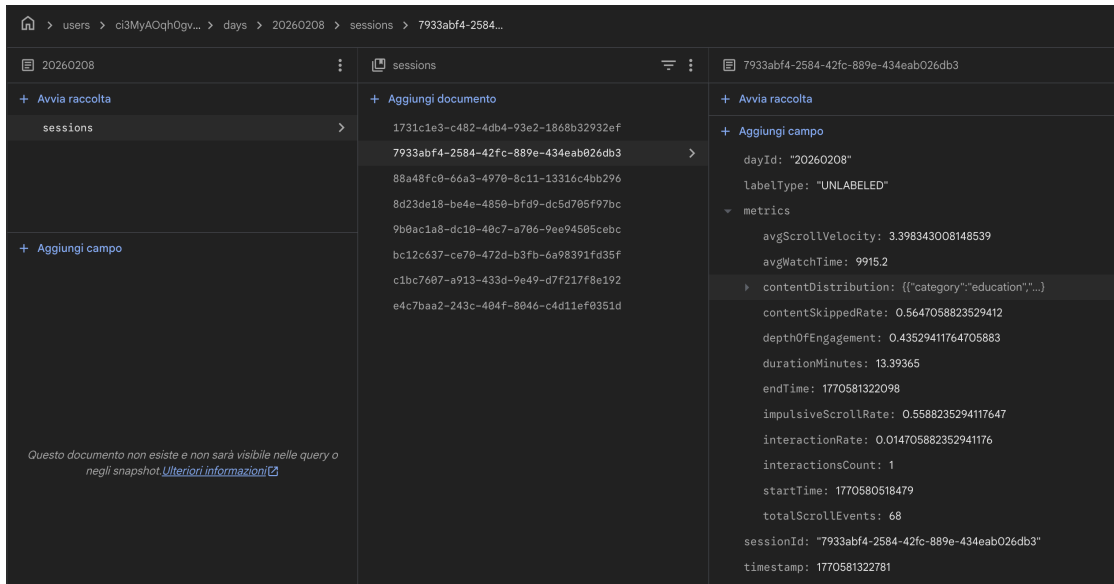


Figure 5.2: Example of a session saved on Firebase for a given user, showing the structure and organization of data

The following data categories were collected as shown in Figure 5.2:

- Session metadata (start time, end time, duration, timestamp);
- Aggregated behavioral metrics (depth of engagement, interaction rate, content skipped rate, average watch time, etc.);
- Content distribution with a list of categories regarding the content viewed
- Subjective session metrics (only for *LABELED* sessions)

Data were securely transmitted from the application to the database and organized at the session level. This structure enabled repeated-measures analysis across multiple sessions for each participant during the 14-day longitudinal period.

The data collection strategy was designed in accordance with principles of data minimization and privacy preservation, ensuring that only information strictly necessary for the research objectives was recorded.

Table 5.1: Session variables collected and stored in the Firestore database.

Variable	Type	Description
Independent Variables (Behavioral Predictors)		
<code>durationMinutes</code>	Continuous	Total duration of the session (in minutes).
<code>totalScrollEvents</code>	Discrete	Total number of scroll events recorded during the session.
<code>avgScrollVelocity</code>	Continuous	Average scrolling velocity computed across the session.
<code>impulsiveScrollRate</code>	Continuous	Proportion of rapid consecutive scroll events within short time intervals.
<code>interactionRate</code>	Continuous	Ratio of active user interactions (e.g., likes, comments) over session duration.
<code>interactionsCount</code>	Discrete	Total number of explicit interactions performed during the session.
<code>contentSkippedRate</code>	Continuous	Proportion of content items skipped after brief exposure.
<code>depthOfEngagement</code>	Continuous	Metric capturing sustained interaction with specific content items.
<code>avgWatchTime</code>	Continuous	Mean viewing time per content item within the session.
<code>contentDistribution</code>	Categorical (Aggregated)	Distribution of viewed content categories within the session.
Dependent Variables (Labeled Sessions only)		
<code>userLabelDoom</code>	Binary	User-reported indication of whether the session was perceived as doomscrolling (Yes/No).
<code>userEstimatedMinutes</code>	Discrete	User-estimated duration of the session (in minutes).

Chapter 6

Results

6.1 Data Preparation

The empirical analysis was conducted on data collected during a *14 day* user study involving a concise but meaningful sample of five participants. Each participant completed a 7 day warm-up phase followed by a 7 day detection phase. During the first phase, the sessions were subjectively labeled by the users as doomscrolling or not-doomscrolling, while during the detection phase, the sessions were automatically classified by the system based on personalized threshold values (as described in 4.5).

All session data were exported from Firestore into three structured JSON files, including: all recorded sessions (≈ 1500 records), sessions detected as doomscrolling during the detection phase (≈ 100 records), and the daily threshold configurations (≈ 50 records). The files collected were, respectively:

- sessions.json
- sessions_doom.json
- thresholds.json

Before the analysis, the dataset was subjected to a cleaning and preparation process to ensure consistency and comparability among the participants.

First, users who did not complete a sufficient portion of the study (at least 12 days) were excluded, ensuring consistent representation and adequate temporal coverage to observe stable behavior patterns and exposure to the system's feedback mechanisms. Second, micro-sessions (i.e., sessions shorter than 3 seconds) were removed from the dataset, excluding from the dataset brief interactions mostly associated with accidental openings or transitional app switches, that did not reflect meaningful usage behavior. Third, the behavioral metrics embedded within each session were extracted and structured into a unified analytical dataset.

The final dataset consisted of all valid sessions in both phases for the included participants, with each session annotated with the following:

- Phase (warm-up or detection)
- Doom label (subjective or system-detected)
- Behavioral metrics (e.g., duration, impulsiveScrollRate, contentSkippedRate, interactionRate, avgWatchTime, avgScrollVelocity)

This preparation process ensured that the following analyses were performed on consistent and meaningful data while minimizing noise and unnecessary records.

6.2 Research Questions Analysis

This section presents the empirical analysis conducted to address the three research questions introduced in chapter 3. This analysis pursues three objectives: first, to examine whether doomscrolling can be characterized through observable and recurring behavioral patterns; second, to evaluate whether awareness feedback influences user behavior over time; and third, to assess the effectiveness of the AI mechanism of identifying correctly doomscrolling sessions.

6.2.1 Characterizing Doomscrolling (RQ1)

To investigate whether doomscrolling can be characterized by recurring behavioral patterns, we conducted a feature activation analysis based on high-intensity behavioral signals. In addition, we performed a cross-user consistency analysis to examine whether the detected behavioral signals were consistent between participants. Specifically, we analyzed how frequently each high-intensity signal was triggered within doomscrolling sessions for different users. The goal of this analysis was to verify whether the same signals consistently characterized doomscrolling among participants, or whether the phenomenon manifested through different patterns depending on the individual user.

Instead of comparing simple averages, the feature activation analysis focuses on the likelihood that a session exhibits extreme behavioral values. A session was considered to activate a signal if its metric exceeded the global 75th percentile (or fell below the 25th percentile for the interaction rate). This approach allows the analysis to capture high-intensity patterns without being overly sensitive to isolated outliers.

For each metric, we then estimated the conditional probability of signal activation separately for doom and non-doom sessions. In particular:

$$P(\text{signal} \mid \text{doom})$$

represents the proportion of doom sessions whose metric crossed the predefined percentile threshold, while

$$P(\text{signal} \mid \text{non-doom})$$

represents the same proportion computed over non-doom sessions.

The comparison between these probabilities was summarized through a risk ratio (RR), defined as:

$$RR = \frac{P(\text{signal} \mid \text{doom})}{P(\text{signal} \mid \text{non-doom})}$$

A risk ratio greater than 1 indicates that the behavioral signal is more likely to occur during doom sessions. Higher values suggest stronger associations between the metric and doomscrolling behavior.

rq1_feature_activation_results

metric	type	P(signal doom)	P(signal non_doom)	risk_ratio
impulsiveScrollRate	high	0,385	0,113	3,407
contentSkippedRate	high	0,385	0,145	2,650
avgScrollVelocity	high	0,385	0,145	2,650
avgWatchTime	high	0,385	0,145	2,650

Figure 6.1: Results derived from the Feature Activation Analysis done for RQ1

Results indicate that doom sessions were:

- **3.4** times more likely to exhibit high impulsiveScrollRate
- **2.6** times more likely to exhibit high contentSkippedRate
- **2.6** times more likely to show high scroll velocity
- **2.6** times more likely to have a high average watch time

These findings demonstrate the presence of consistent and recurring behavioral signatures associated with doomscrolling behavior.

Although feature activation analysis revealed strong behavioral signals associated with doom sessions (with $RR > 2$), cross-user consistency analysis showed variability between participants, suggesting that doomscrolling may manifest through partially subjective behavioral patterns rather than a single universal behavioral pattern.

6.2.2 Behavioral Changes Across Phases (RQ2)

To investigate whether awareness-oriented feedback influenced user behavior over time, we adopted a strategy focused on three main analyses.

Three complementary analytical approaches were adopted: daily doom rate trend analysis, behavioral comparison across early and late detection days, and qualitative inspection of selected participant trends.

Daily Doom Rate During Detection

For each participant, days were ordered chronologically and divided into warm-up (first seven active days) and detection (subsequent days).

For each detection day, a daily doom rate was calculated as:

$$\text{doom rate}_{\text{day}} = \frac{\text{number of detected doom sessions}}{\text{total sessions}}$$

This metric captures the relative frequency of doomscrolling episodes in correlation with total sessions of the day. Linear regression models were fitted individually for each participant using the detection day index as the independent variable and the daily doom rate as the dependent variable. The slope (β) was used to estimate the direction of the behavioral change, while R^2 was used to assess the strength of the linear trend.

The results revealed heterogeneous patterns among the participants. One participant exhibited a moderate positive slope ($\beta = 0.051$), indicating an increase in the doom rate over time. A second participant showed a slightly negative slope ($\beta = -0.023$), suggesting a slight reduction. A third participant showed a slope close to zero ($\beta = 0.024$), indicating no consistent change. The determination coefficients remained relatively low (approximately between 0.01 and 0.18), suggesting that daily doom rate fluctuations were only weakly explained by a linear progression.

At the group level, the aggregated mean daily doom rate showed a small positive slope ($\beta = 0.026$) with moderate explanatory strength ($R^2 \approx 0.32$). This higher explanatory power likely results from the aggregation process, which smooths individual behavioral fluctuations and reduces noise.

In general, these findings do not indicate a systematic reduction in doomscrolling frequency during the detection phase. Instead, behavioral evolution appears to be strongly individual.

Behavioral Intensity Reduction

While the daily doom rate captures changes in the relative frequency of detected episodes, it does not describe how the behavioral *intensity* of sessions evolves over

time. To inspect more subtle adaptations, we compared early and late detection behavior for each participant.

For each user, the detection phase was split into two temporal segments (early vs. late detection days).

- **Early Detection:** first half of detection days
- **Late Detection:** last half of detection days

For each segment, we computed the mean value of key behavioral metrics (session duration, impulsiveScrollRate, contentSkippedRate, interactionRate, avgWatchTime, and avgScrollVelocity). We then calculated the relative change between segments as:

$$\Delta\% = \frac{\mu_{\text{late}} - \mu_{\text{early}}}{\mu_{\text{early}}} \cdot 100$$

The resulting deltas for each user were finally aggregated to obtain a group-level overview of how behavioral intensity shifted over time.

Overall, the analysis revealed a mixed pattern rather than a uniform reduction in intensity. On average, avgScrollVelocity decreased by 19.6% (std = 41.8%) and avgWatchTime decreased by 23.3% (std = 72.1%), suggesting that participants tended to scroll more slowly and spend less time per video during the late detection period. However, other indicators showed opposite dynamics: contentSkippedRate increased by 30.4% (std = 21.0%) and impulsiveScrollRate increased by 27.0% (std = 42.2%), each result pointing to a more fragmented consumption style characterized by frequent skipping and faster decision.

The relatively high standard deviations observed for some metrics indicate substantial variability between participants. This variability is consistent with the individualized nature of doomscrolling behavior, which can manifest itself through different interaction patterns depending on personal habits, content exposure, and engagement levels.

Session duration also increased substantially (+66.3%, std = 99.4%), indicating that the overall time spent per session did not consistently decrease during detection. The very high standard deviation suggests a strong difference between participants and sessions, indicating that changes in session duration were not uniform between users. This variability should therefore be interpreted with caution, as the sample dimension and the length of the longitudinal study may increase noise and dispersion. Finally, interactionRate decreased by 33.2% (std = 23.5%), suggesting fewer explicit interactions (e.g., likes or shares) in late detection sessions.

Taken together, these findings do not support a consistent reduction in behavioral intensity during the detection phase. Instead, changes appear heterogeneous, reinforcing the interpretation that awareness-oriented feedback may affect users in different ways depending on individual habits and different contexts.

rq2_intensity_summary_clean

metric	mean	std
avgScrollVelocity	-19,605	41,753
avgWatchTime	-23,313	72,125
contentSkippedRate	30,389	21,026
durationMinutes	66,303	99,442
impulsiveScrollRate	26,982	42,232
interactionRate	-33,190	23,544

Figure 6.2: Behavioral Intensity Change summary, showing percentage of change and std for each metric

Longitudinal Trajectory Visualization

To further investigate how behavior evolved during the detection phase, we conducted a longitudinal trajectory analysis focusing on two participants with the highest number of valid sessions (participant A and participant B). Rather than relying exclusively on aggregated statistics, this analysis examined the evolution of key behavioral metrics throughout detection.

For each detection day, we computed the daily doom rate and the mean values of selected behavioral indicators, including session duration, *impulsiveScrollRate*, *contentSkippedRate*, and *interactionRate*. Linear regression models were computed for each metric to estimate directional trends, but the primary focus remained on the observed daily trajectories.

Participant A. Across the six detection days, the daily doom rate exhibited a mild upward trend ($\beta = 0.051$, $R^2 = 0.18$), suggesting no reduction in the frequency of detected doom episodes. Session duration also showed a slight increase over time.

However, a different pattern emerged when considering behavioral intensity indicators. *ImpulsiveScrollRate* displayed a clear decreasing trajectory ($\beta = -0.028$, $R^2 = 0.52$), indicating a progressive reduction in impulsive scrolling behavior. *InteractionRate* similarly declined over time ($\beta = -0.027$, $R^2 = 0.32$), while *contentSkippedRate* remained relatively stable.

Taken together, these findings suggest that although the number of detected doom sessions did not decrease, certain aspects of behavioral intensity may have attenuated during detection, potentially reflecting a partial adaptation effect.

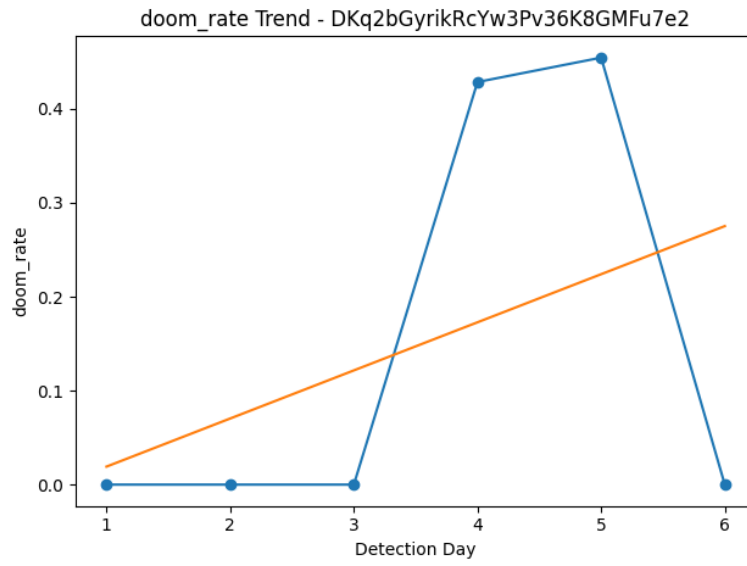


Figure 6.3: Doom Rate trend for participant 1

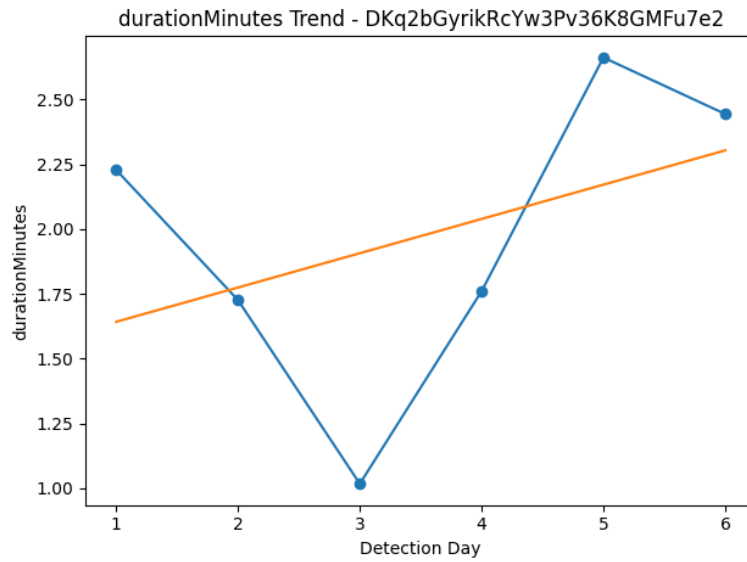


Figure 6.4: Session Duration trend for participant 1

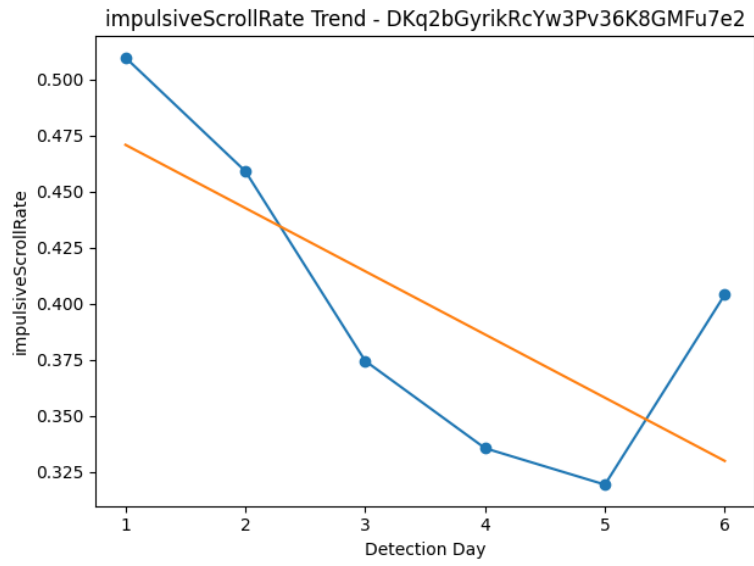


Figure 6.5: Impulsive Scroll Rate trend for participant 1

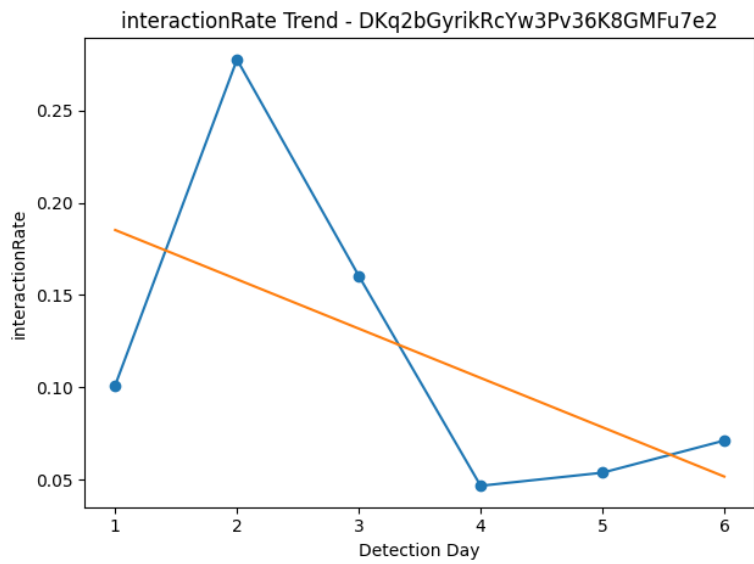


Figure 6.6: Interaction Rate trend for participant 1

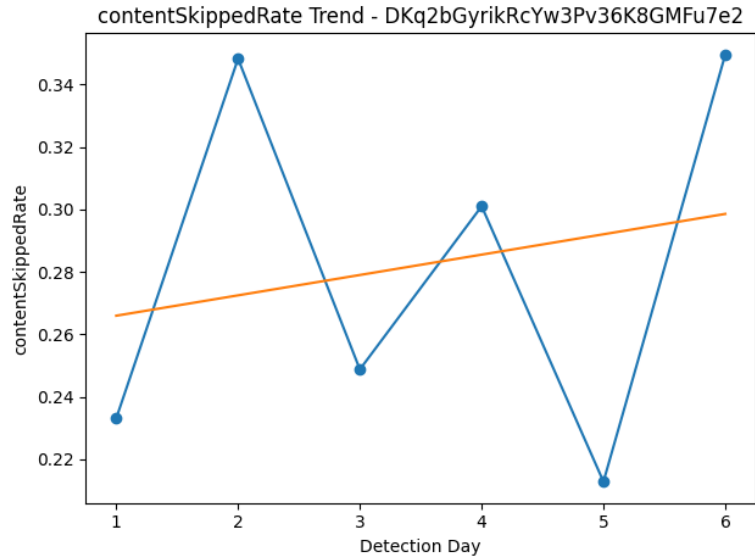


Figure 6.7: Content Skipped rate trend for participant 1

Participant B. In contrast, the second participant exhibited a different trajectory. The doom rate showed only a weak upward trend ($\beta = 0.024$, $R^2 = 0.06$), indicating substantial daily variability without a strong directional pattern.

In addition, other behavioral metrics revealed a more systematic evolution. Session duration increased considerably over time ($\beta = 0.645$, $R^2 = 0.51$), also *impulsiveScrollRate*, which showed a pronounced upward trajectory ($\beta = 0.078$, $R^2 = 0.78$). *ContentSkippedRate* also increased ($\beta = 0.043$, $R^2 = 0.48$), suggesting a progressively more fragmented and rapid consumption pattern. *InteractionRate* remained essentially stable.

This combination of increasing duration and impulsive indicators suggests a trajectory of intensification rather than regulation during the detection phase.

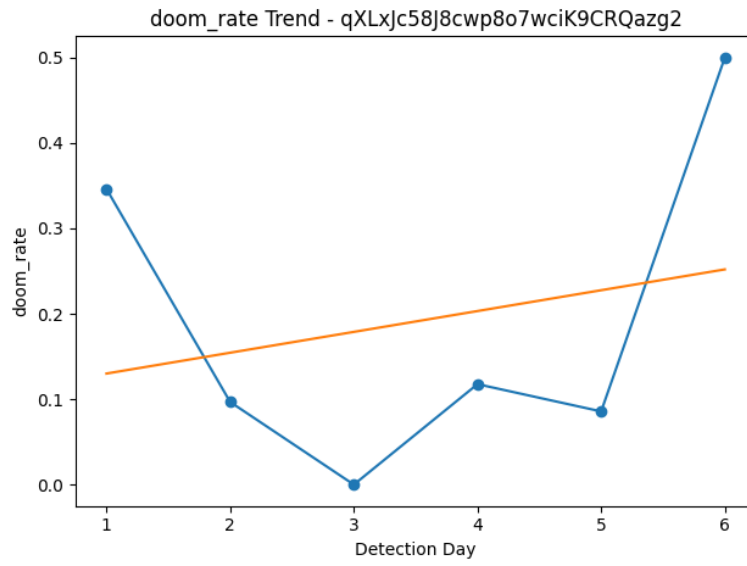


Figure 6.8: Doom Rate trend for participant 2

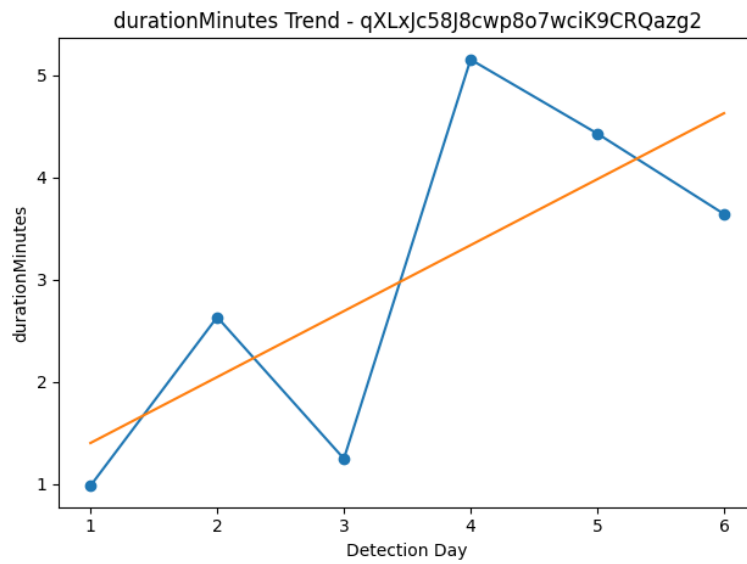


Figure 6.9: Session Duration trend for participant 2

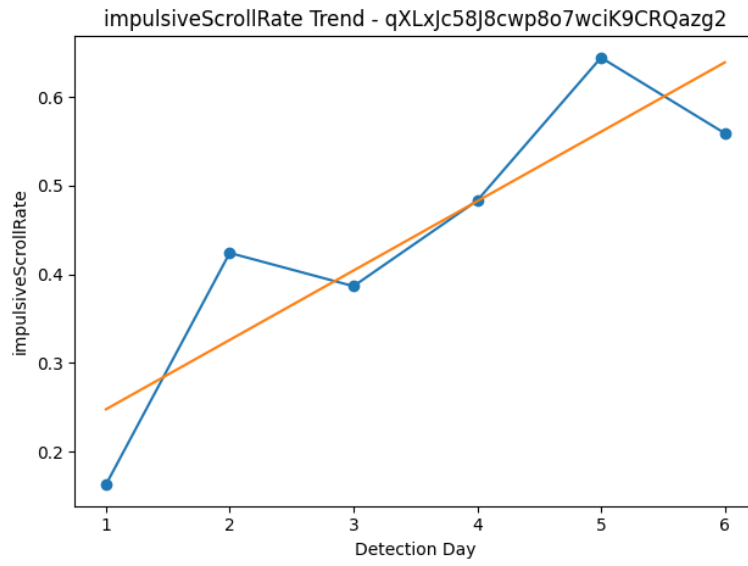


Figure 6.10: Impulsive Scroll Rate trend for participant 2

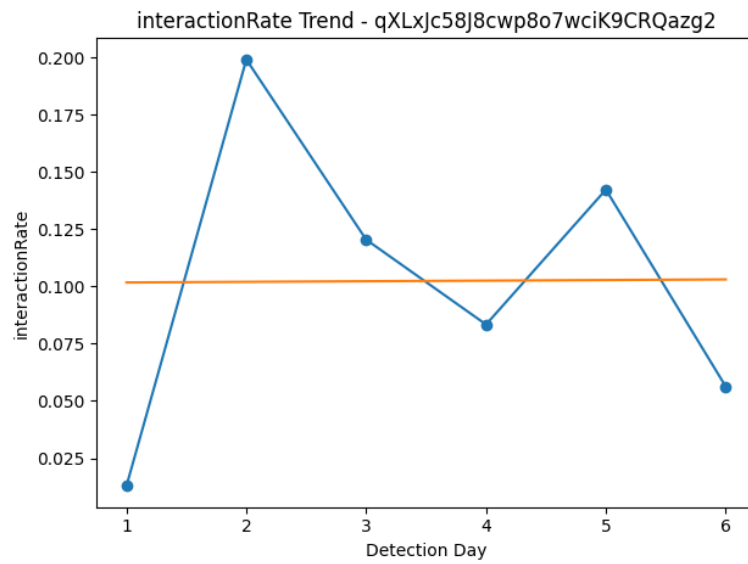


Figure 6.11: Interaction Rate trend for participant 2

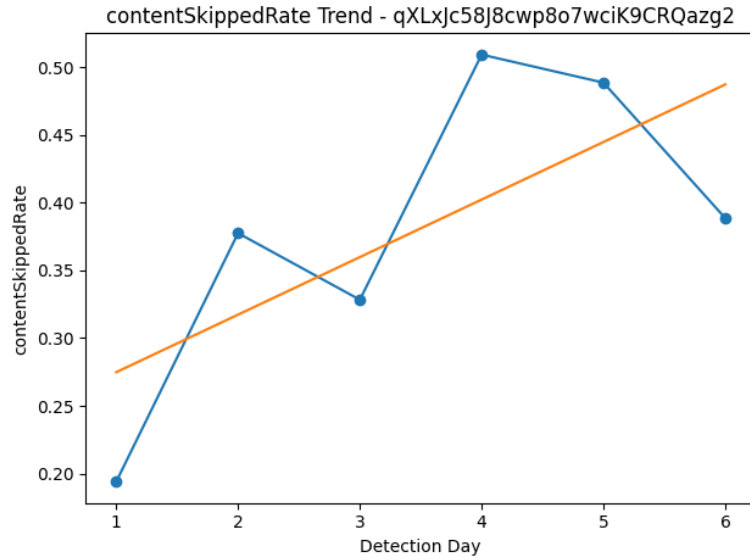


Figure 6.12: Content Skipped rate trend for participant 2

Summary of Trajectory Patterns. The comparison between these two profiles shows the heterogeneous nature of the behavioral change. While one participant exhibited signs of reduced impulsivity despite stable doom frequency, the other showed a consistent intensification of scrolling dynamics.

These different types of trajectories reinforce the interpretation derived from the previous analyses: awareness feedback did not completely produce a uniform behavioral shift among participants. Instead, behavioral adaptation appears to be highly individualized.

In general, this trajectory analysis illustrates how doomscrolling behaviors can evolve in complex and different ways over time.

6.2.3 Evaluation of LLM-Based Thresholding (RQ3)

To address RQ3, we evaluated whether the thresholds generated by the Large Language Model could capture behavioral patterns associated with doomscrolling sessions. The detection logic implemented in the system is based on interpretable threshold values applied to behavioral metrics. Each threshold defines a boundary condition that indicates whether a specific behavioral signal associated with doomscrolling is present within a session.

The thresholds used in the evaluation were extracted from the values dynamically generated by the LLM during the detection phase of the study and stored in the Firestore database. Since thresholds could be updated over time, the values used in the analysis correspond to the mean threshold values computed for each participant

during the study. This design choice reflects the assumption that doomscrolling manifests itself relative to an individual’s baseline behavior rather than as an absolute universal pattern. This aggregation provides a stable estimate of the personalized values inferred by the LLM.

Given a session, we define a set of behavioral conditions corresponding to the six metrics used by the detection system:

- minimum session duration
- minimum impulsive scroll rate
- minimum content skipped rate
- maximum interaction rate
- maximum depth of engagement
- maximum average watch time

Each condition evaluates whether the session metric satisfies the corresponding threshold constraint. Formally, for a session s and a threshold condition i , we define the activation function:

$$I_i(s) = \begin{cases} 1 & \text{if the threshold condition } i \text{ is satisfied} \\ 0 & \text{otherwise} \end{cases}$$

The total number of behavioral signals activated in a session is computed as follows:

$$Signals(s) = \sum_{i=1}^6 I_i(s)$$

where $Signals(s)$ represents the number of threshold conditions satisfied by the session. This value provides an interpretable indicator of how strongly a session exhibits the behavioral characteristics associated with doomscrolling.

The analysis was conducted on 34 sessions explicitly labeled as doomscrolling by the participants during the warm-up phase of the study. After matching sessions with the available personalized thresholds, 31 sessions were included in the evaluation.

Results indicate that doomscrolling sessions activate multiple behavioral signals simultaneously. In particular, the mean number of activated signals across sessions was:

$$\overline{Signals}_{doom} = 2.35$$

while the median number of activated signals was equal to 3. This suggests that doomscrolling sessions are typically characterized by the simultaneous presence of two to three behavioral indicators rather than by a single dominant metric.

To better understand which behavioral signals contributed most frequently to threshold activation, we computed the activation probability of each condition in all analyzed sessions. For a threshold condition i , the activation probability is defined as:

$$P(I_i = 1) = \frac{\text{number of sessions where condition } i \text{ is satisfied}}{\text{total number of evaluated sessions}}$$

The results show that some behavioral signals were consistently more present than others. In particular, low interaction rate and low average watch time emerged as the most frequently activated conditions, with activation probabilities of approximately 61% and 58% respectively. Similarly, high impulsive scrolling behavior was observed in more than half of the sessions analyzed.

In contrast, other signals, such as content skipped rate, were activated less frequently.

To further examine the capability of the generated thresholds, we compared the number of activated signals in sessions labeled as doomscrolling and sessions labeled as non-doomscrolling by participants. For each session, the same signal activation function defined above was applied.

Let $Signals_{doom}$ and $Signals_{nodoom}$ denote the number of activated signals for doom and non-doom sessions respectively. The mean values observed in the dataset were:

$$\overline{Signals}_{doom} = 2.35$$

$$\overline{Signals}_{nodoom} = 2.64$$

Although non-doom sessions exhibited a slightly higher average number of activated signals, this result highlights an important characteristic of the detection framework. The thresholds generated by the LLM do not act as strict classification boundaries but rather as indicators of behavioral tendencies associated with doomscrolling. Individual signals may occasionally appear in normal browsing sessions; however, the presence of multiple simultaneous signals provides a stronger indication of potentially problematic interaction patterns.

In general, the results provide preliminary evidence that the LLM-based threshold generation approach can support the identification of meaningful behavioral boundaries within the observed interaction data.

6.3 Subjective Feedback

To complement the behavioral data collected during the study, participants were asked to complete a post-study questionnaire, to capture their subjective perception of the system and its influence on their social media usage. The questionnaire included a set of Likert-scale questions together with a small number of open-ended questions, allowing participants to rate specific aspects of their experience and provide more detailed qualitative feedback. The questionnaire explored three main aspects: perceived awareness of personal scrolling behavior after the study, perceived usefulness and appropriateness of notifications, and overall utility of the application.

In general, the responses indicate that the application contributed positively to users' awareness of their social media habits.

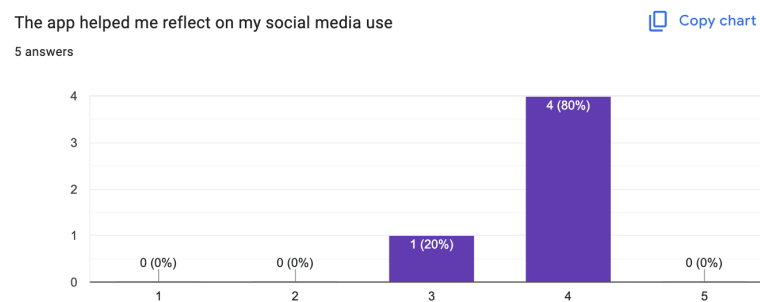


Figure 6.13: Participant responses on whether the application encouraged reflection on their social media usage.

Participants generally reported greater awareness of the time spent on social media platforms and an improved ability to recognize moments where they were engaged in doomscrolling. Most responses to statements related to awareness and reflection were concentrated in the upper-middle range of the five-point Likert scale. For the statement “The app helped me reflect on my social media usage”, 80% of the participants selected a rating of 4, while the remaining 20% selected 3. Similarly, for the statement “I can recognize better when I’m doomscrolling”, 60% of the participants selected 4 and 40% selected 3. These results suggest a moderately positive perceived impact of the system on the awareness of users about their scrolling behavior.

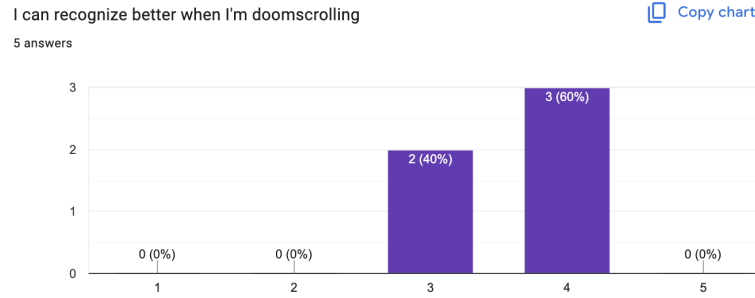


Figure 6.14: Participant responses on whether the application helped them better recognize when they were doomscrolling.

These findings are consistent with the qualitative feedback provided in the open responses. Several participants explicitly mentioned situations in which the system prompted reflection during ongoing usage. For example, a participant noted:

“The app was especially helpful during times when I lost track of time and scrolled endlessly”.

These contexts correspond to moments commonly associated with prolonged or passive consumption of social media content, suggesting that the awareness feedback may be especially relevant during these types of session.

Participants also highlighted situations in which the system helped them interrupt potentially unproductive behavior. One response indicated:

“The application was more useful while studying or working”

suggesting that the notifications sometimes acted as a reminder of ongoing distraction.

Regarding the perceived appropriateness of notifications, responses were homogeneous. The statement evaluating whether notifications appeared at appropriate moments received consistently low ratings, and participants generally selected 2 on the five-point scale. This suggests that even if the notifications were generally understandable and sometimes useful, their timing did not always match the expectations of the users. Several participants pointed out that notifications could occasionally appear too early in a session or during moments that did not clearly correspond to doomscrolling behavior. These observations are consistent with the findings discussed in RQ3, which highlighted the variability in behavior patterns between participants and the difficulty of defining universally accurate thresholds to detect doomscrolling. This feedback reinforces the importance of adaptive and personalized thresholding mechanisms when trying to identify such a complex and individual phenomenon.

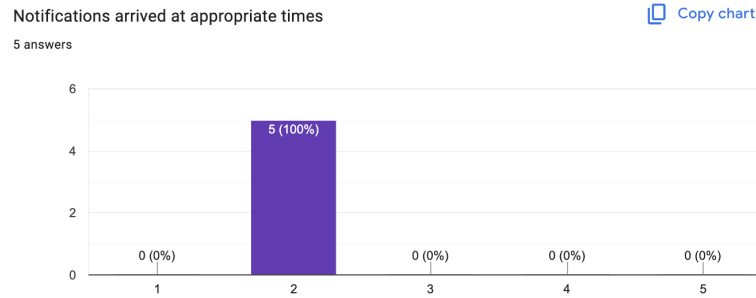


Figure 6.15: Distribution of participant responses regarding whether the notifications were delivered at appropriate moments during usage.

The open responses about suggestions for the application further clarify this aspect. Some participants suggested that the detection could be made more accurate or less sensitive. For example, a participant recommended:

“Avoid notifications after a few seconds of use, and initially incorporate more detailed user analysis to understand patterns and when doomscrolling is really happening.”

Another participant similarly noted that:

“I would probably adjust the detection slightly to make it even more accurate.”

These comments indicate that, while the underlying concept of the system was generally appreciated, users perceived room for improvement in the precision of the detection mechanism.

Despite these concerns about notification timing, participants generally reported that the suggestions provided by the application were understandable. Ratings for the clarity of suggestions ranged from 3 to 5, indicating that users were able to interpret the feedback without difficulty. This suggests that the design of the messages and suggestions successfully communicated the intended purpose of the intervention.

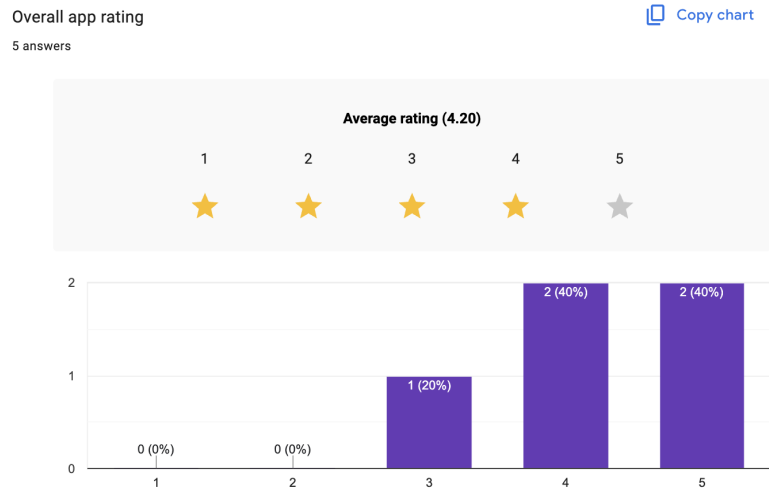


Figure 6.16: Overall user rating of the application on a 5-point scale, showing an average evaluation of 4.2.

Finally, participants were asked to provide an overall evaluation of the application and indicate whether they would consider using a similar tool in the future. The overall ratings ranged from 3 to 5 (with a mean of 4.2), suggesting a generally positive perception of the system. When asked whether they would continue using a similar application, four participants responded positively, while one participant indicated uncertainty. This result suggests that the concept of a system designed to increase awareness of scrolling behavior is acceptable to users, even though improvements in detection accuracy and notification timing could further enhance the user experience.

Chapter 7

Discussion

The analyses presented in the previous chapter provide empirical evidence that allows us to revisit the research questions that guided this study. Rather than focusing on individual metrics alone, the purpose of this chapter is to interpret the observed patterns in relation to the conceptualization of doomscrolling adopted in this work and to evaluate whether the proposed hypotheses are supported by the empirical findings. Because the study was conducted in an in-the-wild setting and relied on natural user sessions rather than controlled experimental manipulations, the results should be interpreted primarily as exploratory observations. Nevertheless, the collected data reveal consistent behavioral trends that contribute to a deeper understanding of how doomscrolling manifests in everyday social media use.

The first research question (**RQ1**) aimed to investigate whether doomscrolling can be characterized through recurring behavioral patterns observable during real-world social media sessions. The results presented in Chapter 6 provide several indications that sessions detected as doomscrolling indeed exhibit distinct interaction characteristics when compared to non-problematic sessions. In particular, feature activation analyses suggest that these sessions tend to have higher watch times and display higher levels of scrolling intensity. Taken together, these patterns align closely with the conceptual definition of doomscrolling adopted in this thesis, which frames the phenomenon as a state of prolonged and passive consumption of content with automatic interaction mechanisms.

Regarding the first hypothesis (**H1**):

Sessions labeled by users as doomscrolling are associated with longer session durations compared to non-doomscrolling sessions.

The findings indicate that sessions labeled by users as doomscrolling are generally associated with longer durations compared to other sessions. This result is consistent with the theoretical understanding of doomscrolling as a behavior in which users remain engaged in content consumption beyond their initial intention or interest.

Rather than representing deliberate exploration, the duration of extended sessions can reflect a form of attention capture, in which the continuous presentation of algorithmically selected content encourages prolonged engagement [12]. In this context, the duration of a session becomes a useful instrument for identifying potential states of reduced engagement, where users continue scrolling without a clear stopping point.

The second hypothesis (**H2**):

Doomscrolling sessions exhibit higher scrolling intensity and more impulsive scrolling patterns than non-doomscrolling sessions.

proposed that doomscrolling sessions would exhibit higher scrolling intensity and more impulsive scrolling patterns. The behavioral metrics analyzed in the previous chapter provide evidence supporting this expectation. Sessions considered problematic were often characterized by faster interaction rhythms and a higher rate of consecutive scroll actions. Such patterns suggest a mode of interaction in which users rapidly move from one content to another without deeply engaging with them. This behavior is consistent with the notion of automatic or habitual interaction regarding the design of the attention capture platform.

Regarding the third hypothesis (**H3**):

Doomscrolling sessions are characterized by lower levels of active interaction compared to non-doomscrolling sessions.

which predicted that doomscrolling sessions would involve lower levels of active interaction, the analysis did not clearly show that sessions detected as doomscrolling typically contain fewer meaningful interactions, such as likes, comments, or other engagement actions, relative to the number of items viewed.

Taken together, these observations offer preliminary evidence consistent with the first research question: doomscrolling appears to manifest through recurring behavioral patterns that can be observed in interaction data collected during real-world usage sessions. Importantly, none of the individual metrics alone is sufficient to fully capture the phenomenon. Rather, it is the combination of multiple behavioral indicators that allows for a more reliable characterization of the phenomenon. This observation supports the design choice adopted in this work, where doomscrolling detection is based on a set of complementary signals rather than on a single metric.

The second research question (**RQ2**) investigated whether awareness-oriented feedback could influence users' reflection on their social media usage during the longitudinal study. This question explored whether the introduction of reflective notifications could be associated with measurable behavioral changes over time. The analyses presented in Chapter 6, including the evolution of the daily doom rate,

the comparison between early and late detection phases, and the inspection of individual behavioral trajectories, revealed heterogeneous patterns among participants. In general, the results did not show a consistent reduction in the frequency of doomscrolling during the detection phase, suggesting that the evolution of behavior during the study was highly individualized.

Regarding hypothesis (**H4**):

Repeated exposure to awareness-oriented feedback increases users' reflection on their social media usage over the course of the longitudinal study.

the empirical results provide only partial support. At the aggregate level, the daily doom rate did not show a clear decreasing trend during the detection phase, and several behavioral indicators did not evolve consistently toward a reduced intensity. In particular, some metrics such as scroll velocity and average watch time slightly decreased, and other indicators such as impulsiveScrollRate and contentSkippedRate increased, suggesting that users sometimes adopted a more fragmented consumption pattern rather than a more reflective one.

However, the longitudinal trajectory analysis highlights that behavioral adaptation can occur in subtle and highly individual ways. For example, a participant showed a gradual reduction in impulsive scrolling despite a stable doom session frequency, suggesting that certain aspects of interaction dynamics may have evolved during the detection phase. At the same time, other participants exhibited opposite trajectories, reinforcing the interpretation that awareness-oriented feedback does not affect all users uniformly.

Several factors may explain the absence of a clear group-level effect. First, the limited number of participants reduces the statistical power of the analysis and makes aggregate trends highly sensitive to variability. Second, the relatively short duration of the longitudinal observation may not have been sufficient to capture more gradual behavioral adaptations. Changes in digital consumption habits often require longer exposure to reflective interventions before becoming observable in interaction data [1, 25].

Together, these findings suggest that the second research question can be addressed at an exploratory level, providing preliminary insights. Although the results do not provide strong empirical evidence for a systematic behavioral reduction in doomscrolling, they indicate that awareness-oriented feedback may influence certain interaction dynamics in ways that are highly dependent on individual habits.

The third research question (**RQ3**) explored whether a Large Language Model could support the definition of threshold values for detecting doomscrolling behavior based on behavioral metrics. The purpose of this analysis was to understand whether the thresholds generated through an LLM could produce interpretable values that reflect meaningful behavioral patterns observed during the study.

Regarding the first hypothesis (**H5**):

Threshold values generated with the support of a Large Language Model are coherent with the distribution of observed behavioral metrics.

the analysis suggests that the generated thresholds were generally consistent with the empirical distribution of the behavioral indicators collected during the study. When applied to sessions labeled as doomscrolling, multiple threshold conditions were frequently activated simultaneously, with an average of approximately 2.35 signals per session and a median of three signals. This indicates that doomscrolling sessions tend to exhibit several behavioral characteristics at the same time rather than being defined by a single dominant metric.

The second hypothesis (**H6**):

AI-supported threshold definitions exhibit meaningful alignment with users' session-level doomscrolling labels.

examined whether threshold signals were consistent with subjective labels provided by participants. Rather than functioning as a strict classification mechanism, the detection framework was designed to use behavioral thresholds as indicators of engagement intensity. Individual signals therefore do not determine the classification of a session, but instead contribute to a pattern of behavior evidence. Within this framework, simultaneous activation of multiple signals represents a stronger indication of potential doomscrolling behavior, while the presence of isolated signals may simply reflect normal variability in user interaction.

Finally, the third hypothesis (**H7**):

The use of an AI-based approach facilitates the adaptability of threshold-based doomscrolling detection in an exploratory setting.

is supported by the nature of the threshold generation process. During the study, the LLM periodically updated the threshold values based on newly collected session data, allowing the system to adapt to evolving behavioral patterns. This adaptive mechanism represents a key advantage compared to traditional manually defined thresholds, which typically remain static and require extensive manual tuning.

In general, these findings provide preliminary evidence that the LLM-based threshold generation approach can help identify behavioral boundaries within the interaction data and support the detection framework, although further validation on larger datasets would be required.

In general, the results discussed in this chapter provide a coherent view of how doomscrolling behavior can be analyzed through behavioral interaction data. Although the analyses do not point to a single pattern of doomscrolling, they

suggest that the phenomenon can be partially characterized through a combination of observable behavioral signals, personalized detection thresholds, and awareness-oriented feedback mechanisms. Taken together, the findings highlight the potential and the complexity of detecting problematic engagement patterns in everyday social media usage.

First, the results suggest that doomscrolling sessions tend to exhibit recognizable behavioral characteristics, consistent with the conceptual definition of the phenomenon adopted in this work.

Second, the analysis of the longitudinal phase indicates that awareness-oriented feedback does not necessarily produce a uniform behavioral change among users. Instead, the observed trajectories suggest that reflective interventions can influence interaction dynamics in more subtle and individualized ways.

Finally, the exploratory integration of Large Language Models for threshold generation demonstrates the potential of AI-supported approaches in the design of behavioral detection systems.

Although the results should be interpreted within the exploratory scope and the limited scale of the study, they still contribute to the broader discussion on how interaction metrics and AI-based techniques can be combined to better understand problematic engagement patterns in modern social media environments.

Chapter 8

Limitations and Future Work

Despite the results discussed in the previous chapters, it is important to recognize several limitations that characterize the present study. These limitations provide the correct context for interpreting the results found and finding potential directions for future research.

8.1 Limitations

An important first limitation refers to the technological constraints related to the operating system on which the system was developed. The proposed application relies on the Android Accessibility Service to capture interaction events such as scrolling gestures, interaction frequency, and session dynamics during TikTok usage. This approach allows for the collection of detailed behavioral data directly from the user's device while maintaining a lightweight monitoring process. However, similar capabilities are currently much more restricted within the iOS ecosystem. Due to Apple's strict privacy and security policies, third-party applications have very limited access to low-level interaction data and gesture detection. As a consequence, extending the proposed system to iOS devices would require alternative technical approaches and would likely provide a significantly reduced set of observable behavioral signals. This platform dependency therefore represents a limitation in terms of cross-platform applicability.

The second limitation relates to the intrinsic complexity of detecting doomscrolling behavior with high precision. Doomscrolling is not a strictly defined behavioral state, but rather a subjective and highly individualized pattern of interaction with digital content. Different users may experience doomscrolling under different conditions, depending on factors such as personal habits, emotional

state, type of content consumed, or time of day. Although the system attempts to address this challenge through personalized thresholds and the integration of user-provided labels, the detection process inevitably remains an approximation of a more complex phenomenon.

Additionally, the size of the participant sample represents another factor that should be considered when interpreting the results. Although the longitudinal study in-the-wild allowed the collection of realistic behavioral data over time, the number of participants remains limited. This restricts the possibility of performing more advanced statistical analyses and limits the generalization of the findings to larger samples. Nevertheless, the exploratory nature of the study provides valuable initial insights that can guide future investigations on larger and more diverse samples.

Another potential limitation relates to the possibility of a Hawthorne effect. Since participants were aware that their interaction behavior was being monitored by the application, their browsing habits may have been partially influenced by this awareness. Even though the application was designed to operate in the background, the knowledge of being observed could have encouraged participants to adopt more deliberate interaction patterns, potentially altering their natural usage behavior during the study period.

A further limitation concerns the fact that the study focused exclusively on a single social media platform, namely TikTok. While TikTok represents one of the most prominent examples of algorithmically curated short-video feeds and infinite scrolling interfaces, doomscrolling behavior can also occur on other platforms such as Instagram and YouTube Shorts. As a consequence, the behavioral patterns identified in this work can be influenced by the specific interaction design and recommendation mechanisms of TikTok, limiting the generalization of the results to other platforms.

8.2 Future Work

The limitations identified above naturally suggest several directions for future work. An important first research direction concerns the extension of the system to a broader set of platforms and applications. Future studies could explore alternative techniques for interaction monitoring that do not rely exclusively on Android accessibility services, potentially enabling compatibility with other mobile operating systems. Moreover, the approach could be generalized beyond TikTok, applying a similar detection mechanism to other social media platforms characterized by infinite scrolling.

Another promising avenue concerns the refinement of doomscrolling detection models. While the current system relies on rule-based thresholds supported by Large

Language Models to personalize detection criteria, future work could investigate the integration of more accurate and precise machine learning models trained exclusively on larger and more specific datasets of behavioral interactions. Such models could potentially capture more complex temporal patterns and interactions between behavioral signals, leading to more robust detection mechanisms.

Finally, future research could also refer to the personalization of feedback and intervention strategies. In the present study, awareness notifications were designed to encourage reflection on ongoing behavior. As shown in the results discussed in chapter 7, where users suggested a more accurate timing for notifications, future systems could explore adaptive interventions that dynamically adjust their timing and content based on the user's interaction patterns, personal preferences, or historical behavioral trends. This could enable more effective digital well-being interventions.

Chapter 9

Conclusions

This thesis addresses the growing concern about problematic engagement patterns on modern social media platforms, with a particular focus on the phenomenon commonly referred to as doomscrolling. Within the broader context of the attention economy, platforms such as TikTok are designed to maximize user engagement through mechanisms such as infinite scrolling and algorithmically personalized content feeds. Although these design strategies contribute to highly engaging user experiences, they can also facilitate prolonged and often unintentional content consumption.

The main objective of this research was to investigate whether behavioral interaction data collected directly from users' devices could be used to identify patterns associated with doomscrolling during everyday social media usage. Rather than focusing exclusively on restricting user behavior, the proposed approach explored the potential of awareness-oriented notifications aimed at encouraging reflection and self-regulation.

To support this investigation, a mobile application was developed for Android devices capable of monitoring interaction metrics during TikTok sessions. The system captured behavioral signals such as scroll dynamics and interaction frequency, which were subsequently used to identify potential doomscrolling episodes. A key component of the proposed framework was the integration of a Large Language Model (GPT-4.1) to support the generation and refinement of personalized behavioral thresholds.

The system was evaluated through a longitudinal study in-the-wild, allowing behavioral data to be collected within realistic everyday conditions.

The results highlighted that doomscrolling episodes can be partially characterized through observable interaction patterns. In particular, the feature activation analysis of behavioral signals showed that these patterns tended to appear more frequently during sessions labeled as doomscrolling, although individual signals were not exclusively associated with problematic sessions.

Furthermore, the longitudinal analysis of the detection phase suggested that awareness-oriented feedback did not lead to a uniform behavioral change between participants. Instead, the observed trajectories indicated heterogeneous responses. These findings suggest that reflective interventions can influence user behavior in subtle and individual ways rather than producing immediate or uniform behavioral changes.

Beyond the specific findings, this work also illustrates the broader potential of combining behavioral data with AI analysis techniques to support digital well-being research. Although there are some limitations bound to RQ3, as shown in chapter 8, the integration of Large Language Models opens new possibilities for building adaptive and personalized behavioral monitoring systems capable of evolving with the user's habits over time.

In conclusion, the proposed system contributes to a growing body of research that seeks to better understand the relationship between interface design, user behavior, and digital well-being. Future work will be essential to address the limitations discussed in this study to refine these approaches, extend them to larger populations, and explore new forms of adaptive interventions capable of supporting healthier and more intentional interactions with digital platforms.

Appendix A

LLM Prompts for Threshold Computation

A.1 File: thresholdDefinerPrompt.txt

```
1 You are an expert behavioral analyst specializing in personalized
   digital wellbeing.
2
3 Your task is to analyze a dataset of user sessions (using Tiktok)
   containing:
4 - objective metrics (scroll velocity, session duration, impulsivity rate,
   avg watch time per tiktok, interactions, etc.)
5 - subjective labels provided by the user ("doomscrolling" vs "not
   doomscrolling")
6 - subjective estimates of perceived duration
7
8 Using ONLY patterns present in this specific dataset, derive
   personalized decision thresholds that best separate DOOM vs NON-DOOM
   sessions ("doomscrolling" vs "not doomscrolling").
9
10 Do NOT use generic assumptions about doomscrolling.
11 Do NOT invent metrics that are not provided.
12 Base your reasoning EXCLUSIVELY on patterns found in the user's data.
13
14 -----
15 ### IN-CONTEXT LEARNING EXAMPLE (follow this logic)
16 Example reasoning:
17 1. Identify the metric that best separates the two classes (usually
   session duration).
```

```
18 2. Determine secondary metrics that consistently increase during DOOM
    sessions.
19 3. Choose stable numeric thresholds that maximize separation.
20 4. Return strict JSON with only the thresholds, no commentary.
21
22 Example output format:
23 {
24     "minSessionDuration": 12,
25     "minImpulsiveScrollRate": 0.4,
26     "minContentSkippedRate": 0.3,
27     "maxDepthOfEngagement": 0.1,
28     "maxInteractionRate": 0.5,
29     "minAvgScrollVelocity": 2.3,
30     "maxAvgWatchTime": 1
31 }
32
33 IMPORTANT: Produce EVERY threshold in the output format example
34
35 -----
36 ### DATASET TO ANALYZE
37 {{DATASET}}
38
39 -----
40
41 ### YOUR TASK
42 Analyze the dataset and output:
43 1. A set of personalized numeric thresholds for doomscrolling detection.
44
45 Return ONLY valid JSON in this exact structure:
46
47 {
48     "minSessionDuration": 12,
49     "minImpulsiveScrollRate": 0.4,
50     "minContentSkippedRate": 0.3,
51     "maxDepthOfEngagement": 0.1,
52     "maxInteractionRate": 0.5,
53     "minAvgScrollVelocity": 2.3,
54     "maxAvgWatchTime": 1
55 }
56
57 IMPORTANT: Produce EVERY threshold in the output format example
```

A.2 File: thresholdUpdaterPrompt.txt

```
1 You are an expert behavioral analyst.
2
3 You will receive:
4 1. The current threshold model.
5 2. A batch of new sessions (some labeled, some unlabeled).
6
7 Your task is to update the thresholds **quite conservatively**:
8 - Labeled sessions can outweigh unlabeled ones.
9 - Thresholds should not drastically change unless multiple sessions
   contradict the current model.
10 - Unlabeled sessions may quite shift thresholds if they show a
    consistent trend.
11 - The primary threshold (usually session duration) must remain the most
    important gatekeeper.
12 - Do NOT use any weights or scoring systems. Only output updated
    thresholds.
13
14 -----
15 ### IN-CONTEXT LEARNING EXAMPLE (follow this pattern)
16
17 Example behavior:
18 - If several newly labeled DOOM sessions exceed a metric, slightly lower
    that metric's threshold.
19 - If several newly labeled NON-DOOM sessions stay below a metric,
    slightly increase threshold.
20 - Ignore isolated outliers.
21
22 Example output:
23 {
24     "minSessionDuration": 12,
25     "minImpulsiveScrollRate": 0.4,
26     "minContentSkippedRate": 0.3,
27     "maxDepthOfEngagement": 0.1,
28     "maxInteractionRate": 0.5,
29     "minAvgScrollVelocity": 2.3,
30     "maxAvgWatchTime": 1
31 }
32 IMPORTAT: Return EVERY threshold in the example, even without changes
    from the the previous ones
33
34 -----
35 ### CURRENT THRESHOLDS
```

```
36 {{CURRENT_THRESHOLDS}}
37
38 ### NEW SESSIONS
39 {{NEW_SESSIONS}}
40
41 -----
42
43 Return ONLY valid JSON in this exact structure:
44 {
45     "minSessionDuration": 12,
46     "minImpulsiveScrollRate": 0.4,
47     "minContentSkippedRate": 0.3,
48     "maxDepthOfEngagement": 0.1,
49     "maxInteractionRate": 0.5,
50     "minAvgScrollVelocity": 2.3,
51     "maxAvgWatchTime": 1
52 }
53 IMPORTANT: Return EVERY threshold in the example, even without changes
    from the the previous ones
```

Appendix B

Participant Questionnaires

PRE-QUESTIONARIO (SCREENING)

Questo questionario ha lo scopo di valutare l'idoneità alla partecipazione a uno studio sull'uso dei social media, in particolare TikTok.

Se interessati a scaricare l'applicazione si prega di inserire la stessa Email utilizzata in PlayStore.

* **Indica una domanda obbligatoria**

1. Email *

Requisiti Tecnici

2. Sistema Operativo Principale *

Contrassegna solo un ovale.

Android

iOS

Altro:
—

3. Versione di Android *

Contrassegna solo un ovale.

Android 11 o superiore

Inferiore

Non so

4. Età *

Contrassegna solo un ovale.

18-24

25-30

31-35

>35

Utilizzo Social

5. Quali social utilizzi regolarmente? *

Seleziona tutte le voci applicabili.

TikTok

Instagram Reels

YouTube Shorts

Altro

Nessuno

6. Quanto tempo passi mediamente al giorno su questi social? *

Contrassegna solo un ovale.

< 30 min

30-60 min

1-2 ore

> 2 ore

Autopercezione

7. Mi capita di scrollare più a lungo di quanto vorrei *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

8. Uso i social in modo automatico, senza pensarci *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

9. Perdo la percezione del tempo mentre scrollo *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

10. Dopo aver usato i social, a volte mi sento insoddisfatto *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

Dati per la beta

La partecipazione allo studio è **volontaria**.

È possibile **ritirarsi in qualsiasi momento**, senza fornire motivazioni, comunicandolo via email all'indirizzo:

s331356@studenti.polito.it

11. Accetti di partecipare allo studio per circa 2 settimane? *

Contrassegna solo un ovale.

Sì

No

12. Se sei curioso di scaricare l'applicazione, inserisci la mail che utilizzi per PlayStore *

Questa email verrà utilizzata esclusivamente per:

- abilitarne l'accesso alla versione beta dell'app tramite Google Play
- ricontattarti con il link per il download dell'app

Non verrà usata per altri scopi né condivisa con terze parti.

Questi contenuti non sono creati né avallati da Google.

Google Moduli

POST-QUESTIONARIO (FINE STUDIO)

Questo questionario serve a valutare la tua esperienza con l'app durante lo studio.

* Indica una domanda obbligatoria

1. Email *

2. Sono più consapevole del tempo che passo sui social *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

3. Riesco a riconoscere meglio quando sto doomscrollando *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

4. Presto più attenzione al modo in cui utilizzo i social *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

5. L'app mi ha aiutato a riflettere sul mio utilizzo dei social *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

6. Le notifiche arrivavano in momenti appropriati *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

7. I suggerimenti erano comprensibili *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

8. Ho trovato l'app invadente *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

9. Mi sono sentito giudicato dall'app *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

10. Le notifiche interrompevano inutilmente l'esperienza *

Contrassegna solo un ovale.

1 2 3 4 5

Per Moltissimo

Valutazione complessiva

11. Valutazione complessiva dell'app *

1 2 3 4 5

12. Continueresti a usare un'app simile? *

Contrassegna solo un ovale.

- Si
- No
- Forse

Domande aperte - Suggerimenti

13. In quali situazioni l'app ti è stata più utile?

14. Cosa miglioreresti o cambieresti?

15. Note aggiuntive (opzionale)

Questi contenuti non sono creati né avallati da Google.

Google Moduli

Bibliography

- [1] Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. «Habits Make Smartphone Use More Pervasive». In: *Personal and Ubiquitous Computing* (2012). DOI: 10.1007/s00779-011-0412-2. URL: <https://dl.acm.org/doi/abs/10.1007/s00779-011-0412-2> (cit. on pp. 1, 2, 10, 15, 19, 85).
- [2] Candice L. Odgers and Michaelaeline R. Jensen. «Annual Research Review: Adolescent Mental Health in the Digital Age: Facts, Fears, and Future Directions». In: *Journal of Child Psychology and Psychiatry* 61.3 (2020), pp. 336–348. DOI: 10.1111/jcpp.13190. URL: <https://pubmed.ncbi.nlm.nih.gov/31951670/> (cit. on pp. 1, 17).
- [3] Amy Orben and Andrew K. Przybylski. «The Association Between Adolescent Well-Being and Digital Technology Use». In: *Nature Human Behaviour* 3 (2019), pp. 173–182. DOI: 10.1038/s41562-018-0506-1. URL: <https://www.nature.com/articles/s41562-018-0506-1> (cit. on p. 1).
- [4] Jon D. Elhai, Robert D. Dvorak, Jason C. Levine, and Brian J. Hall. «Problematic Smartphone Use: A Conceptual Overview and Systematic Review of Relations with Anxiety and Depression Psychopathology». In: *Journal of Affective Disorders* 207 (2017), pp. 251–259. DOI: 10.1016/j.jad.2016.08.030. URL: <https://pubmed.ncbi.nlm.nih.gov/27736736/> (cit. on pp. 1, 3, 22).
- [5] Mariek M. P. Vanden Abeele. «Digital Wellbeing as a Dynamic Construct». In: *Communication Theory* 31.4 (2021), pp. 932–955. DOI: 10.1093/ct/qtaa024. URL: <https://academic.oup.com/ct/article/31/4/932/5927565> (cit. on pp. 1, 10, 12, 16, 18, 20).
- [6] Ulrik Lyngs, Kai Lukoff, Petr Slovak, Reuben Binns, Adam Slack, Michael Inzlicht, Max Van Kleek, and Nigel Shadbolt. «Self-Control in Cyberspace: Applying Dual Systems Theory to a Review of Digital Self-Control Tools». In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, 2019. DOI: 10.1145/3290605.3300361. URL: <https://dl.acm.org/doi/10.1145/3290605.3300361> (cit. on pp. 1, 3, 4, 10, 16, 18, 21).

- [7] Alberto M. Roffarello and Luigi De Russis. «Achieving Digital Wellbeing Through Digital Self-Control Tools: A Systematic Review and Meta-analysis». In: *ACM Transactions on Computer-Human Interaction* (2023). DOI: 10.1145/3571810. URL: <https://dl.acm.org/doi/full/10.1145/3571810> (cit. on pp. 1, 3, 4, 11, 18–20).
- [8] Herbert A. Simon. «Designing Organizations for an Information-Rich World». In: *Computers, Communications, and the Public Interest*. Ed. by Martin Greenberger. Baltimore, MD, USA: The Johns Hopkins Press, 1971, pp. 37–72. URL: <https://atelierdesfuturs.org/wp-content/uploads/2025/07/1971-simon.pdf> (cit. on pp. 1, 8).
- [9] Pablo González de la Torre, Marta Pérez-Verdugo, and Xabier E. Barandiaran. «Attention is all they need: Cognitive science and the (techno)political economy of attention». In: *arXiv* (2024). URL: <https://arxiv.org/abs/2405.06478> (cit. on p. 2).
- [10] Wikipedia contributors. *Attention economy — Wikipedia, The Free Encyclopedia*. [Online; accessed 27-February-2026]. 2026. URL: https://en.wikipedia.org/w/index.php?title=Attention_economy&oldid=1340718634 (cit. on p. 2).
- [11] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. «The Dark (Patterns) Side of UX Design». In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 2018. DOI: 10.1145/3173574.3174108. URL: <https://dl.acm.org/doi/10.1145/3173574.3174108> (cit. on pp. 2, 8, 11, 12, 15).
- [12] Kai Lukoff Alberto Monge Roffarello Luigi De Russis. «Defining and Identifying Attention Capture Deceptive Designs in Digital Interfaces». In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, 2023. DOI: 10.1145/3544548.3580729. URL: <https://dl.acm.org/doi/abs/10.1145/3544548.3580729> (cit. on pp. 2, 11–15, 17–20, 23, 25, 26, 84).
- [13] Christian Montag, Hong Yang, and Jon D. Elhai. «On the Psychology of TikTok Use: A First Glimpse From Empirical Findings». In: *Frontiers in Public Health* 9 (2021), p. 641673. DOI: 10.3389/fpubh.2021.641673. URL: <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2021.641673/full> (cit. on pp. 2, 3, 16).
- [14] Meng Chao, Di Xue, Tianyi Liu, Huan Yang, and Brian J. Hall. «TikTok Use and Psychosocial Factors Among Adolescents: Associations with Depression, Anxiety, and Stress». In: *Journal of Affective Disorders* (2023). Check the exact page range / final DOI suffix in the published version. DOI: 10.1016/

- j.jad.2023.03.OXX. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0165178123001889> (cit. on p. 2).
- [15] DataReportal. *Digital 2025: Global Overview Report*. Accessed: 2026-01-27. 2025. URL: <https://datareportal.com/reports/digital-2025-global-overview-report> (cit. on pp. 2, 24, 25).
- [16] L. Jain et al. «Exploring Problematic TikTok Use and Mental Health Issues: A Systematic Review». In: *Healthcare (Basel) / PMC version* (2025). Please update with full author list, journal name, volume/pages from the final published record. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11924099/> (cit. on p. 2).
- [17] Fanni Bányai, Ágnes Zsila, Orsolya Király, Aniko Maraz, Zsuzsanna Elekes, Mark D. Griffiths, Cecilie Schou Andreassen, and Zsolt Demetrovics. «Problematic Social Media Use: Results from a Large-Scale Nationally Representative Adolescent Sample». In: *PLOS ONE* 12.1 (2017), e0169839. DOI: 10.1371/journal.pone.0169839. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0169839> (cit. on p. 3).
- [18] R.X. Schwartz Alberto Monge Roffarello Luigi De Russis and Panagiotis Apostolellis. «Reducing Risk in Digital Self-Control Tools». In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, 2021. DOI: 10.1145/3411763.3451843. URL: <https://dl.acm.org/doi/10.1145/3411763.3451843> (cit. on pp. 3, 11).
- [19] Luigi De Russis Alberto Monge Roffarello. «Towards Understanding the Dark Patterns That Steal Our Attention». In: (2022). Preprint / Technical report. URL: <https://iris.polito.it/retrieve/handle/11583/2955697/579805/dpimpact.pdf> (cit. on pp. 3, 15, 17).
- [20] Thomas H. Davenport and John C. Beck. *The Attention Economy: Understanding the New Currency of Business*. Harvard Business School Press, 2001. ISBN: 9781578514416 (cit. on pp. 8, 15).
- [21] Tim Wu. *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*. Alfred A. Knopf, 2016. ISBN: 9780307907108 (cit. on pp. 8, 12, 15).
- [22] Shoshana Zuboff. *The Age of Surveillance Capitalism*. PublicAffairs, 2019. ISBN: 9781610395694 (cit. on pp. 8, 12).
- [23] Adrian F. Ward, Kristen Duke, Ayelet Gneezy, and Maarten W. Bos. «Brain Drain: The Mere Presence of One’s Own Smartphone Reduces Available Cognitive Capacity». In: *Journal of the Association for Consumer Research* 2.2 (2017), pp. 140–154. DOI: 10.1086/691462 (cit. on pp. 8, 9, 16, 19).

- [24] Moritz Büchi, Noemi Festic, and Michael Latzer. «Digital Overuse and Subjective Well-Being in a Digitized Society». In: *Social Media + Society* 5.4 (2019). DOI: 10.1177/2056305119886031 (cit. on pp. 9, 10, 12, 15, 17–20).
- [25] Leonard Reinecke and Wilhelm Hofmann. «Slacking Off or Winding Down? Self-Control and Media Use». In: *Communication Research* 45.8 (2018), pp. 1269–1295. DOI: 10.1177/0093650216676386 (cit. on pp. 9, 10, 15, 16, 18, 19, 21, 85).
- [26] Alexander J. A. M. van Deursen and Ellen J. Helsper. «The Third-Level Digital Divide: Who Benefits Most from Being Online?» In: *Communication and Information Technologies Annual* (2015). DOI: 10.1108/S2050-206020150000010002 (cit. on p. 9).
- [27] Andrew K. Przybylski, Kou Murayama, Cody R. DeHaan, and Valerie Gladwell. «Motivational, Emotional, and Behavioral Correlates of Fear of Missing Out». In: *Computers in Human Behavior* 29.4 (2013), pp. 1841–1848. DOI: 10.1016/j.chb.2013.02.014. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0747563213000800> (cit. on p. 9).
- [28] Philippe Verduyn, David S. Lee, Jiyoung Park, Hannah Shablack, Ariana Orvell, Joseph Bayer, Oscar Ybarra, John Jonides, and Ethan Kross. «Passive Facebook Usage Undermines Affective Well-Being: Experimental and Longitudinal Evidence». In: *Journal of Experimental Psychology: General* 144.2 (2015), pp. 480–488. DOI: 10.1037/xge0000057. URL: <https://pubmed.ncbi.nlm.nih.gov/25706656/> (cit. on pp. 10, 16, 22).
- [29] Ethan Kross, Philippe Verduyn, Emre Demiralp, Jiyoung Park, David S. Lee, Natalie Lin, Hannah Shablack, John Jonides, and Oscar Ybarra. «Facebook Use Predicts Declines in Subjective Well-Being in Young Adults». In: *PLOS ONE* 8.8 (2013), e69841. DOI: 10.1371/journal.pone.0069841. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0069841> (cit. on pp. 10, 16).
- [30] Christian Schaffner, Robert Toth, and Leonard Reinecke. «Autoplay and Continuous Watching: How Interface Design Affects Media Consumption». In: *Journal of Broadcasting & Electronic Media* (2022). DOI: 10.1080/08838151.2022.2034538 (cit. on pp. 10, 16, 18, 22, 23).
- [31] Stefan Rixen, Oliver Wetter, and Klaus Opwis. «The Effects of Infinite Scroll on Attention and Content Consumption». In: *Human-Computer Interaction* (2020). DOI: 10.1080/07370024.2020.1734395 (cit. on pp. 10, 15, 18, 19, 25–28).

- [32] Ji Yang, Xue Fu, Xiaoqian Liao, and Yiran Li. «Association of Problematic Smartphone Use with Poor Sleep Quality, Depression, and Anxiety: A Systematic Review and Meta-analysis». In: *Journal of Affective Disorders* 284 (2020), pp. 256–267. DOI: 10.1016/j.jad.2019.12.016. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0165178119317408> (cit. on p. 10).
- [33] Jean M. Twenge, Thomas E. Joiner, Megan L. Rogers, and Gabrielle N. Martin. «Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010». In: *Clinical Psychological Science* 6.1 (2018), pp. 3–17. DOI: 10.1177/2167702617723376 (cit. on p. 10).
- [34] Roy F. Baumeister, Kathleen D. Vohs, and Dianne M. Tice. «The Strength Model of Self-Control». In: *Current Directions in Psychological Science* 16.6 (2007), pp. 351–355. DOI: 10.1111/j.1467-8721.2007.00534.x (cit. on p. 10).
- [35] Matthias Brand, Kimberly S. Young, and Christian Laier. «Prefrontal Control and Internet Addiction». In: *Neuroscience & Biobehavioral Reviews* 104 (2019), pp. 1–14. DOI: 10.1016/j.neubiorev.2019.06.023 (cit. on pp. 10, 15, 17, 19).
- [36] Christian Montag, Bernd Lachmann, Marc Herrlich, and Katharina Zweig. «Addictive Features of Social Media and Video Games». In: *Behavioral Sciences* 9.2 (2019), pp. 1–17. DOI: 10.3390/bs9020026 (cit. on p. 15).
- [37] Patrick Haggard. «Sense of Agency in the Human Brain». In: *Nature Reviews Neuroscience* 18 (2017), pp. 196–207. DOI: 10.1038/nrn.2017.14 (cit. on pp. 16, 17, 20–22).
- [38] James W. Moore. «What Is the Sense of Agency and Why Does It Matter?» In: *Frontiers in Psychology* 7 (2016), p. 1272. DOI: 10.3389/fpsyg.2016.01272 (cit. on pp. 16, 17, 21, 22).
- [39] Albert Bandura. «Social Cognitive Theory: An Agentic Perspective». In: *Annual Review of Psychology* 52 (2001), pp. 1–26. DOI: 10.1146/annurev.psych.52.1.1 (cit. on pp. 16, 17, 20, 21).
- [40] Alexander Ostrovsky and Nele Schwarz. «Short-Video Consumption and Habit Formation». In: *Computers in Human Behavior* 130 (2022), p. 107175. DOI: 10.1016/j.chb.2021.107175 (cit. on pp. 16, 25, 27).
- [41] Aparajita Bhandari and Sharanya Bimo. «Why’s Everyone on TikTok Now? The Algorithmized Self and the Future of Social Media». In: *Social Media + Society* 7.1 (2021). DOI: 10.1177/2056305121998645 (cit. on pp. 16–18, 23, 25).

- [42] Elisa Wegmann, Kai W. Müller, Ofir Turel, and Matthias Brand. «Dissociative Symptoms in Problematic Social Media Use». In: *Addictive Behaviors Reports* 11 (2020), p. 100277. DOI: 10.1016/j.abrep.2020.100277 (cit. on pp. 17–19).
- [43] Matthias Synofzik, Gottfried Vosgerau, and Albert Newen. «Beyond the Comparator Model: A Multifactorial Two-Step Account of Agency». In: *Consciousness and Cognition* 17.1 (2008), pp. 219–239. DOI: 10.1016/j.concog.2007.03.010. URL: <https://www.sciencedirect.com/science/article/pii/S1053810007000520> (cit. on pp. 21, 22).
- [44] Kai Lukoff, Ulrik Lyngs, Himanshu Zade, Morten Rasmussen, Michael Inzlicht, Max Van Kleek, and Nigel Shadbolt. «Addressing Digital Wellbeing by Designing for Ritualized and Intentional Use». In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 2018. DOI: 10.1145/3173574.3173689. URL: <https://dl.acm.org/doi/10.1145/3173574.3173689> (cit. on pp. 22, 23).
- [45] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, and Steven Jacobs. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. 6th. Pearson, 2016. ISBN: 9780134380384 (cit. on p. 22).
- [46] Matthew Didion, Q. Vera Liao, Michael Muller, and Thomas Erickson. «Measuring the Sense of Agency in Human-AI Interaction». In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, 2023. DOI: 10.1145/3544548.3580715. URL: <https://dl.acm.org/doi/10.1145/3544548.3580715> (cit. on p. 23).
- [47] Simon Kemp. *Digital 2023 Deep-Dive: Is social media really dying?* Tech. rep. Data Reportal, 2023. URL: <https://datareportal.com/reports/digital-2023-deep-dive-the-worlds-top-social-media-platforms?rq=time%20spent%20using%20social%20media> (cit. on p. 25).
- [48] David B. V. Kaye, Xiao Chen, and Jing Zeng. «The Co-Evolution of TikTok and Music». In: *International Journal of Communication* 14 (2020), pp. 4984–5006 (cit. on p. 25).
- [49] Yu Feng, Yang-Chen Lin, and Ya-Quin You. «Factors Affecting User Experience Preferences in Short Video Platforms - A Case Study of TikTok». In: *Social Media* (2023). DOI: 10.1109/GCCE59613.2023.10315272. URL: <https://ieeexplore.ieee.org/document/10315272> (cit. on pp. 26, 27).
- [50] Qian Jiang and Liangying Ma. «Swiping more, thinking less: Using TikTok hinders analytic thinking». In: *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 18.3 (June 2024), Article 1. DOI: 10.5817/CP2024-3-1. URL: <https://cyberpsychology.eu/article/view/33099> (cit. on p. 28).

- [51] Yifan Zhang and Wenjing Liu. «The Impact of TikTok Use on User Behavior and Cognitive Engagement». In: *Proceedings of the International Conference on Social Media and Society*. Francis Press, 2023, pp. 45–52. URL: <https://francis-press.com/uploads/papers/KRAJ60ieLedXZuB2rZGVVN4k0qr55NXxWLVeKHGX.pdf> (cit. on p. 28).
- [52] Jan Rixen, Luca Meinhardt, Michael Glöckler, Marius-Lukas Ziegenbein, Anna Schlothauer, Mark Colley, Enrico Rukzio, and Jan Gugenheimer. «The Loop and Reasons to Break It: Investigating Infinite Scrolling Behaviour in Social Media Applications and Reasons to Stop». In: *Proceedings of the ACM on Human-Computer Interaction* 7 (Sept. 2023), pp. 1–22. DOI: 10.1145/3604275 (cit. on p. 29).
- [53] OpenAI. «GPT-4 Technical Report». In: *arXiv preprint arXiv:2303.08774* (2023) (cit. on p. 62).